



**HAL**  
open science

# Développement et application de méthodologies d'évaluation des expositions atmosphériques chroniques aux dioxines et au cadmium dans le cadre d'études épidémiologiques

Thomas Coudon

► **To cite this version:**

Thomas Coudon. Développement et application de méthodologies d'évaluation des expositions atmosphériques chroniques aux dioxines et au cadmium dans le cadre d'études épidémiologiques. Santé. Université de Lyon, 2018. Français. NNT : 2018LYSE1093 . tel-02053064

**HAL Id: tel-02053064**

**<https://theses.hal.science/tel-02053064>**

Submitted on 1 Mar 2019

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

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



N° d'ordre NNT : 2018LYSE1093

**THÈSE de DOCTORAT DE L'UNIVERSITÉ DE LYON**  
opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**École Doctorale N° 205**  
EDISS, École Doctorale Interdisciplinaire Sciences-Santé

**Spécialité de doctorat : Expologie environnementale**

Soutenue publiquement le 01/06/2018, par :  
**Thomas COUDON**

---

**Développement et application de méthodologies d'évaluation  
des expositions atmosphériques chroniques aux dioxines et au  
cadmium dans le cadre d'études épidémiologiques**

---

Devant le jury composé de :

**Pr DEGUEN Séverine**

Professeur, HDR, enseignant chercheur, EHESP Rennes – Rapporteur

**Dr PASCAL Mathilde**

PhD, Chargée de projet scientifique, Santé Publique France Paris– Rapporteur

**Pr VINEIS Paolo**

MD, MRC/PHE Center for Environmental and Health School of Public Health, Imperial College London – Rapporteur

**Pr CHARBOTEL Barbara**

PU-PH2 - Médecin spécialiste en santé au travail, HCL (UMRESTTE), Univ. Lyon 1 – Examinatrice

**Dr SEVERI Gianluca**

PhD, HDR, Directeur de Recherche, Inserm - Centre de Recherche en Épidémiologie et Santé des Populations (CESP, UMR Inserm 1018) Université Paris-Saclay, Gustave Roussy – Examineur

**Pr SOULHAC Lionel**

MCF, HDR, co-directeur du Laboratoire de Mécanique des Fluides et d'Acoustique, Ecole Centrale de Lyon / INSA de Lyon / Université Claude Bernard Lyon 1 / CNRS – Examineur

**Pr FERVERS Béatrice**

MD, HDR, Pr. associé, Coordinatrice Département Cancer Environnement, Centre Léon Bérard, Inserm U1052, CRCL, Université de Lyon – Directrice de thèse

**Dr SALIZZONI Pietro**

MCF, HDR, enseignant chercheur, LMFA, Ecole centrale de Lyon - Co-directeur de thèse



## Résumé en Français

Un certain nombre d'études ont étudié le lien entre cancer du sein et exposition aux dioxines et au cadmium. Toutefois, les résultats de ces études ne sont pas concluants et présentent des limites méthodologiques.

L'objectif principal de cette thèse était de développer un indicateur spatial permettant d'estimer les expositions atmosphériques chroniques aux dioxines et au cadmium des femmes issues de la cohorte E3N en France, entre 1990 et 2008.

Nous avons créé une base de données de 2620 sources de dioxines et 2700 sources de cadmium et estimé et géolocalisé leurs émissions.

L'estimation des émissions à la source et leur localisation ont servi de base pour la construction de l'indicateur spatial. Nous avons identifié une sélection de paramètres permettant d'obtenir un accord « substantiel » entre les classifications des expositions des femmes de la cohorte E3N, localisées à l'adresse de résidence, estimées avec l'indicateur, et avec le modèle gaussien SIRANE.

En utilisant le modèle SIRANE, nous avons également étudié la variabilité spatio-temporelle (1990-2008) des concentrations de dioxines et de cadmium sur la métropole de Lyon. Pour la première fois des concentrations modélisées ont été comparées à des mesures de concentrations de dioxines en air ambiant.

L'indicateur développé dans cette thèse a été utilisé pour estimer le risque de cancer du sein associé à l'exposition atmosphérique aux dioxines dans une étude cas-témoins niché au sein de la cohorte E3N. Il est utilisé actuellement dans une étude sur le lien entre exposition au cadmium et risque de cancer du sein, et pourra être appliqué dans de futures études portant sur d'autres polluants ou d'autres pathologies.

Mots-clés : dioxines, cadmium, exposition atmosphérique, évaluation des expositions, système d'information géographique, modèle gaussien, évaluation des émissions, cancer du sein, perturbateurs endocriniens, xénoestrogènes, métaux lourds

English thesis title

---

**Development and application of assessment methodologies for chronic airborne dioxin and cadmium exposures to be used in epidemiological studies**

---

## English summary

A number of studies have examined the link between breast cancer and exposure to air pollution, including dioxins and cadmium. However, the results of these studies are inconclusive and present a number of methodological limitations.

The main objective of this thesis was to develop a spatial indicator to assess chronic atmospheric exposure to dioxins and cadmium of women from the E3N cohort in France between 1990 and 2008.

We first performed an inventory and created a database of 2620 dioxins and 2700 cadmium emitting sources in France between 1990 and 2008 and estimated and geolocated their annual emissions.

The location of the sources and their estimated emissions were used as the basis for the construction of the indicator. Combination of additional spatial parameters, allowed us to obtain a "substantial" agreement between the dioxin and cadmium exposure classifications of the E3N subjects geolocalised at their residential address, using the estimated by the indicator and exposures estimates derived from the Gaussian model.

We also evaluated the spatial-temporal variability of dioxin and cadmium concentrations over nearly two decades in the Lyon metropolitan area, taking into account a wide variety of source types. This is the first study comparing concentrations predicted by a dispersion model to dioxin concentrations measured in ambient air.

The exposure indicator was used in a case-control study within the E3N cohort to estimate the risk of breast cancer associated with atmospheric exposure to dioxins. It is currently being used in another study on cadmium exposure and breast cancer risk and may be applied in future studies on other pollutants or pathologies.

Keywords: dioxins, cadmium, atmospheric exposure, exposure assessment, geographic information system, gaussian model, emissions assessment, breast cancer, endocrine disrupters, xenoestrogens, heavy metals

# UNIVERSITE CLAUDE BERNARD LYON 1

## **Président de l'Université**

Président du Conseil Académique

Vice-président du Conseil d'Administration

Vice-président du Conseil Formation et Vie Universitaire

Vice-président de la Commission Recherche

Directrice Générale des Services

**M. le Professeur Frédéric FLEURY**

M. le Professeur Hamda BEN HADID

M. le Professeur Didier REVEL

M. le Professeur Philippe CHEVALIER

M. Fabrice VALLÉE

Mme Dominique MARCHAND

## ***COMPOSANTES SANTE***

Faculté de Médecine Lyon Est – Claude Bernard

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Faculté d'Odontologie

Institut des Sciences Pharmaceutiques et Biologiques

Institut des Sciences et Techniques de la Réadaptation

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : M. le Professeur G.RODE

Directeur : Mme la Professeure C. BURILLON

Directeur : M. le Professeur D. BOURGEOIS

Directeur : Mme la Professeure C. VINCIGUERRA

Directeur : M. X. PERROT

Directeur : Mme la Professeure A-M. SCHOTT

## ***COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE***

Faculté des Sciences et Technologies

Département Biologie

Département Chimie Biochimie

Département GEP

Département Informatique

Département Mathématiques

Département Mécanique

Département Physique

UFR Sciences et Techniques des Activités Physiques et Sportives

Observatoire des Sciences de l'Univers de Lyon

Polytech Lyon

Ecole Supérieure de Chimie Physique Electronique

Institut Universitaire de Technologie de Lyon 1

Ecole Supérieure du Professorat et de l'Education

Institut de Science Financière et d'Assurances

Directeur : M. F. DE MARCHI

Directeur : M. le Professeur F. THEVENARD

Directeur : Mme C. FELIX

Directeur : M. Hassan HAMMOURI

Directeur : M. le Professeur S. AKKOUCHE

Directeur : M. le Professeur G. TOMANOV

Directeur : M. le Professeur H. BEN HADID

Directeur : M. le Professeur J-C PLENET

Directeur : M. Y. VANPOULLE

Directeur : M. B. GUIDERDONI

Directeur : M. le Professeur E. PERRIN

Directeur : M. G. PIGNAULT

Directeur : M. le Professeur C. VITON

Directeur : M. le Professeur A. MOUGNIOTTE

Directeur : M. N. LEBOISNE

## **Affiliations of des membres du jury de thèse**

### **Pr CHARBOTEL Barbara**

PU-PH2 Médecin spécialiste en santé au travail,  
HCL (UMRESTTE)  
Université Claude Bernard Lyon 1

### **Pr DEGUEN Séverine**

PU, HDR, enseignant-chercheur à l'Ecole des Hautes Etudes en Santé Publique, Université Sorbonne Paris Cité et Université Bretagne Loire, Rennes

### **Pr FERVERS Béatrice**

Professeur associé, MD, PhD, HDR  
Coordinatrice du Département Cancer et Environnement, Centre Léon Bérard  
Équipe « Signalisation des hormones stéroïdiennes et cancer du sein », UMR Inserm 1052 – CNRS 5286, Centre de Recherche en Cancérologie de Lyon

### **Dr PASCAL Mathilde**

PhD, Chargée de projet scientifique, Santé Publique France Paris

### **Dr SALIZZONI Pietro**

MCF, HDR, enseignant chercheur, Laboratoire de mécanique de fluides et d'acoustique, Ecole centrale de Lyon

### **Dr SEVERI Gianluca**

PhD, HDR, Directeur de Recherche, Inserm - Centre de Recherche en Épidémiologie et Santé des Populations (CESP, UMR Inserm 1018) Université Paris-Saclay, Gustave Roussy

### **Pr SOULHAC Lionel**

MCF, HDR, co-directeur du Laboratoire de Mécanique des Fluides et d'Acoustique, Ecole Centrale de Lyon / INSA de Lyon / Université Claude Bernard Lyon 1 / CNRS

### **Pr VINEIS Paolo**

MD, MPH, FFPH, HDR  
Chair in Environmental Epidemiology  
MRC/PHE Centre for Environment and Health, School of Public Health, Imperial College London



## **Affiliations**

Le travail présenté dans cette thèse a été mené au sein des départements et équipes suivantes:

**Département Cancer et Environnement**

**Centre Léon Bérard**

28 rue Laënnec, 69373 Lyon Cedex 08, France

**Equipe « Atmosphere, Impact & Risk »**

**Laboratoire de mécanique des fluides et d'acoustiques**

**Ecole Centrale de Lyon**

36 Avenue Guy de Collongue, 69134 Écully

**Equipe « Steroid signalling and breast tumor », UMR Inserm 1052, CNRS 5286**

**Centre de recherche contre le cancer de Lyon**

**Université Claude Bernard Lyon 1**

28 rue Laënnec, 69373 Lyon Cedex 08, France

## Remerciements

*Béatrice, merci de m'avoir accueilli dans le Département et d'avoir cru en moi pour réaliser cette thèse. Je souhaite vous remercier pour votre réelle implication dans ce projet, votre sens du détail, votre exigence et pour avoir accepté ma franchise.*

*Pietro, merci pour tes nombreux conseils, tes idées, ton humour, nos discussions et ta disponibilité.*

*Je remercie Mathilde Pascal, Paolo Vineis et Séverine Deguen pour avoir accepté de rapporter ce travail. Je remercie également les autres membres du Jury : Barbara Charbotel, Gianluca Severi et Lionel Soulhac,*

*Je remercie Barbara Charbotel et Lionel Soulhac pour avoir participé à mes comités de suivi de thèse ainsi que pour leurs précieux conseils.*

*Je tiens à remercier Béatrice, Pietro, Aurélie, Elodie, Delphine et Marie-Line pour m'avoir aidé dans la relecture de cette thèse.*

*Je remercie mes collègues, avec qui j'ai travaillé quotidiennement sur les projets GEO3N et XENAIR : Aurélie, Camille, Charlotte C, Delphine, Elodie F, Guillaume, Hassan, Laure, Maxime, Tingting et Xavier et Yue.*

*J'ai une pensée particulière pour les collègues de bureau qui m'ont supporté et tant apporté durant cette thèse : Guillaume (directeur de stage et compagnon d'aventure), Camille (qui m'a ouvert l'esprit), Hassan (on n'oublie jamais les premiers), Xavier (qui m'a appris la science des belles cartes) et Chi-Vuong (toujours disponible pour m'aider avec calme et pédagogie).*

*Après ces 4 années au Département Cancer et Environnement, je souhaite remercier : Alix, Ariane, Aurélie, Axel, Antoine, Alexandre, Audrey, Amina, Baptiste, Cédric, Charlotte, Delphine, Elodie B, Elodie F, Etienne D, Etienne M, Hélène, Jeffrey, Jennifer, Joane, Julien, Katia, Lidia, Marina, Marine, Magali, Margaux, Nora, Nicole, Olivia C, Olivia F, Olivia P, Rémi, Renaud, Rodolf, Solène et Yue.*

*Surtout ne changez rien.*

*Je tiens à remercier mes collègues de l'Ecole Centrale pour m'avoir accueillie et soutenu tout le long de cette thèse : Guillevic, Horacio, Massimo, Patrick, Perrine ; ainsi que tous les doctorants mangeurs de sandwiches : Chi-Vuong, Cristina, Fateh, Hervé, Mateo et Thierry.*

*Je remercie les membres de l'équipe E3N, notamment Françoise Clavet-Chapelon, Francesca Mancini, Pascale Gerboin et Gianluca Severi.*

*Je remercie John Gulliver de l'Imperial College London pour m'avoir accueilli quelques semaines dans son équipe et spécialement Adi pour m'avoir tellement aidé durant ce séjour.*

*Je remercie l'Université Claude Bernard Lyon 1 et l'EDISS pour m'avoir attribué une bourse de thèse et le CLARA pour m'avoir attribué une bourse de mobilité.*

*Je tiens à remercier tous mes amis, qu'ils m'aient soutenus de près ou de loin :*

*Antoine (because we can't, we won't and we don't stop), Maître Bertrand, Guillaume, Marie-Line, Maxime & Bugs bunny, Mon Gronkowski, La meilleure : Aurore, la team barathon, Remy, Pauline, Lucie, Tristan, Manon, l'immense Rompy, Titi, Sochay, Maxime, Flo', Delphine, Elodie, Aurélie, Adi « goodness me », Mateo, Alexandre, mes partenaires de squash responsables de plusieurs raquettes cassées, la bande des chibres qui ne lira jamais ces remerciements (merci Jérôme), les mecs de com-com et les anciens du Master qui manque un peu : Jennyfouille, John, Jeremy, Jack et Xavier (Aighttt !)*

*Je souhaite remercier particulièrement tous les membres de ma famille, qui m'ont immensément soutenu pendant toutes ces années. Sans vous, je ne serais sûrement pas arrivé jusqu'ici.*

*Merci encore à mes parents, mon frère et mes grands-parents.*



## Table des matières

Résumé en Français .....	3
English thesis title .....	4
English summary .....	5
UNIVERSITE CLAUDE BERNARD LYON 1 .....	6
Affiliations of des membres du jury de thèse .....	7
Affiliations .....	8
Remerciements .....	9
Table des matières .....	12
Liste des abréviations .....	15
Liste des figures.....	16
Liste des tableaux .....	17
Liste des publications .....	18
Liste des présentations orales .....	18
Liste des présentations poster .....	19
<b>CHAPITRE I. INTRODUCTION .....</b>	<b>21</b>
1. Contexte scientifique.....	22
1.1. Exposition atmosphérique, perturbateurs endocriniens et risque de cancer.....	22
1.2. Dioxines .....	23
1.2.1. Classification, mécanismes biologiques et propriétés.....	23
1.2.2. Emissions et sources de dioxines .....	24
1.2.3. Expositions humaines aux dioxines .....	27
1.2.4. Dioxine et risque de cancer du sein.....	28
1.3. Cadmium.....	29
1.3.1. Classification, mécanismes biologiques et propriétés.....	29
1.3.2. Emissions et sources de cadmium.....	29
1.3.3. Exposition humaine au cadmium .....	31
1.3.4. Cadmium et risque de cancer du sein.....	32
1.4. Complexité de l'étude du lien entre exposition atmosphérique aux dioxines et au cadmium et risque de cancer du sein .....	33

1.5. Les méthodes de quantification des expositions atmosphériques chroniques.....	36
2. Objectifs et contexte de la thèse.....	44
2.1. Objectifs.....	44
2.2. Contexte.....	46
2.2.1. Département cancer et environnement.....	46
2.2.2. La cohorte E3N.....	46
2.2.3. Equipe AIR (Laboratoire de Mécanique des fluides et d’acoustiques – Ecole centrale de Lyon).....	47
2.2.4. Le modèle SIRANE.....	48
2.2.5. Projet GEO3N.....	50
2.2.6. Projet XENAIR.....	52
<b>CHAPITRE II. EMISSIONS ATMOSPHERIQUES INDUSTRIELLES DE DIOXINE ET DE CADMIUM.....</b>	<b>55</b>
1. Introduction.....	56
2. Article n°1.....	59
3. Extension de la base de données aux émetteurs de cadmium.....	92
4. Conclusion.....	101
<b>CHAPITRE III. EXPOSITION ATMOSPHERIQUE AUX DIOXINES ET AU CADMIUM .</b>	<b>105</b>
1. Introduction.....	106
2. Article n°2.....	109
3. Conclusion.....	143
<b>CHAPITRE IV. EVALUATION DE L’EXPOSITION ATMOSPHERIQUE CHRONIQUE AUX DIOXINES ET CADMIUM EN FRANCE (1990-2008).....</b>	<b>147</b>
1. Introduction.....	148
2. Article n°3.....	150
3. Conclusion.....	186
Application de l’indicateur dans l’étude cas-témoins nichée GEO3N.....	188
<b>CHAPITRE V. DISCUSSION ET PERSPECTIVES.....</b>	<b>192</b>
1. Discussion.....	193
2. Perspectives de recherches.....	198
Le projet XENAIR.....	201
3. Conclusion.....	204
<b>Références.....</b>	<b>206</b>

<b>Annexes .....</b>	<b>217</b>
Annexe 1. Revue de la littérature : liste des équations de recherche bibliographique utilisées pour identifier les paramètres d'intérêts dans la construction de l'indicateur SIG .....	217
Annexe 2. Présentation des 7 articles sélectionnés dans la revue de littérature .....	218
Annexe 3. Danjou et al. 2018. <i>En cours de finalisation. 2018</i> .....	220

## Liste des abréviations

ADEME : Agence de l'environnement et de la maîtrise de l'énergie  
ADMS : Atmospheric Dispersion Modelling System  
AhR : Récepteur d'aryl d'hydrocarbure ou récepteur aux hydrocarbures aromatiques  
ATMO AURA : Association agréée de surveillance et de la qualité de l'air en Auvergne-Rhône-Alpes  
BaP : Benzo(a)pyrene  
C<sub>6</sub>H<sub>6</sub> : Benzène  
CIRC : Centre international de recherche sur le cancer  
CITEPA : Centre interprofessionnel technique d'études de la pollution atmosphérique  
Cnil : Commission nationale de l'informatique et des libertés  
CO : Monoxyde de carbone  
COV : Composé organique volatil  
DCE : Département cancer et environnement  
EMEP : European Monitoring and Evaluation Programme  
ER/PR : Récepteur d'œstrogènes (ER) et de progestérone (PR)  
HAP : Hydrocarbure aromatique polycyclique  
ICPE : Installations Classées pour la Protection de l'Environnement  
LUR : Land use regression  
NO<sub>2</sub> : Dioxyde d'azote  
OMINEA : Organisation et méthodes des inventaires nationaux des émissions atmosphériques  
OMS : Organisation mondiale pour la santé  
O<sub>3</sub> : Ozone  
PCB : Polychlorobiphényles  
PCDD/DF : Polychlorodibenzodioxines/Polychlorodibenzofuranes  
PE : Perturbateurs endocriniens  
PM<sub>10/2.5</sub> : Particules fines d'un diamètre inférieur ou égal à 10/2.5 microns  
PNSE : Plan national santé environnement  
POP : Polluants organique persistant  
SIG : Système d'information géographique  
TCDD : 2,3,7,8-tetrachlorodibenzo-para-dioxin (Tetrachlorodibenzo-p-dioxine)  
TEF : Facteurs équivalents toxiques  
TEQ : Quotient équivalent toxique  
TSP Particules totales en suspension  
UIOM : Unité d'incinération des ordures ménagères  
UNEP : United Nations Environment Programme (Programme des Nations unies pour l'environnement)  
US EPA : United States Environmental Protection Agency (Agence américaine de protection de l'environnement)  
wk : coefficient kappa pondéré



## Liste des figures

Figure 1- Représentation qualitative des niveaux de production de produits chimiques dans les pays occidentaux. L'axe vertical correspond à une échelle arbitraire (Slama et al. 2017).....	22
Figure 2 – Évolution des taux de concentrations de PCDD/DF entre 1850 et 1995, issues d'un herbier de la station expérimentale de Rothamsted (UK) (Kjeller et al. 1996). .....	25
Figure 3 – Estimation des émissions annuelles de dioxines en France entre 1990 et 2016 en g-ITEQ .... (CITEPA 2017) .....	27
Figure 4 - Estimation des émissions annuelles de cadmium (en tonnes) en France entre 1990 et 2016... (CITEPA 2017) .....	31
Figure 5 – Structure de l'écoulement dans une rue canyon.....	48
Figure 6 - Données d'entrée de de sortie de SIRANE .....	49
Figure 7 – Répartition par groupes des sources de dioxines et de cadmium .....	96
Figure 8 – Evolution en pourcentage des émissions annuelles de dioxines par groupes UNEP entre 1990 et 2008 .....	97
Figure 9 – Evolution en pourcentage des émissions annuelles de cadmium par groupe UNEP entre 1990 et 2008 .....	97
Figure 10 – Evolution des émissions total de cadmium (tonnes/an) par groupes UNEP entre 1990 et 2008.....	97
Figure 11 - Carte des zones symposiums définis pour le projet GEO3N (n=223).....	189
Figure 12 - Application de l'indicateur SIG dans l'étude GEO3N. Distribution des sujets (pourcentage sur l'axe verticale et effectifs sur les barres) selon l'exposition environnementale aux dioxines estimé par l'indicateur spatialisé (exposition cumulée) dans l'étude GEO3N (N=9058).....	189
Figure 13 - Application de l'indicateur SIG dans l'étude GEO3N. Répartition des sujets selon leur l'exposition cumulative (1990-2008) aux dioxines (en $\mu\text{g-TEQ}/\text{m}^2$ ) de la moins exposées à la plus exposées. Environ 900 sujets sont considéré comme non-exposées car en dehors de toutes zones d'exposition.....	190
Figure 14 - Exemple de variations des expositions annuelles de 20 sujets, choisis aléatoirement, entre 1990 et 2008, calculée avec l'indicateur SIG (en $\mu\text{g}/\text{m}^3$ ). .....	190

## Liste des tableaux

Tableau 1- Comparaisons entre les mesures et les estimations (basées sur les facteurs d'émissions EMEP et OMINEA) des émissions de cadmium (kg/an) des sources industrielles en Rhône-Alpes .....	94
Tableau 2 – Répartition des nouvelles sources identifiées durant l'inventaire des sources de cadmium .....	95
Tableau 3 – Evolution des émissions annuelles médianes des sources émettrices de cadmium (kg/m <sup>3</sup> ) par groupe entre 1990 et 2008 .....	98
Tableau 4 – Comparaison des émissions de cadmium (kg/an) estimées dans notre étude avec les incertitudes associées en pourcentage ( $\pm$ ) et estimée avec les facteurs d'émission de l'EMEP pour 4 catégories entre 1990 et 2008 .....	99
Tableau 5 - Présentation des 7 articles sélectionnés dans le cadre de la revue de la littérature réalisée pour définir la métrique SIG .....	218

## Liste des publications

Coudon T, Hourani H, Nguyen C, Faure E, Mancini FR, Fervers B, Salizzoni P. 2018. Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon metropolitan area (France). *Environment International* 111:177–190; doi:10.1016/j.envint.2017.11.027.

Coudon T, Danjou AMN, Faure E, Praud D, Mancini FR, Severi G, Salizzoni P, Fervers B. 2018. Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutant emissions from industrial sources. *Accepté Environmental Health (2018)*.

Coudon T, Salizzoni P, Praud D, Faure E, Danjou AMN, Fervers B. 2018. National inventory of historical dioxin air emissions source in France. *Soumis Science of the Total Environment (avril 2018)*.

Danjou AMN, Coudon T, Praud D, Lévêque E, Faure E, Salizzoni P, Le Romancer M, Severi G, Mancini FR, Leffondré K, Dossus L and Fervers B. Airborne Dioxin Exposure and Breast Cancer Risk in a case-control study nested within the French E3N Prospective Cohort: a GIS-based approach. *En cours de finalisation*.

## Liste des présentations orales

Coudon. 2015. Développement et validation d'un score d'exposition aux dioxines pour une application dans des études épidémiologiques. 19ème Journée Scientifique de l'EDISS, Ecole Doctorale Interdisciplinaire Science-Santé, Lyon, France.

Coudon. 2015. Expositions environnementales aux dioxines et cancer du sein dans la cohorte E3N. ADEME : Connaissances des impacts sanitaires et environnementaux de la gestion des déchets, Angers, France.

Coudon. 2016. Reconstructing past and present dioxins atmospheric pollution scenarios in the city of Lyon. Transport & Air pollution congress. Lyon, France

Coudon. 2016. Développement d'un SIG pour l'estimation de l'exposition aux dioxines et la comparaison avec un modèle gaussien (SIRANE). Journée thématique Santé-Environnement, Université Grenoble-Alpes, France.

Coudon. 2016 (*présenté par Béatrice Fervers*). Development of a GIS based exposure metric to assess environmental dioxin exposure and comparison with an urban Gaussian model. 28th Annual Conference International Society for Environmental Epidemiology 2016. Rome, Italie.

Coudon. 2016. Determination of a geographic information system based indicator to assess environmental dioxins exposure in Lyon and through comparisons with an atmospheric dispersion model results. 7eme Congrès national Santé et Environnement : Qualité de l'air et santé, Strasbourg, France.

Coudon 2017. Risque de cancer et pollution de l'air. Soirée débat grand public. Forum de la recherche en cancérologie du CLARA. Lyon, France. [Vidéo](#).

Coudon 2017. Risque de cancer et pollution de l'air. Soirée débat, Maison de l'environnement. Lyon.

Coudon 2018. Présentation des conclusions du projet GEO3N. Journée de valorisation des projets de recherche APR CIDE ADEME. Paris, France.

Coudon 2018. Exposition à la pollution de l'air et risque de cancer. Conférence grand public ERI, Centre Léon Bérard, Lyon, France.

## Liste des présentations poster

Coudon. 2016. Determination of a geographic information system based indicator to assess environmental dioxins exposure in Lyon and through comparisons with an atmospheric dispersion model results. IARC 50th Anniversary Conference 2016. Lyon, France.

Coudon. 2016. Reconstructing past and present dioxins atmospheric pollution scenarios in the city of Lyon. 7eme Congrès national Santé et Environnement : Qualité de l'air et santé. Strasbourg, France

Coudon. 2016. Construction et validation d'un SIG pour l'estimation de l'exposition aux dioxines. Journée médico-scientifique du Centre Léon Bérard. Lyon France.

Coudon. 2017. Determination of a geographic information system based indicator to assess environmental dioxins exposure in Lyon and through comparisons with an atmospheric dispersion model results. Forum de la recherche en cancérologie du CLARA. Lyon, France.

Coudon. 2018. Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutant emissions from industrial sources. European congress of epidemiology, Lyon, France

Coudon. 2018. Development and performance evaluation of a GIS-Based Metric to assess exposure to airborne pollutant emissions from industrial sources. 9<sup>th</sup> International conference on the science of exposure assessment, Manchester, UK.



## **CHAPITRE I. INTRODUCTION**

## 1. Contexte scientifique

### 1.1. Exposition atmosphérique, perturbateurs endocriniens et risque de cancer

La pollution de l'air extérieur est un problème majeur de santé publique entraînant des effets néfastes sur la santé. L'OMS estime qu'elle est responsable de près de 3,1 millions de décès prématurés par an dans le monde (WHO 2013). C'est une problématique ancienne et complexe car la pollution de l'air est un mélange de polluants multiples dont les émissions atmosphériques ont fortement varié au cours du siècle dernier (Figure 1). Parmi ces polluants, un certain nombre sont classés cancérigènes pour l'Homme (Groupe 1) par le Centre International de Recherche sur le Cancer (CIRC), notamment les PCB, le benzo[a]pyrène (B[a]P), le cadmium et la 2,3,7,8-tétrachlorodibenzo-para-dioxine (TCDD) (Baan et al. 2009).

Récemment, le CIRC a classé la pollution de l'air extérieur en tant que mélange et les matières particulaires provenant de la pollution de l'air extérieur, comme cancérigènes pour l'Homme (Groupe 1), avec des preuves suffisantes pour le cancer du poumon (Loomis et al. 2013). La pollution de l'air intérieur, en particulier issue de la combustion, a également été positivement associée à plusieurs cancers (Josyula et al. 2015).

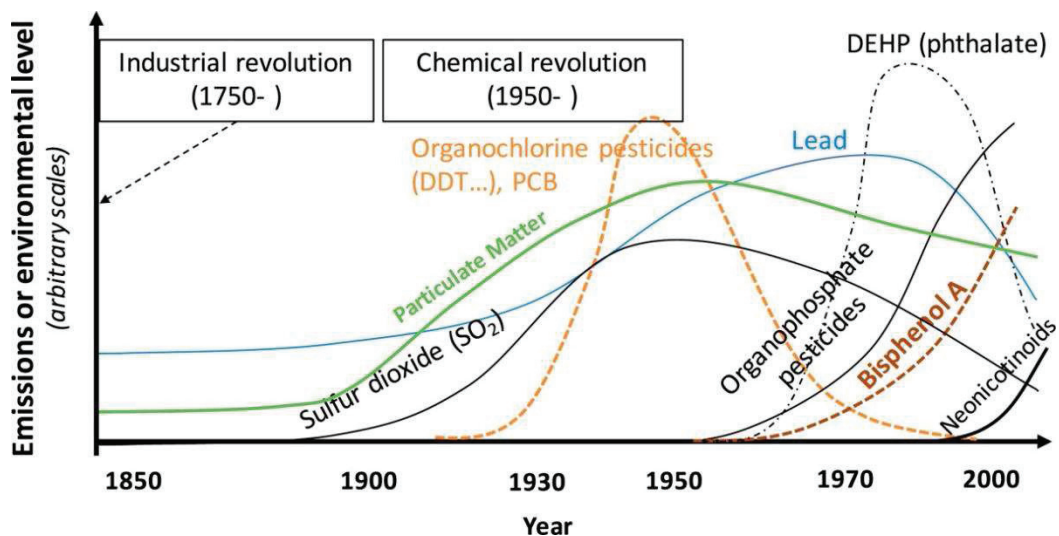


Figure 1- Représentation qualitative des niveaux de production de produits chimiques dans les pays occidentaux. L'axe vertical correspond à une échelle arbitraire (Slama et al. 2017).

Parmi les polluants présents dans l'air ambiant, les études épidémiologiques s'intéressent plus particulièrement aux polluants à effets perturbateurs endocriniens et à leurs impacts sur la santé. L'Environmental Protection Agency (US EPA) définit les perturbateurs endocriniens (PE) comme des « agents exogènes qui interfèrent avec la synthèse, la sécrétion, le transport, le métabolisme, l'action contraignante ou l'élimination des hormones naturelles transmises par le sang qui sont présentes dans le

corps et sont responsables de l'homéostasie, de la reproduction et du processus de développement » (Diamanti-Kandarakis et al. 2009; Kavlock et al. 1996). La famille des PE comprend une large gamme de molécules hétérogènes, incluant des pesticides organochlorés, des produits chimiques industriels, des plastiques, des métaux lourds et de nombreux d'autres produits chimiques. Ils sont libérés dans l'environnement, entre autres lors d'activités industrielles et de combustion, où ils peuvent persister en raison de leur longue demi-vie, notamment les substances chlorées très stables.

Les études épidémiologiques et expérimentales suggèrent que les polluants environnementaux présentant des propriétés de PE peuvent contribuer au développement de cancers hormono-dépendants comme le cancer du sein (Brody et al. 2007; Teitelbaum et al. 2015). Les dioxines et le cadmium sont deux exemples de ces polluants à effets perturbateurs endocriniens.

## **1.2. Dioxines**

### **1.2.1. Classification, mécanismes biologiques et propriétés**

Les polychlorodibenzo-para-dioxines (PCDD) et les polychlorodibenzofuranes (PCDF), communément appelés "dioxines", sont des polluants organiques persistants (POP) avec une longue demi-vie dans l'environnement et chez l'Homme. Ce groupe comprend 210 congénères (75 PCDD et 135 PCDF) qui se distinguent par le nombre et la position des atomes de chlore qu'ils possèdent (entre 4 et 8). Parmi ces congénères, 17 ont été identifiés comme toxiques par l'Organisation mondiale de la santé (OMS) (7 PCDD et 10 PCDF). La TCDD est le congénère le plus toxique et a été classé, par le CIRC, comme cancérigène certain pour l'Homme (Groupe 1) avec des preuves suffisantes pour tous les cancers combinés et avec des preuves limitées pour le cancer du poumon, les sarcomes des tissus mous et les lymphomes non hodgkiniens (IARC 1997).

La TCDD agit comme un perturbateur endocrinien via l'activation ou l'inhibition du récepteur aux hydrocarbures aromatiques (AhR). Le mécanisme d'action principal de la TCDD en cancérogenèse est la promotion du développement tumoral par activation de la réplication cellulaire, l'altération de la sénescence cellulaire et de l'apoptose. Le deuxième mécanisme est l'augmentation du stress oxydatif. La TCDD n'est pas directement génotoxique mais provoque une augmentation des dérivés réactifs de l'oxygène conduisant à des dommages et des mutations de l'ADN (Mandal 2005). Certaines études ont décrit une action antiproliférative de la TCDD, dépendant ou indépendant d'AhR (Yoshioka et al. 2012; Zhang et al. 2009).

La TCDD a une faible solubilité dans l'eau, mais est très lipophile et s'accumule dans les tissus adipeux. Elle a été mesurée chez l'Homme dans les tissus adipeux, le sang et le lait maternel (Arisawa et al. 2003; Sonawane 1995). Sa demi-vie chez l'Homme est estimée à 7,2 ans (Milbrath et al. 2009).



L'exposition humaine aux dioxines est calculée en termes de quotients d'équivalence toxique (TEQ). Un facteur d'équivalence toxique (TEF) est attribué à chacun des 17 congénères toxiques de dioxines, pondéré par rapport à la TCDD. La toxicité d'un mélange de congénères est calculée en multipliant les concentrations individuelles de congénères par leur TEF respectif puis en additionnant les valeurs obtenues pour chaque congénère. Selon les facteurs d'équivalents toxiques utilisés, deux classifications existent : iTEQ (établis par l'OTAN en 1989) et TEQ (créée par l'OMS en 1998 et mise à jour en 2005). L'utilisation de la classification TEQ est maintenant recommandée (US EPA 2010).

### 1.2.2. Emissions et sources de dioxines

Les dioxines n'ont pas d'application, ni d'utilité commerciale connue et sont produites involontairement pendant des processus de combustion impliquant une source de chlore. Le CIRC définit quatre grandes catégories de sources environnementales de dioxines (IARC 2012) :

- Émissions provenant de l'incinération : déchets municipaux, déchets hospitaliers, déchets dangereux et incinération des boues d'épuration ;
- Émissions provenant de la combustion : les fours à ciment, la combustion du bois, le chauffage domestique, le brûlage de câbles, les centrales électriques au charbon et les crématoriums ;
- Émissions provenant de l'activité industrielle : usines de pâtes et papiers, fabrication de produits chimiques et industrie métallurgique ;
- Émissions provenant de processus biochimiques : processus photolytiques, incendies de forêt, éruptions volcaniques. Ces émissions naturelles ponctuelles sont parfois importantes mais représentent une très faible proportion des émissions totales.

Les émissions de dioxines ont évolué au cours des dernières décennies en raison des changements réglementaires encadrant les sources d'émissions. Les dioxines étaient contenues dans des herbicides couramment utilisés dans les années 60 et dans les défoliants utilisés durant la guerre du Vietnam. Les principales émissions résultaient de la phase de fabrication de ces composés chlorés (chlorophénols et herbicides phénoxy chlorés). Ces émissions ont été considérablement réduites en raison de l'évolution de la législation et de la restriction de l'utilisation des pesticides organochlorés dans l'agriculture. Par la suite, la multiplication des usines d'incinération, pour répondre à la problématique de gestion des déchets municipaux, a provoqué une nouvelle augmentation des émissions dans les années 1980 (Figure 2).

Le secteur de l'incinération de déchets est alors devenu la principale source d'émission de dioxines, avant de diminuer suite à la mise en place de réglementation spécifique encadrant les émissions de dioxines, et particulièrement les activités d'incinérations de déchets.

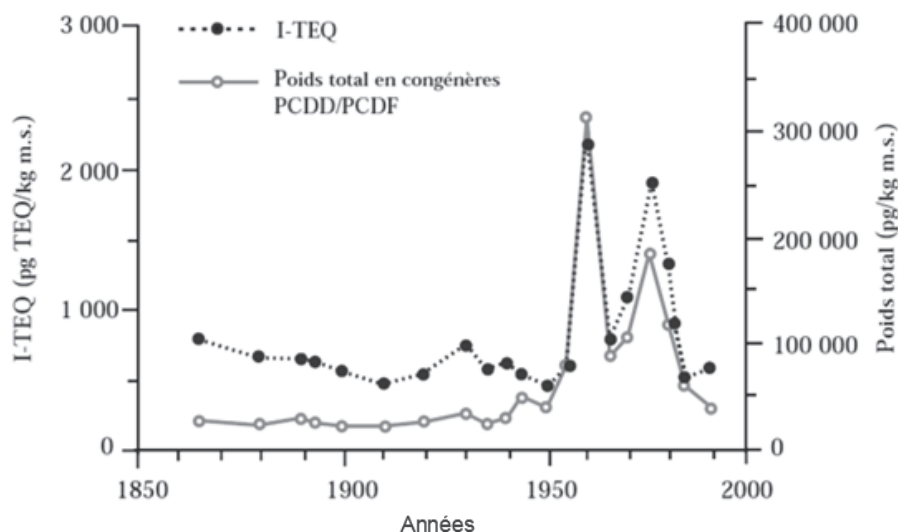


Figure 2 – Évolution des taux de concentrations de PCDD/DF entre 1850 et 1995, issues d'un herbier de la station expérimentale de Rothamsted (UK) (Kjeller et al. 1996).

Dans les années 90, la France possédait le plus grand parc d'incinérateurs de déchets solides municipaux de l'Union européenne. La réglementation française encadrant les émissions de dioxines des Installations Classées pour la Protection de l'Environnement (ICPE) a connu plusieurs évolutions majeures entre 1990 et 2005. Trois arrêtés ministériels progressivement entrés en vigueur ont contribué à la réduction des émissions industrielles de dioxines en France.

Pour les installations d'incinérations de déchets :

- L'arrêté ministériel du 25 janvier 1991, applicable à partir du 1er décembre 2000, et relatif aux installations d'incinérations de résidus urbains (déchets ménagers et provenant d'établissement sanitaires), prescrit un traitement des fumées des incinérateurs. L'arrêté impose des conditions d'incinération en termes de température, de temps de séjour et de taux d'oxygène favorable à la destruction des dioxines en sortie d'incinérateur (maintien des fumées en sortie de combustion à plus de 850°C pendant plus de 2 secondes).
- L'arrêté ministériel du 20 septembre 2002 fixe une valeur limite d'émission de dioxines de 0,1 ng-TEQ/Nm<sup>3</sup>, valeur limite encore utilisée aujourd'hui. La date limite de mise en application fut fixée au 28 décembre 2005. Cet arrêté met également en place une obligation de contrôle des émissions de dioxines 2 fois par an.

Pour les autres installations classées :

- L'arrêté du 02 février 1998 ne propose pas de valeur limite à l'échelle nationale. Néanmoins, à l'échelle locale, de nombreuses installations relevant de cet arrêté ministériel se sont vu fixer des valeurs limites à l'émission par des arrêtés préfectoraux, sur la base des informations contenues dans les « BREF » (documents de référence qui répertorient les meilleures techniques disponibles pour satisfaire les critères de développement durable). C'est le cas pour un grand nombre d'installations métallurgiques.

En plus des arrêtés ministériels, le Plan National Santé Environnement (PNSE) a été mis en place en 2004 pour lutter contre les risques sanitaires dus aux pollutions des milieux de vie. Il contient au total 45 actions, qui recouvrent l'ensemble du champ des relations entre environnement et santé. Son action n°7 portait sur la réduction des émissions aériennes de substances toxiques d'origine industrielle. Cette action visait un objectif de réduction des émissions de 85 % pour les dioxines à l'horizon 2010 (PNSE 2004). Deux autres PNSE ont été publiés en 2009 et 2015 (PNSE 2009, 2015).

Ces différents changements de réglementation ont ainsi entraîné une diminution du nombre d'installations et un meilleur contrôle des émissions (Anzivino-Viricel et al. 2012). Selon le centre technique interprofessionnel d'études sur la pollution atmosphérique (CITEPA), les émissions de dioxines ont diminué de plus de 90 % entre 1990 et 2008, passant de 1776 g-TEQ estimées en 1990 à 143 g-TEQ estimées en 2008 (Figure 3) (CITEPA 2017). Les différents arrêtés ont permis de réduire les émissions des incinérateurs d'ordures ménagères de 99 % et les émissions des incinérateurs de déchets industriels dangereux de 89 %. En revanche, les émissions dues aux transports et au chauffage résidentiel sont restées quasiment constantes sur la même période (CITEPA 2017). A l'heure actuelle, la source d'émission majeure est une activité illégale : le brûlage de câble. Cette source d'émission est officiellement associée au secteur résidentiel tertiaire dans les différents inventaires d'émissions.

### Emissions atmosphériques de dioxine en g I-TEQ en tonnes par secteur en France métropolitaine

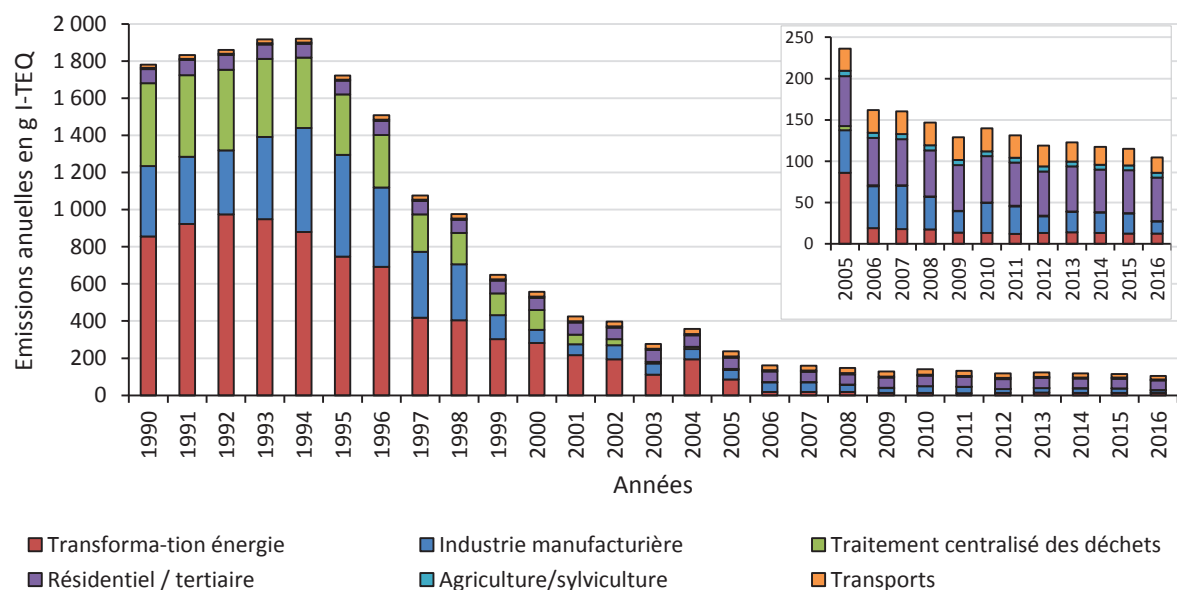


Figure 3 – Estimation des émissions annuelles de dioxines en France entre 1990 et 2016 en g-ITEQ (CITEPA 2017)

#### 1.2.3. Expositions humaines aux dioxines

Les dioxines sont des contaminants ubiquitaire et omniprésents dans l'environnement. En dehors des expositions accidentelles ou professionnelles, les humains sont exposés par voie alimentaire lors d'ingestion de produits riches en matière grasse (œuf, poissons, lait, viandes). Le CIRC a estimé qu'environ 90 % de l'exposition aux dioxines dans la population générale provient de l'ingestion d'aliments contaminés, en raison de l'accumulation de ces composés dans la chaîne alimentaire (IARC 2012). On estime l'exposition journalière aux dioxines via l'alimentation, pour un adulte, à environ 40 pg-TEQ (IARC 2012). L'exposition aux dioxines dans la population générale est considérée comme faible : une étude reprenant les résultats de 109 études publiées entre 1989 et 2010, a estimé que la concentration sanguine moyenne de TCDD dans la population générale était de 1,8 pg/g de lipide (écart type SD = 1,7) (Consonni et al. 2012). Cependant, comme les dioxines sont des polluants persistants, la population générale subit une exposition cumulative à long terme, et les concentrations sanguines de dioxines augmentent avec l'âge (Consonni et al. 2012).

La classification de la TCDD comme cancérigène pour l'Homme par le CIRC repose sur les résultats issus d'études portant sur des populations fortement exposées (aux herbicides contenant des dioxines ou lors d'accidents industriels). Contrairement à la population générale, ces groupes spécifiques ont des niveaux d'exposition beaucoup plus élevés en raison de leur exposition

professionnelle. Parmi ces populations historiquement les plus exposées, on retrouve les ouvriers travaillant à la production d'herbicides et chlorophenols ainsi que les applicateurs d'herbicide utilisant des sprays (IARC 2012). Ces expositions ont mené à des concentrations de TCDD dans les tissus de 10 à 100 fois supérieures aux valeurs de fond de la population générale dans les années 1980 (IARC 1997). De manière générale, due à la baisse des émissions, une baisse des concentrations de TCDD dans les tissus humains (d'un facteur 2 à 3 depuis le milieu des années 1980) a été observée en population générale (IARC 2012).

Les expositions les plus fortes se sont probablement produites lors d'accidents industriels. Le plus connu est l'explosion d'une usine chimique à Seveso (Italie) en 1976. Cet incident a entraîné la libération d'une grande quantité de TCDD, pouvant atteindre selon certaines estimations 30 kg, soit 15 fois les émissions annuelles Française en 1990 (CITEPA 2017; Domenico et al. 1990).

#### 1.2.4. Dioxine et risque de cancer du sein

En raison de leur potentiel de cancérogénicité, de leur durée de demi-vie dans le corps humain, et de leurs forte lipophile, les dioxines sont suspectées d'être impliquées dans le développement du cancer du sein (Brody et al. 2007; Kogevinas 2001). Un certain nombre d'études ont étudié spécifiquement le lien entre exposition aux dioxines et risque de cancer du sein. Dans l'ensemble, ces études épidémiologiques n'ont pas établi de lien entre l'exposition à la dioxine et le risque de cancer du sein chez les femmes. A la suite de l'accident de l'usine SEVESO, les résultats issus de la cohorte, constitué de locaux vivant dans les zones contaminées, n'ont pas montré de lien avec le risque de cancer du sein (Bertazzi et al. 1993; Pesatori et al. 2009). De plus, aucune association n'a été observée entre les taux de TCDD mesurés dans les échantillons sanguins prélevés après l'accident et le risque de cancer du sein chez les femmes âgées de 0 à 40 ans au moment de l'accident et vivant dans les zones les plus contaminées (Warner et al. 2011). Des associations positives ont été observées chez des cohortes d'ouvrières d'usines d'herbicides et d'insecticides (Manuwald et al. 2012) et chez des femmes russes vivant près d'une usine chimique (Revich et al. 2001) ainsi que chez des femmes issues d'une étude écologique française sur les résidents exposés aux émissions de dioxines de plusieurs incinérateurs (Fabre et al. 2008). Viel et al (2008) ont quant à eux observé une diminution statistiquement significative du risque de cancer du sein chez les femmes de plus de 60 ans vivant dans la zone la plus exposée autour d'un incinérateur français (Viel et al. 2008). La plus récente méta-analyse, comprenant dix études réalisées sur l'association entre l'exposition externe à la TCDD (incluant expositions professionnelles, accidentelles et environnementales) et l'incidence de cancers (dont 3 études incluant le cancer du sein) et n'a pas rapporté de risque statistiquement significatif pour le risque tout cancer, ni pour le risque de cancer du sein (Xu et al. 2016).

## 1.3. Cadmium

### 1.3.1. Classification, mécanismes biologiques et propriétés

Le cadmium est un solide métallique naturellement présent dans l'écorce terrestre en faible quantité (0.1-0.2 mg/kg). Ces propriétés naturelles (résistance à la corrosion, faible point de fusion, forte conductivité thermique et électrique, ductilité, etc.) en font un métal très utilisé dans divers secteurs industriels (fabrication de pigments pour les verres, stabilisant dans le PVC, dans les accumulateurs nickel-cadmium, etc.) (IARC 2012). Il peut être dispersé dans l'environnement par entraînement de particules provenant du sol ou lors de phénomènes de combustions anthropique ou naturelles (éruptions volcaniques, incendies, etc.). Le CIRC a classé l'exposition au cadmium et ses composés comme cancérogènes certains pour l'Homme (Groupe 1) avec des preuves suffisantes pour le cancer du poumon et des preuves limitées pour le cancer de la prostate et du rein (Straif et al. 2009).

Un certain nombre de mécanismes biologiques ont été identifiés comme pouvant contribuer à la cancérogénicité du cadmium, notamment la perturbation des protéines réparatrices de l'ADN et des protéines de suppression tumorale (IARC 2012). Plusieurs études ont montré que le cadmium présente des effets œstrogéniques (Brama et al. 2007; Johnson et al. 2003; Nesatyy et al. 2006; Stoica et al. 2000). A faible dose, le cadmium possède des activités œstrogènes et androgènes-like in vitro et in vivo par liaison directe sur les récepteurs aux androgènes et aux œstrogènes (Takiguchi and Yoshihara 2006).

Le cadmium n'est pas essentiel au développement des organismes animaux ou végétaux et ne participe pas au métabolisme cellulaire. En revanche, ses propriétés physiques et chimiques, proches de celles du calcium, lui permettent de traverser les barrières biologiques et de s'accumuler dans les tissus. De par sa toxicité, le cadmium est classé dans la catégorie des métaux lourds et celles des éléments en trace métallique (ETM). L'accumulation s'effectue principalement dans les reins (organe 'cible' pour le cadmium), mais également dans le foie, le pancréas et la thyroïde avec un temps de demi-vie d'environ 20 ans (Järup and Akesson 2009).

### 1.3.2. Emissions et sources de cadmium

Les émissions atmosphériques de cadmium peuvent être d'origine anthropique ou naturelle. Environ 90% des rejets atmosphériques sont dus à l'activité humaine, en particulier lors des activités industrielles comme le raffinage des métaux non-ferreux (plomb, zinc et cuivre), la combustion du charbon et des produits pétroliers, l'incinération d'ordures ménagères et la métallurgie de l'acier. Les sources dominantes sont la production de zinc et l'incinération de déchets. Le cadmium, à l'instar de la plupart des métaux lourds présents dans l'atmosphère, se retrouve principalement incorporé dans les particules

atmosphériques. Contrairement à d'autres métaux lourds tels que le mercure, le cadmium est considéré comme inerte chimiquement, bien qu'il puisse se dissoudre en présence d'eau. Dans l'eau, le cadmium provient de l'érosion naturelle, du lessivage des sols ainsi que des décharges industrielles, du traitement des effluents industriels et des mines.

La Convention sur la Pollution Atmosphérique Transfrontalière à Longue Distance (1979) est le principal cadre de travail pour la protection de l'environnement atmosphérique à l'échelle Européenne. Le protocole d'Aarhus (Danemark), relatif aux métaux lourds, adopté le 24 juin 1998, vise spécifiquement le mercure, le plomb et le cadmium. Il a pour objectif de limiter les rejets d'origine anthropiques de ces composées dans l'atmosphère, grâce à la définition de valeurs limites d'émissions et la préconisation d'utilisation des meilleures techniques disponibles de production et de traitement des effluents.

Depuis 2003, les ICPE sont soumises à l'Arrêté du 24 décembre 2002 relatif à la déclaration annuelle des émissions polluantes. Pour le cadmium, cette déclaration est obligatoire à partir du seuil de 10 kg rejetés par an. Des arrêtés préfectoraux propres à chaque installation adaptent cette limite maximale en fonction des activités. Les déclarations d'émissions de polluants sont à disposition du grand public via le registre iREP.

De manière similaire aux dioxines, l'évolution des différentes législations encadrant les émissions des sources industrielles et l'amélioration des techniques d'épuration des fumées ont permis de faire progressivement diminuer les émissions de cadmium. D'après Olendrzynski et al, les émissions de cadmium en Europe représentaient environ 2400 tonnes en 1955 (et 998 tonnes en France) (Olendrzyński et al. 1996). D'après cette même étude, en 1995 à l'échelle européenne les émissions de cadmium avaient diminué de 58 % (et de 84 % pour la France, soit une émission annuelle de 25 tonnes). Le CITEPA fournit des estimations à l'échelle française depuis 1990 (Serveau et al. 2013). D'après cet inventaire, les émissions atmosphériques françaises de cadmium ont diminué de 87 % entre 1990 et 2015, passant de 20,5 tonnes à 2,7 tonnes (Figure 4). Les émissions liées aux sources industrielles ont qu'en a elles diminuées de 92 % sur la même période.

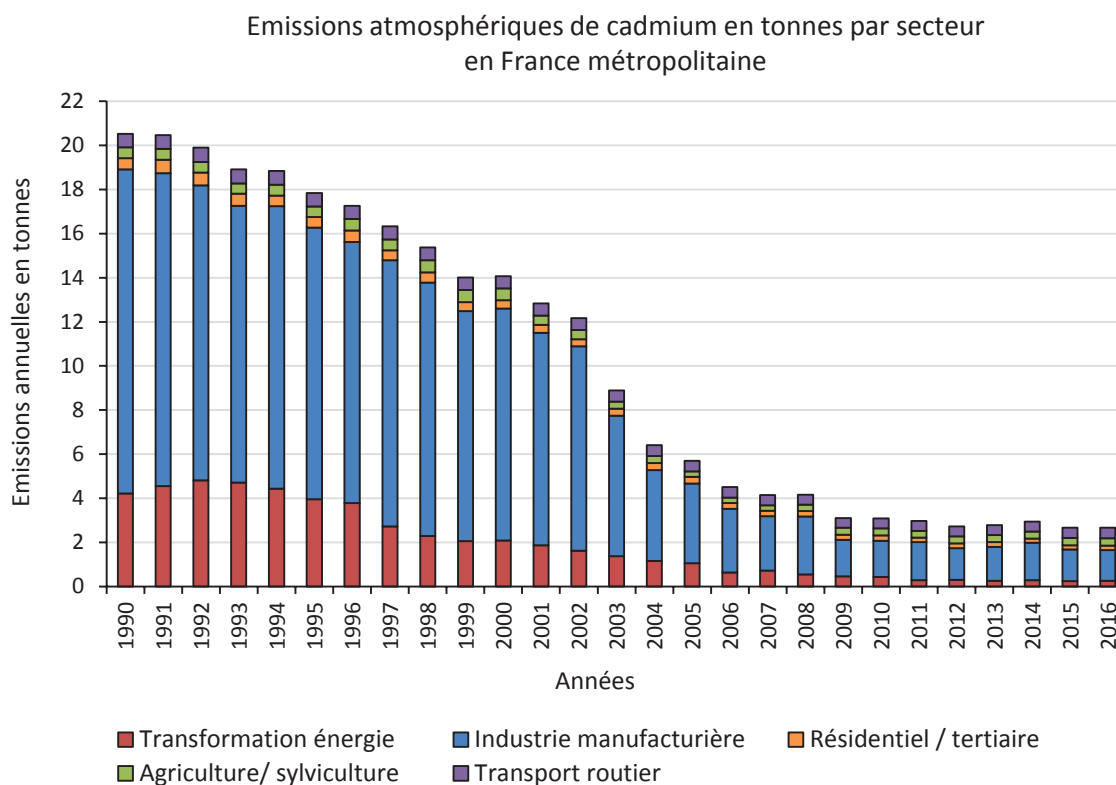


Figure 4 - Estimation des émissions annuelles de cadmium (en tonnes) en France entre 1990 et 2016 (CITEPA 2017)

### 1.3.3. Exposition humaine au cadmium

Le cadmium est présent dans les sols (0.1-0.2 mg/kg), dans l'air ambiant (0.1 ng/m<sup>3</sup>) ainsi que dans les eaux (5-100 ng/L) à de faibles concentrations, mais ces valeurs peuvent devenir plus importantes à proximité de zones industrielles (European Commission 2000). Ces concentrations de fond ne représentent pas une menace, mais la bioaccumulation dans les plantes (dépôt direct, absorption du sol, arrosage, etc.) et dans les animaux peut entraîner une augmentation des concentrations. L'inhalation de fumée de cigarette est une des principales sources d'exposition au cadmium pour la population générale, une cigarette contient en moyenne 2 µg de cadmium et on estime qu'environ 10 % est inhalé par les fumeurs (IARC 2012).

En population générale, chez les non-fumeurs, l'exposition se fait principalement par voie alimentaire, puis de manière secondaire via inhalation de cadmium dans l'air ambiant et ingestion d'eau contaminée. Les aliments tels que les céréales, les légumes et les crustacés sont particulièrement riches en cadmium (Järup and Akesson 2009; WHO 1992). On estime la dose journalière de cadmium via l'alimentation, pour un adulte, à environ 8,5 ng/kg de poids corporel (European Commission 2000).



Dans le cadre d'une exposition professionnelle, la principale voie d'exposition est l'inhalation de cadmium dans l'air (incluant les poussières contaminées). Les concentrations dans le sang (qui permettent de prendre en compte l'exposition cumulative) peuvent atteindre 50 µg/L pour les expositions professionnelles, contre, entre 1 et 4 µg/L pour les fumeurs, et entre 0,4 et 1,0 µg/L pour la population générale (IARC 2012d). De manière similaire aux dioxines, la classification du cadmium en tant que cancérigène certain pour l'Homme repose sur les résultats d'études de cohortes portant sur des populations exposées dans un cadre professionnel. Ces études ont établi un lien entre l'exposition au cadmium et la survenue de cancers pulmonaires (Kazantzis et al. 1988). Une étude menée en population générale, au voisinage de sites industriels, a confirmé un risque accru de cancer du poumon avec des concentrations urinaires en cadmium élevées (Nawrot et al. 2006). En se basant sur les concentrations observées en milieu professionnel dans le passé, l'OMS recommande une concentration moyenne d'exposition au cadmium en air ambiant limite de 5 ng/m<sup>3</sup> (WHO 2000).

#### 1.3.4. Cadmium et risque de cancer du sein

Plusieurs études ont observé une association positive entre exposition au cadmium et risque de cancer du sein (Julin et al. 2012; Liu et al. 2015), en particulier avec les cancers du sein avec des récepteurs aux œstrogènes positifs (Nagata et al. 2013).

Une récente méta-analyse a été publiée sur l'exposition alimentaire au cadmium et le risque de cancer du sein (Lin et al. 2016). Cette méta-analyse a repris les résultats de 6 études, incluant les études de Julien et al. (2012), et de Nagata et al. (2013), considérant le niveau de cadmium dans les urines (qui est considéré comme un biomarqueur de l'exposition vie entière (Lauwerys et al. 1994)), et a suggéré un effet significativement plus important sur le risque de cancer du sein chez les sujets ayant un niveau élevé de cadmium dans les urines. Cette augmentation du risque n'a pas été retrouvée dans la même méta-analyse considérant 6 autres études portant sur l'exposition alimentaire au cadmium et le risque de cancer du sein. Cependant, aucune des études disponibles n'a considéré l'évolution au cours du temps du profil alimentaire des femmes, ni de l'évolution de la contamination des aliments par le cadmium.

À notre connaissance, Liu et al constitue la seule étude menée à ce jour sur l'exposition atmosphérique au cadmium spécifiquement et le risque de cancer du sein chez la femme (Liu et al. 2015). Les auteurs ont estimé l'exposition atmosphérique historique au cadmium des femmes de la « California Teachers Study » en se basant sur des concentrations modélisées par l'US EPA. Dans le cas des expositions les élevées au cadmium, un risque accru de cancers du sein avec des récepteurs aux œstrogènes négatifs a été observé.

#### **1.4. Complexité de l'étude du lien entre exposition atmosphérique aux dioxines et au cadmium et risque de cancer du sein**

Un certain nombre d'études épidémiologiques récentes sur la pollution de l'air ont suggéré une association positive avec le risque de cancer du sein (Bonner et al. 2005; Chen and Bina 2012; Nie et al. 2007; Raaschou-Nielsen et al. 2011; Wei et al. 2012). Comme évoqué dans les sections 1.2.4 et 1.3.4, quelques études ont spécifiquement porté sur le risque de cancer du sein associé à une exposition atmosphérique aux dioxines et au cadmium (Fabre et al. 2008a; Liu et al. 2015; Viel et al. 2008). L'absence de consensus dans la littérature scientifique s'explique en partie par les difficultés et limites rencontrées dans ces études.

*Des expositions en population générale difficiles à quantifier, car chroniques, souvent faibles, et variables dans le temps et l'espace*

La preuve actuelle de la cancérogénicité de la TCDD et du cadmium repose principalement sur des études portant sur des populations exposées dans leur vie professionnelle ou lors d'accidents industriels (IARC 2012a, 2012d). Ces résultats ne peuvent pas être extrapolés à la population générale, qui est exposée de manière chronique à de faibles niveaux de dioxines et de cadmium. Néanmoins, il a été prouvé que certains perturbateurs endocriniens agissaient à faibles doses (Skinner et al. 2011; Vandenberg et al. 2012).

Dans le cas des dioxines et du cadmium, la grande variabilité des émissions au cours des dernières décennies requiert des informations précises sur les sources émettrices pour caractériser leurs émissions au cours du temps (Nzihou et al. 2012). En plus de la variabilité temporelle des émissions, une forte variabilité spatiale de concentrations en air ambiant est observée, due à la domination des sources industrielles qui créent de forts gradients de concentration autour des sources. La caractérisation des émissions de dioxines et de cadmium et l'évaluation des concentrations se heurtent également au très faible nombre de mesures historiques disponibles (à l'émission et en air ambiant), qui sont, de plus, réparties de manière hétérogène dans le temps et l'espace.

Lorsque l'on s'intéresse à des expositions faibles sur les périodes longues, classer correctement les sujets des études épidémiologiques devient complexe et demande la mise en place de méthodes spécifiques de quantification des expositions, pouvant fournir des estimations précises spatialement et avec une qualité constante sur la période d'étude. Ce point est abordé en détails dans la section 1.5.

### *Une grande importance du timing et de la durée d'exposition*

Les effets complexes du timing de l'exposition peuvent contribuer aux résultats non-concluants observés dans les études épidémiologiques. Une période de latence trop courte ne permet pas d'observer les effets de l'exposition sur la survenue de cancer.

Une exposition pendant le stade fœtal est susceptible d'engendrer des effets pathologiques à l'âge adulte. Ainsi, l'exposition précoce aux dioxines a été associée à un retard dans l'initiation du développement du sein chez les femmes (Den Hond et al. 2002), et en particulier chez les personnes vivant près d'incinérateurs de déchets et d'industries métallurgiques (Leijts et al. 2008, 2017). Les études expérimentales ont également démontré que l'exposition prénatale aux dioxines pouvait doubler l'incidence des tumeurs mammaires chez la souris plus tard dans la vie (Fenton and Birnbaum 2015). Les expositions ayant eu lieu dans l'enfance et/ou au cours des fenêtres de sensibilité critique du sein ont également été suggérées comme étant fortement associées à un risque de cancer du sein (Bonner et al. 2005; Cohn et al. 2007; Potischman and Troisi 1999).

Certains perturbateurs endocriniens comme les dioxines ou le cadmium se caractérisent par leurs propriétés de bioaccumulation et leurs persurances dans l'environnement. Le niveau individuel d'exposition est donc également lié à l'historique résidentiel de la population.

La prise en compte d'une exposition sur un temps long implique de maintenir une résolution temporelle fine dans l'évaluation des expositions sur la durée de la période d'étude. Plus le temps d'étude sera long, plus le manque de connaissances sur l'histoire résidentielle et l'exposition historique pourra induire une mauvaise classification des sujets entraînant une incertitude dans les risques estimés (Basagana et al. 2013; Brody et al. 2007; Cordioli et al. 2013).

### *Une méconnaissance des facteurs de risques individuels*

Une des limites majeures des études est l'absence de connaissances sur les facteurs de confusions. En effet, lorsqu'une étude épidémiologique souhaite étudier l'association entre un risque de développer une pathologie et une exposition, il est nécessaire de prendre en compte dans l'analyse des éléments extérieurs pouvant biaiser l'estimation de cette association. Dans le cas du cancer du sein, les facteurs de risques connus pouvant biaiser l'estimation des associations potentielle sont nombreux et variés : antécédents familiaux, facteurs hormonaux et reproductifs (nombre d'enfants, âge à la puberté, allaitement, traitement hormonaux de la ménopause, etc.), tabac, surpoids ou encore le travail de nuit. La prise en compte de ces paramètres semble atténuer les risques observés jusqu'ici (Binachon et al. 2014; Garcia et al. 2015; Raaschou-Nielsen et al. 2011). De plus, la plupart des études précédentes manquaient d'information pour stratifier sur le statut des récepteurs hormonaux du cancer du sein (ER et PR) alors qu'il existe des preuves que l'effet des dioxines et du cadmium peut différer selon le statut de ces récepteurs hormonaux (Garcia et al. 2015; Liu et al. 2015; Raaschou-Nielsen et al. 2005).

*Des expositions multiples souvent insuffisamment connues ou explorées.*

Alors qu'une grande variété de sources peuvent émettre des dioxines ou du cadmium (Lohmann and Jones 1998; Pacyna et al. 2001; Pronk et al. 2013), la majorité des études publiées se concentrent sur une seule source ou un seul type de sources (Ashworth et al. 2013; Cordioli et al. 2013a; Fabre et al. 2008c). Ce choix limite l'interprétation des résultats et peut entraîner une mauvaise classification de l'exposition (Basagana et al. 2013). Le manque de données empêche souvent de considérer des co-expositions (atmosphériques et alimentaires) ou des multi-expositions (plusieurs polluants). Hors, d'autres facteurs environnementaux sont suspectée d'augmenter le risque de cancer du sein (produits chimiques, polluants et substances radioactives). La population générale est donc exposée de façon concomitante à plusieurs facteurs environnementaux.

Les PCB ont été également classés comme cancérogènes pour l'Homme (Groupe 1) par le CIRC en 2013 et certaines associations positives statistiquement significatives ont été observées pour le cancer du sein (Aronson et al. 2000; Demers et al. 2002; Millikan et al. 2000; Moysich et al. 1998, 1999). Plusieurs autres polluants sont également suspectés d'augmenter le risque de cancer du sein, comme l'oxyde d'éthylène (Steenland et al. 2003), les phtalates (López-Carrillo et al. 2009, Chen and Chien 2014) ou encore les HAP (Gammon et al. 2002).

A la lumière de ces éléments, il apparaît que pour limiter au maximum les biais des études épidémiologiques portant sur le lien entre expositions atmosphériques aux dioxines et au cadmium et risque de cancer du sein, il conviendrait :

- D'étudier l'exposition d'une population large (pour obtenir une forte puissance statistique) sur une période longue (pour prendre en compte les temps de latence) ;
- De disposer d'informations précises sur les sujets inclus dans l'étude : exposition professionnelle et alimentaire, histoire résidentielle au cours de la vie, information sur les facteurs de risque (génétiques, hormonaux, reproductifs et comportementaux) et données sur le niveau socio-économique ;
- D'évaluer les expositions au cours de la période d'étude avec une résolution spatio-temporelle élevée et constante sur la période, en prenant en compte les différents types de sources d'émissions afin de ne pas sous-estimer les expositions, et si possible de prendre en compte les multi-expositions ;
- De comparer les résultats obtenus avec des mesures en air ambiant, des mesures à l'émission et des mesures biologiques pour ainsi valider les estimations des expositions ;

## 1.5. Les méthodes de quantification des expositions atmosphériques chroniques

Les émissions atmosphériques de dioxines et cadmium, observées en Europe, ont fortement diminué au cours des 30 dernières années. Le déclin industriel et les contraintes environnementales, de plus en plus fortes, en sont les causes principales. Cette variabilité des émissions est un défi majeur pour les études épidémiologiques qui doivent fournir une évaluation historique précise de l'exposition sur de longues périodes (par exemple depuis le recrutement d'une cohorte). La multiplicité des sources et le temps de latence entre l'exposition et l'apparition du cancer renforcent le besoin de caractériser la variabilité spatio-temporelle des expositions sur de grands espaces et pour de longues périodes.

Les études épidémiologiques publiées s'appuient sur une grande variété de données pour estimer les expositions, en utilisant notamment des données provenant de réseaux de surveillance de la qualité de l'air ou mais aussi d'« indicateur d'expositions » comme la densité du trafic routier autour des sujets ou la proximité résidentielle des sujets avec des sources industrielles. Cependant, de telles données sont souvent insuffisantes pour capter la variabilité spatiale des concentrations de polluants à l'échelle locale, en particulier à l'échelle intra-urbaine où la pollution atmosphérique est à la fois la plus élevée et la plus variable (Wilton et al. 2010). Des techniques plus complexes ont été utilisées dans des études récentes pour représenter de manière adéquate la variation spatiale des concentrations des polluants : des modèles de land use regression (LUR), des modèles de dispersion atmosphérique (gaussien ou chimie-transport), et des modèles systèmes d'information géographique (SIG) (de Hoogh et al. 2014). La méta-analyse de Xu et al. (2016) comprend ainsi (en excluant les études portant sur les expositions professionnelles) des études estimant les expositions externes via des modèles de dispersion (Floret et al. 2006; Viel et al. 2008; Zambon et al. 2007), des mesures de concentrations dans les sols (Consonni et al. 2008; Pesatori et al. 2009; Revich et al. 2001) ou plus simplement des questionnaires (Hardell 1993; Read et al. 2007).

### *Un nombre de données de mesures réduit*

Bien que des réseaux de surveillance de la pollution atmosphérique aient été mis en place en France dans les principales zones urbaines depuis près de 15 ans, il existe un manque de données de mesures des concentrations en air ambiant pour les périodes antérieures aux années 2000 ainsi que pour les communes de taille moyenne (moins de 100 000 habitants).

Dans le cas des dioxines et du cadmium, les mesures de concentration en air ambiant sont extrêmement parcellaires et sont difficilement utilisables pour estimer des expositions. De plus, ces mesures sont réparties inégalement dans le temps et dans l'espace avec un nombre de mesures croissant à partir des années 2000 alors que dans le même temps, les émissions baissent fortement. Le même constat s'applique pour les mesures réalisées à l'émission par les industriels. En effet, une seule base recense les émissions industrielles et depuis 2003 et pour les plus gros émetteurs seulement. De plus, ces mesures sont effectuées de manière ponctuelle (une fois par an dans les meilleurs cas).

En l'absence de données de mesure en air ambiant pour les dioxines ou le cadmium, les études épidémiologiques se sont basées sur des mesures de concentration dans les sols (moins chère à effectuer et donc plus nombreuses), ou sur des mesures biologiques (sang ou urines) pour estimer les expositions historiques (Dai and Oyana 2008a; Nagata et al. 2013; Nawrot et al. 2006). Étant donné la persistance dans l'environnement des dioxines et du cadmium, ces données de mesures dans les sols ou dans le corps humain ne renseignent pas sur la variabilité temporelle des concentrations, et témoignent principalement de l'exposition cumulative toutes voies d'expositions confondues (alimentaire, atmosphérique, cutanée). De plus, ces mesures sont très rarement disponibles sur des larges populations ou de grande zone d'étude. En revanche, ces mesures peuvent être utilisées pour renforcer le degré de validité des estimations réalisées avec d'autres méthodologies. Floret et al (2006), a ainsi comparé les estimations de concentrations de dioxines en air ambiant réalisées par un modèle gaussien avec des mesures dans les sols sur une zone d'étude réduite autour d'un incinérateur (Floret et al. 2006).

#### *Les modèles de Land Use Regression (LUR)*

Les modèles LUR utilisent des données géoréférencées (flux de trafic, densité de population, land use) pour construire des indicateurs de l'exposition aux polluants atmosphériques implémenté dans des SIG. Les LUR permettent de développer des modèles d'exposition paramétriques en associant ces données habituellement utilisées comme « indicateurs » des activités génératrices d'émission polluantes (intensité du trafic, taille des routes, ...) avec des mesures de concentration en air ambiant. Les expositions estimées sont principalement celles liées du trafic routier (Pedersen et al. 2013). La modélisation LUR s'est révélé être une approche robuste pour estimer les expositions à la pollution de l'air extérieur dans des zones avec de fortes variabilités spatiales des concentrations (Briggs et al. 2000; Hoek et al. 2008), ainsi que sur différentes échelles spatiales (Vienneau et al. 2013) et temporelles (Eeftens et al. 2012; Gulliver et al. 2013). Ces modèles demandent une faible capacité de calcul et sont calibrés en utilisant les données de mesures. Ils ne nécessitent pas d'inventaire détaillé des émissions ou de données météorologiques, contrairement aux modèles de dispersion atmosphérique.

Les modèles LUR sont couramment utilisés dans des études épidémiologiques portant sur la pollution atmosphérique. La plupart ont été développés pour couvrir des villes, ou des régions dans le cadre d'études de cohorte. Des modèles ont été développés dans l'étude ESCAPE pour Paris (étude sur les PM<sub>10</sub>, PM<sub>2.5</sub>, NO<sub>2</sub>), Grenoble, Lyon et Marseille (Pedersen et al. 2013). Il existe également des modèles pour les PM<sub>10</sub> et le NO<sub>2</sub> pour l'Europe de l'Ouest mais ceux-ci n'ont pas obtenus de bon résultats au niveau national (Vienneau et al. 2013).

Une des limites des méthodes LUR est le besoin important en termes de données de mesures (au minimum une trentaine de points pour une zone d'étude locale). Cette faiblesse s'accroît lors de l'étude d'expositions historiques. Pour pallier au manque de données de mesures dans le passé, des méthodes dites de « back-extrapolation » des modèles LUR ont été développées et ont obtenu des estimations valides sur des périodes jusqu'à 18 ans (Eeftens et al. 2012; Gulliver et al. 2013). Même si dans la plupart des situations, le nombre de sites de mesure dans le passé est moindre, la back extrapolation des modèles LUR à l'échelle nationale est possible en utilisant un nombre réduit de points de mesures, localisés en zones rurales ou reculées (Gulliver et al. 2013).

Dans le cas des dioxines et du cadmium, le manque de données de mesures en air ambiant combiné à la forte baisse des concentrations (Serveau et al. 2013) limite la possibilité d'utiliser des modèles LUR pour évaluer les expositions de dioxines et de cadmium. De plus, les émissions de dioxines et de cadmium sont historiquement dominées par les sources industrielles (sources ponctuelles), pour lesquelles, usuellement, les LUR sont peu ou pas utilisés car moins performants.

#### *Les modèles de dispersion atmosphérique gaussiens*

Les modèles de dispersion atmosphérique gaussiens décrivent la dispersion de polluants non-réactifs transportés par le vent et diffusés par la turbulence atmosphérique sur des zones allant jusqu'à quelques dizaines de kilomètres. L'approche gaussienne englobe les modèles de panache gaussien et les modèles à bouffées gaussiennes. La dispersion est calculée en fonction des conditions initiales d'émissions et des conditions météorologiques. Les conditions météorologiques définissent les classes de stabilité et donc les degrés de turbulences atmosphériques. Sur chaque pas de temps, on considère la météorologie stationnaire. Ces modèles demandent assez peu de ressources de calculs et sont priorisés en cas de nécessité de simulations rapides. Les champs de concentrations obtenus peuvent aller jusqu'à quelques mètres de résolution. Ils sont particulièrement utilisés pour le calcul de dispersion de rejets canalisés type cheminée industrielle. Comme les autres types de modèles de dispersion atmosphérique, ils sont limités par leurs besoins en données d'entrée : émissions des sources, paramètres techniques (hauteur, vitesse, débit) ou données météorologiques (température, nébulosité, direction du vent...). De plus les modèles gaussiens sont adaptés pour des études à l'échelle locale : à proximité des sources, sur des terrains non accidentés, et avec des vitesses de vent supérieures à 1m/s.

Un certain nombre d'études ont pu établir un lien entre exposition à des polluants atmosphériques et risque de cancer sein en se basant sur les résultats de modèles gaussiens (Fabre et al. 2008; Liu et al. 2015). Ils sont couramment utilisés dans les études épidémiologiques pour estimer les expositions aux dioxines et au cadmium (Fabre et al. 2008; Floret et al. 2006; Liu et al. 2015, 2015; Ranzi et al. 2011; Viel et al. 2008; Zambon et al. 2007). Ces études se concentrent le plus souvent sur des zones réduites en simulant les émissions d'une ou plusieurs sources industrielles.

Floret et al. (2003) a modélisé les concentrations de dioxines au niveau du sol, dues aux rejets d'un incinérateur en 1999, dans l'agglomération de Besançon (114 000 habitants), en utilisant un modèle gaussien. L'exposition de 222 cas de lymphome non-Hodgkinien a été comparée à celle de témoins sélectionnés aléatoirement. Les sujets étaient répartis en 4 groupes d'exposition. Les résultats ont montré que le risque de développer un lymphome non-Hodgkinien était 2.3 fois plus important chez les sujets vivant dans les zones les plus exposées. Par la suite, Floret et al. (2006) a mené une étude de validation. En l'absence de données de mesure en air ambiant, des mesures de concentrations dans le sol ont été utilisées. Les résultats obtenus étaient contradictoires, avec une association significative entre les concentrations modélisées et le logarithme des concentrations dans les sols au niveau de la zone Nord-Est de l'incinérateur (zone plate). En revanche, dans la zone sud (zone la plus exposée et terrain accidenté), les valeurs modélisées de concentrations surestimaient les concentrations dans les sols.

Dans la même zone, Viel et al. (2008) a étudié le lien entre les émissions de dioxines de l'incinérateur et le risque de cancer du sein. Les expositions étaient estimées en utilisant un modèle gaussien pour simuler les concentrations au niveau du sol. En comparant les expositions de 434 cas de cancer du sein avec 2170 témoins (réparti en 4 catégories d'exposition), les résultats n'ont pas montré d'augmentation du risque, mais une diminution du risque de cancer du sein pour les femmes de plus de 60 ans vivant dans la zone la plus exposée.

En modélisant la dispersion des panaches de fumée de 16 incinérateurs français entre 1990 et 1999, Fabre et al. (2008) a réalisé une étude écologique sur l'incidence des cancers dans une population répartie sur 4 départements français (135 123 cas de cancer pour une population d'environ 2.5 million de personnes). Un important travail a été mené dans la caractérisation des émissions des incinérateurs (recueil d'informations techniques, dates de fonctionnement et d'arrêt, mesures à l'émission, etc.). Comme indicateur de l'exposition, un mélange de dioxines, furanes et PCB a été estimé et modélisé avec le modèle gaussien ADMS (Atmospheric Dispersion Modelling System). Plusieurs relations positives statistiquement significatives ont été observées, notamment pour le risque tout cancer chez la femme, le risque de cancer du sein chez la femme, le risque de lymphome non hodgkiniens (femmes et hommes) et de myélomes multiples chez l'homme. Cependant, cette étude souffre de limites inhérentes aux études écologiques comme l'absence d'information sur les facteurs de risque individuels (l'exposition professionnelle, le tabagisme, la consommation d'alcool, des traitements médicaux, etc.) et la méconnaissance de l'histoire résidentielle. Ces limites ne permettent pas d'établir un lien de causalité entre l'exposition aux rejets des incinérateurs et l'incidence des cancers (Fabre et al. 2008).



Dans le cadre d'une étude cas-témoin incluant 172 cas de sarcome sur une population d'environ 400 000 sujets, Zambon et al. (2007) ont estimé l'exposition aux rejets atmosphérique de dioxines issues de sources industrielles (incinérateurs des déchets, incinérateurs hospitaliers, raffinerie, fonderie...) en utilisant le modèle de dispersion gaussien ISC3. L'étude portait sur Venise et son agglomération sur la période 1960-1996. Comme pour l'étude de Fabre et al (2008), les émissions ont été estimées grâce à un recueil historique d'informations techniques sur les sources industrielles. Les auteurs ont observé une augmentation significative du risque de sarcome pour les deux sexes, chez les sujets plus exposés.

Cordier et al. (2010) ont étudié le risque de malformations congénitales de l'appareil urinaire chez 304 sujets en Région Rhône-Alpes, liée à une exposition aux dioxines durant la grossesse entre janvier 2001 et décembre 2003. Dans cette étude, les dioxines étaient considérées comme un « traceur » de l'exposition aux fumées émis par les incinérateurs. L'exposition a été estimée en simulant la dispersion des émissions de 21 incinérateurs et en calculant les concentrations aux domiciles des sujets de l'étude, à l'aide du modèle gaussien ADMS. Les résultats ont montré une augmentation significative du risque de malformation congénitale chez les femmes exposées aux dioxines en début de grossesse. La prise en compte des facteurs de confusion dans un sous-groupe de sujets pour lesquels les facteurs de confusion (caractéristiques sociales et démographiques, antécédents obstétricaux, antécédents médicaux et familiaux, situation professionnel, lieux de résidence pendant la grossesse, consommation de tabac et d'alcool et consommation d'aliments cultivés localement) étaient disponibles et vivant dans les zones les moins exposées a entraîné une diminution de l'association, qui est devenue non-significative (Cordier et al. 2010).

Ranzi et al. (2011) a étudié l'exposition atmosphérique au cadmium d'une population de 31 347 sujets répartis dans un rayon de 3.5 km autour de 2 incinérateurs, entre 1990 et 2003, en utilisant le modèle de dispersion gaussien ISC3. Un certain nombre de polluants ont été modélisés (NO<sub>2</sub>, COV, CO, TSP, C<sub>6</sub>H<sub>6</sub>, HCl et métaux lourds). Les concentrations annuelles de métaux lourds ont été considérées comme des indicateurs de l'exposition aux rejets des incinérateurs, tandis que les concentrations de NO<sub>2</sub> étaient considérés comme représentatives de l'exposition aux autres sources prise en compte (trafic et chauffage). Les auteurs ont observé une augmentation significative du risque de cancer du sein pour les femmes et de risque du cancer de la plèvre pour les hommes, pour les sujets plus exposées aux métaux lourds.

En utilisant les concentrations de cadmium modélisées par l'US EPA en 2002 à l'échelle du « census tract » (équivalent de l'IRIS en France) comme proxy de l'exposition, Liu et al. (2015) ont observé un risque plus élevé de cancer du sein hormonaux négatifs chez les femmes les plus exposées. Ces résultats ont été observés sur la population de la cohorte « California Teachers Study » qui regroupe 112 379 femmes dont 5 361 cas de cancer du sein. Ces concentrations modélisées par l'US EPA, ont été comparées à des mesures en air ambiant. Dans le cas du cadmium, cette comparaison a montré une sous-estimation des concentrations par la modélisation en comparaison aux mesures avec un ratio entre les valeurs modélisées et mesurées inférieur à 0.2 (Garcia et al. 2014).

Hormis Liu et al. (2015), aucune des études citées ci-dessus n'a comparé les données modélisées avec les modèles de dispersion avec des mesures en air ambiant. De plus, dans le cas de dioxines, la majorité des études ont pris en compte uniquement les émissions provenant d'usines d'incinérations.

### *Les modèles de Chimie-Transport*

Les modèles de chimie-transport permettent de modéliser des concentrations sur des échelles régionales ou nationales. Les applications de ce type de modèles sont multiples : comprendre les processus physico-chimiques de l'atmosphère (Bessagnet et al. 2010; Hodzic et al. 2010), évaluer les scénarios de réduction des émissions (Kiesewetter et al. 2014), connaître les tendances passées et futures de la pollution à une échelle régionale. Ces modèles ont l'avantage de couvrir de larges territoires (nationaux voir continentaux) et permettent d'estimer les concentrations dues à des sources en dehors d'un domaine d'étude locale et indiquées plus génériquement comme « pollution de fond » (Soulhac et al. 2003). L'estimation de cette concentration représente une des problématiques clefs dans l'étude de l'exposition à la pollution atmosphérique.

Plusieurs simulations ont été réalisées en Europe au cours de la dernière décennie mais la résolution était grossière (environ 25 km) et l'évaluation reposait sur un ensemble limité de stations de mesure (Pay et al. 2012). La validation de modèles pour des résolutions plus fine (environ 10 km) a été uniquement réalisée sur des domaines limités (Flaounas et al. 2009; Hodzic et al. 2005; Tombette and Sportisse 2007). En 2015, Terrenoire et al, ont réalisé, avec le modèle Chimère une simulation pour une série de polluants (NO<sub>2</sub>, O<sub>3</sub>, PM<sub>10</sub>, PM<sub>2.5</sub>,) sur l'Europe à une résolution de 7 km (Terrenoire et al. 2015).

Pour les polluants comme le cadmium, ou les dioxines, seul Guzev et al. ont réalisé une simulation à l'échelle européenne des concentrations en dioxines, et PCBs mais à une résolution de 50 km (Gusev et al. 2016). Un modèle avec une telle résolution aurait une faible capacité discriminatoire des sujets selon expositions, particulièrement dans les zones urbaines, et ne serait pas adapté aux besoins des études épidémiologiques.

### *Les indicateurs basés sur des Systèmes d'information géographique (SIG)*

Depuis les années 2000, les SIG sont de plus en plus utilisés dans les études épidémiologiques. Ces outils permettent notamment de répondre à l'absence de données de mesure, en utilisant des données spatiales, comme la proximité résidentielle à différents types de sources d'émission (usines, routes, etc.) et sont ainsi intégrés comme un indicateur/proxy de l'exposition des sujets (Nuckols et al. 2004).

Ces indicateurs peuvent être basiques et prendre en compte uniquement la proximité géographique avec des sources, en considérant ou non les intensités d'émissions. Pour investiguer le risque associé à une exposition atmosphérique, un certain nombre d'études se sont basées principalement la distance sujet-source comme proxy de l'exposition (e.g. Langlois et al. 2009; Pronk et al. 2013; Visser et al. 2004). L'étude la plus complète sur l'exposition atmosphérique aux dioxines basée sur un indicateur SIG simple a été réalisée par Pronk et al. (2013). Cette étude menée aux Etats Unis a pris en compte la distance sujet-source et l'intensité des émissions des sources. Elle se distingue par l'utilisation d'un inventaire historique des émissions de dioxines (1987-2000) ainsi qu'une approche multi-sources (incinérateur, cimenterie, métallurgie, etc.). En étudiant l'exposition de 969 cas atteints de lymphome non-hodgkinien et 749 témoins, aucune augmentation du risque lié à l'exposition aux dioxines n'a été observée, mais un risque significativement plus élevé a été mis en évidence pour un sous-groupe de sujets vivant à proximité de cimenteries.

En 2013, Cordioli et al. (2013) ont réalisé une revue de la littérature sur les études épidémiologiques portant sur l'impact sanitaire des incinérateurs. Sur les 41 études recensées, 11 utilisaient des modèles de dispersion atmosphérique, 19 évaluaient les expositions en se basant sur la distance sujet-source et les 11 restantes utilisaient des indicateurs qualitatifs (présence/absence de sources à proximité des sujets). Pour étudier l'impact des différentes méthodologies sur l'estimation des expositions, Cordioli et al ont comparé les performances de ces indicateurs observés dans la littérature. Pour cela, ces différents indicateurs ont été appliqués à un même scénario : estimer l'exposition aux PM<sub>10</sub> des habitants de la ville de Parme résidant à moins de 4 km d'un incinérateur (n= 31000). Une valeur d'exposition de référence a été établie avec un modèle gaussien (ADMS) puis les estimations de chaque indicateur ont été comparées à cette référence. Les sujets étaient répartis en quintiles d'expositions. La concordance la plus faible a été observée pour un indicateur quantitatif (présence/absence de l'incinérateur dans un rayon de 800m) : seulement 25% des sujets étaient classés dans les mêmes catégories d'expositions qu'avec la référence (et 35% des sujets étaient classés avec deux catégories de différences par rapport à la référence). Les indicateurs utilisant la distance sujet/source obtenaient des meilleurs résultats avec 40% des sujets correctement classés et 15% de sujets classés avec plus d'une catégorie de différence avec la référence. Les auteurs ont donc montré que l'utilisation d'indicateurs SIG basiques induisait d'importantes erreurs de classification des sujets (en prenant comme référence un modèle gaussien). Les résultats de cette étude ont également souligné l'importance d'une localisation précise des sujets (taux de bonne classification diminuant de 25% si les sujets étaient positionnés à l'IRIS plutôt qu'à l'adresse).

Néanmoins, les outils SIG permettent également le développement d'indicateurs d'exposition plus complexes, applicable à des échelles spatio-temporelles fines et sur de zones géographiques larges pour estimer les expositions environnementales au niveau de l'adresse individuelle et en intégrant des paramètres influençant la dispersion des polluants comme la météorologiques ou la topographiques (Gulliver and Briggs 2011; Han et al. 2013; Hoek et al. 2001; White et al. 2009; Zou et al. 2009). Ainsi, les indicateurs développés dans ces études prennent en compte des paramètres comme la direction du vent (Yu et al. 2006), la direction et la vitesse du vent combinées (Gulliver and Briggs 2011; Vienneau et al. 2013; White et al. 2009) ou encore la pollution de fond (Hoek et al. 2001).

Ces indicateurs plus complexes, portent sur des polluants comme le NO<sub>2</sub> (Vienneau et al. 2009), les PM<sub>10</sub> (Gulliver and Briggs 2011) ou les fumées noires (Hoek et al. 2001), pour lesquels, à la différence des dioxines ou du cadmium, un grand nombre de mesures est disponible. Ces mesures permettent de calibrer les indicateurs et d'évaluer leurs performances (Gulliver and Briggs 2011; Vienneau et al. 2009). Ainsi, en comparant les estimations de leur indicateur SIG avec des concentrations annuelles de PM<sub>10</sub> à Londres, Gulliver and Briggs (2011) ont obtenu une corrélation R<sup>2</sup> de 0.67-0.77, et un résultat similaire en comparant leur estimation avec celle réalisée par un modèle gaussien (R<sup>2</sup>: 0.71-0.77). Vienneau et al (2009) ont obtenu un R<sup>2</sup> de 0.60, en comparant les estimations de leur indicateur SIG avec les concentrations de NO<sub>2</sub> en Europe. Néanmoins, les estimations avaient une précision spatiale limitée (1x1 km<sup>2</sup>) (Vienneau et al. 2009).

La taille de la zone d'application de ces indicateurs peut être variable (d'une ville à plusieurs pays) et selon les sources de polluants étudiées, la dimension de la zone d'exposition (buffer) prise en compte varie de plusieurs kilomètres pour les rejets industriels (Pronk et al. 2013; Vienneau et al. 2013; Yu et al. 2006) à quelques centaines de mètres pour les études sur les rejets liés au trafic (Gulliver and Briggs 2011)). Enfin, la précision de la localisation géographique des sujets et des sources est un élément clef pour minimiser les erreurs de classification liées au gradient spatial des concentrations (Faure et al. 2017; Han et al. 2013; Jacquemin et al. 2013).

Malgré les progrès réalisés dans le développement des modèles LUR et l'utilisation croissante de modèles déterministes (gaussien, chimie-transport, ...) considérés comme plus fiable (Cordioli et al. 2013), leur applicabilité pour l'étude de l'exposition aux dioxines et au cadmium, pour lesquels le nombre de mesures est extrêmement réduit dans le temps et dans l'espace et pour lesquels il n'existe pas d'inventaire exhaustif des émissions historiques, est limitée. En conséquence, l'utilisation des SIG ouvre une perspective pour la caractérisation des expositions atmosphériques à ces polluants dans les études épidémiologiques.

## 2. Objectifs et contexte de la thèse

### 2.1. Objectifs

Dans ce premier chapitre, j'ai détaillé les études portant sur l'exposition atmosphérique aux dioxines et au cadmium et le risque de cancer (cancer du sein inclus) ainsi que les différentes méthodologies d'évaluation des expositions mise en œuvre. Il apparaît important, dans l'optique d'étudier ce lien, de quantifier la variabilité spatio-temporelle des expositions avec une résolution fine sur de longues périodes.

Dans le cas des dioxines et du cadmium, il est nécessaire de prendre en compte de manière exhaustive les émissions atmosphérique des sources de dioxines et de cadmium au niveau national (Pronk et al. 2013). Cependant, le nombre de mesures à l'émission est extrêmement réduit et les inventaires de sources sont quasi inexistantes (ou limitées temporellement ainsi qu'en terme de représentativité des sources). Un travail d'inventaire est donc à mener avec rigueur, car il joue un rôle central dans l'évaluation des expositions, et ce quelles que soient les méthodologies utilisées. En l'absence de mesures, cet inventaire historique des émissions de dioxines et de cadmium pourra reposer sur un recueil historique complet des technologies mise en place par les industriels pour réduire leurs émissions, pour ensuite estimer les intensités d'émissions et leurs variations dans le temps (Cordier et al. 2010; Fabre et al. 2008; Ranzi et al. 2011; Zambon et al. 2007).

Au vu des données disponible, l'utilisation d'une métrique implémentée dans un SIG est l'approche la plus pertinente (Pronk et al. 2013). En revanche Il semble nécessaire, pour améliorer l'estimation des expositions, de prendre en compte les paramètres (météorologique, techniques, etc.) connus pour influencer la dispersion des polluants dans l'atmosphère (White et al. 2009; Zou et al. 2009). En effet, Cordioli et al. (2013) ont montré que les indicateurs SIG simplistes (par exemple basés uniquement sur la distance sujet-source) pouvaient générer des erreurs de classification des sujets. En l'absence de mesures en air ambiant, la calibration et la validation de cette métrique peuvent être réalisées en comparaison avec un modèle gaussien dans des zones restreinte comme montré par Gulliver et Briggs (Gulliver and Briggs 2011). Dans le cadre de cette thèse, cette métrique devra être applicable pour estimer les expositions d'une population répartie sur l'ensemble du territoire Français pour laquelle des données précises sont disponibles (exposition professionnelle, histoire résidentielle, information sur les facteurs de risque individuels et des données sur le niveau socio-économique) (Liu et al. 2015) en prenant en compte les expositions liées aux rejets de plusieurs milliers de sources industrielles sur une période de presque 20 ans. A défaut de pouvoir utiliser un modèle déterministe pour évaluer les expositions sur la France entière, un modèle gaussien peut permettre d'identifier les déterminants des expositions chroniques sur des zones réduites (Ranzi et al. 2011). Une comparaison avec des mesures de concentrations en air ambiant permettrait de quantifier la qualité des estimations (Liu et al. 2015).

Dans ce contexte, les objectifs de mon travail de recherche étaient d'étudier l'exposition atmosphérique aux dioxines et au cadmium, des femmes de la cohorte E3N (section 2.2.2) sur une période de 18 ans (1990-2008). Plus précisément, l'objectif de mon travail doctoral était de développer et valider une méthodologie d'évaluation de l'exposition chronique aux dioxines et au cadmium, due aux rejets de plusieurs milliers de sources industrielles sur le territoire Français. Elle devait être applicable dans le cadre d'études épidémiologiques comprenant une dizaine de milliers de sujets et témoins répartie sur le territoire Français. Par ailleurs, un des objectifs était également d'étudier les déterminants d'une exposition locale en milieu urbain via l'utilisation d'un modèle gaussien.

Les objectifs de recherche détaillés étaient les suivants :

- Développer une base de données recensant l'ensemble des sources industrielles (ou assimilées) émettrices de dioxines et de cadmium en France entre 1990 et 2008, recueillir les paramètres techniques sur le fonctionnement des sources, connaître les évolutions technologiques spécifiques de chaque source et leur taux d'activité, ainsi que d'estimer les émissions annuelles (CHAPITRE II). Les résultats obtenus ont servi de donnée d'entrée à la métrique d'exposition utilisée pour étudier la relation entre exposition aux dioxines et risque de cancer du sein.

- Etudier les déterminants de l'exposition chronique (1990-2008) aux dioxines et au cadmium au sein de la métropole de Lyon, ainsi que la variation de la contribution des différentes sources d'émission à l'exposition sur la période d'étude (CHAPITRE III).

- Développer, calibrer et valider une métrique d'exposition des émissions industrielles basée sur un système d'information géographique (SIG) applicable pour des études épidémiologiques évaluant les expositions atmosphérique individuelle aux dioxines et au cadmium chez les femmes de la cohorte E3N en France entre 1990 et 2008 (CHAPITRE IV).

À partir de ces objectifs de recherche, quatre articles scientifiques ont été produits, publiés ou soumis à des revues internationales (voir listes des publications). Le premier, présenté dans le CHAPITRE II, détaille les résultats de l'inventaire et de la caractérisation des sources émettrices de dioxines en France entre 1990 et 2008 en appliquant une méthode standardisée proposée par l'UNEP (article soumis, *Science of total environment*). Le deuxième article, présenté dans le CHAPITRE III, porte sur l'évaluation de l'exposition chronique aux dioxines et au cadmium dans la métropole de Lyon à l'aide d'une approche multi-sources et la comparaison des résultats obtenus avec différentes mesures (Coudon et al. 2018). Le troisième article, présenté au CHAPITRE IV, décrit la méthodologie SIG développée pour évaluer l'exposition individuelle aux dioxines et cadmium par des sources industrielles et sa validation avec un modèle gaussien (article accepté, *Environmental Health*). Enfin, un quatrième article, écrit par A. Danjou et présenté en annexe, détaille l'application de l'indicateur en région Rhône-Alpes ainsi que les résultats épidémiologiques observés.

## 2.2. Contexte

### 2.2.1. Département cancer et environnement

Le Centre Léon Bérard (CLB) à Lyon est un centre régional français de lutte contre le cancer (établissement de santé privé d'intérêt collectif – ESPIC), qui s'organise autour de trois axes principaux : les soins, la recherche et l'enseignement. Le CLB fait partie du groupe hospitalier Unicancer, exclusivement dédié à la lutte contre le cancer. En 2009, un programme interdisciplinaire portant sur les liens entre cancer, environnement et nutrition a été mis en place au sein du CLB. Ce programme a conduit à la création du département cancer et environnement qui a pour mission d'effectuer un travail de recherche sur deux thématiques majeures :

- L'étude des facteurs de risque environnementaux et professionnels associés au cancer du sein, au cancer des testicules et au cancer de l'ovaire ;
- L'évaluation et la gestion nutritionnelle des patients atteints de cancer du sein en surpoids ou à risque et l'introduction d'un programme d'activité physique adaptée à tous les patients cancéreux ;

Le département cancer et environnement (DCE) participe également à l'amélioration du diagnostic et à la déclaration des cancers professionnels grâce à des consultations médicales. Un site web a été développé (<http://cancer-environnement.fr>) pour faciliter la diffusion et à l'accès des professionnels de la santé, des patients et du public à des informations scientifiques basées sur les publications scientifiques les plus récentes. Le département est affilié à l'équipe «Signalisation des stéroïdes et tumeurs du sein» au Centre de recherche sur le cancer de Lyon (CRCL, UMR Inserm 1052 CNRS 5286, Centre Léon Bérard).

### 2.2.2. La cohorte E3N

La cohorte E3N est une étude prospective visant à identifier les causes environnementales des principales maladies chroniques chez les femmes, et en particulier du cancer (Clavel-Chapelon et al. 1997; Clavel-Chapelon 2015). La cohorte est conduite au sein de l'équipe "Mode de vie, gènes et santé" (CESP, Inserm U1018) à l'Institut Gustave Roussy. Elle a été dirigée de 1990 à 2015, par Françoise Clavel-Chapelon, puis par Marie-Christine Boutron-Ruault, et depuis 2017, par Gianluca Severi.

L'étude E3N a été lancée en 1989-1991 et comprend près de 100 000 femmes, nées entre 1925 et 1950. Ces femmes étaient assurées par la MGEN, un régime national d'assurance maladie couvrant principalement les enseignants.

Les participantes ont également signé un formulaire de consentement éclairé qui a permis de recueillir des informations supplémentaires sur les changements d'adresse et les remboursements de frais médicaux de la MGEN. Le protocole de l'étude E3N a été approuvé par la Commission nationale française des données informatisées et des libertés individuelles (CNIL, Commission Nationale Informatique et Libertés).

Depuis 1990, les participants ont complété des questionnaires tous les 2 à 3 ans sur des facteurs liés au mode de vie, comme la nutrition, l'activité physique, les caractéristiques anthropométriques, la prise d'un traitement hormonal, les facteurs de reproduction et les antécédents médicaux et familiaux. L'histoire résidentielle des femmes de la cohorte E3N a été reconstruite à l'aide des informations obtenues lors des différents questionnaires. Les adresses des domiciles des participantes ont été recueillies lors du premier questionnaire puis du cinquième au onzième questionnaire.

De 1994 à 1999, des échantillons de sang ont été recueillis chez près de 25 000 des femmes de la cohorte. Cette banque de matériel biologique a ensuite été complétée en 2009-2011 avec des échantillons de salive provenant d'environ 47 000 femmes qui n'avaient pas fourni d'échantillon de sang.

### 2.2.3. Equipe AIR (Laboratoire de Mécanique des fluides et d'acoustiques – Ecole centrale de Lyon)

L'équipe de recherche AIR (Atmosphere, Impact & Risk) est une équipe du Laboratoire de mécanique des fluides et d'acoustiques (LMFA) de l'Ecole Centrale de Lyon. Le LMFA axe ses activités de recherche autour de quatre thématiques : acoustique, fluides complexes et transferts, turbomachines et turbulence et stabilité.

L'équipe AIR, inclut dans le groupe fluides complexe et transferts, se concentre sur la modélisation de flux et la dispersion des polluants dans la couche limite atmosphérique. Cette équipe assure l'interface entre chercheurs universitaires et industriels. Les compétences de recherche couvrent les domaines de mécanique des fluides et de transport de polluants dans l'environnement.

Elle dispose d'un ensemble d'installations expérimentales :

- Une grande soufflerie pour l'étude de la dispersion atmosphérique sur les sites réels principalement utilisés pour des contrats de recherche avec le secteur privé ;
- Une soufflerie pour l'étude de la dispersion dans des configurations idéalisées, plus adaptée aux activités de recherche fondamentale sur la dispersion turbulente ;
- Un tunnel à petite échelle pour l'étude de la propagation des rejets de gaz légers, tels que les fumées chaudes, dans les tunnels ventilés ;



Au cours des dernières années, l'équipe AIR a réalisé différents types d'études allant de l'étude de la qualité de l'air et des pollutions chroniques en passant par l'évaluation de risque de pollution atmosphérique accidentelle. Elle collabore avec le département cancer et environnement du Centre Léon Bérard depuis 2013 dans le cadre du projet GEO3N puis XENAIR.

#### 2.2.4. Le modèle SIRANE

SIRANE est un modèle opérationnel de dispersion atmosphérique des polluants en milieu urbain, adapté à l'échelle d'une agglomération et développé par l'équipe AIR (Soulhac et al. 2012, 2017, 2011). Pour simuler la dispersion des polluants dans l'atmosphère urbaine, le modèle SIRANE décompose le domaine d'étude en deux parties : la canopée urbaine et la couche limite atmosphériques se développant au-dessus. La canopée urbaine est représentée comme un réseau de rue connectée (à titre d'exemple le réseau décrivant la canopée lyonnaise comprend environ 20 000 brins de rues). Pour simuler les polluants dans ce réseau de rues, le modèle se base sur 3 principaux phénomènes de transport des polluants dans la canopée urbaine : le confinement des polluants entre les bâtiments (rue-canyon), les échanges des polluants avec l'atmosphère extérieure et les échange des polluants au niveau des intersections de rues (Figure 5). Le phénomène de confinement peut jouer un rôle important dans les villes avec une forte densité de hauts bâtiments, qui peuvent piéger les polluants émis au niveau du sol (principalement liée au trafic) et qui vont s'accumuler au fil de temps en induisant une augmentation des concentrations au sol.

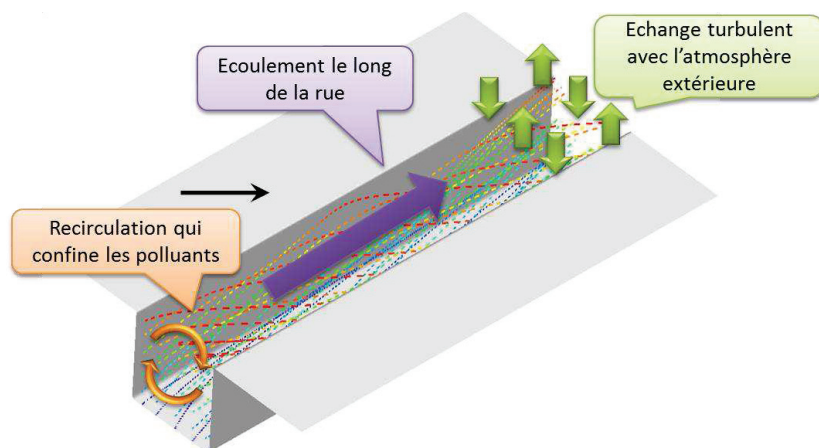


Figure 5 – Structure de l'écoulement dans une rue canyon

En plus de simuler la dispersion dans la canopée urbaine, SIRANE prend en compte le transport des polluants au-dessus du niveau des toits dans la couche limite atmosphérique. Les polluants présents au-dessus du niveau des toits proviennent soit de sources situées au-dessus des bâtiments (cheminées), soit de la diffusion turbulente, à travers l'interface rue-atmosphère, des polluants présents dans les rues. Le transport des polluants dans l'atmosphère extérieure est représenté à l'aide d'un modèle de panache gaussien. Chaque intersection et chaque rue deviennent alors la source d'un panache responsable d'une distribution de concentration. Selon les polluants étudiés, les phénomènes physico-chimiques comme le cycle de Chapman, les dépôts humides (lessivage) et secs peuvent être pris en compte.

SIRANE peut fournir des cartographies des concentrations avec un maillage fin (10m) ainsi que l'évolution temporelle des concentrations en certains récepteurs d'intérêt (stations de mesure ou lieux de résidence des sujets des études) (Figure 6).

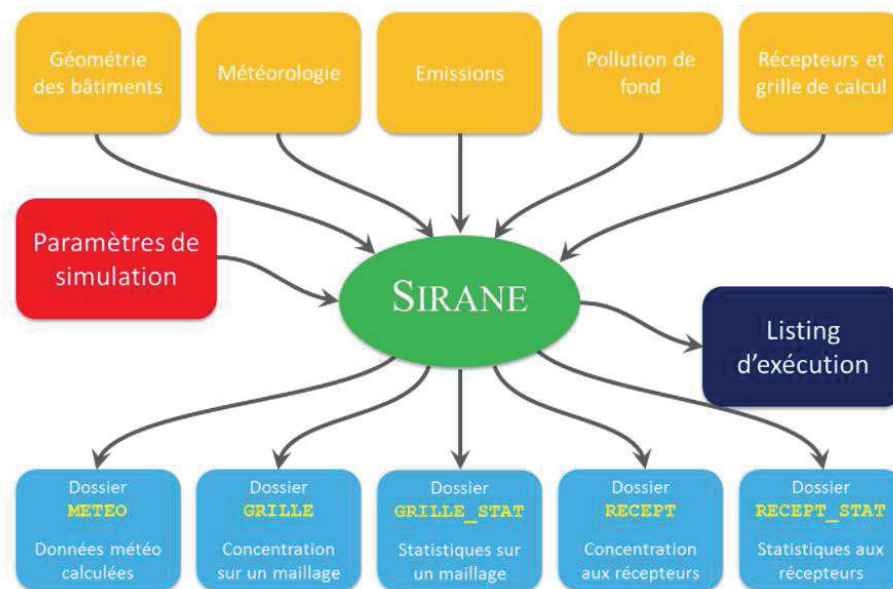


Figure 6 - Données d'entrée et de sortie de SIRANE

À ce jour, le modèle SIRANE a été validé au moyen d'expériences en soufflerie (Ben Salem et al. 2015; Carpentieri et al. 2012), par comparaison avec des simulations numériques détaillées et en utilisant des mesures de terrain pour plusieurs polluants ( $\text{NO}_2$  et  $\text{PM}_{10}$ ) (Soulhac et al. 2012, 2017). La validation la plus complète (Soulhac et al. 2017) a été réalisée sur la ville de Lyon en modélisant les concentrations de  $\text{NO}_2$  pour l'année 2008 et en comparant les résultats avec des stations de mesures continues et avec un grand nombre de mesures ponctuelles.

### 2.2.5. Projet GEO3N

En 2008, l'Institut national de la santé et de la recherche médicale (Inserm) et l'Agence française de l'alimentation, de l'environnement et de la santé et sécurité au travail (AFSSET) ont conclu dans un rapport à la nécessité de mener de nouveaux travaux de recherche sur les expositions environnementales et le risque de cancer du sein (Inserm and Afsset 2008).

En 2011, le DCE, avec la collaboration de l'Observatoire Régional de Santé (ORS) en Rhône-Alpes, a procédé à une revue de la littérature sur l'impact de la gestion des déchets ménagers et assimilés (collecte et triage, compostage, incinération et stockage) sur la santé des travailleurs et de la population locale (Anzivino-Viricel et al. 2012). Les résultats ont suggéré une augmentation possible de l'incidence du cancer du sein chez les femmes vivant près des installations, et donc une association potentielle entre l'apparition de cancer du sein et l'exposition aux fumées des incinérateurs de déchets ménagers avant la mise en place des normes européennes.

Dans ce contexte, le projet de recherche GEO3N a été mis en place par le DCE en collaboration avec l'équipe "Mode de vie, gènes et santé" (Inserm U1018, Institut Gustave Roussy). Il étudie l'hypothèse d'une association entre le risque de cancer du sein chez les femmes de la cohorte E3N et leur exposition aux dioxines.

Le projet a été construit pour surmonter les limites des études existantes, en utilisant :

- Des données précises sur les femmes recrutées sur tout le territoire national par la cohorte E3N: histoire résidentielle depuis 1990, principaux facteurs de risque de cancer (facteurs hormonaux, reproductifs et liés au mode de vie), statut des récepteurs hormonaux liés au cancer du sein et données alimentaires ;
- Un inventaire historique (1990-2008) des sources industrielles pour l'évaluation de l'exposition atmosphérique aux dioxines basée sur une approche multi-source (incinérateur, industries du métal, industries manufacturière, chimie, etc.) avec une connaissance de l'évolution des caractéristiques techniques des installations industrielles ;
- En l'absence de mesures d'émission, une méthodologie normalisée pour estimer les émissions de dioxines des installations, en utilisant le Toolkit pour l'identification et la quantification des rejets de dioxines, de furannes et autres POP du Programme des Nations Unies pour l'environnement (UNEP 2013) ;

- Le développement d'une métrique intégrée dans un SIG et utilisant des paramètres météorologiques et des paramètres techniques des sources de dioxines, pour évaluer l'exposition atmosphérique au niveau de l'adresse des femmes de la cohorte E3N entre 1990 et 2008 ;
- L'évaluation de l'exposition alimentaire aux dioxines, en combinant les données d'un questionnaire alimentaire envoyé aux femmes de la cohorte E3N avec des données sur la contamination des aliments par les dioxines ;

Le projet de recherche GEO3N comprend tous les cas de cancer du sein identifiés et validés dans la cohorte E3N entre 1990 et 2008 (N = 5 455). L'exposition professionnelle aux dioxines est considérée comme négligeable car les femmes de la cohorte E3N sont principalement des enseignantes et des travailleurs du secteur de l'éducation.

#### *L'équipe projet*

Le projet implique une collaboration pluridisciplinaire avec une expertise en oncologie, épidémiologie, géographie, sciences de l'environnement et biologie.

Les membres de l'équipe sont: Pr. Béatrice Fervers (MD, HDR, professeur associé), Laure Dossus (PhD, épidémiologiste, 2011-2015), Elodie Faure (géomaticienne, 2012-aujourd'hui), Aurélie Danjou (doctorante, épidémiologiste, 2013-2017), Thomas Coudon (doctorant en expologie environnementale, 2014-aujourd'hui), Lucie Anzivino-Viricel (PhD, épidémiologiste, 2012-2013), Guillaume Harel, Camille Denis et Hassan Hourani (ingénieurs en environnement 2014-2015, 2015-2016 et 2016-2017), Charlotte Carretero et Maxime Guillou (techniciens de géocodage, 2015-2016 et 2016-2017), Delphine Praud (PhD, post-doctorant épidémiologiste, 2016-aujourd'hui), et Xavier Morelli (post-doctorant PhD, épidémiologiste, 2016-2017).

#### *Les partenaires du projet*

- Françoise Clavel-Chapelon (DR, épidémiologiste) de l'équipe Équipe «Mode de vie, gènes et santé», Inserm U1018, Centre de recherche en épidémiologie et santé de la population (CESP), Université Paris-Sud;

- Pietro Salizzoni (MCU, ingénieur en mécanique des fluides) et Chi-Vuong Nguyen (docteur en mécanique des fluides), du Laboratoire de Mécanique et Acoustique des Fluides, UMR 5509, Ecole Centrale de Lyon

- Gianluca Severi (Directeur de Recherche Inserm) et Francesca Mancini (post-doctorante en épidémiologie nutritionnelle), de l'équipe "Mode de vie, gènes et santé" (CESP, Inserm U1018) à l'Institut Gustave Roussy

### *Le comité scientifique*

Un comité scientifique a été constitué pour suivre le projet dans son développement méthodologie. Les membres sont l'Agence Santé Publique France (anciennement Institut de Veille Sanitaire), le Bureau de Recherche Géologique et Minière (BRGM), le Centre National de la Sécurité Industrielle et de la Protection de l'Environnement (INERIS), le Centre Technique Interprofessionnel d'Etudes sur la Pollution Atmosphérique (CITEPA), l'Agence de l'Environnement et de la Maîtrise de l'Energie (ADEME), la Direction Régionale de l'Environnement, de l'Aménagement et du Logement Auvergne-Rhône-Alpes (DREAL RA), l'ANSES, l'Association Agréée Surveillance Qualité de l'Air (ATMO-AURA) et INSAValor.

### *Financement*

Le financement du projet comprend la subvention nationale de l'ADEME et la subvention régionale de la Ligue Contre le Cancer 71. L'étude menée dans la région Rhône-Alpes a été financée par le Programme Oncostarter du Cancéropôle Lyon Auvergne Rhône-Alpes (CLARA) et la Région Auvergne-Rhône-Alpes. Deux étudiants en doctorat (Aurélié Danjou et moi-même) ont reçu une bourse universitaire de l'Ecole doctorale interdisciplinaire science-santé (EDISS) de l'Université Lyon 1.

#### 2.2.6. Projet XENAIR

En parallèle et à la suite du projet GEO3N, le projet XENAIR vise à évaluer l'hypothèse d'un lien entre exposition atmosphérique chronique à faible dose aux polluants xénoestrogènes (ayant un effet œstrogénique sur les êtres vivants) et risque de cancer du sein dans la cohorte E3N. Ce projet comporte plusieurs objectifs spécifiques :

1. Estimer l'exposition des femmes de la cohorte E3N, entre 1990 et 2010, aux polluants suivants : PM<sub>10</sub>, PM<sub>2,5</sub>, NO<sub>2</sub>, O<sub>3</sub>, BaP, dioxines, PCB et cadmium. Cette évaluation de l'exposition sera basée sur des méthodologies complémentaires à différentes résolutions spatiales.

2. Etudier le rôle potentiel de la méthylation de l'ADN comme marqueur intermédiaire du risque dans l'association entre l'exposition aux polluants xénoestrogènes et risque de cancer du sein.

3. Estimer l'interaction entre l'exposition à long terme à de faibles doses de polluants atmosphériques et polymorphismes génétiques pour étudier l'hypothèse selon laquelle le risque de cancer du sein dépend de la susceptibilité génétique (polymorphismes) des femmes.

4. Si un lien est identifié, analyser la fraction de cancer du sein attribuable à la pollution atmosphérique en France et quantifier les coûts associés et les actions à mener en termes de santé publique

Ce projet de recherche interdisciplinaire comporte 6 équipes, avec une expertise en épidémiologie, évaluation de l'exposition, géographie et biostatistique :

- Département Cancer et Environnement, Centre Léon Bérard, Lyon - Responsable: Béatrice Fervers
- Gustave Roussy, Equipe Générations et santé - Inserm UMR 1018 (données de la cohorte E3N). Responsables: Gianluca Severi, Francesca Mancini
- Equipe AIR, LMFA, Ecole Centre Lyon, Ecully, France - Responsable: Pietro Salizzoni
- INERIS, Verneuil-en-Halatte, Oise, France – Responsable : Bertrand Bessagnet
- Imperial College of London, UK – Responsable : John Gulliver
- ISPED, Université de Bordeaux, France – Responsable : Karen Leffondré

#### *Financement*

Le projet XENAIR bénéficie d'un financement de l'appel à projet CANC' AIR de la Fondation ARC pour la recherche sur le cancer en 2015. Ce financement est réparti sur 4 ans (2016-2020).

Dans le cadre du projet XENAIR, j'ai reçu une bourse de mobilité du CLARA qui m'a permis de, me former sur les méthodes de Land Use Regression (LUR), dans l'équipe de John Gulliver (Impérial College London). Ces méthodes seront appliquées dans le projet XENAIR.



## **CHAPITRE II. EMISSIONS ATMOSPHERIQUES INDUSTRIELLES DE DIOXINE ET DE CADMIUM**

*Inventaire et Géolocalisation des sources et estimation des rejets atmosphériques de dioxines et cadmium entre 1990 et 2008*



## 1. Introduction

Les dioxines sont une famille de polluants chimiques ayant des effets délétères sur la santé, et pouvant perturber l'équilibre hormonal des êtres humains (Yang et al. 2015). Ces polluants sont produits de manière non-intentionnelle lors de processus de combustion et dans un certain nombre de secteurs industriels. Les dioxines font partis de polluants organiques persistants pour lesquels les nations signataires de la convention de Stockholm (2001) se sont engagées à réduire leurs émissions (Zhang et al. 2011). De cette convention est née la nécessité de mieux maîtriser et comprendre les différentes sources de dioxines d'origine anthropiques. Depuis cette date, de nombreux inventaires des émissions de dioxines ont été réalisés à travers le monde (Cheruiyot et al. 2016). A l'échelle européenne, plusieurs études se sont intéressées à quantifier les émissions historiques de dioxines (Pacyna et al. 2001; Quaß et al. 2004) ou les émissions à un instant t (Kok et al. 2005). Ces études avaient pour objectifs principaux de déterminer les principales sources de dioxines, d'identifier les tendances au niveau national et d'évaluer l'efficacité des politiques de réduction des émissions. D'une manière générale, les résultats ont montré une forte baisse des émissions de dioxines et particulièrement dans les secteurs identifiés comme fortement émetteurs. Ainsi, au niveau Français, le CITEPA estime que les émissions totales en France ont diminué de -94% entre 1990 et 2015 (CITEPA 2017). Il est à noter que cette baisse ne s'est pas produite à la même vitesse, ni avec la même intensité pour les différents secteurs émetteurs de dioxines.

Néanmoins, ces inventaires n'ont pas été construits dans un objectif d'évaluation des expositions et ne sont pas forcément adaptés pour des études épidémiologiques. Pour évaluer les effets des dioxines sur le risque de cancer, des inventaires exhaustifs des émissions doivent être développés avec une connaissance des émissions à l'échelle des sources et de leurs localisations précises (les études mentionnées ci-dessus se limitent à une échelle nationale). Malgré le nombre de sources considérées dans ces études, certaines sources locales importantes (comme les incinérateurs hospitaliers) peuvent être manquantes ou mal caractérisées (Pulles et al. 2006; Quass et al. 2004). De plus, les variations des émissions au cours du temps doivent être connues, car il est nécessaire de couvrir une large plage temporelle pour tenir compte du temps de latence entre l'exposition et le développement du cancer (Ries et al. 2006). Un inventaire incomplet aboutirait à une estimation biaisée de l'exposition et à une mauvaise classification des sujets lors des études épidémiologiques.

Toutes ces contraintes prennent une autre ampleur dans le cas des dioxines. Les mesures à l'émission sont coûteuses et sont, par conséquent, rares. Ils en résultent 3 problématiques : une couverture spatiale faible, des mesures effectuées de manière ponctuelle (1 fois par an dans le meilleur des cas), uniquement auprès des plus gros émetteurs, et une quasi-absence de mesures pour les années 1990-2000.

De plus, la diminution hétérogène des émissions selon les secteurs émetteurs sur les 20 dernières années empêche d'extrapoler les émissions à partir d'inventaires récents plus précis. A l'échelle individuelle, les émissions des sources industrielles de dioxines varient fortement selon la technologie mise en place. Hors, les précédents inventaires ont été construits en se basant sur des émissions estimées à partir de statistiques nationales et ne prennent donc pas en compte les sources de manière individuelle. Dans la période 1990-2005, des mesures ponctuelles sur certains incinérateurs ont illustrées cette variation des émissions à l'échelle individuelles, avec des valeurs mesurées pouvant dépassant plusieurs dizaines de fois les normes en vigueur.

### *Objectifs et méthodes*

Le but de ce travail est de construire une base de données incluant toutes les sources industrielles (ou assimilé) de dioxines sur le territoire français entre 1990 et 2008, ainsi que leurs localisations précises, leurs émissions en dioxines, les mesures éventuelles associées à l'émission et les caractéristiques de chaque source ainsi que leurs évolutions au cours du temps.

Il n'existe pas de base de données des sources de dioxines en France couvrant la période de 1990 à 2008. La première étape de cette thèse était donc la construction d'un inventaire des sources de dioxines, permettant de couvrir des types de sources plus large au-delà des seuls incinérateurs. Nous nous sommes basés sur la méthodologie d'inventaire proposé par l'UNEP via l'utilisation de son « Toolkit for Identifications and quantifications of dioxins and furans releases » (UNEP 2013). Le Toolkit distingue 10 groupes de sources de dioxines, chacun sous-divisé en plusieurs catégories. Pour chaque catégorie, en fonction des caractéristiques technique et procédés mis en place pour le traitement des fumées, des facteurs d'émission sont proposés. Les différents facteurs d'émissions disponibles reflètent l'impact des technologies d'épurations des fumées sur les émissions. Ces facteurs d'émission sont basés sur des données de la littérature et des résultats de mesures. Ils ont fait l'objet d'un consensus au niveau européen.

A partir du classement en catégories des activités potentiellement émettrices de dioxine, proposé par le Toolkit, nous avons, en utilisant plusieurs bases de données officielles, identifié les catégories de sources émettrices de dioxines en France. Nous avons inventorié les sources en activité sur la période 1990-2008 pour chaque catégorie. Puis, nous avons recueilli, pour chaque source, des données technologiques nécessaires à la classification ainsi qu'un certain nombre de métadonnées (hauteur de cheminée, vitesse d'éjection des fumées mesures, localisation géographique, taux d'activité) via un contact direct auprès des industriels ou en utilisant plusieurs bases de données.

La connaissance de la technologie à l'échelle individuelle pour chaque source recensée, ainsi que le changement technologique effectué au cours du temps entre 1990 et 2008, a permis d'estimer les émissions annuelles de chaque source. En combinant les données sur les taux d'activité des sources et les facteurs d'émission correspondants, nous avons estimé les rejets annuels en dioxine (g-TEQ/an). De plus, grâce aux informations recueillies, nous avons pu estimer les incertitudes sur les émissions liées à la qualité des données. Chaque sources a été localisée le plus précisément possible (cheminée, parcelle, rue, etc.).

Nous avons effectué ensuite des comparaisons à plusieurs niveaux, pour évaluer la qualité de nos estimations avec le Toolkit :

- au niveau des installations en utilisant les mesures disponibles (pour environ 2% des sources recensés) ;
- à l'échelle départementale par rapport aux estimations de l'inventaire national spatialisé de l'INERIS ;
- au niveau national, à partir des données de l'European Monitoring and Evaluation Programme (EMEP), nous avons comparé les tendances d'émissions sur la période 1990-2008 avec nos estimations ;

## 2. Article n°1

### National inventory of historical dioxin air emissions source in France

*Thomas Coudon<sup>1,2</sup>, Pietro Salizzon<sup>3</sup>, Delphine Praud<sup>1,4</sup>, Elodie Faure<sup>1</sup> Aurélie Danjou<sup>1,2</sup> and Beatrice Fervers<sup>1,2,4</sup>*

<sup>1</sup> Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France

<sup>2</sup> Université Claude Bernard Lyon 1, Villeurbanne 69100, France

<sup>3</sup> Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully 69134, France

<sup>4</sup> Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.

#### *Reference:*

Coudon T, Salizzoni P, Praud D, Faure E, Danjou AMN, Fervers B. 2018. National inventory of historical dioxin air emissions source in France. Soumis, Science of the Total Environment (avril 2018).

# A NATIONAL INVENTORY OF HISTORICAL DIOXIN AIR EMISSIONS SOURCES IN FRANCE

*T. Coudon<sup>1,2</sup>, P. Salizzoni<sup>3</sup>, D. Praud<sup>1,4</sup>, AMN. Danjou<sup>1</sup>, L. Dossus<sup>1,4,5</sup>, E. Faure<sup>1</sup>, B. Fervers<sup>1,2,6</sup>*

*<sup>1</sup>Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France*

*<sup>2</sup>Université Claude Bernard Lyon 1, Villeurbanne 69100, France*

*<sup>3</sup>Laboratoire de Mécanique des Fluides et d'Acoustique, UMR CNRS 5509, University of Lyon – École Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, 36, Avenue Guy de Collongue, 69134 Écully CEDEX - France.*

*<sup>4</sup>Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS UVSQ, Gustave Roussy, 114 rue Edouard-Vaillant, 94805 Villejuif Cedex, France*

*<sup>5</sup>Nutrition and Metabolism Section, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France*

*<sup>6</sup>INSERM 1052, CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, France.*

Keywords:

dioxins, emission, inventory, UNEP Toolkit, exposure assessment, endocrine disruptor

## Abstract

Polychlorinated dibenzodioxins (PCDD) and dibenzofurans (PCDF), commonly referred to as 'dioxins', are environmental and persistent organic pollutants with endocrine disrupting properties and are classified as human carcinogens. In France, since 1994, due to the reduction in industrial emissions, total dioxin emissions have decreased by more than 90%. To better characterise historical atmospheric dioxin exposure in an epidemiological study, we performed a comprehensive inventory of all dioxin-emitting sources between 1990 and 2008 at a fine spatial and temporal resolution, using regional and national databases and institutional information sources, according to the methodology of the "Standardized Toolkit for the Identification of Quantification of Dioxin and Furan Releases" by the United Nations Environment Program (UNEP).

Overall, we identified 2620 distinct dioxin-emitting sources over the French territory. The majority of the sources were geolocated either at the stack (43%) or at the building (27%). The final database contains the technical characteristics and activity rates for each source over the study period. These data served to estimate total annual dioxin emissions (g-TEQ/year) with the UNEP Toolkit.

We compared our estimates with both measurements (available only from 2003 onwards for 2.3% of the inventoried sources) and inventories at both departmental and national level. The discrepancies in the comparisons between the Toolkit estimates, the dioxin measurements available, and the inventories at the national and departmental level remained within the uncertainty range of the UNEP Toolkit estimates. The differences observed were due to a lack of data in particular for earlier years and to the large number of small incinerators with large dioxin emissions inventoried by our study. Due to better access to relevant data and a general decrease in emissions, the UNEP Toolkit estimates get closer to measurements and estimates of others inventories for more recent years.

We created the first database containing the annual atmospheric dioxin emissions for a large number of emission sources at the facility scale for an entire country. This database will be used to assess exposure to dioxins in an epidemiological study on breast cancer and data collected will allow the performance of emission estimates for other pollutants emitted by these sources to be assessed.

## Introduction

Polychlorinated dibenzodioxins (PCDD) and dibenzofurans (PCDF), commonly referred to as ‘dioxins’, are toxic chemical pollutants, with endocrine properties, produced unintentionally by the combustion of chlorinated products in industrial processes, such as waste incineration, metallurgy, etc., as well as by accidental fire events (e.g. forest fires). The 2,3,7,8-TCDD, the most toxic dioxin congener, is classified as a group 1 carcinogen by the International Agency for Research on Cancer (IARC) with a causal relationship for all cancer types combined and a positive association for lung cancer, soft tissue sarcoma and non-Hodgkin lymphoma (Baan et al. 2009). Evidence of its association with other cancers such as breast cancer remains inconsistent (Dai and Oyana 2008; Fabre et al. 2008; Cordioli et al. 2013; Huang and Batterman 2000; Pronk et al. 2013; Viel et al. 2008) and motivates further studies to investigate the effects of atmospheric dioxin exposure on human health.

The multiplicity of exposure sources and the latency between exposure and cancer occurrence represent major challenges in such studies and require an accurate characterisation of the spatial-temporal variability of exposures over large areas and long time periods. In numerous studies, the lack of past residential history and historical air pollutant exposure assessment at a fine spatial and temporal scale may have resulted in non-differential exposure misclassification and, hence, is likely to have contributed to imprecise risk estimates (Basagana et al. 2013; Brody et al. 2007; Cordioli et al. 2013).

As confirmed by the few available long-term measurements of atmospheric concentrations of dioxin (Augusto et al. 2015; Coutinho et al. 2015; Katsoyiannis et al. 2010), a strong decrease in industrial emissions over the last two decades has been observed in Western Europe, consequent to the implementation of specific abatement technologies. However, this overall decrease in industrial emissions varies significantly across countries and, most importantly, is dependent on the sector (waste incineration, metallurgy, etc.) (Fiedler 2015). While numerous facilities, including metal industries and cement kilns are likely to emit dioxins (Coudon et al. 2018; Lohmann and Jones 1998; Pacyna et al. 2001; Pronk et al. 2013), the majority of published studies have restricted exposure assessment to incinerators (Ashworth et al. 2013; Cordioli et al. 2013; Fabre et al. 2008b). This stresses the need to consider the exposure from multiple sources in epidemiological studies.

While the spatial coverage of ambient air quality monitoring networks has steadily increased in recent years, monitoring data for dioxins are extremely sparse, making it impossible to rely on the monitoring station networks to assess human exposure to atmospheric dioxin emissions.

To characterise dioxin exposure of study subjects (Cordioli et al. 2013), epidemiological studies used exposure indicators based on geographic information systems or deterministic models such as Gaussian dispersion models (Coudon et al. 2018, Liu et al. 2015, Ranzi et al. 2011). All these exposure assessment methodologies rely on an inventory of dioxin emissions, and their evolution over time, performed at a fine spatial and temporal scale and covering large periods, to take into account the latency of cancer development (Ries et al. 2006). Several historical inventories of industrial and non-industrial dioxin emissions have been performed in Europe (Kok et al. 2005; Pacyna et al. 2003; Quass et al. 2004; Quina et al. 2011; Saral et al. 2014) following the Stockholm Convention (Zhang et al. 2011). The aim of these studies was to inventory major dioxin-emitting sources, identify their temporal trends and assess the effectiveness of emission reduction policies at national or European level.

The inventories by Pacyna et al. (2003) and Quass et al. (2004) covered time periods of more than 10 years (1970-1995 and 1985-2005, respectively). These were built with data reported from each country and dioxin releases were estimated with emission factors from the literature (or based on national reported emissions and official reported activity data). They were an important starting point for the assessment of dioxin emissions, and their environmental and health impacts. They present however three major limitations. Firstly, these results are generally shown with a coarse spatial resolution and are mostly based on values averaged over each single nation. Yet, in epidemiological studies, detailed emission estimates by source at a fine temporal and spatial scale are important. Secondly, despite the number of sources covered by these inventories, some important local sources (like hospital waste incinerators) may be missing or are not fully characterised (Quass et al. 2004). Thirdly, the lack of detailed information implies large uncertainties in emission estimates, especially in the 1990s (Webster and Connett 1998). Even in more recent inventories, uncertainties over emissions were high, varying by about a factor of three (Pulles et al. 2006).

This study aims to build a comprehensive dioxin emission inventory, to include all relevant sources, for the French metropolitan territory over a long time period, i.e. between 1990 and 2008, with a high spatial resolution, (i.e. location of dioxin-emitting facilities at the stack) and at a fine temporal scale (i.e. annual dioxin emissions) to be used in an epidemiological study on dioxin exposure and risk of cancer. The objective is to provide annual estimates of dioxin emissions, associated uncertainties using historical information on technical abatement measures and activity rate for each dioxin source. Unlike previous studies, measurements of source emissions will be used to assess the accuracy of our estimates.



## Materials and Methods

The present inventory was based on the methodology proposed by the "Standardized Toolkit for the Identification of Quantification of Dioxin and Furan Releases" developed by the United Nations Environment Program (UNEP; <http://toolkit.pops.int/>). More particularly, our inventory used the update of the Toolkit, published in 2013 (UNEP 2013).

Furthermore, the application of the methodology of the UNEP Toolkit to the French context was submitted and approved by a national group of experts from the French Environment and Energy Management Agency (ADEME); the Regional Directorate for the Environment, Planning and Housing (DREAL) of the Rhône-Alpes Region; and industry associations.

The UNEP Toolkit methodology suggests a five-step inventory (UNEP 2013):

- Use of the UNEP Toolkit's list of activities likely to release PCDD/PCDF into the air to identify the main group of sources in the country
- Production of an inventory of all PCDD/PCDF sources active during the study period (1990-2008) and distribution of these source into the appropriate categories
- Collection of detailed information on technical process and activity characteristics of the sources
- Identification of the suitable emission factor for each source, using the extensive UNEP list of emission factors, according to the data collected
- Quantification of annual source emissions using emission factors

### **Identification of dioxin sources from 1990 to 2008**

The UNEP Toolkit provides a comprehensive list of dioxin sources divided into ten groups:

- Group 1: Waste Incineration, including seven categories of waste incinerator;
- Group 2: Ferrous and Non-Ferrous Metal Production, including twelve categories for the production of metals and metal alloys including recycling operations;
- Group 3: Heat and Power Generation, including five categories of large and small installations using fossil fuels, biomass or gas;
- Group 4: Production of Mineral Products, including seven categories of processes to manufacture mineral products;
- Group 5: Transport, including four categories (among which are road and ship transport);
- Group 6: Open Burning Processes, including two categories of burning biomass or waste

without technical equipment;

- Group 7: Production and Use of Chemicals and Consumer Goods, including eight categories of various industrial activities;
- Group 8: Miscellaneous Sources, including an array of five source categories that do not match the description of any other source group;
- Group 9: Disposal and Landfill, comprising five categories of waste disposal;
- Group 10: Contaminated Sites and Hotspots, including thirteen categories.

Since the study focused on air emissions, source groups 9 and 10, as well as some source categories of groups 7 and 8, were excluded due to the absence of air emission factors. Also, traffic emissions (group 5) were disregarded due to their weak contribution to local dioxin exposure (Coudon et al. 2018). Group 6, covering open burning processes, was omitted due to the impossibility of collecting consistent information to reliably characterise and locate emissions. Table 1 provides the detailed list of UNEP source groups and associated source categories inventoried in this study.

Relevant sources between 1990 and 2008 were identified through a combination of regional and national databases, in particular the Annual Emission Reporting on Pollutant and Greenhouse Gas Releases (GEREP), the French National Registry of Pollutant Emissions (IREP) and the Inventory of French Classified Facilities (ICPE). The databases are detailed in Table S1.

### **Collection of technological characteristics**

The information collected for the characterisation of the facilities includes the following items: name of the facility; name of the operator; address; geographical coordinates; start and end dates of the activity; annual activity rate; and technical characteristics (including process, incoming material, smoke treatment system, physical parameters of flue gases, etc.).

The information was collected through the above-mentioned databases (Table S1) and complemented by institutional information sources, contact with industrial unions and other nationally recognized associations. Furthermore, a structured questionnaire was sent to the facilities operating company or group to collect additional information on their technical characteristics and activity rate. Technological changes impacting dioxin releases were registered in the database on an annual basis.

## **Default air emission factors and emission estimates**

The Toolkit provides a comprehensive range of emission factors based on the international literature, experimental data, and governmental and institutional reports. Within each source category, the UNEP Toolkit distinguishes several classes and provides a default emission factor specific to each class. The default emission factors are expressed as TEQ (toxic equivalent) per unit of activity (e.g.,  $\mu\text{g-TEQ}$  per ton of material produced, per ton of fuel burned, etc.). Based on the information collected on technology, process characteristics, and technical abatement measures and operating practices, each source is assigned to one of the several classes and corresponding default emission factor for this class.

Classification and assignment of the corresponding default emission factor, performed for each calendar year, was carried out independently by two engineers blinded to each other's classification. Disagreement in classification was resolved by discussion to reach consensus.

The estimate of the annual dioxin emission for each source in grams TEQ per year was calculated as the product of the air emission factor assigned and the annual activity rate of the source.

## **Estimates for emission uncertainty**

A confidence level of the annual emission intensity was established by source based on the quality of the information on technical parameters and activity rate for each source. The quality of the information collected on technical parameters was classified as High (in-depth knowledge of technological parameters of interest), Medium (approximate knowledge of parameters of interest) or Low (parameters of interest are only known in a summary way). Similarly, the quality of the information on the activity rate for each source was classified as High (activity rate known), Medium (activity rate known for at least one year of the source) or Low (activity rate unknown: approximated using an average activity rate for similar sources). The overall confidence level of the annual emission intensity of each source was then derived by combining the confidence levels for the activity rate and the emission factor; this distinguished four categories (High, Medium+, Medium-, and Low) corresponding to the uncertainty in the annual emission estimates ( $\pm 30\%$ ;  $\pm 50\%$ ;  $\pm 75\%$  and  $\pm 100\%$ ) for each inventoried source (Coudon et al. 2018).

### **Specific methodology for medical waste incineration**

A specific process was put in place for medical waste incineration due to the large number of sources over the French territory between 1990 and 2008, the absence of a specific database and difficulties in obtaining reliable historical information on technical characteristics and activity rates for each medical waste incinerator. According to Tabeaud and Hamez (2000), in the early 1990s, 54% of the 2676 medical centres in France were equipped with an internal incinerator. These incinerators were small (< 6 t/h) and most of them had no air pollution control system and emission abatement (APC). Without any available information on the technologies, emissions from medical waste incinerators were estimated using a unique emission factor provided by the UNEP Toolkit for incinerators with “discontinuous operation” and “no or minimal APC”.

The tonnage of medical waste incineration was estimated based on the number of beds in the hospital. An average annual activity ratio of 0.38 t/bed was considered according to a study on medical waste incinerators in the French Rhône-Alpes Region (Vela 2010). Given these assumptions, the lowest confidence level was assigned to medical waste incineration emissions.

Based on the French decree of January 1991 (restricting waste from incineration plants with a nominal capacity < 6t/h), small incinerators in hospital centres, local hospitals and dialysis centres were considered as inactive by the end of 1995.

### **Geolocation of dioxin-emitting sources**

All identified dioxin-emitting facilities were manually geolocated using ARCGIS software (ESRI, ArcMap 10.1). Whenever possible, facilities were positioned at the chimney stack using aerial photography. When several stacks were identified on the aerial photograph, the facility was positioned at equal distance from each of the chimneys. When no chimney could be identified on the aerial photograph, the facility was positioned at the centroid of the building. For facilities that had been closed for several years, we used historical aerial photography available from the Geographical National Institute (<https://remonterletemps.ign.fr/>). When neither a chimney nor the building could be identified, facilities were positioned, in descending order of precision, at the centroid of the parcel, the town hall of the municipality of the facility.

For hospital waste incinerators, a specific geolocation by the French Ministry of Health was used (<https://www.atlasante.fr/accueil>) locating these facilities at the address plate or street segment. The precision of the geolocation was specified in the source database and has been directly integrated in our GIS.

## **Comparison with emission measurements and other emission inventories**

The emission estimates provided by the UNEP Toolkit were compared with the dioxin declaration from the IREP database, available since 2003. Since 2003, the IREP registers mandatory declarations of facilities exceeding the reporting threshold of annual dioxin emissions of 0.1 g-TEQ/year. We compared each registration in the IREP database with the UNEP Toolkit estimate for the corresponding source and activity.

Furthermore, we compared the emission estimates provided by the UNEP Toolkit with existing emission inventories available at departmental and national level.

At departmental level, we compared our estimates to the data of the French National Spatial Inventory (INS) from the French National Institute for Industrial Environment and Risk (INERIS). Dioxin emissions were estimated as the product of the activity rate and a specific emission factor (established for French emissions according to measurement data). These data were publicly available only for 2004, 2007 and 2012 (<http://emissions-air.developpement-durable.gouv.fr>).

At the national level, we compared our estimates with the annual emission inventory for France by the European Monitoring and Evaluation Program (EMEP; [www.ceip.at](http://www.ceip.at)). Since they used a different source classification than the UNEP Toolkit, we could only compare estimates for a limited number of source groups and categories matching the UNEP classification, i.e. group 2 (ferrous and non-ferrous metal production), group 4 (production of mineral products) and group 8 (miscellaneous), as well as categories 1.b (hazardous waste incineration) and 1.c (medical waste incineration).

The EMEP distinguishes municipal solid waste incineration (MSWI) with heat recovery (classified in the heat production sector) and MSWI without heat recovery (classified in the incineration sector). As we were unable to perform this distinction due to lack of data, we compared the sum of the EMEP emissions estimates for MSWI with heat recovery, and MSWI without heat recovery and heat production to the sum of the UNEP Toolkit estimates for category 1.a (municipal solid waste incineration) and group 3 (heat and power generation).

## Results

### National inventory of dioxin sources from 1990 to 2008

Overall, 2620 dioxin sources were identified over the French territory between 1990 and 2008 (Table 1). Waste incineration (n=1913), and ferrous and non-ferrous metal production (n=241) represent, respectively, 73% and 9% of the sources inventoried. The remaining 18% correspond to heat and power generation (6%, n=165), production of mineral products (6%, n=156), production and use of chemicals and consumers goods (1%, n=13), and miscellaneous sources (5%, n=131). Medical waste incineration represents the most frequent source category (category 1.c, n=1442), followed by municipal waste incineration (1.a, n=377), fossil fuel plants (3.a, n=147) and crematoria (8.b, n=131). The categories of iron and steel production (2.ci, n=46; 2.c.ii, n=60), and aluminium production (2.e, n=54) constitute further categories comprising numerous sources. Setting aside medical waste incinerators, the total annual number of dioxin sources only slightly decreased over time from 905 in 1996 to 853 in 2008 (Table S2). Of the 2620 sources identified, only 920 (35%) were active in 2008.

The overall response rate to the questionnaire sent to facilities companies or operating groups was 32% despite mail and phone reminders. The highest response rates were observed for cement production (4.a, 100%); ferrous and non-ferrous metal production (80%); lime production (4.b, 80%); sewage sludge incineration (1.e, 78%); and hazardous waste incineration (1.b, 70%). It should be noted that sole confirmation of information through phone contact with operators was not accounted for in the response rate. Some companies explicitly refused to respond. The main reasons for refusal were lack of time, information being difficult to access or lost, in particular for facilities which were no longer active, and mistrust in the study.

### Geolocations of the inventoried sources

Figure 1 shows the geolocations of the inventoried sources. The location accuracy of the hospital waste incinerators (n=1444) is as follows: 62% are located at the address plate, 18% at the town hall, 18% at the track and 2% nearby. Of the remaining sources types (n=1176), 82% are located at the stack, 12% at the centroid of the building, 5% at the address plate and 1% at the town hall.

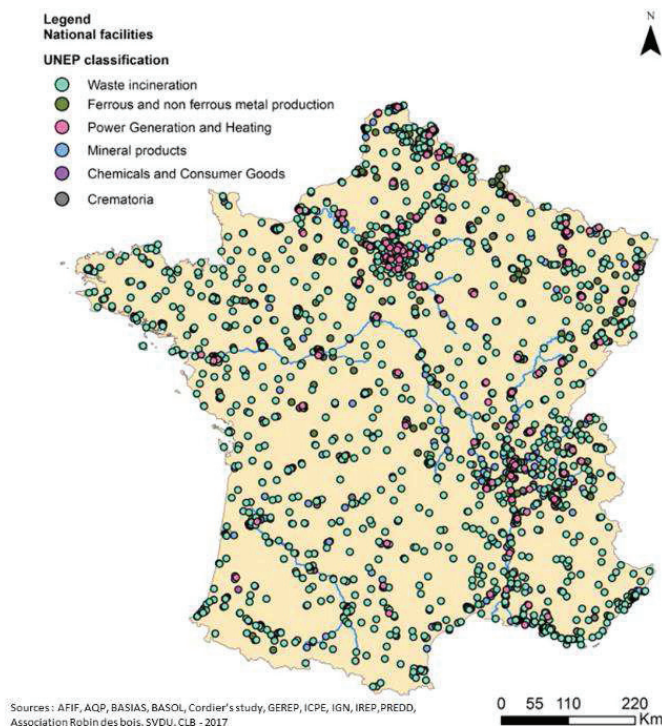


Figure 1 – Geolocations of the dioxin sources between 1990 and 2008 in France

### Emission intensity estimates

The distribution of emission intensities by the UNEP Toolkit group and activity category is shown in Table 1. The median annual emission per source for waste incineration (0.402 g-TEQ/year) largely exceeds that of other groups, i.e. ferrous and non-ferrous metal production (0.017 g-TEQ/year), heat and power generation (0.001 g-TEQ/year), production of mineral products (0.008 g-TEQ/year), production and use of chemicals and consumer goods (0.017 g-TEQ/year) and miscellaneous sources (0.008 g-TEQ/year).

Municipal waste incineration (1.a) and iron ore sintering (2.a) are the categories with the most important median intensity per source of 3.85g TEQ/year and 1.3 g-TEQ/year, respectively. The median annual emissions of two additional categories are above 0.1 g-TEQ/year: medical waste incineration (1.c; 0.323 g-TEQ/year) and cement production (4.a; 0.147 g-TEQ/year). A further six categories (Pulp and paper production; Hazardous waste incineration; Iron and steel making; Aluminium production; Biomass power plants and Waste wood and biomass incineration) display median emissions close to 0.5 g-TEQ/year (Table 1).

The estimated uncertainties of dioxin emission estimates vary across categories ranging from  $\pm 30\%$  (1.g, 2.a, 3.b, 7.a and 7.f) to  $\pm 100\%$  (1.c, 4.f and 8.b) (Table 1). At the national level, we observe an average uncertainty of  $\pm 62\%$  in the total emission estimates. Setting aside hospital waste incinerators (1.c, n=1442), characterised by the lowest confidence level, the uncertainty in the total emission estimates is  $\pm 53\%$ .

Table 1 – Median annual intensity of emissions (g-TEQ/year) by UNEP Toolkit group and category

Group	Nb of sources	Median (g-TEQ/year)	IQR	Uncertainty (%)	
1	Waste incineration	1913	0.402	0.180 – 0.981	96%
2	Metal production	241	0.017	0.002 - 0.105	53%
3	Heat and power generation	165	0.001	0.0002 - 0.007	40%
4	Production of mineral products	156	0.008	0.001-0.051	63%
7	Chemicals and consumer goods	14	0.048	0.009-0.080	29%
8	Miscellaneous	131	0.008	0.006-0.014	100%
Category	Nb of sources	Median (g-TEQ/year)	IQR	Uncertainty (%)	
1.a	Municipal solid waste incineration	377	3.850	1.750-8.467	49%
1.b	Hazardous waste incineration	67	0.055	0.010-0.229	63%
1.c	Medical waste incineration	1442	0.323	0.174-0.630	100%
1.e	Sewage sludge incineration	21	0.018	0.0040.029	45%
1.f	Waste wood and biomass incineration	1	0.04	NA	75%
1.g	Destruction of animal carcasses	5	0.011	0.045-0.146	30%
2.a	Iron ore sintering	7	1.31	0.603-6.000	30%
2.b	Coke production	5	0.018	0.010-0.042	75%
2.c.i	Iron and Steel Making	46	0.05	0.010-0.419	51%
2.c.ii	Iron Foundries	60	0.009	0.002-0.026	56%
2.c.iii	Hot-dip Galvanizing Plants	9	0.008	0.001-0.013	32%
2.d	Copper production	23	0.001	0.0001-0.056	75%
2.e	Aluminium production	54	0.05	0.014-0.179	59%
2.f	Lead production	14	0.01	0.002-0.020	42%
2.g	Zinc production	7	0.002	0.001-0.020	50%
2.h	Brass and bronze production	2	0.021	0.011-0.032	50%
2.j	Others non-ferrous metal production	14	0.003	0.002-0.022	36%
3.a	Fossil fuel power plant	147	0.001	0.0002-0.005	37%
3.b	Biomass power plants	9	0.05	0.028-0.059	30%
3.c	Landfill Biogas combustion	9	0.001	0.001-0.001	75%
4.a	Cement production	39	0.147	0.043-0.285	58%
4.b	Lime production	36	0.008	0.004-0.031	66%
4.c	Brick production	5	0.003	0.002-0.003	32%
4.d	Glass production	30	0.029	0.006-0.040	37%
4.f	Asphalt mixing	46	0.001	0.0001-0.001	100%
7.a	Pulp and paper production	13	0.075	0.012-0.080	30%
7.f	Chlorinated inorganic chemicals	1	0.007	NA	30%
8.b	Crematoria	131	0.008	0.006-0.014	100%

### Emission trends (1990-2008)

Results show a 98.4% decrease in total dioxin emissions between 1990 and 2008 (Figure 2). Waste incineration (UNEP Toolkit Group 1) constitutes the main contributor to total annual dioxin emissions until 2005, representing 95%, 85% and 75% of the total emissions respectively in 1990, 2000 and 2005. Over the last three years of the study period, 2006 - 2008, emissions from ferrous and non-ferrous metal production (Group 2) instead are dominant. Emissions from waste incineration decreased by 99.5% over the whole period and two sharp decreases are observed in 1996 and 2005, subsequent to changes of regulatory standards in January 1991 and September 2002 involving the implementation of new abatement technologies for incinerators and the closing of hospital waste incinerators (Table



S3). A similar decrease is observed over the whole period for ferrous and non-ferrous metal production, and for the production of mineral products (57% and 74%, respectively) (Table S4). The trend is different for the production and use of chemicals and consumer goods, and miscellaneous groups, whose emissions increased over the entire period by 31% and 100%, respectively (Table S5). However, the annual emissions of these two groups are low (1.27 and 1.26 g-TEQ/year, respectively, in 2008), representing less than 3% each of total emissions. Emissions from the heat and power generation group vary slightly over the study period (-8%) and remain negligible at 4% of total emissions. Overall uncertainties in emissions decrease over the years, from 67% in 1990 to 49% in 2008 (Table S2). Total uncertainties in emissions represent almost 3 kg-TEQ/year in 1990, whereas in 2008, they represented 35 g-TEQ/year.

Except for Iron and steel making (category 2.c.i), all categories with emissions above 10 g-TEQ/year in 1990 (1.a, 1.b, 1.c, 2.a, 1.e and 4.a), strongly decrease over the 18-year period by between 57% and 99% (Tables S3, S4 and S5). We observe an increasing trend in emissions for 8 categories out of 28. For crematoria (category 8.b), this increase in emissions is due to an almost three-fold rise of the number of crematoria from 41 in 1990 to 120 in 2008. For other categories, displaying a reduced number of sources ( $n < 10$ ) and overall weak emissions, an increased sensitivity to emission variations of a single source is observed. As for uncertainties in group emissions, we observe a decrease in the uncertainties by category (Tables S3, S4 and S5). Except for hospital waste incineration, the emission estimates for the 8 categories with a median emission  $\geq 0.05$  g-TEQ/year over the study period were associated with relative low uncertainties, ranging from 30% to 63%.

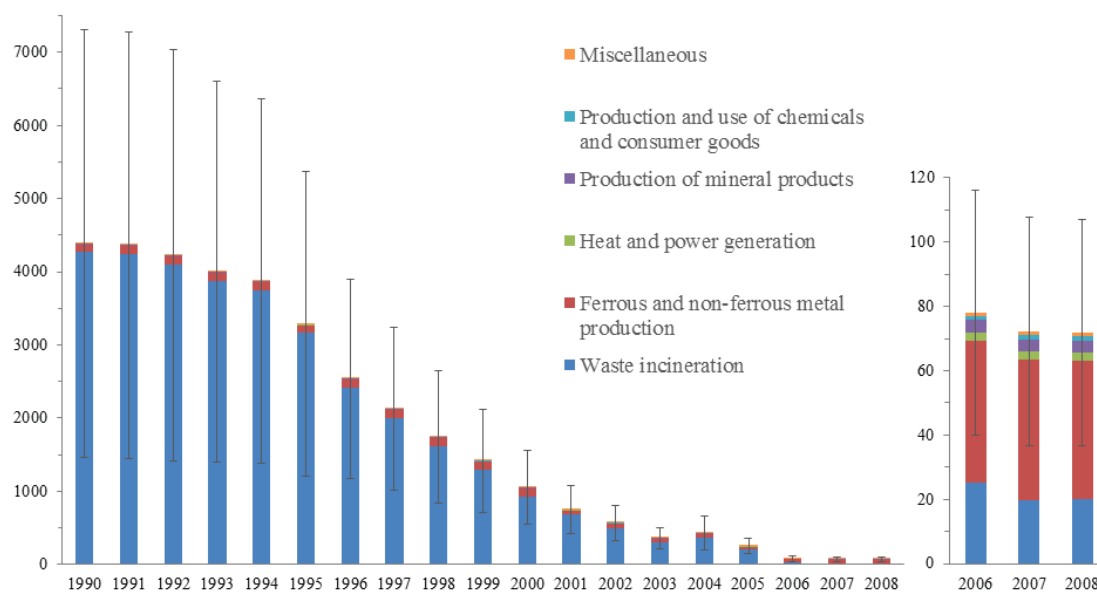


Figure 2 - Trends in dioxin emission estimates (g-TEQ/year) by UNEP group in France from 1990 to

2008

The decrease in emissions varies spatially (Figure 3). In 1990, we observe high spatial differences in emissions over the French territory, with the highest emissions in the North and strong emissions in the Rhône-Alpes region. While in 1990, 33 French departments (34%) have an annual emission above 50 g-TEQ (including 4 departments with emissions above 150 g-TEQ/year), there are only 12 (13%) with an annual emission above 50g-TEQ in 1996 (including two departments with emissions above 150 g-TEQ/year) and the large majority of the French departments have emissions between 0 and 25 g-TEQ/year. In 2002, the emissions of the majority of the departments are below 5 g-TEQ/year. Some high emission values remain in the most populated areas (Paris, Lyon and Bordeaux), often due to a single major source. In 2008, almost all French departments have annual emissions below 5 g-TEQ/year (Figure 3).

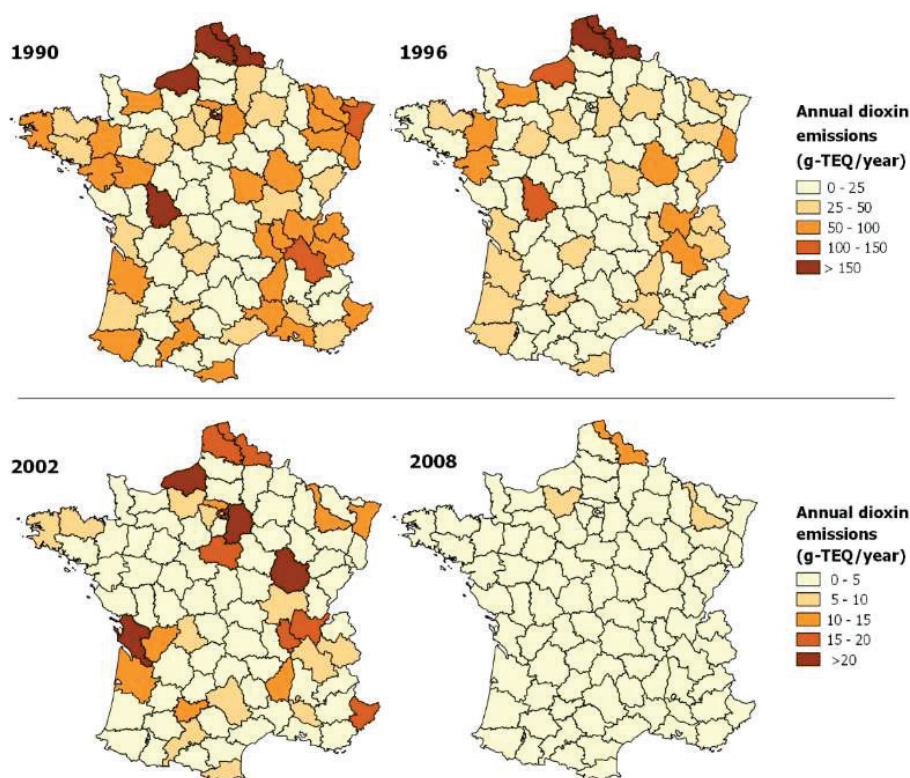


Figure 3 - UNEP Toolkit estimates of dioxin emissions by department in France from 1990 to 2008

### Comparison of the estimates with available measurements (2003-2008)

We compared annual UNEP Toolkit emission estimates and dioxin emissions declared in the IREP database, i.e. dioxin emissions exceeding the reporting threshold of annual dioxin emissions of 0.1 g-TEQ/year in place since 2003 (Figure 4 and Table S6). Between 2003 and 2008, we identified 128 dioxin air emission declarations, corresponding to 59 distinct sources (2.3% of the inventoried sources over the whole period). Overall, UNEP Toolkit estimates of

annual averaged emissions exceed emission measurements by 34%, while median source annual emissions underestimate measurements by 40% (Median (IQR): 1.65 (0.16-3.64) g-TEQ/year; 2.76 (1.27-5.51) g-TEQ/year). For waste incineration (63 declarations from 36 distinct sources), the median emission estimates (median (IQR): 2.51 (0.13-5.40) g-TEQ/year) are close to the median emission measurements declared in the IREP (2.80 (1.64-4.73) g-TEQ/year), with a difference below 10% (Table S6). For ferrous and non-ferrous metal production (65 declarations from 23 distinct sources), median source annual emissions underestimate median emission measurements declared in the IREP by 32% (1.65 (0.38-2.76) versus 2.41 (0.93-6.54) g-TEQ, respectively). Yet, the declared measurements remain within the range of the estimated uncertainty in the UNEP Toolkit estimates. Overall, the majority (52%) of our estimates are within  $\pm 200\%$  of the IREP data for the corresponding facilities and activity period (Figure 4) and the Spearman's rank correlation coefficient shows a “moderate” correlation ( $r_s = 0.42$ ).

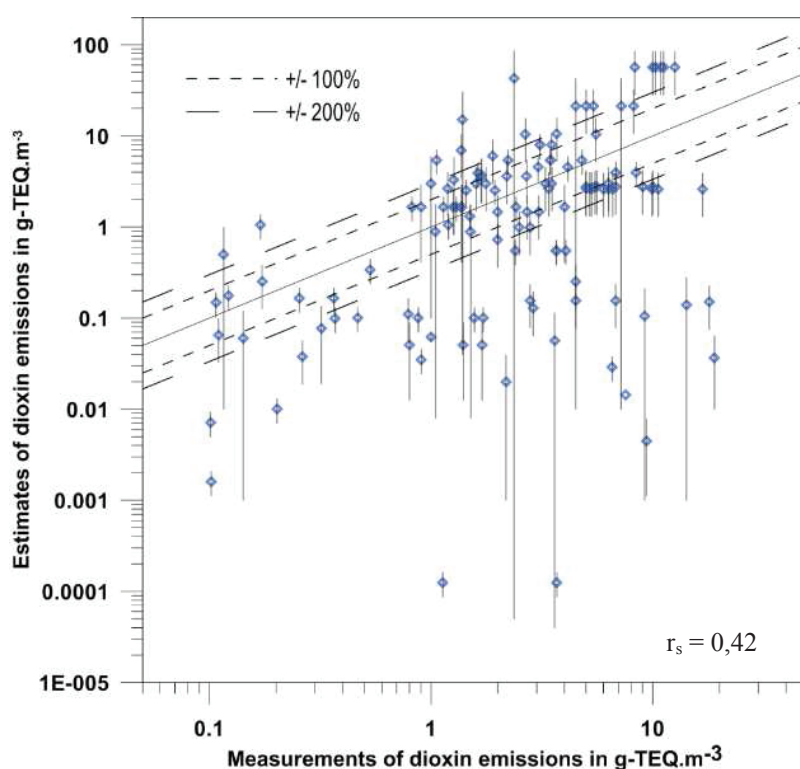


Figure 4 - Comparison between annual average estimates of dioxin emissions and declared dioxin measurements exceeding the reporting threshold of annual dioxin emissions (0.1 g-TEQ/year) (IREP database) in France between 2003 and 2008

## Comparison at the departmental level

At the departmental level, the summed annual UNEP Toolkit estimates overestimate those from the INS emission inventory by 63% in 2004 (425.5 g-TEQ vs 260.6 g-TEQ, respectively) and 29% in 2007 (71.8 g-TEQ vs 55.8 g-TEQ) (Figure 5). In both cases, INS values are included in the uncertainties ranges of the UNEP Toolkit estimates.

We observe a decrease in the difference in average emission by department (from +58% to +27%, respectively) from 2004 to 2007, with a constant coefficient of determination (0.69 in 2004 and 0.67 in 2007). The difference in median emission by department remains small (approximately 0.1 g-TEQ/year). The number of departments with a difference between the Toolkit estimates and the INS inventory exceeding  $\pm 0.5$  g-TEQ/year and  $\pm 2$  g-TEQ/year decreases, respectively, from 53 to 18 and from 32 to 4 departments between 2004 and 2007 (Figure 5).

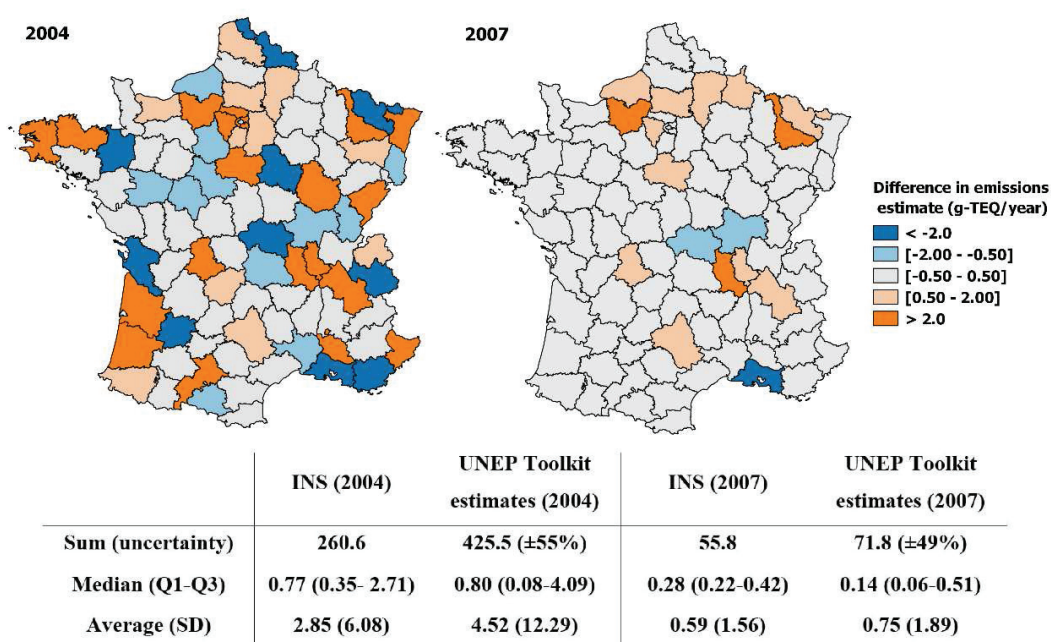


Figure 5 – Comparison at the departmental level between emission estimates (g-TEQ/year) from the INS inventory and the UNEP Toolkit estimates inventory for 2004 and 2007

## Comparison at the national level

By using the EMEP emission inventory, we performed comparisons for the following groups and categories: ferrous and non-ferrous metal production (group 2), production of mineral products (group 4), miscellaneous sources (group 8), hazardous waste incineration (category 1.b), medical waste incineration (1.c), and municipal solid waste incineration and heat and power generation combined (group 3 and category 1.a) (see Table S7).

We observe discrepancies between the two inventories in emission estimates and historical trends for the different groups and categories. The UNEP Toolkit total emission estimates from 1990 to 2008 underestimate the EMEP estimates for ferrous and non-ferrous metal production and hazardous waste incineration (by 37% and 60%, respectively) and overestimate the EMEP estimates for production of mineral products, crematoria and medical waste incineration (by approximately 130%, 1200% and 1800%). For group 8, the relative difference is high but the absolute total difference is low (0.86 g-TEQ/year on average).

The largest absolute difference is observed in 1990 for the group including both municipal solid waste incineration and heat and power generation, with an overestimation of the EMEP annual dioxin emissions by the UNEP Toolkit of nearly 2.4 kg (+200%) (Table S7). This difference progressively decreases over time: 1467 g-TEQ in 1995, 564 g-TEQ in 2000, 167 g-TEQ in 2004 and 12 g-TEQ in 2008. From 1990 to 2008, the difference between the EMEP estimates and UNEP Toolkit estimates falls by more than 95% for ferrous and non-ferrous metal production, municipal solid waste incineration and heat and power generation, production of mineral products groups and medical waste incineration. For crematoria and hazardous waste incineration, the difference increases over the years by 100% (0.6 g-TEQ) and 184% (22.5 g-TEQ), respectively.

## Discussion

This study reports a national inventory to estimate annual dioxin emissions for France at the source level, from 1990 to 2008 as a basis for future use in an epidemiological study. To our knowledge the present inventory is the first historical inventory to include a systematic geolocation of the sources with a high spatial resolution taking into account a source's technological characteristic and the activity rates of a large number of dioxin-emitting sources (n=2620) including an uncertainty evaluation of emission estimates. We further compared the UNEP Toolkit estimates with emissions measurements, as well as inventories at the departmental (INS) and national level (EMEP).

Our results showed a large decrease in atmospheric dioxin emissions in France between 1990 and 2008 (-98.4%). Similar decreases over the same period in emissions from controlled sources have been described by others: for Europe (Quass et al. 2004), the United States (Dwyer and Themelis 2015) and Japan (Government of Japan 2012). The observed reduction of emissions in our study and contribution of different dioxin source groups to total dioxin emissions was heterogeneous over the French territory over time.

In the absence of dioxin atmospheric monitoring data, the use of technical parameters to reconstruct historical dioxin atmospheric emissions has been shown to be a robust methodology (Cordier et al. 2010; Coudon et al. 2018; Fabre et al. 2008b). The UNEP Toolkit provides a series of emission factors that may differ by a factor of three to ten for different technologies in many categories. The reliability of the information used to derive the emission factor and subsequent choice of the emission factor based on technical characteristics, as well as uncertainties in the activity rate may lead to large differences in the emission estimates for a specific source or group of sources. Uncertainty estimates are therefore an essential element of an emissions inventory. However, little is known about the accuracy of existing emission estimates and few emission inventories report uncertainties explicitly. Our observation that dioxin emission estimates are associated with large uncertainties, that vary across categories, are consistent with observations from the literature (Breivik et al 2006). Pulles et al., using Monte Carlo modelling, showed that the annual total emission estimates with the UNEP Toolkit for 2000, in 13 Central and Eastern European countries, was situated in a range of 2-3 times lower and 3–5 times higher than the mean value. Quina et al. (2011), using the UNEP Toolkit, estimated that the total amount of PCDD/PCDF emissions was situated in a range of 2 times lower to 2.5 times higher the most likely value of the total annual amount of PCDD/PCDF released in Portugal in 2006. An inventory of PCDD/PCDF releases estimates total annual emissions from 15 countries to be 10500 g I-TEQ/year for 1995, with an inferior value 1.3 times lower and an upper value approximately 3.5 times higher (Fiedler 1999).

Similarly to Pulles et al., the largest contribution to the overall uncertainty was due to the uncertainty in emissions from hospital waste incinerators (Pulles et al. 2006). The large number of sources and the poor quality of available information strongly affected the average confidence level and subsequent uncertainty estimates. Our comprehensive data search that included multiple databases and use of historical information sources as well as direct contact with operators allowed for numerous facilities to obtain and confirm information on technical characteristics and activity rates. However, the overall response rate to the questionnaire sent to facilities companies or operating groups was low (32%) despite mail and phone reminders. Similar to other studies, the lack of information on technical characteristics and activity rates of dioxin sources especially in the early 1990s constitutes an important limitation in our study.

In response to this and to assess the reliability of the UNEP Toolkit estimates, we compared our estimates with both measurements at the individual source level and two others inventories providing global annual average dioxin emission estimates at the departmental (INS) and national (EMEP) scales.

Measurements of dioxin emissions are scarce (Coudon et al 2018). The comparison of our annual UNEP Toolkit emission estimates to 128 dioxin emissions measures from 59 sources, declared since 2003 in the IREP database, showed a “moderate” correlation overall. The median measurements were situated within 1.5 times the median values of the annual Toolkit estimates. The discrepancies observed can be explained by several factors. The one-off measurements declared in the IREP database by the facility operators once a year have been taken for another purpose and may not be representative of the annual average emission. Furthermore, the variability at a single plant can be an order of magnitude between different sampling periods (Pacyna 2003). Also, measurements were available only for municipal waste incinerators and the metal industry, and the number of sources was small.

The comparison at the departmental scale showed that the median estimates of the INS inventory were between 1.03 times lower (2004) and 2 times higher (2007) than the median values of the UNEP Toolkit estimates. The comparison at the national scale showed that the annual estimates from the EMEP emission database were up to 3.5 times lower compared to the annual UNEP Toolkit estimates for municipal solid waste incineration and heat and power generation, and production of mineral products. The largest differences were found for medical waste incineration. Conversely, EMEP emission estimates exceeded the UNEP Toolkit estimates for ferrous and non-ferrous metal production and hazardous waste incineration. The major differences were observed between 1990 and 1999 corresponding to the period with the highest dioxin emissions. The UNEP Toolkit estimates tend to get closer to the departmental and national inventory estimates for more recent years, probably due to better access to technical source data but also to the overall decline in emissions. These differences can be explained by differences in the emission factors and activity rates used to estimate emissions, the lack of technical data on facilities, and the absence of information on the real number of dioxin sources, such as hospital waste incinerators (Quass et al. 2004). For the latter, the factor of 10 difference is likely to be due to the large number of sources in our inventory (n=1442) with low-quality information collected for this specific category. For the crematoria, the discrepancies are due to differences in the emission factors varying by a factor of 10, and to the low emissions of this category that favours large relative differences. Also, several studies suggest that for the period 1990-2000, dioxin emissions were underestimated due to a poor characterisation of emission sources (Pacyna et al. 2003; Webster and Connett 1998; Brzuzy and Hites 1996). In our study regarding municipal solid waste incineration, we inventoried numerous small MSWI in rural areas; most of them did not have an abatement system and were closed before 1996 due to their high emissions. These sources may not have been included in the EMEP inventory. For example, using specific databases from the Rhône-Alpes region, we inventoried numerous small waste incinerators located in the mountainous area.

Industrial emissions today are low in Europe and further refinement should incorporate other punctual and non-industrial sources that can emit relatively high amounts of dioxins (Coudon et al. 2018; Lee et al. 1999; Serveau et al. 2013) locally on an ad hoc basis over time, such as biomass fires, the burning of manufactured goods and cable, and outdoor burning and illegal landfills in the early 1990s. Yet, these poor or uncontrolled emission sources could not be considered here due to the lack of reliable activity data and estimates of such emissions, in particular retrospectively, and the difficulty of their geolocation. Furthermore, considering only air emissions from controlled sources may also lead to an underestimation of the total amount of PCDD/PCDF released (Quina et al 2011; Brzuzny and Hites 1996).

## **Conclusion**

In the framework of an epidemiological study on the long-term effects of dioxin exposure on cancer risk, we created the first database containing annual dioxin emissions at the facility scale for 2620 facilities in France between 1990 and 2008, including geolocation at the facilities' stack and the technical information needed to characterise emissions over time. Our study shows that the UNEP Toolkit estimates provide reliable estimates for some comparison and that the main difficulty is the availability of data and subsequent uncertainties in the emission estimates. The differences observed in the comparisons at the source, and departmental and national scales were within the estimated uncertainty ranges from our study and as reported in the literature (Breivik et al. 2006).

The final database contains further complementary information on the sources' technical parameters (namely stack height, smoke temperature and velocity) and was used to provide emission estimates and technical information for the modelling of dioxin exposure between 1990 and 2008 (Coudon et al. 2018). The estimates will further be used in an epidemiological study on historical exposure to dioxins and risk of breast cancer. Our database could be useful in the assessment of human exposure in future studies and forms a basis to estimate emissions of other pollutants such as heavy metals (Coudon et al. 2018).



### **Competing interests**

The authors declare to have no competing interests.

### **Acknowledgement**

This work was supported by public funding from the French Environment and Energy Management Agency (ADEME, Grant N° 1306C0031), the Cancéropôle Lyon Auvergne Rhône-Alpes (CLARA), and the Regional Committee of the French League against Cancer of the Saone and Loire Region, and was carried out in partnership with the ARC Foundations for Cancer Research. Thomas Coudon is supported by a Doctoral Fellowship of University Lyon 1. We gratefully acknowledge Harel. G, Denis. C and Hourani. H for their work on the dioxin source inventory and source classification, and Carretero. C and Guillou. M for the geolocation of the dioxin sources.

## Bibliography

1. Ashworth DC, Fuller GW, Toledano MB, Font A, Elliott P, Hansell AL et al. 2013. Comparative assessment of particulate air pollution exposure from municipal solid waste incinerator emissions. *J Environ Public Health* 2013:560342; doi:10.1155/2013/560342.
2. Augusto S, Pinho P, Santos A, Botelho MJ, Palma-Oliveira J, Branquinho C. 2015. Declining trends of PCDD/Fs in lichens over a decade in a Mediterranean area with multiple pollution sources. *Sci Total Environ* 508:95–100; doi:10.1016/j.scitotenv.2014.11.065.
3. Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. 2009. A review of human carcinogens—Part F: Chemical agents and related occupations. *Lancet Oncol* 10:1143–1144; doi:10.1016/S1470-2045(09)70358-4.
4. Basagana X, Aguilera I, Rivera M, Agis D, Foraster M, Marrugat J, et al. 2013. Measurement Error in Epidemiologic Studies of Air Pollution Based on Land-Use Regression Models. *Am J Epidemiol* 178:1342–1346; doi:10.1093/aje/kwt127.
5. Breivik K, Vestreng V, Rozovskaya O, Pacyna JM. 2006. Atmospheric emissions of some POPs in Europe: a discussion of existing inventories and data needs. *Environ Sci Policy* 9:663–674; doi:10.1016/j.envsci.2006.09.001.
6. Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. 2007. Environmental pollutants and breast cancer - Epidemiologic studies. *Cancer* 109:2667–2711; doi:10.1002/cncr.22655.
7. Brzuzy LP, Hites RA. 1996. Global mass balance for polychlorinated dibenzo-p-dioxins and dibenzofurans. *Environ Sci Technol* 30:1797–1804; doi:10.1021/es950714n.
8. Cordier S, Lehébel A, Amar E, Anzivino-Viricel L, Hours M, Monfort C, et al. 2010. Maternal residence near municipal waste incinerators and the risk of urinary tract birth defects. *Occup Environ Med* 67:493–499; doi:10.1136/oem.2009.052456.
9. Cordioli M, Ranzi A, De Leo GA, Lauriola P. 2013. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators. *J Environ Public Health*; doi:10.1155/2013/129470.
10. Coudon T, Hourani H, Nguyen C, Faure E, Mancini FR, Fervers B, et al. 2018. Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon metropolitan area (France). *Environ Int* 111:177–190; doi:10.1016/j.envint.2017.11.027.
11. Coutinho M, Albuquerque M, Silva AP, Rodrigues J, Borrego C. 2015. Long-time monitoring of polychlorinated dibenzo-p-dioxins and dibenzofurans over a decade in the ambient air of Porto, Portugal. *Chemosphere* 137:207–213; doi:10.1016/j.chemosphere.2015.07.052.
12. Dai D, Oyana TJ. 2008. Spatial variations in the incidence of breast cancer and potential risks associated with soil dioxin contamination in Midland, Saginaw, and Bay Counties, Michigan, USA. *Environ Health* 7:49; doi:10.1186/1476-069X-7-49.
13. Dwyer H, Themelis NJ. 2015. Inventory of U.S. 2012 dioxin emissions to atmosphere. *Waste Manag* 46:242–246; doi:10.1016/j.wasman.2015.08.009.

14. Fabre P, Daniau C, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. 2008a. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. Synthèse St-Maurice InVS.
15. Fabre P, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. 2008b. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. Synthèse St-Maurice InVS.
16. Fiedler H. 1999. National and regional dioxin and furan inventories: national and regional emissions of PCDD/PCDF.
17. Fiedler H. 2015. Release Inventories of Polychlorinated Dibenzo-p-Dioxins and Polychlorinated Dibenzofurans. In: *Dioxin and Related Compounds. The Handbook of Environmental Chemistry*. Springer, Cham. 1–27.
18. Government of Japan. 2012. Dioxins.
19. Huang YL, Batterman S. 2000. Residence location as a measure of environmental exposure: a review of air pollution epidemiology studies. *J Expo Anal Environ Epidemiol* 10: 66–85.
20. Katsoyiannis A, Gioia R, Sweetman AJ, Jones KC. 2010. Continuous Monitoring of PCDD/Fs in the UK Atmosphere: 1991-2008. *Environ Sci Technol* 44:5735–5740; doi:10.1021/es1009828.
21. Kok H, Pulles T, Quass U. 2005. Dioxin emission estimates and reduction options for Central and Eastern Europe. In: *Environmental Health Risk III* (C.A. Brebbia, V. Popov, and D. Fayzieva, eds). Vol. 9 of. Wit Press:Southampton. 189–197.
22. Lee RGM, Green NJL, Lohmann R, Jones KC. 1999. Seasonal, Anthropogenic, Air Mass, and Meteorological Influences on the Atmospheric Concentrations of Polychlorinated Dibenzo-p-dioxins and Dibenzofurans (PCDD/Fs): Evidence for the Importance of Diffuse Combustion Sources. *Environ Sci Technol* 33:2864–2871; doi:10.1021/es981323m.
23. Lohmann R, Jones KC. 1998. Dioxins and furans in air and deposition: a review of levels, behaviour and processes. *Sci Total Environ* 219: 53–81.
24. Pacyna EG, Pacyna JM, Pirrone N. 2001. European emissions of atmospheric mercury from anthropogenic sources in 1995. *Atmos Environ* 35:2987–2996; doi:10.1016/S1352-2310(01)00102-9.
25. Pacyna JM, Breivik K, Munch J, Fudala J. 2003. European atmospheric emissions of selected persistent organic pollutants, 1970-1995. *Atmos Environ* 37:S119–S131; doi:10.1016/S1352-2310(03)00240-1.
26. Pronk A, Nuckols JR, De Roos AJ, Airola M, Colt JS, Cerhan JR, et al. 2013. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case-control study. *Environ Health* 12:20; doi:10.1186/1476-069X-12-20.
27. Pulles T, Kok H, Quass U. 2006. Application of the emission inventory model TEAM: Uncertainties in dioxin emission estimates for central Europe. *Atmos Environ* 40:2321–2332; doi:10.1016/j.atmosenv.2005.12.015.
28. Quass U, Fermann M, Broker G. 2004. The European dioxin air emission inventory

- project - Final results. *Chemosphere* 54:1319–1327; doi:10.1016/S0045-6535(03)00251-0.
29. Quina MJ, Pedro RS, Gando-Ferreira LM, Quinta-Ferreira RM. 2011. A national inventory to estimate release of polychlorinated dibenzo-p-dioxins and dibenzofurans in Portugal. *Chemosphere* 85:1749–1758; doi:10.1016/j.chemosphere.2011.09.028.
  30. Ries L, Harkins D, Krapcho M, Mariotto A, Miller B, Feuer E, et al. 2006. SEER Cancer Statistics Review, 1975-2003. Public Health Fac Publ.
  31. Saral A, Gunes G, Demir S. 2014. Atmospheric PCDD/PCDF emission inventory for Turkey. *Atmospheric Pollut Res* 5:24–33; doi:10.5094/APR.2014.004.
  32. Serveau L, Boutang J, Allemand N, Fontelle J-P, Vincent J. 2013. source CITEPA / format SECTEN – INVENTAIRE DES EMISSIONS DE POLLUANTS ATMOSPHERIQUES ET DE GAZ A EFFET DE SERRE EN FRANCE SERIES SECTORIELLES ET ANALYSES ETENDUES.
  33. Tabeaud M, Hamez G. 2000. *Les métamorphoses du déchet*. Publications de la Sorbonne.
  34. UNEP. 2013. United Nations Environment Programme | Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs. Available: [http://toolkit.pops.int/Publish/Main/01\\_Index.html](http://toolkit.pops.int/Publish/Main/01_Index.html) [accessed 22 August 2016].
  35. Vela G. 2010. ARMEN Vague 2 Segment Gestion des déchets Livrable long.
  36. Viel J-F, Clement M-C, Hagi M, Grandjean S, Challier B, Danzon A. 2008. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int J Health Geogr* 7:4; doi:10.1186/1476-072X-7-4.
  37. Webster T, Connett P. 1998. Dioxin emission inventories and trends: the importance of large point sources. *Chemosphere* 37:2105–2118; doi:10.1016/S0045-6535(98)00273-2.
  38. Zhang T, Fiedler H, Yu G, Solorzano Ochoa G, Carroll WF, Gullett BK, et al. 2011. Emissions of unintentional persistent organic pollutants from open burning of municipal solid waste from developing countries. *Chemosphere* 84:994–1001; doi:10.1016/j.chemosphere.2011.04.070.

## Supplementary material

# **A NATIONAL INVENTORY OF HISTORICAL DIOXIN AIR EMISSIONS SOURCES IN FRANCE**

*T. Coudon<sup>1,2</sup>, P. Salizzon<sup>3</sup>, D. Praud<sup>1,4</sup>, AMN. Danjou<sup>1</sup>, L. Dossus<sup>1,4,5</sup>, E. Faure<sup>1</sup>, B. Fervers<sup>1,2,6</sup>*

*<sup>1</sup>Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France*

*<sup>2</sup>Université Claude Bernard Lyon 1, Villeurbanne 69100, France*

*<sup>3</sup>Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully, France*

*<sup>4</sup>Centre de Recherche en Epidémiologie et Santé des Populations (CESP, Inserm U1018),  
Facultés de Médecine, Université Paris-Saclay, UPS UVSQ, Gustave Roussy, 114 rue  
Edouard-Vaillant, 94805 Villejuif Cedex, France*

*<sup>5</sup>Nutrition and Metabolism Section, International Agency for Research on Cancer, 150 cours  
Albert Thomas, 69372 Lyon Cedex 08, France*

*<sup>6</sup>INSERM 1052, CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, France.*

Keywords:

dioxins, emission, inventory, UNEP toolkit, exposure assessment, endocrine disruptor

Table S1 - List of public database used to build the BD source

Database	UNEP Categories covered	Spatial coverage	Temporal coverage*	Description	Limitations
<b>AFIF</b> - French Funeral Association	8.b	National	Since 1992	Address ; Activity rate ; Flue gas treatment systems	Lack of information on activity period and activity rate
<b>AQP</b> - Association Weighing Quality	4.f	National	Since 1998	Address ; technical characteristic	Restricted to some plants
<b>BASIAS</b> - Historical inventory of industrial sites and service activities	All	National	Since 1978	Address ; geographical coordinates; activity history	Lack of information on technical characteristics
<b>BASOL</b> - Inventory of polluted sites and soils	All	National	Since 1994	Address ; cadastral plans; geographical coordinates; historical information	Data completeness and quality varies by region
<b>Cordier</b> – Study by Cordier et al., 2004	1.a	Regional (Rhône-Alpes)	1988-1997	Address ; activity rate ; technical characteristics	Lack of information on technical characteristics
<b>GEREP</b> - Annual reporting of pollutant and waste emissions and transfers	All	National	Since 2003	Address ; geographical coordinates; activity rate; technical characteristic; measured emissions	No data before 2003
<b>Base ICPE</b> – Database of classified installations for environmental protection	All	National	Since 2003	Prefectural Orders; inspection reports; authorization dates; technical metadata	Current sources only
<b>IREP</b> - French pollutant emissions register	All	National	Since 2003	Address ; geographical coordinates; tonnage treated waste / products; measured emissions	No data before 2003
<b>ITOM</b> - Household garbage treatment facilities survey	1.a	National	Since 1975	Historical technological characterization; activity rate	Lack of information on technical characteristics
<b>PREDD</b> - Regional Hazardous Waste Disposal Plans	1.b	National	Since 2010	Address ; date of first activity ; activity rate	Lack of information on technical characteristics; Source active on 2010 only
<b>Association Robin des Bois</b> (Association for the Protection of Man and the Environment)	1.a	Régional	1966-2003	Address ; date of beginning and end of activity; activity rate	Quantity and quality of information varying across regions; Lack of data on technical characteristics
<b>SVDU</b> - National Union for the Treatment of Urban and Related Waste	1.a	National	Since approx. 1970	Address ; technical characteristics; activity rate	Current source only
<b>Via Seva</b> - Heat networks	3	National	Since 2000	Plant type; activity rate	Current source only
<b>Internet documents</b>	All	National	Any dates	Any kind of information	Quality of information variable

\*do not imply full completeness on the period.

Table S2 - Number of sources, average and median annual emissions estimates (in g-TEQ/year), standard deviation, interquartile range and average uncertainty from 1990 to 2008

Years	Number of sources	Average annual emissions (SD) g-TEQ/year	Median annual emissions (IQR) g-TEQ/year	uncertainty
1990	2323	1.89 (8.68)	0.288 (0.132-0.83)	66%
1991	2326	1.88 (8.65)	0.287 (0.132-0.82)	66%
1992	2338	1.81 (8.56)	0.281 (0.129-0.80)	66%
1993	2346	1.70 (7.73)	0.279 (0.129-0.79)	65%
1994	2348	1.65 (7.63)	0.273 (0.126-0.77)	64%
1995	2348	1.40 (5.22)	0.269 (0.126-0.76)	64%
1996	905	2.80 (8.13)	0.050 (0.0049-2.12)	54%
1997	906	2.35 (7.43)	0.047 (0.0045-1.68)	52%
1998	902	1.93 (5.28)	0.040 (0.0041-1.41)	52%
1999	892	1.59 (4.81)	0.030 (0.0036-0.85)	50%
2000	876	1.20 (3.95)	0.022 (0.0029-0.30)	48%
2001	867	0.87 (2.94)	0.018 (0.0025-0.15)	44%
2002	849	0.69 (2.64)	0.015 (0.0022-0.10)	41%
2003	836	0.43 (2.19)	0.012 (0.0018-0.05)	40%
2004	845	0.50 (4.06)	0.012 (0.0017-0.05)	56%
2005	852	0.30 (1.47)	0.011 (0.0016-0.05)	40%
2006	847	0.10 (0.44)	0.010 (0.0015-0.04)	40%
2007	858	0.08 (0.43)	0.009 (0.0014-0.04)	50%
2008	853	0.08 (0.43)	0.009 (0.0013-0.04)	50%

Table S3 – Annual UNEP toolkit emission estimates and associated uncertainties by category for waste incineration group from 1990 to 2008 in g-TEQ/year

Years	Municipal solid waste incineration	Hazardous waste incineration	Medical waste incineration	Sewage sludge incineration	Waste wood and waste biomass incineration	Destruction of animal carcasses					
1990	3587.66	61%	34.90	75%	642.17	100%	1.13	61%	ND	0.01	30%
1991	3566.27	61%	35.39	74%	642.17	100%	1.13	61%	ND	0.01	30%
1992	3440.40	61%	20.53	50%	642.17	100%	1.13	61%	ND	0.01	30%
1993	3202.10	59%	31.25	43%	642.17	100%	1.13	61%	ND	0.01	30%
1994	3075.23	58%	32.04	44%	641.80	100%	1.13	61%	ND	0.01	30%
1995	2487.14	55%	32.34	44%	641.80	100%	0.69	81%	ND	0.01	30%
1996	2376.42	54%	32.40	44%	2.06	100%	0.69	81%	ND	0.01	30%
1997	1975.79	52%	23.72	49%	2.06	100%	0.71	80%	ND	0.01	30%
1998	1585.93	52%	24.08	49%	2.06	100%	0.39	65%	ND	0.01	30%
1999	1267.41	49%	21.67	47%	1.16	100%	0.40	65%	0.04	0.15	98%
2000	923.07	47%	8.65	59%	0.60	100%	0.40	65%	0.04	1.60	100%
2001	665.83	43%	9.36	62%	0.60	100%	0.43	63%	0.04	1.60	100%
2002	482.91	40%	8.36	59%	0.60	100%	0.40	60%	0.04	1.60	100%
2003	286.77	37%	6.27	58%	0.60	100%	0.33	51%	0.04	1.60	100%
2004	352.00	56%	5.90	57%	0.02	30%	0.34	51%	0.04	1.61	99%
2005	192.50	39%	5.91	46%	0.02	30%	0.22	51%	0.04	1.61	99%
2006	17.66	48%	5.75	45%	0.02	30%	0.18	42%	0.04	1.61	99%
2007	12.33	53%	5.68	44%	0.02	30%	0.13	47%	0.04	1.61	99%
2008	12.58	52%	5.61	44%	0.02	30%	0.13	47%	0.04	1.61	99%



Table S4 – Annual UNEP toolkit emission estimates (in g-TEQ/year) and associated uncertainties (in %) by category for ferrous and non-ferrous metal production group from 1990 to 2008 in g-TEQ/year

Years	Iron ore sintering	Coke production	Iron and Steel Making	Iron Foundries	Hot-dip Galvanizing Plants	Copper production	Aluminum production	Lead production	Zinc production	Brass and bronze production	Other non-ferrous metal production											
1990	58.44	47%	2.22	50%	16.35	46%	3.68	42%	0.06	75%	5.74	71%	13.31	63%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1991	58.44	47%	2.22	50%	16.35	46%	3.84	42%	0.06	75%	5.74	71%	13.49	63%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1992	59.75	47%	2.22	50%	15.73	47%	4.06	44%	0.06	75%	5.74	71%	13.55	63%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1993	59.75	47%	2.22	50%	15.73	47%	4.06	44%	0.06	75%	5.74	71%	16.86	57%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1994	59.75	47%	2.22	50%	15.79	47%	3.84	45%	0.06	75%	5.74	71%	16.86	57%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1995	59.75	47%	2.22	50%	19.09	47%	3.84	45%	0.06	75%	5.74	71%	16.86	57%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1996	59.75	47%	2.22	50%	20.71	46%	2.88	50%	0.06	75%	6.24	69%	16.91	57%	1.06	54%	0.01	88%	0.04	50%	0.26	52%
1997	59.75	47%	2.22	50%	20.71	46%	2.88	50%	0.06	75%	6.24	69%	16.91	57%	1.06	54%	0.01	88%	0.04	50%	0.28	52%
1998	59.75	47%	2.22	50%	20.71	46%	2.83	49%	0.06	75%	6.24	69%	20.92	52%	1.06	54%	0.11	98%	0.04	50%	0.28	52%
1999	59.75	47%	2.22	50%	20.71	46%	2.82	49%	0.06	75%	6.34	69%	16.79	53%	1.06	54%	0.11	98%	0.04	50%	2.78	50%
2000	59.75	47%	2.22	50%	20.71	46%	2.81	49%	0.06	75%	3.78	89%	14.44	57%	1.06	54%	0.11	98%	0.04	50%	2.76	50%
2001	15.75	30%	2.21	50%	20.63	46%	2.78	49%	0.06	75%	3.78	89%	14.37	57%	1.06	54%	0.11	98%	0.04	50%	2.76	50%
2002	15.75	30%	2.19	50%	20.87	46%	2.13	67%	0.06	75%	3.78	89%	14.37	57%	0.65	56%	0.11	98%	0.04	50%	2.76	50%
2003	15.75	30%	2.19	50%	19.80	47%	1.75	66%	0.06	75%	3.78	89%	9.10	66%	0.62	57%	0.14	88%	0.04	50%	2.76	50%
2004	15.75	30%	2.21	50%	19.73	47%	1.71	67%	0.06	75%	3.78	89%	9.12	66%	0.62	57%	0.14	88%	0.04	50%	2.76	50%
2005	14.54	30%	0.11	51%	16.54	46%	1.62	65%	0.06	75%	3.78	89%	6.33	52%	0.59	58%	0.14	88%	0.04	50%	0.31	52%
2006	14.54	30%	0.11	51%	16.54	46%	1.41	65%	0.06	75%	3.78	89%	6.46	52%	0.59	58%	0.14	88%	0.04	50%	0.31	52%
2007	14.76	31%	0.11	51%	16.40	46%	1.25	60%	0.06	74%	3.79	89%	6.22	51%	0.59	58%	0.14	88%	0.04	50%	0.31	52%
2008	14.76	31%	0.11	51%	16.40	46%	1.22	61%	0.06	74%	3.78	89%	5.71	47%	0.59	58%	0.14	88%	0.00	100%	0.31	52%

Table S5 - Annual UNEP toolkit emission estimates and associated uncertainties by category for Heat and Power Generation, Production of Mineral Products, Chemicals and Consumer Goods and Miscellaneous groups from 1990 to 2008 in g-TEQ/year

Years	Fossil fuel power plant	Biomass power plant	Landfill biogas production	Cement production	Lime production	Brick production	Glass production	Asphalt mixing	Pulp and paper production	Chlorinated inorganic chemicals	Crematoria											
1990	2.40	47%	0.50	100%	ND	9.99	71%	2.97	52%	0.006	39%	0.88	35%	0.01	100%	0.96	34%	0.01	30%	0.61	100%	
1991	2.26	48%	0.50	100%	ND	9.99	71%	2.97	52%	0.006	39%	0.88	35%	0.01	100%	0.96	34%	0.01	30%	0.56	100%	
1992	2.26	48%	0.50	100%	ND	9.99	71%	2.38	47%	0.006	39%	0.88	35%	0.01	100%	1.26	33%	0.01	30%	0.66	100%	
1993	2.25	47%	0.50	100%	ND	9.85	71%	1.64	45%	0.006	39%	0.88	35%	0.01	100%	1.26	33%	0.01	30%	0.67	100%	
1994	2.25	47%	0.50	100%	ND	9.46	70%	1.64	45%	0.006	39%	0.88	35%	0.02	100%	1.26	33%	0.01	30%	0.68	100%	
1995	2.25	47%	0.53	97%	ND	8.98	68%	1.64	45%	0.010	36%	0.92	35%	0.02	100%	1.26	33%	0.01	30%	0.72	100%	
1996	2.25	47%	0.53	97%	ND	8.98	68%	1.64	45%	0.010	36%	0.92	35%	0.02	100%	1.26	33%	0.01	30%	0.74	100%	
1997	2.26	47%	0.53	97%	0.012	58%	8.38	66%	1.64	45%	0.010	36%	0.92	35%	0.02	100%	1.26	33%	0.01	30%	0.75	100%
1998	2.26	47%	0.58	92%	0.012	58%	6.07	64%	1.65	45%	0.010	36%	0.92	35%	0.02	100%	1.26	33%	0.01	30%	0.78	100%
1999	2.26	47%	0.74	80%	0.012	58%	5.10	67%	1.65	45%	0.010	36%	0.93	34%	0.02	100%	1.26	33%	0.01	30%	0.81	100%
2000	2.26	47%	0.74	80%	0.012	58%	4.61	66%	1.64	45%	0.010	36%	0.93	34%	0.02	100%	1.26	33%	0.01	30%	0.92	100%
2001	2.25	47%	0.74	80%	0.012	58%	2.64	57%	1.63	45%	0.010	36%	0.89	34%	0.03	100%	1.26	33%	0.01	30%	1.01	100%
2002	2.25	47%	0.74	80%	0.012	58%	2.64	57%	1.63	45%	0.010	36%	0.89	34%	0.03	100%	1.26	33%	0.01	30%	1.08	100%
2003	2.24	47%	0.74	80%	0.012	58%	2.24	57%	1.22	49%	0.013	34%	0.80	33%	0.03	100%	1.26	33%	0.01	30%	1.11	100%
2004	2.24	47%	0.74	80%	0.012	58%	2.24	57%	1.22	49%	0.013	34%	0.74	33%	0.03	100%	1.26	33%	0.01	30%	1.18	100%
2005	2.09	44%	0.87	74%	0.012	58%	2.24	57%	1.22	49%	0.013	34%	0.68	36%	0.03	100%	1.26	33%	0.01	30%	1.19	100%
2006	1.82	41%	0.90	73%	0.012	58%	1.82	51%	1.22	49%	0.015	34%	0.62	37%	0.03	100%	1.26	33%	0.01	30%	1.20	100%
2007	1.72	38%	0.95	71%	0.012	58%	1.76	50%	1.22	49%	0.015	34%	0.53	38%	0.03	100%	1.26	33%	0.01	30%	1.26	100%
2008	1.70	38%	0.95	71%	0.012	58%	1.76	50%	1.22	49%	0.015	34%	0.53	38%	0.03	100%	1.26	33%	0.01	30%	1.26	100%

Table S6 – Comparison between dioxin emission above the mandatory declaration threshold of 0.1 g-TEQ/year (from the iREP database) and UNEP toolkit estimates between 2003 and 2008 in g-TEQ/year

Years	Waste incineration			Ferrous and non-Ferrous metal production				
	n	Median of emissions declaration	Median of UNEP estimates	uncertainty of UNEP estimates	n	Median of emissions declaration	Median of UNEP estimates	uncertainty of UNEP estimates
2003	19	3.61	2.76	40%	3	10.3	3.00	44%
2004	20	3.05	3.85	37%	11	2.5	2.62	45%
2005	16	2.80	3.19	36%	12	3.4	1.65	40%
2006	2	4.63	0.03	40%	8	4.4	2.62	35%
2007	3	1.00	0.05	33%	15	0.9	1.10	33%
2008	3	0.17	0.06	33%	16	0.9	0.80	33%
Years	n	Median (IQR) of emissions declaration	Median (IQR) of UNEP estimates	uncertainty of UNEP estimates	n	Median (IQR) of emissions declaration	Median (IQR) of UNEP estimates	uncertainty of UNEP estimates
2003-2008	63	2.80 (1.64-4.73)	2.51 (0.13-5.40)	38%	65	2.41 (0.93-6.54)	1.65 (0.38-2.76)	40%

Table S7– Comparison of the emission annual estimates (g-TEQ/year) from EMEP and UNEP toolkit from 1990 and 2008

Years	Municipal solid waste incineration & Heat and power generation		Hazardous waste incineration		Medical waste incineration		Ferrous and non-ferrous metal production		Production of mineral products		Crematoria	
	EMEP	UNEP toolkit estimates	EMEP	UNEP toolkit estimates	EMEP	UNEP toolkit estimates	EMEP	UNEP toolkit estimates	EMEP	UNEP toolkit estimates	EMEP	UNEP toolkit estimates
1990	1202.46	3590.56	48.21	36.03	68.16	642.17	374.36	101.19	4.27	14.83	0.02	0.61
1991	1281.38	3569.04	48.26	36.52	51.95	642.17	354.87	101.52	4.29	14.83	0.02	0.56
1992	1339.00	3443.16	48.39	21.66	40.11	642.17	338.82	102.48	4.22	14.54	0.02	0.66
1993	1310.60	3204.85	48.33	32.37	31.50	642.17	436.39	105.80	4.25	13.66	0.03	0.67
1994	1209.37	3077.98	48.69	33.16	25.05	641.80	553.09	105.64	4.44	13.28	0.03	0.68
1995	1023.00	2489.92	48.62	33.02	22.99	641.80	541.22	108.94	3.85	12.84	0.03	0.72
1996	928.19	2379.20	48.82	33.08	19.46	2.06	421.76	110.15	4.28	12.84	0.04	0.74
1997	578.49	1978.59	47.95	24.43	15.75	2.06	349.90	110.16	4.32	12.23	0.04	0.75
1998	540.59	1588.78	47.39	24.47	7.81	2.06	294.16	114.22	4.26	9.93	0.04	0.78
1999	389.11	1270.43	46.71	22.07	6.15	1.16	124.35	112.67	4.05	8.97	0.05	0.81
2000	362.56	926.09	45.97	9.05	3.94	0.00	64.81	107.76	3.93	8.47	0.05	0.92
2001	246.21	668.84	44.78	9.78	2.81	0.60	52.06	63.56	3.60	6.46	0.05	1.01
2002	204.34	485.91	43.64	8.76	1.72	0.60	72.23	62.71	3.48	6.47	0.06	1.08
2003	102.18	289.77	42.61	6.60	0.61	0.60	55.06	56.00	3.52	5.56	0.06	1.11
2004	188.01	355.00	42.32	6.24	0.20	0.02	50.59	55.92	3.50	5.51	0.06	1.18
2005	74.66	195.47	41.84	6.13	0.00	0.02	45.05	44.07	3.68	5.45	0.07	1.19
2006	4.62	20.39	41.15	5.93	0.00	0.02	45.30	43.99	3.92	4.98	0.07	1.20
2007	2.97	15.01	40.43	5.81	0.00	0.02	46.11	43.70	4.36	4.83	0.07	1.26
2008	2.68	15.24	40.41	5.74	0.00	0.02	32.87	43.10	4.57	4.83	0.08	1.26

### 3. Extension de la base de données aux émetteurs de cadmium

Les sources de dioxines sont également émettrices d'un certain nombre d'autres polluants cancérigènes parmi lesquels on peut citer plusieurs métaux lourds dont le cadmium (Nawrot et al. 2006). Les dioxines ont parfois été utilisées dans la littérature comme polluant « traceur » de l'exposition globale liée aux émissions des incinérateurs (Fabre et al. 2008; Ranzi et al. 2014). Le travail méthodique et étendu en terme de récolte de données, réalisé au sein du projet GEO3N, a permis d'aboutir à une base comprenant les données (émission, localisation, technologie de réduction des émissions, etc.) de 2620 sources de dioxines entre 1990 et 2008. Dans le cadre du projet XENAIR (chapitre 1), nous avons souhaité exploiter cette base pour estimer l'exposition atmosphérique au cadmium liée aux sources industrielles. La méthodologie retenue pour construire cette base nous offrait plusieurs avantages :

- Pour chaque source, nous disposions d'un recueil de données technologiques complètes qui ne se concentrent pas uniquement sur les systèmes d'épuration des fumées spécifiques aux dioxines ;
- L'estimation des émissions est basée sur l'utilisation combinée de facteurs d'émission et de taux d'activité (tonnages annuels), et permet, en changeant les facteurs émission, d'extrapoler les émissions de cadmium ;
- Une (probable) couverture très élevée des sources industrielles de cadmium et par conséquent une quantité de travail réduit pour compléter la base des sources émettrices ;
- Les estimations des incertitudes sur les émissions de cadmium peuvent se baser sur les données déjà recueillies sur les dioxines ;

Avant de « convertir » cette base de données de sources de dioxines en sources de cadmium, il était nécessaire de compléter deux points : vérifier que la base couvrait bien la majorité des sources de cadmium et identifier un jeu de facteur d'émissions, si possible unique, pour estimer de manière correcte les estimations de cadmium. Pour ce faire un travail préliminaire a été lancé en 2016 sur la région Rhône-Alpes et a fait l'objet d'un stage de Master 2, réalisée par Hassan HOURANI, que j'ai supervisé.

D'après les données de la littérature, les secteurs émetteurs de cadmium sont semblables aux secteurs émetteurs de dioxines (CITEPA 2017). Nous avons, dans un premier temps, identifié les principaux secteurs et activités pouvant émettre du cadmium en recensant toutes les déclarations inscrites au registre iREP en Rhône-Alpes sur la période 2003-2012. En couvrant une large période, nous nous assurons de considérer la majorité des secteurs émetteurs de cadmium. Il est apparu que sur la trentaine de sources dépassant le seuil de déclaration de rejets de cadmium en Rhône-Alpes, seulement quatre n'étaient pas déjà dans la base des émetteurs de dioxines.

Durant le projet GEO3N, devant le temps extrêmement long consacré à l'inventaire des sources, une priorisation des sources a été effectuée pour ne pas conserver les activités très peu émettrices liées à certains secteurs spécifiques (chimie, production de produits minéraux, retraitement des batteries usagées...), ce qui explique ces sources manquantes. Nous pouvons donc nous appuyer sur la base « dioxine » et la compléter en recensant les sources manquantes dans les secteurs identifiés, au cas par cas, afin de pouvoir considérer tous les émetteurs de cadmium.

Pour estimer les rejets en dioxines des installations industrielles, le projet GEO3N se basait sur le jeu de facteur d'émission fourni par le Toolkit de l'UNEP (UNEP 2013). Le choix de cet outil fut évident car abondamment utilisé dans la littérature scientifique et repris par des inventaires européens (Kok et al. 2005; Momeniha et al. 2017; Pulles et al. 2006; Quina et al. 2011; Saral et al. 2014). Pour estimer les émissions de cadmium, nous avons identifié deux jeux de facteurs d'émissions différents pouvant être utilisés : le Guide book de l'EMEP et la base OMINEA (Organisation et méthodes des inventaires nationaux des émissions atmosphériques) du CITEPA. Pour choisir la base appropriée (même si, pour certains secteurs, les facteurs d'émission sont communs) nous avons estimé les rejets annuels des sources en Rhône-Alpes avec ces deux outils et avons comparé les résultats avec les mesures disponibles. Il ressort de ces comparaisons que les facteurs d'émission d'OMINEA sont plus fiables que le GuideBook EMEP (Tableau 1) mais également que les estimations du GuideBook s'éloignent fortement des données mesurées avant 2005.

Ce résultat pourrait s'expliquer par le fait que le jeu de facteurs d'émission OMINEA est issu d'une institution nationale (CITEPA) et est donc potentiellement plus adaptée pour estimer les rejets des industries françaises que le Guidebook de l'EMEP qui est un outil européen. De plus, le CITEPA adapte annuellement les facteurs d'émission utilisés grâce à une approche statistique basée sur les mesures réalisées sur les émissions industrielles. La spécificité nationale se traduit donc par le détail des informations et leurs représentativités plutôt que par l'approche méthodologique proprement dite.

Les analyses comparatives ont donc permis d'identifier l'OMINEA du CITEPA comme base de référence pour estimer les rejets de cadmium. Le croisement de la base OMINEA avec le jeu de facteurs d'émission du Toolkit a par ailleurs permis de vérifier qu'aucun secteur émetteur de cadmium n'avait été absent du précédent recensement des sources industrielles.

Tableau 1- Comparaisons entre les mesures et les estimations (basées sur les facteurs d'émissions EMEP et OMINEA) des émissions de cadmium (kg/an) des sources industrielles en Rhône-Alpes

	Nombre de mesures	Emissions moyennes (kg/an)	Estimation EMEP (kg/an)	Estimation OMINEA (kg/an)	Ecart (%) Mesure / EMEP	Ecart (%) Mesure / OMINEA
Source 1	3	22.33	11.90	17.27	-47%	-23%
Source 2	1	11.00	4.76	8.20	-57%	-25%
Source 3	2	15.00	4.05	7.44	-73%	-50%
Source 4	1	23.00	3.02	13.82	-87%	-40%
Source 5	1	11.00	3.19	3.36	-71%	-69%
Source 6	4	31.75	8.50	19.28	-73%	-39%
Source 7	1	11.00	3.45	6.75	-69%	-39%
Source 8	2	15.50	5.77	18.83	-63%	21%
Source 9	3	28.33	7.83	21.53	-72%	-24%
Source 10	2	17.50	3.86	10.84	-78%	-38%
Source 11	1	15.00	3.07	6.01	-80%	-60%
Source 12	1	21.00	4.85	22.16	-77%	6%
Source 13	2	12.00	2.92	5.37	-76%	-55%
Source 14	2	11.00	3.40	11.10	-69%	1%
Source 15	3	15.67	4.59	12.62	-71%	-19%
Source 16	1	13.00	3.40	6.66	-74%	-49%
Source 17	1	13.00	11.77	9.70	-9%	-25%
Source 18	4	13.75	6.12	12.63	-55%	-8%
Source 19	2	18.50	4.25	7.82	-77%	-58%
Source 20	1	16.00	3.13	6.12	-80%	-62%
Source 21	5	49.60	24.04	47.96	-52%	-3%
Source 22	2	19.00	6.45	20.30	-66%	7%
Source 23	2	11.00	2.99	5.51	-73%	-50%
Source 24	2	12.50	7.20	20.24	-42%	62%
Source 25	5	26.40	21.01	42.54	-20%	61%
Source 26	6	19.33	8.53	15.36	-56%	-21%
Source 27	3	17.67	6.11	14.96	-65%	-15%

## Résultats au niveau national

Le recensement au niveau national a permis d'identifier 80 sources de cadmium non présentes dans la base des sources de dioxines (Tableau 2). Comme détaillé dans l'article n°1, certaines de ces sources avaient été négligées car trop faiblement émettrices de dioxines (notamment groupes 4 et 7 et catégories 1d et 3a). Le nombre de sources de cadmium sur la période 1990-2008 s'élève donc à 2700.

Tableau 2 – Répartition des nouvelles sources identifiées durant l'inventaire des sources de cadmium

Groupes		Sources
1	Incinération de déchets	7
2	Production de métaux ferreux et non-ferreux	8
3	Production de chaleur et d'électricité	12
4	Production de produits minéraux	35
7	Chimie	11
8	Divers	0
Catégories		Sources
1.d	Incinération de déchets de déchiqueteuse à fraction légère	7
2.c.i	Fabrication de Fer et d'aciers	3
2.c.ii	Fonderies	2
2.e	Production d'aluminium	3
3.a	Centrale à combustible fossile	12
4.a	Cimenterie	2
4.c	Production de briques	7
4.d	Production de verre	8
4.e	Production de céramique	18
7.a	Production de papier	2
7.b	Produits chimiques inorganiques chlorés	9
7.f	Industrie pétrolière	7
	Total	80

Chacune de ces nouvelles sources a été géolocalisée et a fait l'objet d'un recueil de données techniques pour appliquer le facteur d'émission approprié et estimer les rejets en cadmium. La répartition en nombre total de sources est peu affectée par ces nouvelles sources, la seule évolution notable étant le nombre de sources du groupe « Production de produits minéraux » qui passe de 156 sources à 191 sources (Figure 7).



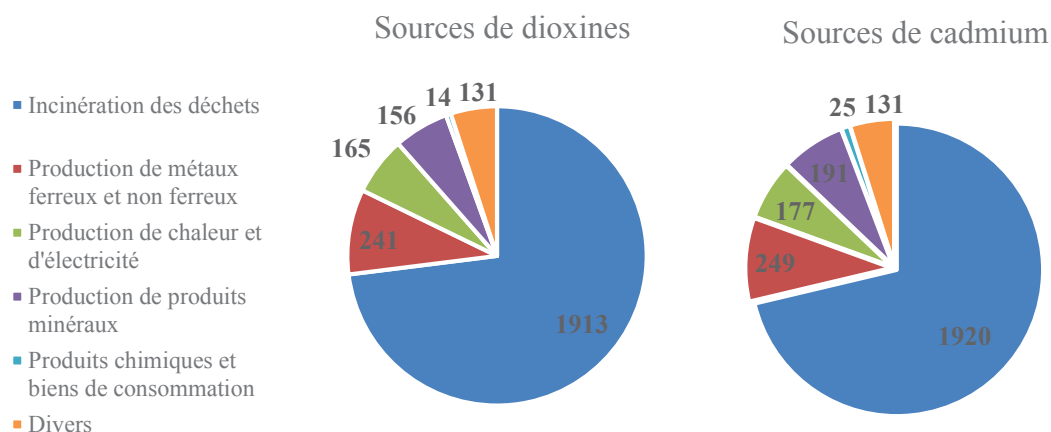


Figure 7 – Répartition par groupes des sources de dioxines et de cadmium

Si la répartition en termes de nombre de sources par polluant reste peu modifiée, la répartition des émissions par groupes est considérablement modifiée. Dans le cas des dioxines, le groupe « incinération des déchets » était très largement le secteur le plus émetteur et, après une baisse importante entre 2005 et 2006, il est dépassé par le groupe « production de métaux ferreux et non ferreux » (Figure 8). Pour le cadmium, la répartition est différente (Figure 9). Le groupe « production de métaux » reste le contributeur principal aux émissions totales sur toute la période (variant de 56% à 74% entre 1990 et 2008). La contribution du groupe « Incinération des déchets » décroît sur la période, de 34% en 1990 à 6% en 2008, contrairement aux groupes « divers » et « production de chaleur » qui augmentent de 1% en 1990 jusqu'à 5% et 7% respectivement en 2008. Le groupe « production de produit minéraux » devient le deuxième contributeur à la fin de la période d'étude avec une contribution relative aux émissions total de 11% en 2008.

De manière similaire aux dioxines, on observe, une baisse du total des émissions de cadmium de 25255 kg en 1990 à 9340 kg en 2008, soit une baisse de 63% (Figure 10). Les incertitudes concernant nos estimations annuelles, évalué tel que décrit dans l'article 1, diminue très légèrement entre 1990 et 2008, passant de 63% à 60%.

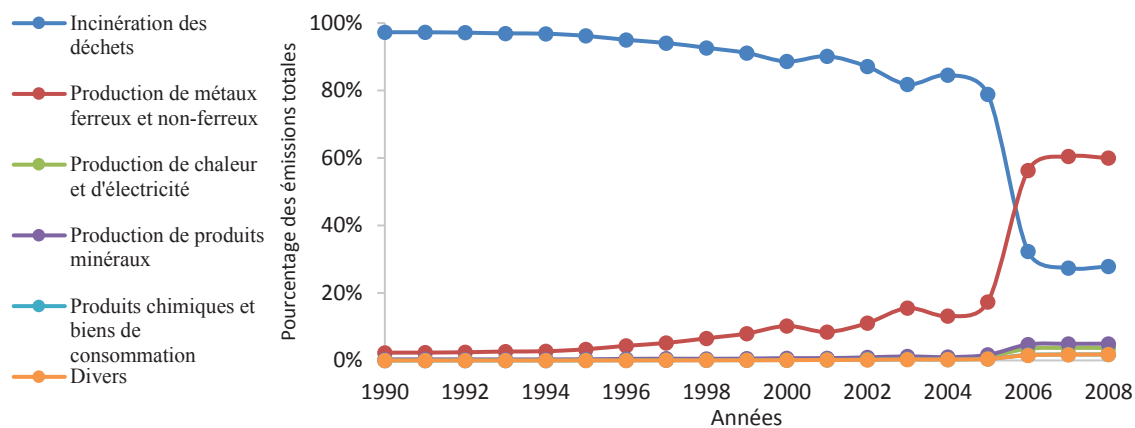


Figure 8 – Evolution en pourcentage des émissions annuelles de dioxines par groupes UNEP entre 1990 et 2008

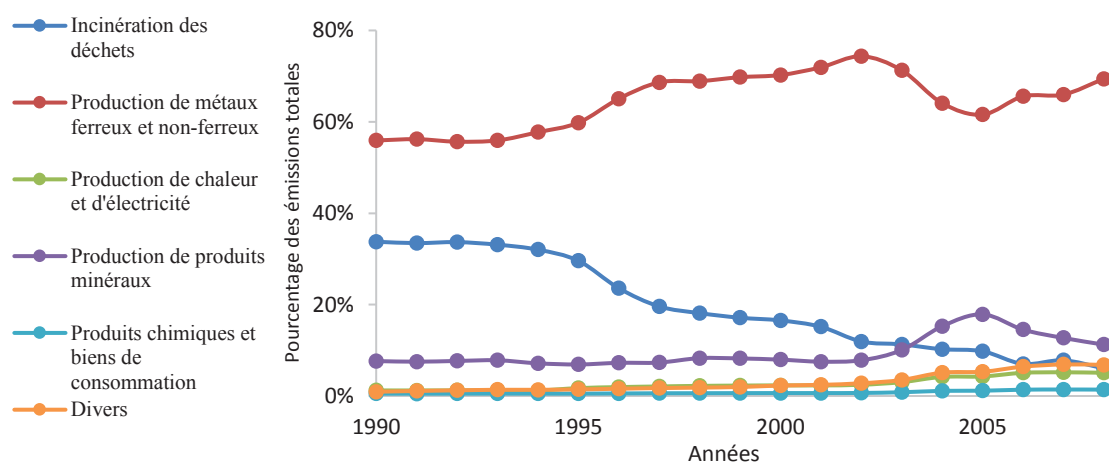


Figure 9 – Evolution en pourcentage des émissions annuelles de cadmium par groupe UNEP entre 1990 et 2008

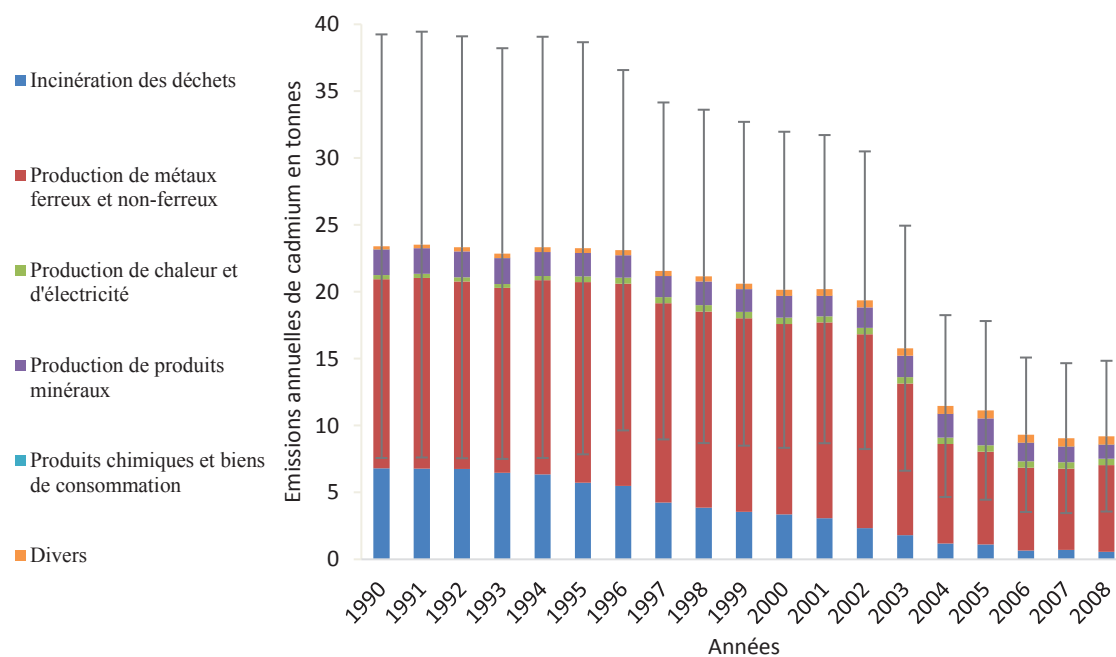


Figure 10 – Evolution des émissions total de cadmium (tonnes/an) par groupes UNEP entre 1990 et 2008

Dans le cadre d'une étude épidémiologique, il est intéressant d'étudier la variation des émissions des sources à l'échelle individuelle. On observe dans le tableau 3 qu'en 1990 malgré une domination du groupe « production des métaux » en termes d'émission globale, les valeurs médianes d'émission de cadmium par source sont plus élevées pour les groupes de l'incinération des déchets, de la production de produits minéraux et « divers ». Ces résultats illustrent l'intérêt de créer une base de données multi-sources avec une caractérisation à l'échelle individuelle. Dans le cas des groupes incinération des déchets, production de métaux ferreux et non ferreux et production de produit minéraux, ces valeurs médianes diminuent fortement et se stabilisent autour de 1.5 kg/an par sources en 2008. La baisse des émissions est plus légère pour le groupe « divers » (crématorium), passant de 6.7 à 4.1 kg/an par sources car la baisse des émissions s'est conjuguée avec une augmentation du nombre de crémations annuelle. Il est important de noter que les valeurs de ce groupe sont à prendre avec précaution due à la très forte incertitude liée à la collecte de données (nombre de crémations souvent non communiqué et difficulté d'obtenir des informations sur les techniques d'épuration des fumées mis en place) et à l'incertitude sur la valeur des facteurs d'émissions (très forte variabilité observée dans la littérature). La technologie assez rudimentaire utilisée dans les crématoriums implique l'application de facteurs d'émissions élevés (CITEPA 2017). Après 2008, de nouvelles réglementations plus strictes ont été mises en place pour les crématoriums (facteur 10 d'écart entre les facteurs d'émission avant et après mise en place de filtres).

Tableau 3 – Evolution des émissions annuelles médianes des sources émettrices de cadmium (kg/m<sup>3</sup>) par groupe entre 1990 et 2008

	Incinération des déchets	Incinérateur hospitalier	Production de métaux ferreux et non ferreux	Production de chaleur et d'électricité	Production de produits minéraux	Produits chimiques et biens de consommation	Divers
1990	5.21	0.86	3.86	0.18	5.98	1.07	6.78
1991	5.29	0.86	3.78	0.18	5.45	1.07	6.05
1992	5.60	0.86	3.72	0.18	5.00	0.01	5.28
1993	5.51	0.86	3.53	0.19	4.69	0.01	5.28
1994	5.48	0.86	3.69	0.11	3.88	0.01	5.12
1995	5.07	0.86	3.73	0.19	3.67	0.01	5.12
1996	5.24	1.94	3.64	0.19	3.57	0.01	5.10
1997	4.53	1.94	3.50	0.19	3.38	0.01	5.11
1998	4.58	1.94	3.50	0.18	3.17	0.01	5.10
1999	4.78	3.09	3.50	0.18	3.08	0.01	5.11
2000	5.26	1.17	3.50	0.16	2.95	0.01	4.75
2001	5.46	1.17	3.45	0.14	2.75	0.01	4.63
2002	4.42	1.17	3.50	0.15	2.72	0.01	4.69
2003	4.24	1.17	3.09	0.14	2.76	0.01	4.63
2004	3.20	1.17	2.35	0.14	2.37	0.01	4.59
2005	3.05	0.56	2.04	0.14	1.89	0.01	4.57
2006	1.77	0.72	1.74	0.09	1.87	0.01	4.37
2007	2.04	0.46	2.04	0.08	1.67	0.01	4.10
2008	1.53	0.46	1.68	0.08	1.67	0.01	4.10

Une comparaison de nos estimations avec un échantillon de mesures à l'émission (409 mesures réparties sur environ 100 sources soit 4% du nombre total des sources de cadmium recensées) a été réalisée et est détaillée dans l'article n°2 (chapitre 3). Les résultats sont meilleurs que pour les dioxines, avec une valeur moyenne estimée (OMINEA) par source de 47 kg/an pour une moyenne des valeurs mesurées de 49 kg/an. Environ 60% des estimations ont un écart avec les émissions mesurées compris entre  $\pm 100\%$ . Sur les 409 mesures, seulement 20 sont surestimées de plus d'un facteur 3. Néanmoins la corrélation mesure/estimation reste faible ( $r_s = 0,41$ ) et similaire à celle observé pour les dioxines (article n°1).

Nous avons observé, en début de chapitre, que l'écart entre émissions mesurées et émissions estimées était plus important lorsque les facteurs d'émissions de l'EMEP étaient utilisés. Néanmoins, les mesures de cadmium, sont limitées, réparties sur quelques sources et disponibles sur une courte période temporelle. On ne peut donc pas considérer qu'elles sont représentatives de la qualité des estimations en particulier pour la période 1990-2000, qui est la période où les émissions sont les plus fortes. La problématique est similaire pour les dioxines. C'est pourquoi, nous avons tout de même comparé nos estimations sur la période 1990-2008 avec les estimations fournies par l'inventaire des émissions EMEP pour la France, pour quelques catégories parmi les plus émettrices (Tableau 4).

Tableau 4 – Comparaison des émissions de cadmium (kg/an) estimées dans notre étude avec les incertitudes associées en pourcentage ( $\pm$ ) et estimée avec les facteurs d'émission de l'EMEP pour 4 catégories entre 1990 et 2008.

Année	Incinérateur de déchets municipaux			Incinérateurs de déchets dangereux			Incinérateurs de déchets médicaux			Production de fer et d'acier et fonderies		
	UNEP	$\pm$	EMEP	UNEP	$\pm$	EMEP	UNEP	$\pm$	EMEP	UNEP	$\pm$	EMEP
1990	6357	60%	1734	456	63%	460	1712	100%	2181	9055	61%	5594
1991	6300	61%	1795	478	61%	471	1712	100%	1662	8977	61%	5630
1992	6234	60%	1844	531	63%	494	1712	100%	1284	8881	61%	5293
1993	5931	57%	1839	537	62%	487	1712	100%	1008	8798	61%	5024
1994	5782	57%	1690	569	62%	548	1710	100%	801	9646	62%	5689
1995	5175	52%	1456	544	63%	539	1710	100%	736	10245	61%	5844
1996	4951	51%	1281	531	63%	536	4	100%	623	10451	60%	5751
1997	3794	49%	947	456	63%	459	4	100%	504	10343	60%	6571
1998	3414	49%	780	443	62%	438	4	100%	269	10176	61%	6501
1999	3148	47%	548	400	62%	404	3	100%	241	10062	61%	6194
2000	3009	45%	520	344	64%	369	2	100%	192	9958	61%	6556
2001	2795	43%	281	285	63%	315	2	100%	144	9838	61%	6198
2002	2066	41%	190	257	58%	261	2	100%	96	9866	61%	6113
2003	1582	41%	113	215	54%	191	2	100%	62	6799	63%	4326
2004	1085	42%	45	100	53%	132	1	100%	37	3403	65%	2147
2005	986	41%	37	116	49%	157	1	100%	4	2998	66%	1641
2006	571	39%	17	92	47%	152	1	100%	4	2702	67%	1221
2007	657	38%	20	60	50%	121	1	100%	2	2510	67%	1063
2008	498	40%	12	67	47%	144	1	100%	1	2514	67%	1327

Le plus gros écart est observé pour la catégorie des incinérateurs de déchets municipaux avec 4600 kg de différence en 1990 (soit une erreur d'un facteur 3.5). Cette différence se réduit progressivement au fil des ans : 2600 kg en 1999, 1040 kg en 2004 et 486 kg en 2008. C'est d'ailleurs le cas pour les 4 catégories présentés dans le Tableau 4.

Pour la catégorie des incinérateurs de déchets industriels, les estimations des deux inventaires sont proches avec des différences ne dépassant pas les 100 kg sur la période. Cette meilleure corrélation entre les deux inventaires pour la catégorie des incinérateurs de déchets industriels par rapport à celle des incinérateurs de déchets municipaux a plusieurs explications. Tout d'abord, nous avons inventorié un grand nombre de petits incinérateurs de déchets municipaux vétustes, localisés principalement dans les campagnes. La plupart n'étaient pas équipés de systèmes d'épuration et fonctionnaient de manière discontinue sans réel contrôle des déchets en entrée. Les facteurs d'émissions recommandés, par le Toolkit dans ces situations, conduisent à des estimations de fortes émissions même avec des taux d'activité faibles. De par leurs tonnages réduits et leurs faibles durées de fonctionnement (pour la plupart fermés avant 1996) ils ont pu échapper (ou être négligés) par les inventaires nationaux précédents. De plus, le manque de données techniques et les volumes traités par ces incinérateurs génèrent de très fortes incertitudes concernant les émissions. Contrairement aux incinérateurs municipaux, les incinérateurs de déchets dangereux sont identifiés depuis très longtemps comme des sources potentielles de pollution et sont bien mieux référencés, encadrés et équipés. Ce qui explique la proximité des valeurs des deux inventaires.

Pour la catégorie de l'incinération de déchets hospitaliers, le problème est similaire : un nombre de sources inventoriées probablement supérieures aux autres inventaires (n=1442 en 1990), et un manque de données pour estimer les rejets avec une meilleure précision. Les valeurs des deux inventaires sont assez proches en début de période puis à partir de 1995, nos estimations baissent brutalement car nous avons considéré que ces sources fermaient suite à la mise en application de l'arrêté du 21 janvier 1991. À partir de 2005, les valeurs des deux inventaires redeviennent très proches. Cette problématique de la contribution des incinérateurs hospitaliers aux émissions sur la période 1990-2000 est évoquée comme une limite importante par Quass et al (2005) dans son inventaire des émissions de dioxine entre 1985 et 2005.

Enfin pour la catégorie de la production de fer et d'acier et fonderies, qui est la catégorie la plus émettrice de cadmium, nos estimations sont environ 2 fois supérieures aux estimations EMEP. Cet écart relatif reste constant sur la période, mais la différence absolue diminue (due à la baisse des émissions). Cet écart est probablement dû à la différence entre les facteurs d'émissions utilisés. De plus, le groupe de la production de métaux ferreux et non-ferreux est le groupe avec les émissions les plus difficiles à quantifier due à la complexité et à la variété des processus mis en place par les industriels. À titre d'exemple la qualité des métaux en entrée de processus peut faire varier les facteurs d'émission d'un facteur 20 (UNEP 2013).

## 4. Conclusion

Nous avons créé la première base de données française contenant les émissions annuelles des sources de dioxines et de cadmium à l'échelle individuelle des sources en France entre 1990 et 2008, ainsi que des informations techniques et leurs localisations (le plus souvent à la cheminée). Cette base de données sera utilisée dans des études épidémiologiques portant sur le risque de cancer du sein associé à l'exposition à la dioxine et au cadmium, mais pourra être étendu à d'autres polluants. Ce travail illustre les qualités et les défauts d'un travail de reconstitution des émissions historiques dans le cadre d'une étude épidémiologique.

En utilisant le Toolkit de l'UNEP et la base OMINEA, nous avons pu inventorier 2620 et 2700 sources d'émission de dioxines et de cadmium respectivement. Le travail de recueil de données techniques a permis de caractériser l'évolution des technologies d'épuration des fumées pour chaque source sur la période 1990-2008 et ainsi d'estimer les émissions annuelles et les incertitudes associées. La comparaison avec un faible échantillon de mesures de l'émission montre une corrélation modérée avec les estimations ( $r_s$  autour de 0.40). Pour les deux polluants, l'émission médiane estimée par sources est inférieure à l'émission médiane mesurée (-32% pour le cadmium, -40% pour les dioxines). Les différences entre nos estimations et les mesures peuvent s'expliquer par un certain nombre de facteurs. Premièrement, les mesures proviennent de la base de données de l'IREP qui recueille les déclarations d'émissions des industriels. Deuxièmement, les mesures de dioxines n'ont été effectuées qu'une fois par an (au mieux) et peuvent ne pas être représentatives des émissions moyennes annuelles. Pour ces raisons, les mesures d'émissions ne peuvent pas être considérées comme pleinement représentatives des émissions réelles survenant sur une année entière. De plus, les mesures sont disponibles sur la période 2003-2008 (qui correspond aux plus faibles émissions sur la période d'étude) et ont été quasi exclusivement effectuées sur des incinérateurs de déchets municipaux et des installations métallurgiques, qui ne sont pas représentatives de la variété des sources inventoriées. C'est pourquoi il est difficile d'extrapoler ces résultats pour estimer la qualité des estimations, en particulier pour la décennie 1990-2000.

On constate que, pour ces deux polluants, les plus gros écarts entre nos estimations et les estimations d'autres inventaires, sont observés pour le groupe de l'incinération des déchets et particulièrement sur la période 1990-2000. En comparant deux inventaires des émissions de dioxine portant sur une période similaire, Breivik et al. (2006) a observé des écarts de plusieurs ordres de grandeur, similaire à nos résultats lors de la comparaison avec l'inventaire EMEP. Globalement, les différences observées durant les comparaisons aux échelles départementale et nationale se situaient dans la gamme d'incertitude. Les estimations de l'incertitude sont un élément essentiel des inventaires des émissions, mais peu d'étude reporte explicitement ces incertitudes. Les ordres de grandeur des incertitudes quantifiée dans les autres études de la littérature sont similaires aux nôtres (Breivik et al. 2006; Fiedler 1999; Pulles et al. 2006; Quina et al. 2011).

Les différences observées lors de la comparaison avec d'autres inventaires sont dues au manque de données sur les sources, à la différence sur le facteur d'émission utilisé et à l'absence de connaissances sur le nombre précis de sources. Les nombreux petits incinérateurs avec émissions élevées, ainsi que les nombreux incinérateurs de déchets hospitaliers, ont peut-être échappé aux inventaires antérieurs. Il faut également noter que sur cette période (1990-2000), un certain nombre de sources furent fermées en raison du non-respect des normes en vigueur (affaire des œufs contaminés dans le Nord-Pas de Calais en 1998, incinérateur de Gilly sur Isère en 2000, usine Recytech en 1998, etc.). Ces cas illustrent les incertitudes sur les estimations des émissions à cette période (Webster and Connett 1998). Dans le cas des dioxines, la problématique est encore plus complexe car les émissions sont particulièrement liées à l'installation de systèmes spécifiques de réduction des émissions de dioxines. Selon la mise en place, ou non, d'une telle installation, les émissions peuvent varier d'un facteur 100. D'une manière générale, les incertitudes diminuent mais restent fortes dans ces inventaires ainsi que dans les études récentes. Une étude sur les émissions de dioxines en Europe centrale pour l'année 2000 a montré que les estimations du total des émissions, avec un intervalle de confiance à 90%, variaient autour de la valeur estimée d'un facteur 2-3 inférieur et un facteur 4 supérieurs (Pulles et al. 2006).

Nos résultats montrent une très forte variabilité des émissions à l'échelle des sources, même pour des sources issues d'une même catégorie, ainsi qu'une forte variabilité temporelle. Ceci illustre l'intérêt de la reconstitution des émissions historiques en se basant sur les technologies mises en place à l'échelle individuelle. Les émissions établies en se basant sur les mesures directement effectuées auprès des sources peuvent induire un biais car elles sont effectuées sur un sous-échantillon non représentatif de l'ensemble de sources.

Nous avons réussi à évaluer les rejets industriels avec une très grande précision spatiale (70% au minimum à la parcelle dont 43% à la cheminée), mais un des défauts majeurs de ce travail, outre le manque de données, reste la considération d'une émission continue dans l'année. De nombreuses études ont montré que les arrêts et redémarrages successifs des industries entraînent de fortes émissions ponctuelles (Cheruiyot et al. 2016). S'il est envisageable d'appliquer une modulation annuelle pour les sources liées au chauffage, il est extrêmement difficile voire impossible de caractériser les rythmes de fonctionnement de chaque usine. Certaines fonctionnaient en continu, d'autres uniquement en semaines et un certain nombre ont subi des arrêts pour maintenance pendant de longues périodes. Ces paramètres viennent agrandir les incertitudes sur les émissions. Dans le cas d'études futures sur des pathologies pouvant être liée à des expositions sur des fenêtres temporelles plus réduites (expositions mensuelles ou journalières), obtenir une plus grande finesse temporelle dans l'évaluation des émissions de dioxine et de cadmium sera probablement nécessaire, mais difficilement réalisable.

La sollicitation par questionnaires des industrielles a joué un rôle clef dans la caractérisation des sources grâce à l'obtention d'informations précises et détaillées sur le fonctionnement des sources. Néanmoins, le faible taux de retour (environ 550 questionnaires envoyés par mail ou courrier pour 175 réponses) pose problème. Un certain nombre d'industriels ont refusé de nous répondre pour diverses raisons : manque de temps, données difficile d'accès, données perdues, manque de confiance dans l'étude, remise en question de la méthodologie, etc. De plus, entre et au sein des trois groupes qui représentent 95% des questionnaires envoyé (incinération de déchets, production de métaux, production de produits minéraux) les taux réponse sont très variable (respectivement 20%, 80% et 25%). Néanmoins, au sein de ces groupes, certaines catégories de sources ont des taux réponse élevées (incinération de déchets industriels spéciaux : 70%, incinération de boues d'épuration : 78%, production de ciment : 100% et production de chaux : 80%). Il est à noter que, si seule une confirmation été nécessaire, les industriels étaient contacté par téléphone (ce qui n'apparaît pas dans les statistiques). Enfin sur les plus de 2600 sources recensées, environ 1700 n'étaient plus en activités après 2008, ce qui a compliqué fortement la récolte d'informations qui a commencé en 2014.

Nos résultats illustrent la réussite des politiques publiques visant à réduire les émissions atmosphériques, les efforts réalisés par les industriels pour réduire leurs émissions devant être salués. Il est à noter que cette diminution des émissions des sources industrielles durant les décennies 1990 et 2000 implique de changer les méthodologies et de ne plus prendre en compte uniquement les sources industrielles comme ce fut le cas par le passé dans de précédentes études épidémiologiques (Cordier et al. 2010; Fabre et al. 2008; Floret et al. 2003; Pronk et al. 2013; Viel et al. 2008; Zambon et al. 2007). La contribution de certaines sources non-industrielles (secteur tertiaire/résidentielle, trafic routier) est devenue non négligeable sur les périodes les plus récentes et doit être prise en compte. La non-prise en compte de ces sources pourrait entraîner des biais de classifications non-différentielles des sujets dans les études épidémiologiques. Pour quantifier les déterminants de l'exposition historique à la dioxine et au cadmium et leurs évolutions, le chapitre 3 présentera les résultats d'un travail de modélisation des émissions multi-sources (industrielles, trafic, chauffage, divers) réalisé sur le Grand Lyon entre 1990 et 2008, grâce au modèle SIRANE.





### **CHAPITRE III. EXPOSITION ATMOSPHERIQUE AUX DIOXINES ET AU CADMIUM**

*Evaluation de l'exposition atmosphérique chronique aux dioxines et au cadmium dans la métropole du Grand Lyon entre 1990 et 2008.*

## 1. Introduction

La pollution de l'air en milieu urbain est un mélange complexe de divers polluants provenant d'une grande variété de sources. Plusieurs études épidémiologiques récentes ont analysé l'association entre la pollution de l'air extérieur et le risque de cancer du sein, mais les résultats restent incohérents. Pour les polluants atmosphériques liés au trafic (oxydes d'azote, particules et hydrocarbures aromatiques polycycliques), des études cas-témoins ont mis en évidence des associations positives (Bonner et al. 2005; Hystad et al. 2015; Mordukhovich et al. 2016; Nie et al. 2007) alors que les études de cohortes prospectives n'ont pas rapporté d'associations significatives (Andersen et al. 2017; Hart et al. 2016; Raaschou-Nielsen et al. 2011; Reding et al. 2015).

Pour les polluants à effets perturbateur endocrinien, certaines études ont également observé des associations positives avec le risque de cancer du sein, notamment pour les PCB (Lauby-Secretan et al. 2013), la 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (Dai and Oyana 2008b; Fabre et al. 2008c), le Benzo[a]pyrene (B[a]P) (Nie et al. 2007; Salehi et al. 2008), ou encore le cadmium (Julin et al. 2012; Liu et al. 2015b; Salehi et al. 2008).

L'incohérence des résultats de ces études pourrait s'expliquer par des limites méthodologiques qui ne permettent pas d'évaluer de manière complète l'impact de l'exposition atmosphérique. Outre les limitations liées à l'absence de connaissances sur les facteurs de confusions du cancer du sein (IMC, éducation, alcool, etc.) ou sur l'histoire résidentielle des sujets, les incertitudes liées à l'évaluation de l'exposition historique aux polluants atmosphériques reste un frein majeur à ce type d'étude et peuvent induire une mauvaise classification des sujets et donc une incertitude dans les risques estimés (Basagana et al. 2013; Cordioli et al. 2013c; Huang and Batterman 2000).

Lorsque l'on s'intéresse aux polluants que l'on pourrait qualifier de «classiques», mesurés par un grand nombre de stations de surveillance de la qualité de l'air, comme le NO<sub>2</sub> ou les particules fines, l'évaluation des performances des modèles est réalisée en comparant les résultats avec les données de mesure de concentration en air ambiant. C'est le cas par exemple des travaux de Vienneau et al, de Gulliver and Briggs et de Hoek et al (Gulliver and Briggs 2011; Hoek et al. 2001; Vienneau et al. 2009). Notons toutefois que, même dans ces études, les auteurs ne sont pas en mesure d'estimer les incertitudes liées aux émissions de sources, car elles sont fournies par des inventaires nationaux d'émissions. Lorsque les études se concentrent sur des polluants comme les dioxines, les PCB et les métaux lourds, l'évaluation des modèles d'expositions est plus complexe dû au manque de données de mesures de concentrations (à émissions et à l'air ambiant) et aux incertitudes sur la quantification des émissions.

En raison de ce manque de données de surveillance pour les dioxines, aucune étude n'a réalisé de comparaison entre des concentrations modélisées et mesurées en air ambiant dans une zone urbaine. Les études antérieures portant sur l'exposition atmosphérique aux dioxines et le risque de cancer ont principalement utilisées des indicateurs d'exposition simple basés sur la distance source-sujet (Bidgoli et al. 2010; Pronk et al. 2013; Visser et al. 2004). Comme montré par Cordioli et al. (2013), ces indicateurs, considérés comme un moyen acceptable de pallier au manque de données de mesure, sont cependant associés à des erreurs dans le classement des sujets (Cordioli et al. 2013). Ce n'est que récemment que des études ont commencé à utiliser des modèles de dispersion (Cordier et al. 2010; Fabre et al. 2008; Ranzi et al. 2011; Viel et al. 2008).

Pour le cadmium, à notre connaissance, il n'existe pas de comparaisons entre les concentrations modélisées et mesurées dans l'air ambiant publiées dans la littérature, à l'exception de l'étude de Liu et al. (2015) qui se base sur les résultats des modélisations réalisées par l'US EPA pour estimer les expositions des sujets de la cohorte (Liu et al. 2015). Néanmoins, selon Garcia et al (2013), les résultats des modélisations de l'EPA, utilisée par Liu et al, ont montré des sous-estimations importantes par rapport aux mesures de métaux lourds en air ambiant et notamment de cadmium (Garcia et al. 2014).

### *Objectif et méthodes*

L'objectif de ce chapitre est de comprendre comment les multiples contributeurs à la pollution atmosphérique aux dioxines et au cadmium affectent l'exposition humaine, et comment les niveaux d'exposition ont évolué dans un contexte urbain au cours des vingt dernières années.

Dans un premier temps, nous avons déterminé les émissions atmosphériques de dioxines et de cadmium des différents types de sources (industrielle, liées chauffage, liées au trafic de la circulation ou diffuse), ainsi que les variations de ces émissions entre 1990 et 2008 dans la métropole lyonnaise. Les émissions des sources industrielles ont été estimées en utilisant la méthodologie développée durant le travail d'inventaire et de caractérisation effectuée précédemment (chapitre 2). Pour les dioxines, nous avons utilisé le Toolkit créé par l'UNEP qui fournit un jeu de facteurs d'émission validés au niveau européen et déjà utilisés dans plusieurs inventaires européens (Quina et al. 2011; Saral et al. 2014). Concernant le cadmium, nous nous sommes basé sur le jeu de facteurs d'émission de l'OMINEA qui a été créé spécialement par le CITEPA pour évaluer les rejets français (CITEPA 2017). Concernant les autres sources d'émission (trafic, chauffage, diffuse), nous avons tiré profit d'un ensemble de données d'émissions fourni par l'ATMO AURA et issus de leur cadastre des émissions (Atmo Auvergne-Rhône-Alpes 2016).

Dans un second temps, nous avons estimé la variabilité temporelle et spatiale des niveaux de concentrations à l'adresse d'un sous-échantillon de femmes de la cohorte E3N vivant dans cette zone. A cette fin, nous avons utilisé le modèle SIRANE pour reconstruire les champs de concentration au niveau du sol et les niveaux d'exposition des sujets résidant dans le Grand Lyon (zone de 1020 km<sup>2</sup>), sur une période de 18 ans avec une résolution spatiale (25 m) et temporelle (1h) fine. Contrairement aux études précédentes (Fabre et al. 2008; Ranzi et al. 2011), nous disposions de mesures pour les émissions des sources et pour les concentrations dans l'air ambiant. Pour évaluer la performance du modèle SIRANE, les résultats ont été comparés aux critères définis par Hanna & Chang pour un modèle de dispersion en milieu urbain (Hanna and Chang 2012). De plus, grâce au travail mené précédemment (Chap. 2), nous disposions des incertitudes sur les estimations des émissions des sources industrielles et nous avons pu évaluer leurs impacts sur les concentrations estimées.

## 2. Article n°2

### **Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon metropolitan area (France)**

Thomas Coudon<sup>1,2</sup>, Hassan Hourani<sup>1</sup>, Chi-Vuong Nguyen<sup>3</sup>, Elodie Faure<sup>1</sup>, Francesca Mancini<sup>4</sup>, Béatrice Fervers<sup>1</sup> and Pietro Salizzoni<sup>3</sup>

<sup>1</sup> Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France

<sup>2</sup> Université Claude Bernard Lyon 1, Villeurbanne 69100, France

<sup>3</sup> Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully 69134, France

<sup>4</sup> Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.

#### *Reference:*

Coudon T, Hourani H, Nguyen C, Faure E, Mancini FR, Fervers B, Salizzoni P. 2018. Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon metropolitan area (France). *Environment International* 111:177–190; doi:10.1016/j.envint.2017.11.027.

# **Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon Metropolitan Area (France)**

*T. Coudon<sup>1,2</sup>, H. Hourani<sup>1</sup>, C. Nguyen<sup>3</sup>, E. Faure<sup>1</sup>, FR.Mancini<sup>4</sup>, B. Fervers<sup>1</sup>, P. Salizzoni<sup>3</sup>*

*<sup>1</sup> Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France*

*<sup>2</sup> Université Claude Bernard Lyon 1, Villeurbanne 69100, France*

*<sup>3</sup> Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully 69134, France*

*<sup>4</sup> Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.*

## Abstract

Identifying the long-term effects of airborne pollutants requires the computation of the spatial and temporal variability of their concentration in air to estimate the exposure of the population. To estimate exposure levels of individuals in a breast cancer case control study nested in a national prospective cohort, we determine here the differential impact of a variety of cadmium and dioxin sources on urban air over a large urban area and over a period of almost 20 years.

To that end, we couple an emission model, to estimate dioxin and cadmium atmospheric annual releases, with an urban dispersion model in order to compute pollutant concentration fields at a fine temporal (1 hour) and spatial (25 meters) resolution. The reliability of the modelling chain is compared to two types of measurement: i) localized industrial emissions and ii) dioxin and cadmium air monitoring data (from 2007 to 2008), collected at a fixed station, placed in the city centre, as well as at three mobile short-term dioxin monitoring stations, located in the suburbs (the latter providing dioxin data, only). Comparisons between measured and estimated emissions show non-negligible difference, with a correlations for dioxin ( $r_s = 0.42$ ) and cadmium ( $r_s = 0.41$ ). Despite this, mean values between estimated emissions and emission measurements are close to each other, in particular for cadmium. Weekly average modelled concentrations show an overall good agreement with weekly average measured concentrations in spring and summer but are generally lower than monitored data in winter due to peak concentrations from diffuse sources representing an important proportion of emissions in 2007/2008. The model provides better results for cadmium than for dioxin. Despite the relevant errors in the model predictions, the model meets the validation criteria, defined by Chang and Hanna for an urban dispersion model.

Simulation scenarios of air pollutant concentrations, reconstructed over the last 20 years, show the effects of the variability of the pollutant sources over time with decreasing levels of dioxin and cadmium concentrations in air. This is primarily due to the reduction in localized industrial releases, which results in a general trend of homogenization of the exposure of the population.

The model further allows us to dissociate the contribution of different types of pollutant sources on the population exposure. The impact on local concentrations due to industrial emissions, which were originally responsible for the major impact on air quality, is shown to drop over the years by 99% and 92% for dioxin and cadmium, respectively. Today, the major contributions are due to diffuse miscellaneous sources in the case of dioxin and to traffic-related emissions for cadmium.

Average modelled concentrations at the study subjects' residential locations range from 0.10 to 5.75 fg-TEQ/m<sup>3</sup> for dioxin and 0.10 to 1.6 ng/m<sup>3</sup> for cadmium and are comparable with data from the literature.

The study results will be essential to increase the accuracy of the assessment of long-term airborne dioxin and cadmium exposure and improve the results of epidemiological studies.



# 1. Introduction

Long-term exposure to air pollution has been consistently associated with adverse health effects and reduced life expectancy (Pope et al., 2009). Health problems related to atmospheric pollution increase in urban areas, due to the large number of sources of pollutant. Urban air pollution is a mixture of multiple pollutants originating from a large variety of sources and includes various substances classified by the International Agency for Research on Cancer (IARC) as carcinogenic or probably carcinogenic to humans (Loomis et al., 2013). Among these, pollutants with endocrine-disrupting properties and heavy metals have been specifically investigated due to their adverse outcomes (Tchounwou et al., 2012; Kabir et al., 2015). Typical examples of these kinds of pollutant are dioxins and cadmium.

Dioxins are environmental and persistent organic pollutants, with lipophilic properties, which can contaminate food and accumulate in adipose tissues (Arisawa et al., 2003). Cadmium is a heavy metal with endocrine-disrupting properties. The most toxic dioxin congener, 2,3,7,8-PCDD/F, and cadmium have been classified as human carcinogens (Group 1) by the IARC (Baan et al., 2009; Nawrot et al., 2006). Assessment of cadmium and dioxin exposures is therefore warranted, although it is complicated by the latency in the disease development, especially for cancer (Ries et al., 2006).

Breast cancer has been positively linked with exposure to IARC Group 1 carcinogens with endocrine-disrupting properties, such as 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (Dai and Oyana, 2008; Fabre et al., 2008), and cadmium (Cd) (Salehi et al., 2008; Liu et al., 2015; Julin et al., 2012), but current evidence remains inconsistent. Some of the inconsistencies observed in epidemiological studies may be due to methodological limitations, such as a lack of statistical power due to small study populations, or a failure to control for individual breast cancer risk factors. Furthermore, the lack of a past residential history for the study subjects and use of historical exposure estimates may have resulted in exposure misclassification, which is likely to contribute to imprecise risk estimates.

The accurate assessment of airborne environmental pollutant exposures and their variation in space and time represents a major challenge in epidemiological studies. Despite the presence of some natural sources, such as forest fires and volcanic eruptions, the large majority of dioxin and cadmium atmospheric releases are due to anthropogenic emissions (Quaß et al., 2000; Nriagu, 1989). These are generally related to incomplete combustion that occurs in industrial processes (chemicals manufacturing, combustion processes, metallurgical processes, paper and pulp processing), domestic heating and traffic (Pacyna and Pacyna, 2001; Lohmann and Jones, 1998).

Measurements of airborne dioxin and cadmium concentrations, as well as the flow rates of their atmospheric emissions, are sparse, especially before the year 2000. Even though the amount of monitoring data has progressively increased in the last few years (Quaß et al., 2004), it remains largely insufficient for estimating the spatial variability of the concentrations. Therefore, previous studies investigating the impact of dioxin exposure on health mainly relied on the use of subject-source

distance as the key parameter for the exposure evaluation (e.g. Pronk et al., 2013; Langlois et al., 2009; Zambon et al., 2007). However, the use of these kinds of proxies, seen as a means to overcome the lack of measurement data, has been shown to be subject to substantial misclassification in epidemiology studies (Cordioli et al., 2013). Therefore dispersion modelling is nowadays considered as a more pertinent approach for exposure assessment with a high spatial resolution. Among the studies based on dispersion modeling exposure assessment, only few focused on the link between dioxin or cadmium exposure and breast cancer (Ranzi et al., 2011). Fabre et al. (2008) suggested that exposure to dioxins is associated with an increased risk of breast cancer, and Liu et al. (2015) observed an elevated risk of hormone receptor-negative tumors with higher exposure to cadmium compounds. However, current evidence remains inconsistent and poor exposure assessment has been advanced to explain the inconsistency of results in published studies (Cordioli et al., 2013). With the exception of the study by Liu et al (2015), no previous studies have evaluated the agreement of modeled data with air monitored data. One major reason for this is the lack of measurement data for dioxin and cadmium emissions, in particular for earlier years and, as mentioned above, the lack of monitoring data in ambient air.

To inform the exposure assessment in a case-control study nested in the French national prospective women's cohort of teaching and allied professionals (E3N) on the effects of long-term air pollution exposure (including airborne dioxin and cadmium exposure) on breast cancer risk (Clavel-Chapelon et al., 1997), the present study aims to i) determine the contribution of different source categories (industrial, heating, traffic and diffuse) on air concentrations of dioxin and cadmium, as well as their temporal variation over the 1991-2008 period in the Lyon metropolitan area (LMA) and ii) estimate the temporal and spatial variability of the exposure levels of the E3N women living in this area.

For this purpose, following usual methods (e.g. Fabre et al., (2008); Ranzi et al., 2011), we combine an emission model with an atmospheric dispersion model. In doing so, we take advantage of i) a complete emission set and ii) available measurements of both pollutant emissions from industrial sources and measurements of pollutant concentrations in ambient air. The latter allows us to evaluate the reliability of the two models through comparison with measurement data and to estimate the uncertainty of the industrial emissions.

The modelling chain is subsequently used to reconstruct i) historical trends of pollutant concentration fields at ground level and ii) the relative exposure levels in a large urban area (1020 km<sup>2</sup>), over a period of 18 years at a fine temporal (1h) and spatial (25m) resolution.

## **2. Materials and Methods**

The assessment of dioxin and cadmium concentrations within a selected area and over a specific period (2.1) requires two main steps: i) evaluation of the spatial and temporal variability of the intensity of the pollutant sources (2.2), and ii) simulation of the effect of atmospheric turbulence on pollutant dispersion (2.3).

### **2.1. Study population, area and period**

The E3N study (Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale) is an ongoing prospective cohort involving 98,995 French female volunteers, born between 1925 and 1950, who are members of a national teachers' health insurance plan (Clavel-Chapelon et al., 2015). E3N participants were enrolled in 1990 and they completed self-administered questionnaires every 2 to 3 years on cancer risk factors and residential history. For the present study, we selected 421 participants who lived in a 30km x 34km area in the LMA in 1990. The LMA is the urban area with the second largest population in France (1.4 million people) around Lyon, the third largest city in the country. Figure 1 shows the geocoded residential addresses of the 421 women residing in the LMA in 1990 using ESRI ArcGis software (Faure et al., 2017) and the population density of the LMA. For the purpose of the present study, we consider the residential addresses to remain unchanged over the whole study period. The choice of the LMA is principally based on the availability of measurement of pollutant concentrations in ambient air, notably from a fixed station providing data on cadmium and dioxin concentrations over two years (2007-2008), and three mobile stations, providing 8 dioxin measurements per year. In France, this kind of measurement is exclusively available for the LMA, thus explaining the choice of the study area. Another point of interest is given by the historical variety of source categories in this area, and namely by the presence of several industrial sources.

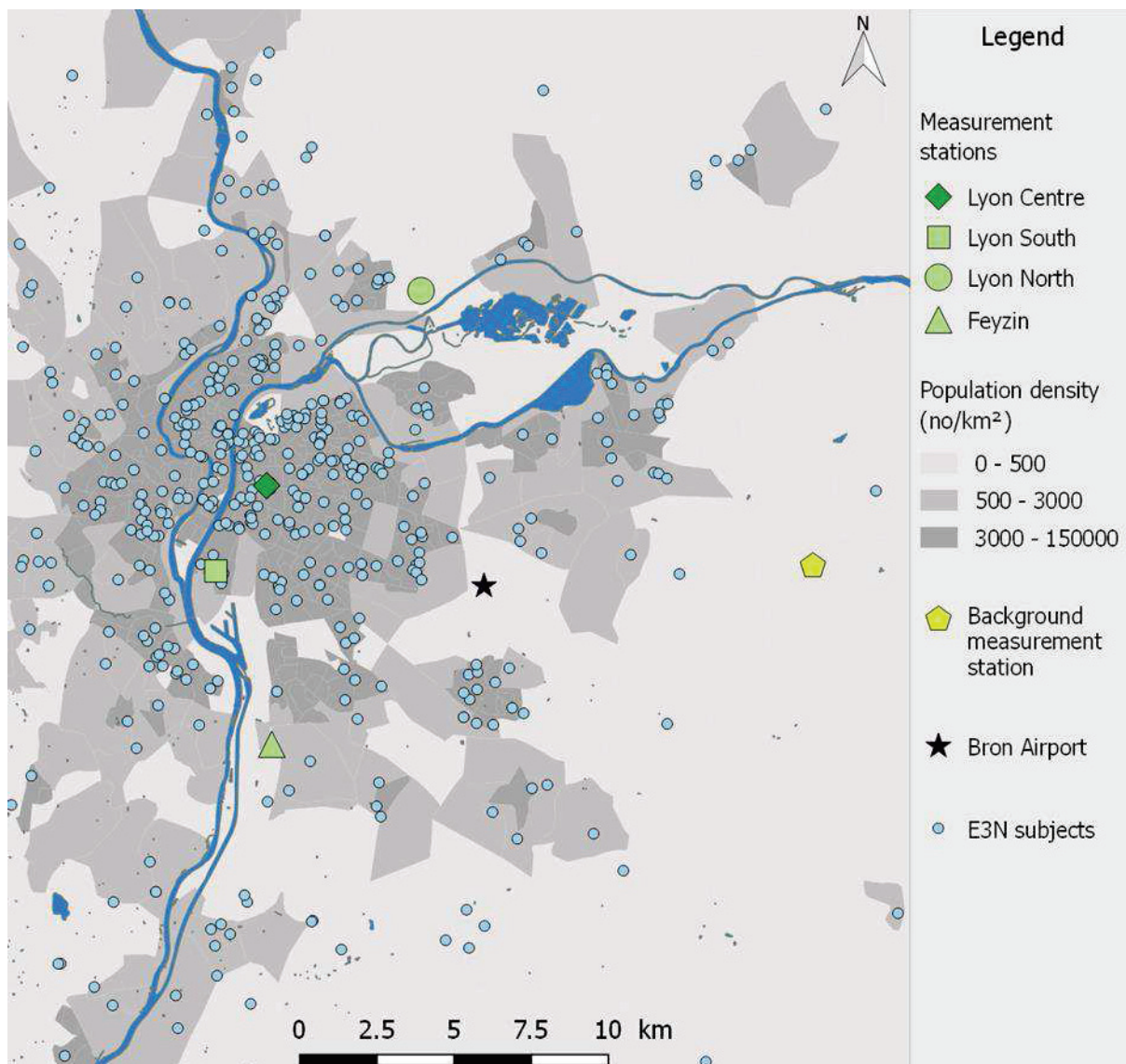


Figure 1 – The Lyon metropolitan area (LMA) showing the distribution of the E3N subjects, the population density (no/km<sup>2</sup>) in 2008 and locations of the measurement stations for long-term dioxin and cadmium concentrations (Lyon Centre), short-term dioxin concentrations (Lyon South, Lyon North and Feyzin), background PM10 concentrations and meteorological data (Bron airport).

## 2.2 Dioxin and cadmium emissions

To assess the historical dioxin and cadmium concentrations, the dispersion model requires a complete set of emission data (localization, intensity, type) for each year. We distinguish three categories of sources: industrial sources (2.2.1), miscellaneous diffuse sources (2.2.2) and traffic sources (2.2.3). Measurements are available only for industrial emissions and for a limited period (2.2.4).

### 2.2.1 Industrial sources

Industrial dioxin and cadmium sources are identified through regional and national databases, namely the GEREP (annual emission report on pollutant and greenhouse gas releases), IREP (French National Registry of Pollutant Emissions) and the ICPE (Inventory of French classified facilities). This information is complemented with information from institutions, industrial unions and nationally recognized associations. Structured questionnaires were sent to the facilities identified to collect additional information on their technical characteristics and activity rate. Given the absence of direct measurement of dioxin and cadmium emissions during the early period of our study and in order to achieve a standardized estimate of the source contributions, we used emission factors provided by the Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases from UNEP (<https://toolkit.pops.int/>) for dioxin emissions and the OMINEA (Organization and Methods of the National Inventories of the Atmospheric Releases in France) for cadmium.

The Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases from UNEP was used to characterize and quantify dioxin releases, in a wide range of industrial sectors, based on their activity and technical characteristics (input materials, air pollution control system, and flue gas or chimney height) as well as their evolution over time. The emission factor, assessed for each facility, was coupled with the activity rate of the industry to obtain an estimate of annual dioxin release (g-TEQ/year) for each industrial source.

Cadmium emissions were estimated using emission factors provided by the French database, OMINEA, from the Inter-professional Technical Centre for studies of Air Pollution (CITEPA). This database uses the Selected Nomenclature for Air Pollution (SNAP) European classification. To take into account activity sectors not included in OMINEA, we used the EMEP database.

In order to assess uncertainties of industrial emissions, the information collected for each industrial source during the inventory was rated according to the completeness of data on technical parameters (in-depth information; approximate information; summary information) and activity rate (known; known for at least one year of the activity period; approximated using an average activity rate from similar sources) for each year of the study period. The four categories of confidence level (High, Medium+, Medium-, and Low) of the emission estimates take into account both the technical parameters and the activity rate and corresponds to the uncertainty of the emission estimates ( $\pm 30\%$ ;  $\pm 50\%$ ;  $\pm 75\%$  and  $\pm 100\%$ ) for each inventoried industrial source. All industrial sources were geocoded at the position of the stack, based on aerial photographs and Google Street View using ESRI ArcGis software. The methodology used to geocode study subjects and industrial sources is described in detail elsewhere (Faure et al., 2017). Industrial emissions were modelled as elevated point sources with an intensity that varies with an annual time step.

### 2.2.2 Diffuse and small sources

The estimate of all other kinds of dioxin and cadmium sources was provided by the Association of Air Quality Surveillance in the Auvergne-Rhone-Alpes Region (ATMO-AURA) cadaster, developed since 2000 for cadmium and 2002 for dioxin, and included releases for both anthropic and natural sources. The cadaster was built with a top-down (disaggregation of national data to the city level) and a bottom-up approach (aggregation of local data at the city-scale level) combined with the use of emissions factors. The cadaster also takes into account the emissions of small facilities, traffic, domestic heating and other miscellaneous sources. Sources were classified, based on the activity sectors using the SNAP classification (Atmo Auvergne-Rhône-Alpes., 2016). The inventory at the city level was distributed and spatially aggregated on a regular 1x1km grid. Each cell contains contributions of the 11 SNAP source categories. In order to avoid double counting the industrial sources in the 1x1km cadaster cells, SNAP categories corresponding to facilities already inventoried (2.2.1) were excluded from the cadaster.

To consider the different contributions to dioxin and cadmium air concentrations, the cadaster is split into three parts according to the SNAP classification: traffic, heating and miscellaneous emissions. Traffic emissions are projected into the street network (Sect. 2.2.3), and heating emissions are considered as spatially distributed sources with their intensity on the 1x1km grid depending on the population density. The temporal variability of heating emissions is modulated according to heating activity as a function of the external temperature (using previous year's correlation). Miscellaneous sources include all other kinds of releases, such as diffuse source or burning of stolen copper cable. In the absence of information about their variability, their emissions are considered as constant over the 2007-2008 period. Note that a significant contribution to emissions from miscellaneous sources is provided by the combustion of the plastic sheath covering copper cables: the sheath being burned off after the cables have been stolen. This activity appears to be the most important source of dioxin in France for the most recent years of the study period (Serveau et al., 2013). Given the illegal nature of this activity, these releases are extremely difficult to localize in space and time, and the total amount of these releases can only be estimated at national level (based on the total amount of copper stolen over the years). Their disaggregation across the national territory at the local scale is then simply based on the local density of the population.

As is often the case in atmospheric pollutant dispersion studies at the urban scale (e.g. Soulhac et al., 2017, Meneses et al., 2004), the intensity of the distributed and traffic emissions is provided by local environmental agencies. For the LMA, the distribution of emissions intensity is provided by ATMO AURA, based on data and information which are not in public access. In this context, it is therefore impossible to quantify appropriate uncertainty levels and sensitivity analyses can only be based on rough variation of emissions around the reference value (e.g. Soulhac 2012).

### **2.2.3 Traffic sources**

Traffic is generally considered as a weak contributor to dioxin and cadmium emissions in France (Serveau et al., 2013) and in Europe (Quaß et al., 2004). However, traffic emissions may be relevant in cities with a high building density, since they occur at ground level and within deep street canyons that are characterized by low ventilation rates (Salizzoni et al., 2009; Buccolieri et al., 2015), which leads to an accumulation of airborne pollutants. The traffic network of the LMA is split into 21 922 segments, each of them representing a street between two intersections, characterized by a specific pollutant emission rate, with an hourly time step. For a detailed description of the methodology used to estimate pollutant emissions from traffic fluxes the reader is referred to Soulhac et al. (2012). The non-quantification of the uncertainty levels of traffic sources is due to the same reasons discussed above for diffuse and small sources (see 2.2.2).

Direct information on traffic emissions was only available for 2007-2008. Before 2007, a spatially averaged emission over the whole agglomeration provided by the cadaster was available (2.2.3). This latter average is used to modulate the emissions over the whole traffic network (which is assumed to be the same as for 2007-2008), keeping the same ratio of emissions between the different street segments. The temporal variability of traffic emissions is modulated day by day according to previous traffic flow observations (see Soulhac et al. (2012), for details).

### **2.2.4 Emission measurements**

Measurements of emissions are only available for industrial sources and are provided by the iREP database (French pollutant release register). IREP data are provided by industrialists since 2003. For France as a whole, 128 and 409 measurements were available for dioxin and cadmium, respectively. Concerning the LMA, the dataset contains emission measurement data for 5 sources of cadmium and 3 sources of dioxin. Due to the scarcity of measurements available for the LMA area, we perform a comparison of our estimates against emissions measurements for both the LMA area and the whole of France (Figure 2).

## **2.3 Atmospheric dispersion**

### **2.3.1 The SIRANE Model**

The numerical simulations are performed with SIRANE, a model conceived to simulate atmospheric pollutant dispersion in a built environment. Its specificity relies on the parameterisations of the pollutant transfers within the urban canopy (Soulhac et al., 2011), namely the channeling and the retention of pollutant within the street canyons (Soulhac et al., 2008; Salizzoni et al., 2009;

Soulhac et al., 2013) and the mass transfer within street intersections (Soulhac et al., 2009). To date SIRANE has been validated by means of wind tunnel experiments (Carpintieri et al., 2012; Ben Salem et al., 2015) and open field measurement data (Soulhac et al., 2012). The most comprehensive validation in a real urban area was recently published by Soulhac et al. (2017), based on the dispersion of nitrogen oxides for the whole year 2008 over the same domain as in the present study taking into account all kinds of pollutant sources, i.e. traffic emissions, domestic heating, and industrial emissions.

The meteorological data (wind speed and direction, temperature, cloud cover and precipitation intensity) used in the model were collected by Météo France at the Bron Airport (Figure 1). These are needed to estimate the parameter governing the pollutant dispersion, as well as their dry and wet deposition, with an hourly time-step. In the model, we consider different source typologies: industrial emissions (Sect 2.2.1) are represented as elevated point sources (taking into account the effects related to buoyancy and momentum fluxes at the emissions), traffic emissions are line sources distributed on a road network, and miscellaneous diffuse sources are represented as ground level area sources. Without any direct information about the size distribution of the pollutant particles, according to previous studies (Sugita et al., 2004; Fabre et al., 2008; European Commission, 2000; Ángeles González et al., 2012) we assume these to have a diameter of  $1\mu\text{m}$ . We performed simulations, with a fine resolution (25 meters), over five years for each pollutant, namely in 1991, 1996, 2002, 2007 and 2008 for dioxin and in 1991, 1996, 2000, 2007 and 2008 for cadmium. The earliest year taken into account is 1991, due to the low quality of the meteorological data for 1990. The choice of two different years, i.e. 2000 for cadmium and 2002 for dioxin, is due to the fact that these were the earliest years for which information from the source cadaster was available. Before 2000 we only had estimates for the industrial sources, and no cadaster data were available. We therefore assumed that the contributions of all other sources are those provided by the earliest data available in the emission cadaster, i.e. 2000 for cadmium and 2002 for dioxin. With this assumption, dioxin and cadmium dispersion is modelled for 1991 and 1996.

### **2.3.2 Background concentrations**

Pollutant concentrations within a given domain result from both local emissions and emissions from sources located outside the domain. The contribution of the latter is usually referred to as ‘background concentration’. The characterization of the background concentrations is a key issue in pollutant dispersion modelling at the local scale (Beelen et al., 2009). These background concentrations are generally estimated by means of nested model simulations or through direct measurements. In the first approach (nested models), the results provided by a regional model are used as boundary conditions for the simulation at the local scale (Soulhac et al., 2003). In the second approach, the data provided by stations placed in rural areas (i.e. located away from the pollutant



sources) surrounding the urban agglomeration are assumed to be representative of the background values (Soulhac et al., 2012). For the pollutants of interest in our study, both estimates are missing since we do not have any large-scale simulations or any measurements in the surrounding rural area. The only possible way to estimate a plausible value of background concentration is therefore to refer to data from the literature.

A first survey of typical values of dioxin concentration was provided by Lohmann and Jones (1998) who reported annual average values of 100-400 fg TEQ.m<sup>-3</sup> in urban areas, 20-50 fg TEQm<sup>-3</sup> in rural areas and close to 10 fg TEQ.m<sup>-3</sup> in remote rural areas. Despite the general decrease of the intensity of dioxin emissions over the last few years, recent measurements collected throughout the French territory (Table 1) generally confirm these early estimates by Lohmann and Jones (1998). These values are similar to the background values reported by the U.S. Environmental Protection Agency (2013), for the National Dioxin Air Monitoring Network (NDAMN) for the period 1998-2004, showing TEQ concentrations of 15.9 fg TEQ.m<sup>-3</sup> for urban and 13.9 fg TEQ.m<sup>-3</sup> for rural areas, with high values during the fall and winter months, compared to the spring and summer months. Today, different estimates of background concentrations collected in different regions of France indicate minimal dioxin concentrations in the range 5-40 fg TEQ.m<sup>-3</sup> and values in remote urban areas that are around 10 fg TEQ.m<sup>-3</sup>. This latter value was assumed here as an annual average for the dioxin background concentrations.

A review from the European Commission (2000) indicate usual concentrations of cadmium ranging from 2 to 20 ng.m<sup>-3</sup> in industrial areas, 0.2 to 2.5 ng.m<sup>-3</sup> in urban areas and 0.1 to 0.4 ng.m<sup>-3</sup> in rural areas. More recent data from the European air quality database AirBase indicate cadmium concentrations of 0.1-3.0 ng.m<sup>-3</sup> for both urban traffic and background levels (Strincone et al., 2013). For the present study a background concentration of 0.1 ng.m<sup>-3</sup> is considered for cadmium.

	Nb. of sites	fg-TEQ.m <sup>-3</sup>	Type	Years of measurement	Reference
Rural	1 site	10-90	Weekly average	2004	AirParif, 2005
Urban	9 sites	20-3410			
Urban	1 site	5-20	Weekly average	2006	Atmo Poitou-Charentes, 2006
Industrials	1 site	9-28			
Urban	1 site	39	Annual average	2010	Atmo Nord-Pas de Calais, 2010
Industrial	1 site	104			
Industrial	1 site	18-140	Annual average	1999-2012	Atmo Languedoc-Roussillon, 2013
Industrial	15 sites	21-129	Annual average	2008-2012	Atmo Auvergne-Rhône-Alpes, 2014
Urban	1 site	24-37			

Table 1 – Details on dioxin measurements by French Air Quality Associations (AASQA).

Several studies show a relevant seasonal variability in the background concentrations for both pollutants, with higher values in winter than in summer (Wai Ting Tung et al., 2005; E.Ahmedv et al., 2016; Duan et al., 2012; EPA 2013). Accordingly, dioxin and cadmium background weekly concentrations are modulated over time, in line with the variation of the weekly average PM10 concentrations (Umlauf et al., 2010; Castro-Jiménez et al., 2012), as measured from the reference background monitoring station (Figure 1) of air quality studies in Lyon (Soulhac et al., 2017).

### **2.3.3 Concentration measurements**

The monitoring of airborne dioxin and cadmium has been performed by the air quality association of the Auvergne-Rhône-Alpes region (ATMO-AURA) in the LMA since 2007 (see Figure 1). For the 2007-2008 period, 102 and 94 weekly average measurements for dioxin and cadmium, respectively, are available, from a fixed station, referred to here as 'Lyon Centre'. For this same period, three mobile stations, referred to as 'Lyon South', 'Lyon Nord' and 'Feyzin', provide short-term dioxin measurements. The 'Lyon South' station, placed within the urban area, and the 'Lyon North' station, placed in a sub-urban area, provided 8 weekly measurements only (in 2007 for Lyon South; in 2008 for Lyon North). The 'Feyzin' station (industrialised area in the south of Lyon) collected 16 measurements (evenly distributed in 2007 and 2008). To compare modelled concentrations and monitoring data from the Lyon Centre station, we use weekly average concentrations of dioxin (ITEQ-fg.m-3) and cadmium (ng.m-3). We further use statistical indices and acceptance criteria (Chang and Hanna, 2012). Since the short-term period measurements were not performed exactly on the same periods, we could not include these data in the statistical analysis.

## **3. Results and Discussion**

The reliability of the modelling chain is evaluated by a systematic comparison of the model predictions with direct measurements of pollutant emissions (3.1.1) and concentrations (3.2.1). The emission model and the dispersion model is then used to simulate the historical trends over the 20-year period considered in this study (3.2.2). Finally, we provide a classification of the E3N subjects on the LMA, depending on their exposure to concentrations of dioxin and cadmium in air over the 1991-2008 period (3.3).

### 3.1 Emissions

#### 3.1.1 Comparison with available measurement

In the LMA, between 1990 and 2008, we identify 31 industrial dioxin and cadmium sources. At the national level, 2620 dioxin sources and 2690 cadmium sources are inventoried in the same period. Given the low number of measurements in the LMA (Sect. 2.2.4), we perform a comparison between emission estimates and measurements considering two samples: the emissions in France as a whole and those in the LMA (Figure 2 and Table 2).

Dioxin	France (n=128)		LMA (n=7)	
	Mean (SD)	Median (Min-Max)	Mean (SD)	Median (Min-Max)
Measurements (g/year)	4.0 (3.8)	2.8 (0.1-19.0)	2.5 (1.4)	3.1 (0.1-3.5)
Estimations (g-TEQ/year)	5.8 (12.7)	1.7 (0.0001-56.6)	5.1 (3.3)	5.4 (0.1-7.9)
Cadmium	France (n=409)		LMA (n=17)	
	Mean (SD)	Median (Min-Max)	Mean (SD)	Median (Min-Max)
Measurements (kg/year)	47 (105)	23 (1.8-1240)	12 (14)	10 (0.1-50)
Estimations (kg/year)	49 (197)	16 (0.1-3007)	11 (10)	7.0 (0.2-41)

Table 2 – Statistical description of measured and estimated emissions for dioxin and cadmium for France as a whole and the LMA.

For both, the national level and the LMA, mean values of estimated emissions were close to emission measurements for cadmium and generally overestimated dioxin measurement values (Table 2), with similar Spearman's rank correlations for dioxin ( $r_s = 0.42$ ) and cadmium ( $r_s = 0.41$ ). The comparison between our estimate (and their relative uncertainties) and measurement of dioxin and cadmium emission is represented in Figure 2. The majority of estimated emissions are within  $\pm 100\%$  of measured emissions for cadmium and  $\pm 200\%$  for dioxin (Figure 2). For both pollutants, the number of sites for which emissions are overestimated by more than 200% is small (5% for cadmium and 12% for dioxin). For both pollutants, we obtain a similar value of the Spearman's rank correlation coefficient. When considering the LMA, we observe similar results. On average, estimates of emissions tend to overestimate the measured dioxin emissions. For cadmium, the averages of our estimates and of the measurements are very similar. On Figure 2, for emission estimates placed in the LMA (red dot), 2 out of 7 dioxin emission measurements and 4 out of 17 cadmium emissions measurements are over/underestimated by more than 200%.

The uncertainties of dioxin and cadmium industrial source emission estimates range from 30% to 100%. At the national level, we observe average uncertainties of  $\pm 65\%$  and  $\pm 68\%$  for dioxin and cadmium emission estimates, respectively. These national averaged values are highly influenced by the large number of hospital waste incinerators (n=1442), characterised by a low confidence level of

the technical and activity information collected. Compared to the national level, average uncertainties in the LMA are lower for both dioxin ( $\pm 38\%$ ) and cadmium emissions ( $\pm 35\%$ ) due to the relatively low number of hospital waste incinerators ( $n=3$ ).

This comparison between model estimates of emissions and measurements cannot however be considered as a “validation” of the emission model. The emission measurements from industrial sources represent only a small portion of the pollutant sources, whose intensity is reasonably expected to be characterized by relevant fluctuations over time. For this reason the emission measurements cannot be considered as fully representative of actual emissions occurring over a whole year, and have therefore to be treated with caution.

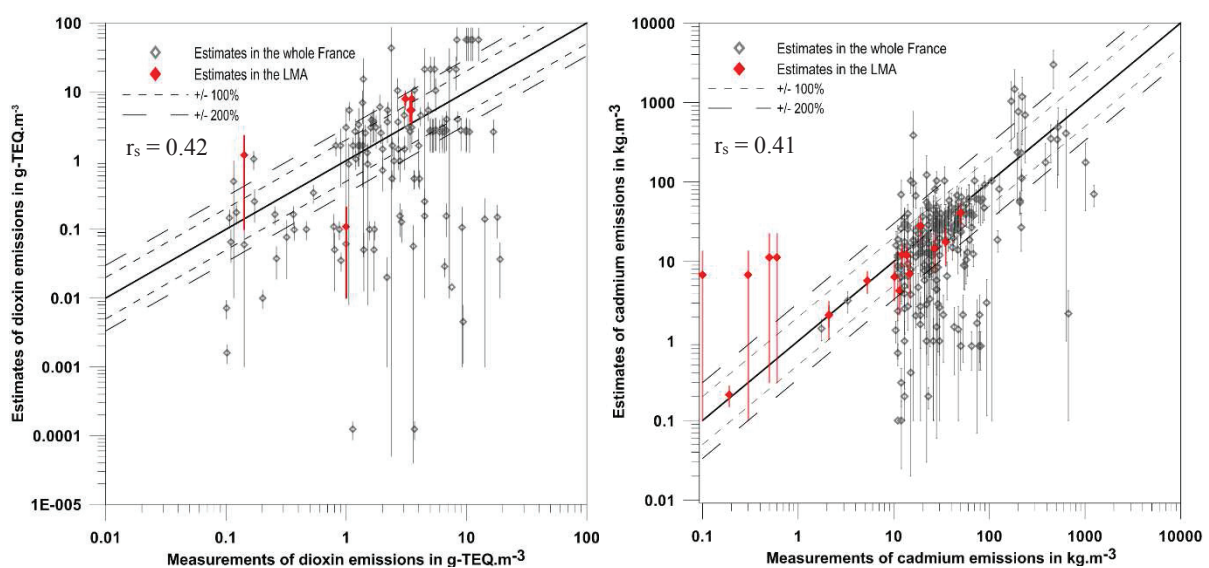


Figure 2 – Comparison between annual average estimates of dioxin (a) and cadmium (b) industrial emissions and measurements in the LMA area and in France as a whole between 2003 and 2008

### 3.1.2 Historical trends

Estimates of industrial emissions significantly decrease in the period 1990-2008, by 99% and 85%, for dioxins and cadmium, respectively. For both pollutants, maximal industrial releases occur before 1994. Dioxin emissions exhibit two clear drops in 1996 and in 2006, whereas cadmium emissions decrease gradually over the period (Figures 3a and 3b). The variability of the emissions is due to regulatory constraints (Nzihou et al., 2012) and technological evolutions. Dioxin emissions show the effects of prescription on municipal solid waste incinerator flue gas treatment systems in 1996, and the implementation of a new dioxin emission limit value in 2005. Over time, emission uncertainties of industrial sources in the LMA vary slightly from  $\pm 37\%$  in 1990 to  $\pm 33\%$  in 2008, except for 2 year ( $\pm 68\%$  in 2006 and  $\pm 69\%$  in 2007), due to a major dioxin source with a low confidence level.

The data presented in Table 3 show that the global reduction in emissions is associated with a major change in the relative contribution of the different sources to total emissions. Industry is the main source of dioxin emissions for 1991, 1996 and 2002. Between 2002 and 2008, industrial emissions of dioxin decrease and diffuse sources become the major contributor (56% of total dioxin emissions). Despite a decrease over the years, industrial sources remain the main source of cadmium emissions for the whole period. The proportion of cadmium emissions from traffic increases from 4.3% to 11.9% between 2002 and 2008.

These general tendencies are in agreement with those identified by CITEPA, reporting that the total emissions of dioxin and cadmium in France decreased by 95 % and 80%, respectively, between 1990 and 2008 (Serveau et al., 2013).

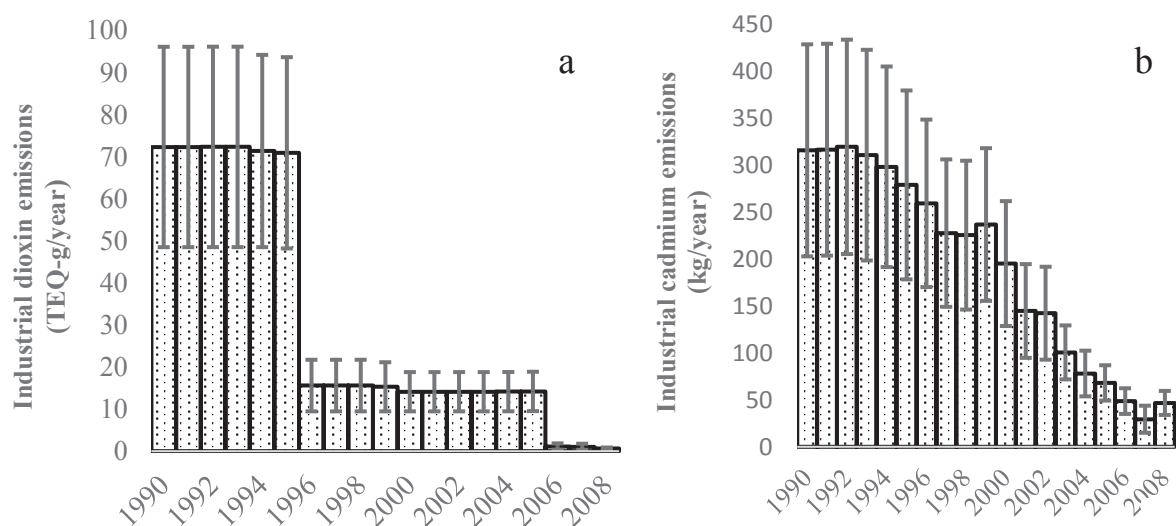


Figure 3 - Estimates of industrial dioxin (a) and cadmium (b) releases in the LMA (1990-2008)

	Dioxin emissions				Total (g/year)	Cadmium emissions				Total (kg/year)
	Industrial	Diffuse	Heating	Traffic		Industrial	Diffuse	Heating	Traffic	
1991	98.2%	1.4%	0.3%	0.1%	73.7	91.1%	5.2%	1.0%	2.8%	347.2
1996	92.1%	5.9%	1.4%	0.5%	16.9	89.3%	6.2%	1.1%	3.3%	290.2
2002	91.4%	6.5%	1.6%	0.6%	15.5	86.3%	8.0%	1.5%	4.3%	164.8
2008	31.4%	56.0%	8.9%	3.7%	2.0	78.7%	4.6%	4.8%	11.9%	59.3

Table 3 – Contribution to total dioxin (a) and cadmium (b) emissions in the LMA from industrial, diffuse, heating and traffic sources

In 2007 and 2008, i.e. the period for which we have monitoring data, dioxin emissions from diffuse sources are the major contributor to total dioxin emissions and releases due to the burning of stolen cables are considered to represent the main component of diffuse emissions at the national level (Serveau et al., 2013).

## **3.2 Ambient air concentrations**

### **3.2.1 Comparison with available measurements**

The comparison between simulated ground-level concentrations and measured weekly-average concentrations registered at the fixed monitoring station (Lyon Centre) for 2007-2008 is presented in Figure 4 and 5.

Weekly average measurements show a clear seasonal trend, with higher concentrations in winter and lower in summer and spring. Winter dioxin peaks attain concentration levels up to 10 times higher than summer values; cadmium peaks are instead only four times the minimal summer values. Besides the winter peaks, we observe the presence of other peaks, which appear almost randomly during the year and seem not to be correlated to the occurrence of peculiar meteorological conditions. When compared to minimal summer values, these distributed peaks are larger for dioxin than for cadmium. These unusually high measured values are therefore likely to be due to miscellaneous uncontrolled sources, and they mainly concern dioxin emissions. As shown in Sect 3.1 (Table 3), in 2007-2008, a large proportion of dioxin emissions are due to miscellaneous and diffuse releases, such as the burning of stolen copper cables. These kinds of releases are not predictable in an emission scenario and can produce very intense emissions, localized both in time and in space that can in turn affect local measured concentrations.

Modelled dioxin concentrations (Figure 4) are overall in good agreement with those registered at the Lyon Centre station in spring and summer. However the trend of the modeled data is not able to reproduce the seasonal variations observed in the measurements. Non-negligible differences between simulations and experimental data are therefore observed during the winter, when the modeled concentrations are ten times lower than the peaks observed experimentally. As previously discussed, these peaks are likely due to uncontrolled and unpredictable diffuse emissions of dioxin, which are not taken into account in our emission scenarios. We therefore expect these systematic underestimates to be due to the uncertainties in our emission scenarios rather than to uncertainties in the dispersion models. Other errors in the numerical estimates of dioxin concentration can be attributed to the lack of information on plume depletion mechanisms such as the degradation due to the reaction with hydroxyl radicals (Cleverly et al., 2007), direct photolysis, or uncertainties in dioxin- vapor/particle partitioning.

In a similar way, modelled cadmium concentrations show a general good agreement with measured concentrations in spring and summer and significant discrepancies in winter, a period during which the model generally tends to underestimate monitored values. For cadmium, the model results generally reproduce the seasonal trend better than for dioxin, even though they fail to predict some winter peaks during the winter. The discrepancies between model and measured data are however significantly lower than those observed in case of dioxin data. This reduced uncertainty reflects the higher accuracy of the cadmium emission scenario (compared to that of dioxin). Indeed, the emission scenario for cadmium is not affected by the large uncertainties related to diffuse sources, whose contribution in the emission scenario is significantly lower than that of industrial sources (Tab. 3).

Based on these observations, we can assert that the emissions related to the burning of copper cables (and others unpredictable releases) are a major source of uncertainty in modelled dioxin concentrations. In order to have an estimate of the occurrence of these (unpredictable) releases and their impact on the air pollutant concentrations, we examined the ratio between measured and modelled dioxin concentrations (median: 1.47) and the ratio between measured concentrations of dioxin and cadmium (median:  $1.25 \times 10^{-4}$ ). For the weeks with a measured/modelled dioxin ratio in the range 0.5-2, the median dioxin/cadmium ratio is close to  $1 \times 10^{-4}$  ( $1.09 \times 10^{-4}$ ). However, in the 14 weeks during which the model exhibits relevant under-prediction of the dioxin concentrations (measured/modelled dioxin ratio above twice the median ratio), the median dioxin/cadmium concentration ratio is equal to  $2.39 \times 10^{-4}$  (nearly 2 time the median dioxin/cadmium ratio). This decoupling between dioxin and cadmium concentration levels is likely to be associated with the occurrence of unpredictable intense dioxin releases in the study area. Note that the impact of these ‘accidental’ releases become more and more important as the general concentration level tended to decrease over time.

For both pollutants, the systematic under-prediction of the measured concentrations may be, at least partially, attributed to the role of re-suspension of particles, which can have a major role in the dispersion of particles in a highly urbanized area (ATMO Auvergne-Rhône-Alpes., 2014). However, this physical phenomenon is not taken into account in the SIRANE model.

Since background concentrations is a major contributor to air pollution levels over the whole agglomeration (this is also the case for other pollutant as NO<sub>x</sub> and PM, see Soulhac et al. (2012, 2017)), we have plotted its time-dependency in Figure 4. This clearly shows that the evolutions of the concentration levels are not only related to the modulation of the background. Notably, the model is able to predict cadmium concentration peaks, as determined by particular meteorological conditions (Figure 4b). Even in the case of dioxins, some of the seasonal peaks can actually be predicted by the model. The variability associated to these peaks is however small when observed with a concentration scale that is imposed by the large winter peaks (Figure 4a).

Despite the different levels of agreement between modelled and measured dioxin and cadmium concentrations, the two time series produce similar scatter plots, with similar  $R^2$  coefficients ( $R^2 = 0.37$  for dioxin and  $0.39$  for cadmium). Yet, for cadmium the dispersion of the data is reduced and the value of  $R^2$  increases from  $0.39$  to  $0.57$  after the exclusion of two outliers. Despite the differences observed between the time series, the model satisfies the validation criteria for both pollutants, as defined by Chang and Hanna (2012), for an urban dispersion model (Table 4).

The 2007 and 2008 average modelled dioxin concentrations ( $17.7\text{fg-TEQ/m}^3$ ) are significantly lower than the monitored concentrations ( $35.5\text{fg-TEQ/m}^3$ ). For cadmium, the comparison between modelled ( $0.18\text{ng/m}^3$ ) and measurement data ( $0.23\text{ng/m}^3$ ) show a better agreement between the two.

The results for the mobile short-term dioxin stations (namely North Lyon, South Lyon and Feyzin) in the 2007-2008 (Figure 5a) are similar to those for the fixed station, with a clear tendency to underestimate measured data. A detailed analysis of the data suggests however different tendencies at the three stations. At the urban and suburban stations (South Lyon and North Lyon), we observe an average modelled concentration of  $15.70\text{ fg-TEQ/m}^3$  (measured:  $25.19\text{ fg-TEQ/m}^3$ ) and  $11.32\text{fg-TEQ/m}^3$  (measured:  $24.76\text{ fg-TEQ/m}^3$ ), respectively. For these two station, the range of concentrations is reduced, the model underestimates the dioxin concentration, by a factor of two, for only 2 and 3 measurement concentrations (over 8). At the industrial station (Feyzin), the average measured concentrations are higher than for the other stations ( $39.68\text{ fg-TEQ/m}^3$ ), and the average modelled concentration is equal to  $14.42\text{ fg/m}^3$ . At this location, the model seems less performant, and underestimates the dioxin concentrations by a factor of 2 for 11 weeks (out of 16). Unfortunately, the limited number of these short-term measurements prevents us to include these in a more comprehensive statistical analysis of the model performance.

The analyses of the sensitivity of our modelled concentrations to the uncertainties in industrial emissions shows a variability of the total average concentration (at the Lyon center monitoring station) of  $\pm 0.5\%$  ( $\pm 0.09\text{ fg-TEQ/m}^3$ ) for dioxin and  $\pm 1.4\%$  ( $\pm 0.003\text{ ng/m}^3$ ) for cadmium, respectively (Figure 4a and 4b). These variations turn out to be negligible compared to the absolute concentration values. The same results are observed for the short term measurement stations. This is of course due to the small contribution of industrials sources to the total concentration in this particular period (Table 5 and 7). Therefore, the uncertainties in the industrial emissions cannot reveal the reasons for the systematic underestimate of the average concentrations measured at the monitoring stations.



### 3.2.2 Historical trend and source sector contributions

The background concentration, whose annual average is assumed to be constant over the whole period, is the major contributor to the average concentration over the area, for both cadmium and dioxin (Tables 6 & 8). On the other hand, contributions from the sources in the LMA vary during the study period.

#### **Dioxin**

In the 1991-2008 period, modelled annual average dioxin concentrations on the LMA domain decrease by 36%. This reduction mainly occurs before 1996, whereas between 2002 and 2008, the average annual dioxin concentration decreases only slightly (from 12.56fg TEQ/m<sup>3</sup> to 11.28fg TEQ/m<sup>3</sup>) (Table 5). Between 1996 and 2008, we observe, however, a relevant change in the spatial concentration patterns, and the local concentration peaks exhibit a reduction, which is much larger than that of the mean values.

This general trend implies two main features: i) dioxin concentrations decrease in the most exposed areas and ii) the difference between the most exposed and the least exposed areas is also reduced. It is worth noting that in 1991, the 1% most exposed area to industrial emissions was 100 times more exposed than in 2008. As a result, the concentration field appears more homogeneous in 2008 compared to that in the previous years. These variations are mainly due to the fall by 99% in the contribution of industrial sources. Furthermore, uncertainties in dioxin emission (that are due to industrial emissions) in the LMA decrease from  $\pm 12.5\%$  in 1991 to  $\pm 0.3\%$  in 2008. Between 2002 and 2008, the contribution from industrial sources becomes negligible whereas the contribution of all other sources (diffuse sources, heating and traffic) remains stable (Table 6).

In each modelled scenario, dioxin concentrations are higher in the city centre, due to the higher density of pollutant sources. Highest concentrations are however located downwind close to industrial sources (Figure 6).

#### **Cadmium**

Considering the whole 18-year period, the average annual cadmium concentration and the 99<sup>th</sup> percentile declined by 26% and 51%, respectively. This decrease is mainly due to a strong reduction in the intensity of industrial and diffuse sources (92% and 82%, respectively) and is also associated with a decrease of uncertainties in the total concentration over the LMA, due to industrial cadmium emission (from  $\pm 6.3\%$  to  $\pm 0.6\%$ ). Between 1991 and 2000, we observe slight differences in average concentrations (from 0.16 to 1.49ng.m<sup>-3</sup>) and in the concentration patterns, with hot spots above 1.6 ng/m<sup>3</sup> occurring close to industrial facilities and high concentrations (between 0.5 and 1ng/m<sup>3</sup>) on

major road (as show on Figure 7). In the 1990-2000 period, the highest concentrations of cadmium always occur close to industrial sources.

The main decrease in the average annual cadmium concentration is observed between 2000 and 2008 (from 0.149ng/m<sup>3</sup> to 0.119ng/m<sup>3</sup>). In this 8-years period, the contribution due to traffic and domestic heating to the modelled cadmium concentrations remains stable (around 8% and 4% respectively) while the contribution due to industrial and diffuse sources decrease (from 10% to around 2%). Therefore, compared to the year 2000 scenario, the concentration field in 2008 appears to be significantly altered. Among the emissions occurring within the studied domain, traffic has become the major contributor to the average modelled cadmium concentrations (Table 7 and Table 8), and this results in concentration fields (Figure 7d) in which the maximal values take place along the major roads. Note that, when occurring within street canyons, the impact of these emissions is further amplified, and requires specific modelling tools to be predicted.

### **3.3 Dioxin and cadmium exposure of study subjects in the study area**

To evaluate the differences in population exposure we considered annual average concentrations, computed at the residential address location of the 421 E3N subject, for four references years (1991, 1996, 2002, and 2008).

#### **3.3.1 Dioxin**

The modelled median dioxin concentration at the address locations decrease by 31%, i.e. from 19.50fg TEQ/m<sup>3</sup> in 1991 to 13.45fg TEQ/m<sup>3</sup> in 2008, and maximal concentrations decrease by 76%, from 82.1fg TEQ/m<sup>3</sup> in 1999 to 20.1fg TEQ/m<sup>3</sup> in 2008. Whereas in 1991, 75% of the study subjects were exposed above 16fg TEQ/m<sup>3</sup>, only 25% are exposed to values exceeding 16fg TEQ/m<sup>3</sup> in 2008 (Figure8a). Similar to the decrease of modelled concentrations over the entire LMA, the observed decrease mainly occurs before 1996 and is associated with the decrease of industrial dioxin emissions. After 1996 dioxin average concentrations vary slightly (from 15.61fg TEQ/m<sup>3</sup> to 13.45fg TEQ/m<sup>3</sup>) with an average decrease of 14%.

Uncertainties due to industrial dioxin emissions decrease over time and are associated with small variations of the total concentration at the E3N subjects' residence location (from ±10.6% in 1991 to ±0.3% in 2008). These relative variations remain constant over space even in the most exposed area of the LMA.

During this 18 year period, background levels of exposure are the major contribution to average dioxin concentrations in the LMA (Table 9). However, at some E3N subject locations, background contribute to less than 20% to the exposure. Their exposure is instead mainly linked to industrial sources in 1991 and to diffuse sources in the subsequent period. The contribution from traffic and

heating to average concentration levels remains weak (around 3% and 1%, respectively). Overall, annual average dioxin concentrations, at the subjects residence location, remain below the WHO indicator, i.e.  $< 300 \text{ fg-TEQ/m}^3$  (WHO, 2000), even when considering uncertainties of the industrial emission estimates. Note however that a small area of the LMA was above this limit in 1991.

Summarizing, over the study period, average dioxin concentrations at the study subjects residential locations due to industrial sources range from  $0.10$  to  $5.75 \text{ fg TEQ/m}^3$ . These values are similar to those provided by previous modelling studies (which however took into account industrial sources only), namely Cordier et al. (2010), Zambon et al. (2007) and Floret et al. (2006), who found average dioxin concentrations of  $3 \text{ fg TEQ/m}^3$ , between  $4$  and  $6 \text{ fg TEQ/m}^3$  and between  $0.1$  and  $1.6 \text{ fg TEQ/m}^3$ , respectively.

### 3.3.2 Cadmium

Over the 18-year period, median cadmium concentrations at the E3N volunteers' locations decrease by 33%, from  $0.21 \text{ ng/m}^3$  to  $0.14 \text{ ng/m}^3$  and the maximum concentration decline by 27%, from  $1.60 \text{ ng/m}^3$  to  $1.17 \text{ ng/m}^3$ .

In 1991, 75% of the E3N subjects are exposed to values above  $0.17 \text{ ng/m}^3$  whereas in 2008, only 5% are exposed to concentrations exceeding this value (Figure 8b). At the same time, the contribution of industrial and diffuse sources on the median cadmium concentrations decreases by 91% and 87%, respectively. The main reduction in concentration occurs between 2000 and 2008. This is mainly due to the huge decrease in releases due to industrial and diffuse sources (-85% and -86%). The contribution of domestic heating to cadmium concentration remains weak and stable during the whole period. As for dioxin, uncertainties due to industrials emissions decrease over the years and have a small impact on the total concentration at the E3N location (from  $\pm 5.9\%$  in 1991 to  $\pm 0.8\%$  in 2008).

Differently from the case of dioxin, traffic is one of the major contributors (around 19%) of cadmium concentrations at the E3N location, over the whole period. In 2008, traffic is by far the main contributor to median cadmium concentrations. The highest concentration value ( $> 1 \text{ ng/m}^3$ ) are therefore observed for E3N subjects (Table 10) located in street canyons. Furthermore, these locations contribute in amplifying the impact of traffic-related emissions, due to the reduced ventilation of these peculiar urban environments. Overall, the observed annual average cadmium concentrations remain below the WHO indicator for cadmium air concentration ( $< 5 \text{ ng/m}^3$ ) (WHO 2000), even after considering uncertainties of the estimates.

Average cadmium concentrations at the study subject's residential location range from  $0.10 \text{ ng.m}^{-3}$  to  $1.6 \text{ ng.m}^{-3}$ . These levels are similar to those reported by Liu et al. (2015), who modelled the airborne cadmium exposure of more than 100.000 women from the California Teachers Study cohort, and estimated values in the range  $0.04\text{-}12 \text{ ng.m}^{-3}$ .

## Conclusions

In the framework of an epidemiological study on the long-term effects of dioxin and cadmium exposure on breast cancer risk, we have herein presented the results of a modelling chain, simulating the evolution in space and time of measured and modelled ground-level air concentrations of these pollutants over a large urban area (1020km<sup>2</sup>) and over a period of 18 years at a fine temporal (1h) and spatial (25m) resolution. To our knowledge, this is the first study to assess the spatiotemporal variability of dioxin and cadmium concentrations over almost 2 decades, comparing modelled and measured concentrations, and taking into account information on a large variety of pollutant sources (industrial, traffic, diffuse and heating) and background concentrations for the whole study area and at individual residence locations. Previous studies mainly focused on emissions from incinerators and a majority of studies used simplified exposure measures, subject to misclassification bias. Furthermore, this is the first study comparing model-predicted concentrations to ambient air concentrations for dioxin.

The study relies on a modelling chain made up of two modules: one dedicated to the estimate of dioxin and cadmium releases and the other to the simulation of their atmospheric dispersion. The performance of these two modules has been evaluated against industrial emissions measurements (both at a national and local level) and to air pollution concentrations.

Results obtained for pollutant emission estimates show that these are affected by non-negligible differences compared to emission measurements that are generally within the range of  $\pm 100\%$  for cadmium and  $\pm 200\%$  for dioxin. Comparisons between modelled and monitored concentrations, registered at a fixed position (located in the city centre) over two years (2007-2008), are also characterized by some non-negligible differences. In particular the model is unable to reproduce dioxin peaks observed during the winter. The model instead captures better the cadmium variations, including some of the peaks observed during the winter. Despite these differences observed in the time series, our model satisfied the validation criteria for both pollutants, as defined by Chang and Hanna (2012). The trend observed in the dioxins data at the fixed monitoring station is supported by the short term measurement made at the three mobile sites (at urban, suburban and industrial sites).

The sensitivity analysis shows that, in 2007-2008, uncertainties associated with industrial emissions are responsible of negligible variations of ground level concentrations. Evaluated at the fixed monitoring station these are of  $\pm 1\%$  for dioxin and  $\pm 4\%$  cadmium. Using the model to reconstruct past air pollution scenarios over two decades, we observed a decrease of dioxin and cadmium concentration. Their spatial average is reduced by 36% for dioxin and 26% for cadmium, whereas the 99<sup>th</sup> percentile of the concentration is reduced by 75% for dioxin and -51% for cadmium. This reduction is mainly associated with the progressive reduction of the intensity of localized industrial sources and results in concentration fields that tend to become more homogeneous.

The comparison with published results shows that our estimates of dioxin and cadmium concentrations at the E3N ‘study subjects’ locations are similar to those provided by other studies (Liu et al., 2015; Cordier et al., 2010; Zambon et al., 2007 and Floret et al., 2006). Median concentrations at these locations decrease by 31% and 33% for dioxin and cadmium, respectively, from 1991 and 2008. The contributions of the industrial sources to the median concentration drop by 98% for dioxin and 91% for cadmium over the years. Note that variations of the concentration at the E3N subject location due to uncertainties in industrial releases also decrease during the study period (from  $\pm 10.6\%$  to  $\pm 0.3\%$  for dioxin concentration; from  $\pm 5.9\%$  to  $\pm 0.8\%$  for cadmium concentration). Notably, for the most recent scenario, i.e. 2008, we observe that the main contributors to dioxin and cadmium concentrations are due to miscellaneous diffuse sources and to traffic, respectively. During the study period, background levels are the major contributors to annual average dioxin and cadmium concentrations in the LMA. This contribution is however variable according to the year and the subject location and ranges from 20% to 98%.

A major limitation in our study is the lack of direct measurements or emission data for diffuse and small sources and traffic included in our modelling chain. Contrary to industrial emissions, we were not able to quantify precisely the uncertainties associated with this type of emissions. This should be investigated in future research that should also assess the implications on the misclassification of individuals’ exposure. The information provided by our numerical results about the contribution of the different sources on the exposure of the E3N cohort subjects (and their variation over time) will be essential to increase the accuracy of the assessment of long-term airborne dioxin and cadmium exposure and improve the results of epidemiological studies.

### **Competing interests**

The authors declare that they have no competing interests.

### **Acknowledgments**

This work was supported by public funding from the French Environment and Energy Management Agency (ADEME, Grant N° 1306C0031), the Cancéropôle Lyon Auvergne Rhône-Alpes (CLARA), the Regional Committee of the French League Against Cancer of the Saone et Loire Region, and was carried out in partnership with the ARC Foundations for Cancer Research. The E3N study is financially supported by the French League Against Cancer, the Mutuelle Générale de l’Education Nationale, the Institut Gustave Roussy, the Institut National de la Santé et de la Recherche Médicale. Thomas Coudon is supported by a doctoral fellowship of University Lyon 1. Chi-Vuong Nguyen was supported by a doctoral fellowship of the Rhone-Alpes Region. We gratefully acknowledge the ATMO Auvergne Rhone-Alpes for providing air measurement and emissions data and the E3N cohort participants for providing data.

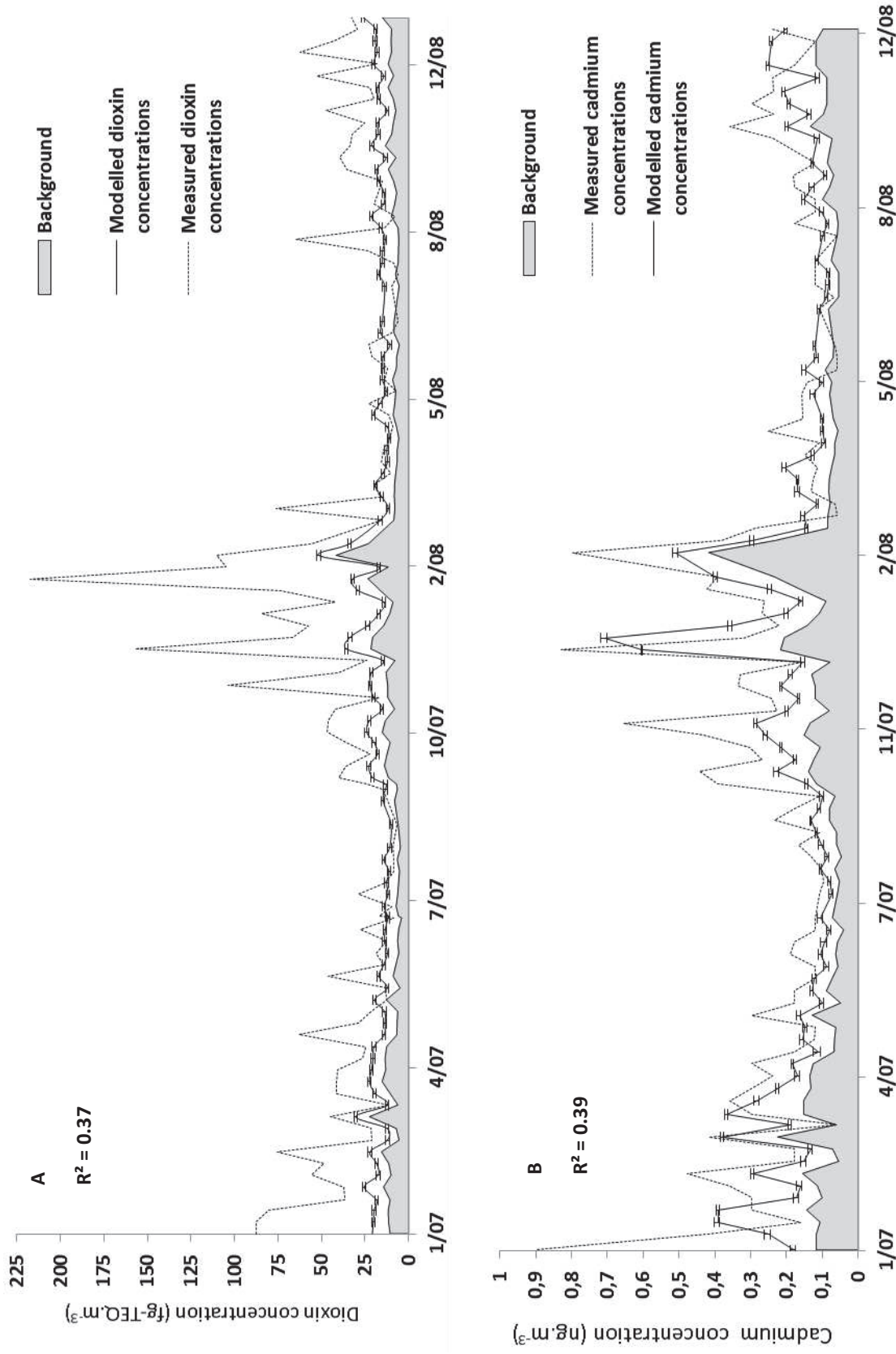


Figure 4 - Comparison between modelled and measured weekly average (a) dioxin and (b) cadmium concentrations at the Lyon centre station between January 2007 and December 2008. Error bars refer to uncertainties due to industrial sources.

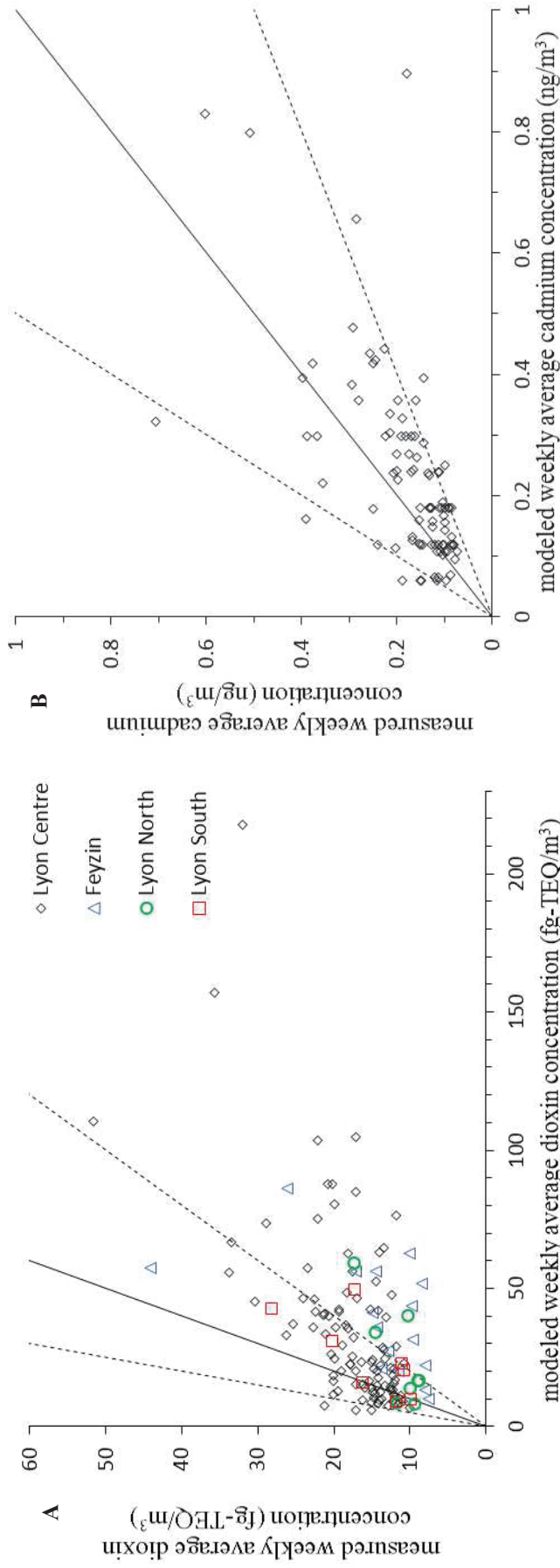


Figure 5 – Scatter plot of the modelled and measured weekly average of (a) dioxin and (fig-TEQ m<sup>-3</sup>) (b) cadmium (ng m<sup>-3</sup>) in 2007 and 2008.

	Fractional mean bias	Normalized mean-square error	Fraction of C <sub>p</sub> within a factor of two of C <sub>m</sub>	Normalized absolute difference
Definition	$FB = \frac{2(\overline{C_m - C_p})}{(\overline{C_m + C_p})}$	$NMSE = \frac{(\overline{C_m - C_p})^2}{\overline{C_m C_p}}$	$FAC2 = 0.5 \leq \frac{C_p}{C_m} \leq 2$	$NAD = \left( \frac{ C_m - C_p }{(C_m + C_p)} \right)$
Optimal value	0	0	1	0
Objectives	$\leq 0.67$	$\leq 6$	$\geq 0.3$	$\leq 0.5$
Dioxin	0.66	1.87	0.67	0.38
Cadmium	0.45	0.35	0.96	0.23

Table 4 - Validation criteria for an urban model result at the continuous measurement station in the city centre according to Chang and

Hanna (2013); C<sub>p</sub> refers to 'predicted' weekly mean concentrations and C<sub>m</sub> to 'measured' weekly mean concentrations.

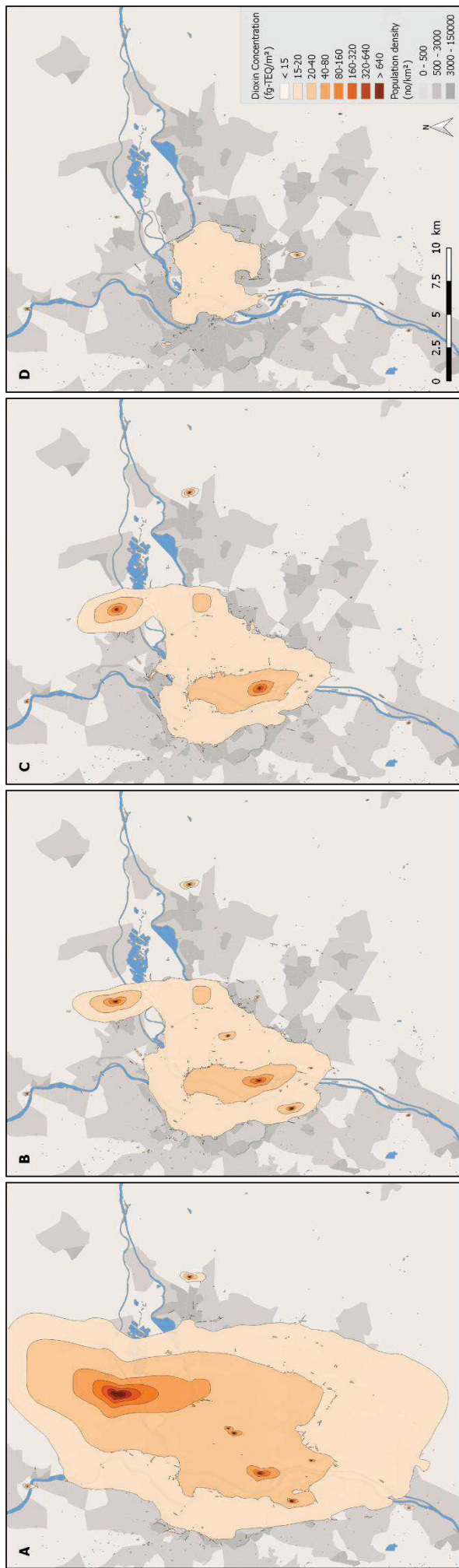


Figure 6 - Evolution over time of annual dioxin concentrations, with a 25x25m resolution, in (a) 1991, (b) 1996, (c) 2002 and (d) 2008 in the LMA

Average concentrations (*) in the LMA in fg TEQ/m <sup>3</sup> (1 <sup>st</sup> - 99 <sup>th</sup> percentile)	Average concentrations (fg TEQ/m <sup>3</sup> ) in the LMA due to			
	Industrial sources (1 <sup>st</sup> - 99 <sup>th</sup> )	Heating sources (1 <sup>st</sup> - 99 <sup>th</sup> )	Diffuse sources (01 <sup>st</sup> - 99 <sup>th</sup> )	Traffic (01 <sup>st</sup> - 99 <sup>th</sup> )
1991	6.35 (0.32-54.56)	0.23 (0.04-0.68)	1.06 (0.10-5.76)	0.09 (0.002-0.697)
1996	1.27 (0.07-9.93)	0.22 (0.04-0.64)	1.05 (0.10-5.62)	0.09 (0.002-0.653)
2002	1.29 (0.08-10.52)	0.24 (0.04-0.70)	0.94 (0.08-5.49)	0.09 (0.003-0.693)
2008	0.08 (0.003-0.55)	0.17 (0.02-0.53)	0.95 (0.07-5.87)	0.08 (0.001-0.548)

Table 5 - Annual average dioxin concentrations (fg TEQ/m<sup>3</sup>) in 1991, 1996, 2002 and 2008 in the LMA and contribution of industrial, diffuse, heating and traffic sources in fg TEQ/m<sup>3</sup> (\*) annual dioxin background is assumed to be constant over the study period and is equal to 10 fg/m<sup>3</sup>

Background (01 <sup>st</sup> - 99 <sup>th</sup> )	Annual average dioxin concentrations (fg TEQ/m <sup>3</sup> ) in 1991, 1996, 2002 and 2008 in the LMA and contribution of industrial, diffuse, heating and traffic sources in fg TEQ/m <sup>3</sup>			
	Industrial sources	Heating sources	Diffuse sources	Traffic
1991	35.8%	1.3%	6.0%	0.5%
1996	10.1%	1.7%	8.3%	0.7%
2002	10.3%	1.9%	7.5%	0.7%
2008	0.7%	1.5%	8.4%	0.7%

Table 6 - Evolution over time of the contribution to the average dioxin concentrations of background, industrial, diffuse, heating and traffic sources



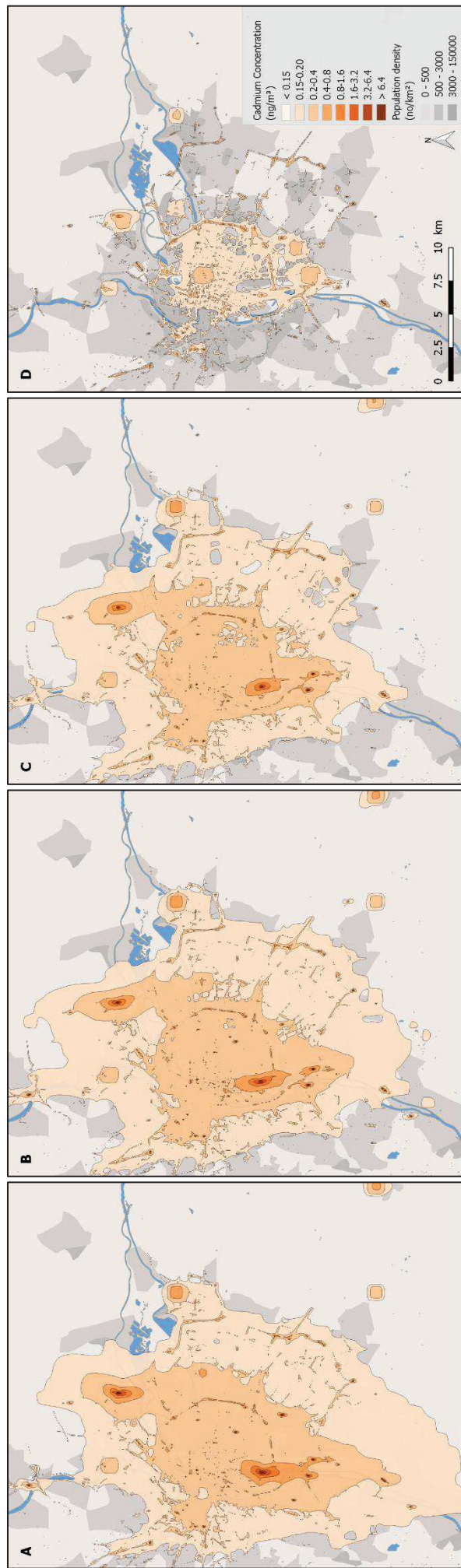


Figure 7 - Evolution over time of annual cadmium concentration, with a 25x25m resolution, in (a) 1991, (b) 1996, (c) 2000 and (d) 2008 in the LMA

Average concentration on the grid in		Average concentration (ng/m <sup>3</sup> ) on the area due to			
Year	ng/m <sup>3</sup> (01 <sup>st</sup> – 99 <sup>th</sup> ) (*)	Industrial sources (01 <sup>st</sup> – 99 <sup>th</sup> )	Heating sources (01 <sup>st</sup> – 99 <sup>th</sup> )	Diffuse sources (01 <sup>st</sup> – 99 <sup>th</sup> )	Traffic (01 <sup>st</sup> – 99 <sup>th</sup> )
1991	0.160 (0.105 – 0.445)	0.025 (0.0011-0.214)	0.006 (0.0009-0.017)	0.017 (0.0021-0.062)	0.012 (0.0003-0.0998)
1996	0.155 (0.105 - 0.402)	0.020 (0.0010-0.171)	0.006 (0.0009-0.015)	0.017 (0.0023-0.060)	0.012 (0.0003-0.0984)
2000	0.149 (0.105 – 0.365)	0.015 (0.0094-0.129)	0.006 (0.0010-0.017)	0.016 (0.0021-0.057)	0.012 (0.0003-0.0986)
2008	0.119 (0.102 – 0.220)	0.002 (0.0002-0.015)	0.004 (0.0005-0.021)	0.003 (0.0001-0.015)	0.010 (0.0003-0.0773)

Table 7 - Annual average cadmium concentration (ng/m<sup>3</sup>) in 1991, 1996, 2002 and 2008 in the LMA and contribution of industrial, diffuse, heating and traffic sources in ng/m<sup>3</sup> (\*): annual cadmium background is assumed to be constant over the study period and is equal to 0.1 ng/m<sup>3</sup>

Year	Background (01 <sup>st</sup> – 99 <sup>th</sup> )	Industrial sources	Heating sources	Diffuse sources	Traffic
1991	62.5% (95% – 22%)	15.6%	3.8%	10.6%	7.5%
1996	64.5% (95% - 25%)	12.9%	3.9%	11.0%	7.7%
2000	67.1% (95% - 28%)	10.1%	4.0%	10.7%	8.1%
2008	84.0% (98% - 45%)	1.7%	3.4%	2.5%	8.4%

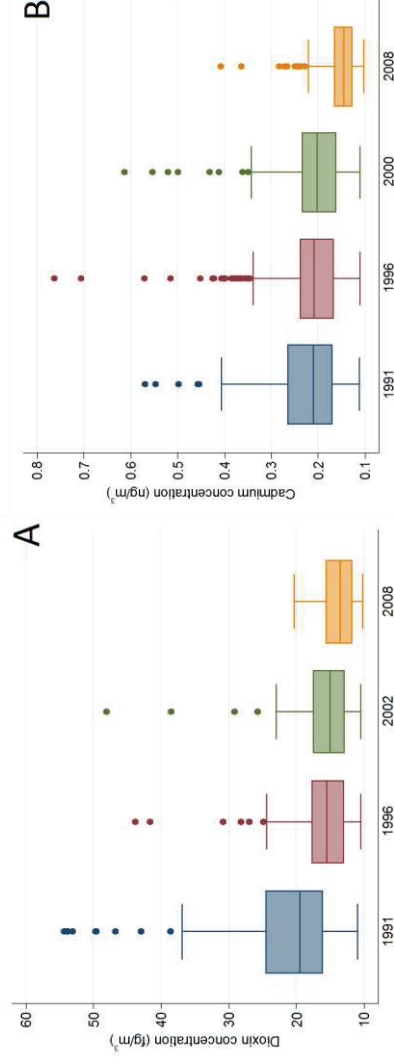
Table 8 – Evolution over time of the contribution to the average cadmium concentrations of background, industrial, diffuse, heating and traffic sources

Median (min-max) dioxin concentrations (fg TEQ/m <sup>3</sup> ) at E3N location	Median (min-max) dioxin concentration (fg TEQ/m <sup>3</sup> ) at the study subjects' residential location of				Relative contribution to average dioxin concentrations at the residential location of study subjects				
	Industrial	Heating	Diffuse sources	Traffic	Background	Industrial	Heating	Diffuse	Traffic
1991 19.50 (10.9 - 82.1)	5.75 (0.51-70.25)	0.47 (0.05-1.23)	2.93 (0.13-11.24)	0.22 (0.003-4.56)	51.2%	37.6%	2.4%	16.0%	1.3%
1996 15.61 (10.4 - 43.8)	1.53 (0.12-31.38)	0.45 (0.05-1.17)	2.88 (0.18-09.15)	0.22 (0.005-1.43)	63.9%	14.2%	2.9%	19.9%	1.5%
2002 15.11 (10.4 - 48.0)	1.52 (0.15-35.71)	0.49 (0.05-1.22)	2.65 (0.05-09.11)	0.23 (0.005-3.27)	63.6%	12.9%	3.1%	18.6%	1.6%
2008 13.45 (10.2 - 20.1)	0.10 (0.01-01.24)	0.37 (0.04-0.67)	2.80 (0.12-07.06)	0.18 (0.002-2.89)	72.4%	0.8%	2.6%	22.6%	1.5%

Table 9 - Median dioxin concentration (fg TEQ/m<sup>3</sup>) in 1991, 1996, 2002 and 2008 at the E3N location and contribution of background (10fg/m<sup>3</sup>), industrial, diffuse, heating and traffic sources

Median (min-max) cadmium concentration (ng/m <sup>3</sup> ) at E3N location	Median (min-max) cadmium concentration (ng/m <sup>3</sup> ) at the study subjects' residential location of				Relative contribution to average dioxin concentrations at the residential location of study subjects				
	Industrials sources	Heating sources	Diffuse sources	Traffic	Background	Industrials	Heating	Diffuse	Traffic
1991 0.21 (0.11 - 1.60)	0.032 (0.0014-0.393)	0.011 (0.002-0.062)	0.031 (0.0039-0.327)	0.031 (0.001-1.42)	44.5%	17.5%	5.2%	14.8%	18.1%
1996 0.21 (0.11 - 1.55)	0.026 (0.0017-0.578)	0.011 (0.002-0.060)	0.029 (0.0046-0.326)	0.032 (0.001-1.36)	45.6%	16.0%	4.9%	14.9%	18.6%
2000 0.20 (0.11 - 1.51)	0.020 (0.0016-0.433)	0.011 (0.002-0.064)	0.029 (0.0029-0.285)	0.032 (0.001-1.33)	47.8%	12.2%	5.5%	15.0%	19.5%
2008 0.14 (0.10 - 1.17)	0.003 (0.0003-0.023)	0.010 (0.001-0.124)	0.004 (0.0002-0.078)	0.026 (0.001-1.04)	65.4%	2.0%	8.1%	3.4%	21.1%

Table 10 - Median cadmium concentration (ng/m<sup>3</sup>) in 1991, 1996, 2002 and 2008 at the E3N location and contribution of background (0.1ng/m<sup>3</sup>), industrial, diffuse, heating and traffic sources



\*1 value above 60 fg-TEQ/m<sup>3</sup> (in 1991) removed from of the graphics

\*\*1 value above 1 ng/m<sup>3</sup> (one for each year) were removed from of the graphics

Figure 8 – Boxplots\* of dioxin (a) and cadmium concentrations (b) at the study subjects' residential location

## Bibliography

1. Ahmed, E., Kim, K.-H., Kim, J.-O., Park, J.-K., Chambers, S., Feng, X., Sohn, J.R., Jeon, E.-C., 2016. Pollution of airborne metallic species in Seoul, Korea from 1998 to 2010. *Atmospheric Environment* 124, Part A, 85–94.
2. AirParif, 2005. Campagnes de mesure des dioxines dans l'air ambiant francilien. ATMO France, Paris.
3. Angeles Gonzalez, M., Vivanco, M.G., Palomino, I., Luis Garrido, J., Santiago, M., Bessagnet, B., 2012. Modelling Some Heavy Metals Air Concentration in Europe. *Water Air Soil Pollut.* 223, 5227–5242.
4. Arisawa, K., Matsumura, T., Tohyama, C., Saito, H., Satoh, H., Nagai, M., Morita, M., Suzuki, T., 2003. Fish intake, plasma  $\omega$ -3 polyunsaturated fatty acids, and polychlorinated dibenzo-p-dioxins/polychlorinated dibenzo-furans and co-planar polychlorinated biphenyls in the blood of the Japanese population. *Int Arch Occup Environ Health* 76, 205–215.
5. Atmo Auvergne-Rhône-Alpes, 2016. Méthode d'élaboration de l'inventaire des émissions atmosphériques en Rhône-Alpes. ATMO France, Bron, France.
6. Atmo Auvergne-Rhône-Alpes, 2014. Programme de surveillance des dioxines, furanes et métaux lourds - Inventaire des émissions & synthèse des mesures (INVENTAIRE DES EMISSIONS & SYNTHESE DES MESURES – 2012). ATMO France, Bron, France.
7. Atmo Auvergne-Rhône-Alpes, 2014. Transferts de dioxines, furanes et PCB entre le sol et l'atmosphère. ATMO France, Bron.
8. Atmo Languedoc-Roussillon, 2013. Surveillance permanente de la qualité de l'air - Bilan 2012. ATMO France.
9. Atmo Nord-Pas de Calais, 2010. Mesure des dioxines, furanes et PCB DL sur le secteur de Bousbecque (No. N° 03-2010-AA). ATMO France.
10. Atmo Poitou-Charentes, 2006. Rapport annuel sur le prix et la qualité du service public d'élimination des déchets. ATMO France.
11. Baan, R., Grosse, Y., Straif, K., Secretan, B., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., Cogliano, V., 2009. A review of human carcinogens—Part F: Chemical agents and related occupations. *The Lancet Oncology* 10, 1143–1144.
12. Beelen, R., Hoek, G., Pebesma, E., Vienneau, D., de Hoogh, K., Briggs, D.J., 2009. Mapping of background air pollution at a fine spatial scale across the European Union. *Sci. Total Environ.* 407, 1852–1867
13. Ben Salem, N., Garbero, V., Salizzoni, P., Lamaison, G., Soulhac, L., 2015. Modelling Pollutant Dispersion in a Street Network. *Bound.-Layer Meteor.* 155, 157–187.
14. Buccolieri, R., Salizzoni, P., Soulhac, L., Garbero, V., Di Sabatino, S., 2015. The breathability of compact cities. *Urban Climate* 13, 73–93.

15. Carpentieri, M., Salizzoni, P., Robins, A., Soulhac, L., 2012. Evaluation of a neighbourhood scale, street network dispersion model through comparison with wind tunnel data. *Environ. Modell. Softw.* 37, 110–124.
16. Castro-Jimenez, J., Eisenreich, S.J., Mariani, G., Skejo, H., Umlauf, G., 2012. Monitoring atmospheric levels and deposition of dioxin-like pollutants in sub-alpine Northern Italy. *Atmos. Environ.* 56, 194–202.
17. Clavel-Chapelon F, van Liere MJ, Giubout C, Niravong MY, Goulard H, Le CC, et al. E3N, a French cohort study on cancer risk factors. E3N Group. *Etude Epidémiologique auprès de femmes de l'Education Nationale. Eur J Cancer*
18. Clavel-Chapelon, F., 2015. Cohort Profile: The French E3N Cohort Study. *Int J Epidemiol* 44, 801–809.
19. Cleverly, D., Ferrario, J., Byrne, C., Riggs, K., Joseph, D., Hartford, P., 2007. A General Indication of the Contemporary Background Levels of PCDDs, PCDFs, and Coplanar PCBs in the Ambient Air over Rural and Remote Areas of the United States. *Environ. Sci. Technol.* 41, 1537–1544.
20. Cordier, S., Lehébel, A., Amar, E., Anzivino-Viricel, L., Hours, M., Monfort, C., Chevrier, C., Chiron, M., Robert-Gnansia, E., 2010. Maternal residence near municipal waste incinerators and the risk of urinary tract birth defects. *Occup Environ Med* 67, 493–499.
21. Cordioli, M., Ranzi, A., De Leo, G.A., Lauriola, P., 2013. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators [WWW Document]. *Journal of Environmental and Public Health*.
22. Dai, D., Oyana, T.J., 2008. Spatial variations in the incidence of breast cancer and potential risks associated with soil dioxin contamination in Midland, Saginaw, and Bay Counties, Michigan, USA. *Environmental Health* 7, 49.
23. Duan, J., Tan, J., Wang, S., Hao, J., Chai, F., 2012. Size distributions and sources of elements in particulate matter at curbside, urban and rural sites in Beijing. *J Environ Sci (China)* 24, 87–94.
24. European Commission, 2000. Ambient air pollution by AS, CD and NI compounds. European Commission.
25. Fabre, P., Daniau, C., Gorla, S., de Crouy-Chanel, P., Empeur-Bissonnet, P., 2008. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. Synthèse. Saint-Maurice: InVS.
26. Faure E, Danjou AMN, Clavel-Chapelon F, Boutron-Ruault M-C, Dossus L, Fervers B. Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environmental Health*. 2017;16:15.
27. Floret, N., Viel, J.-F., Lucot, E., Dudermel, P.-M., Cahn, J.-Y., Badot, P.-M., Mauny, F., 2006. Dispersion Modeling as a Dioxin Exposure Indicator in the Vicinity of a Municipal Solid Waste Incinerator: A Validation Study. *Environ. Sci. Technol.* 40, 2149–2155.
28. Hanna, S., Chang, J., 2012. Acceptance criteria for urban dispersion model evaluation. *Meteorology and Atmospheric Physics* 116, 133–146.
29. Julin, B., Wolk, A., Bergkvist, L., Bottai, M., Akesson, A., 2012. Dietary cadmium exposure and risk of postmenopausal breast cancer: a population-based prospective cohort study. *Cancer Res.* 72, 1459–1466.

30. Kabir, E.R., Rahman, M.S., Rahman, I., 2015. A review on endocrine disruptors and their possible impacts on human health. *Environmental Toxicology and Pharmacology* 40, 241–258.
31. Langlois, P.H., Brender, J.D., Suarez, L., Zhan, F.B., Mistry, J.H., Scheuerle, A., Moody, K., 2009. Maternal residential proximity to waste sites and industrial facilities and conotruncal heart defects in offspring. *Paediatr. Perinat. Epidemiol.* 23, 321–331.
32. Liu, R., Nelson, D.O., Hurley, S., Hertz, A., Reynolds, P., 2015. Residential exposure to estrogen disrupting hazardous air pollutants and breast cancer risk: the California Teachers Study. *Epidemiology* 26, 365–373.
33. Lohmann, R., Jones, K.C., 1998. Dioxins and furans in air and deposition: a review of levels, behaviour and processes. *Sci. Total Environ.* 219, 53–81.
34. Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, Guha N, Baan R, Mattock H, Straif K; International Agency for Research on Cancer Monograph Working Group IARC. The carcinogenicity of outdoor air pollution. *Lancet Oncol.* 2013 Dec;14(13):1262-3
35. Meneses, M., Schuhmacher, M., Domingo, J.L., 2004. Health risk assessment of emissions of dioxins and furans from a municipal waste incinerator: comparison with other emission sources. *Environment International* 30, 481–489. doi:10.1016/j.envint.2003.10.001
36. Nawrot, T., Plusquin, M., Hogervorst, J., Roels, H.A., Celis, H., Thijs, L., Vangronsveld, J., Van Hecke, E., Staessen, J.A., 2006. Environmental exposure to cadmium and risk of cancer: a prospective population-based study. *Lancet Oncol.* 7, 119–126.
37. Nriagu, J.O., 1989. A global assessment of natural sources of atmospheric trace metals. *Nature* 338, 47–49.
38. Nzihou, A., Themelis, N.J., Kemiha, M., Benhamou, Y., 2012. Dioxin emissions from municipal solid waste incinerators (MSWIs) in France. *Waste Management* 32, 2273–2277.
39. Organization, W.H., Safety, I.P. on C., 1992. Cadmium : environmental aspects.
40. Pacyna, J.M., Pacyna, E.G., 2001. An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide. *Environ. Rev.* 9, 269–298.
41. Pope, C.A.I., Ezzati, M., Dockery, D.W., 2009. Fine-Particulate Air Pollution and Life Expectancy in the United States. *New England Journal of Medicine* 360, 376–386.
42. Pronk, A., Nuckols, J.R., De Roos, A.J., Airola, M., Colt, J.S., Cerhan, J.R., Morton, L., Cozen, W., Severson, R., Blair, A., Cleverly, D., Ward, M.H., 2013. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case-control study. *Environ. Health* 12, 20.
43. Quaß, U., Fermann, M.W., Bröker, G., 2004. The European Dioxin Air Emission Inventory Project - Final Results (PDF Download Available). *Chemosphere* 54, 1319–1327.
44. Quaß, U., Fermann, M.W., Bröker, G., 2000. Steps towards a European dioxin emission inventory. *Chemosphere* 40, 1125–1129.

45. Ranzi, A., Fano, V., Erspamer, L., Lauriola, P., Perucci, C.A., Forastiere, F., 2011. Mortality and morbidity among people living close to incinerators: a cohort study based on dispersion modeling for exposure assessment. *Environ. Health* 10, 22.
46. Ries, L., Harkins, D., Krapcho, M., Mariotto, A., Miller, B., Feuer, E., Clegg, L., Eisner, M., Horner, M.-J., Howlander, N., Hayat, M., Hankey, B., Edwards, B., 2006. SEER Cancer Statistics Review, 1975-2003. Public Health Faculty Publications.
47. Salehi, F., Turner, M.C., Phillips, K.P., Wigle, D.T., Krewski, D., Aronson, K.J., 2008. Review of the etiology of breast cancer with special attention to organochlorines as potential endocrine disruptors. *J Toxicol Environ Health B Crit Rev* 11, 276–300.
48. Salizzoni, P., Soulhac, L., Mejean, P., 2009. Street canyon ventilation and atmospheric turbulence. *Atmos. Environ.* 43, 5056–5067.
49. Serveau, L., Boutang, J., Allemand, N., Fontelle, J.-P., Vincent, J., 2013. Source CITEPA / format SECTEN – Inventaire des émissions de polluants atmosphériques et de gaz à effet de serre en France séries sectorielles et analyses étendues. CITEPA, Paris.
50. Soulhac, L., Chi Vuong Nguyen, P., Volta, P., Salizzoni, P., 2017. The model SIRANE for atmospheric urban pollutant dispersion. PART III: Validation against NO<sub>2</sub> yearly concentration measurements in a large urban agglomeration *Atmospheric Environment*. 167, 377-388.
51. Soulhac, L., Garbero, V., Salizzoni, P., Mejean, P., Perkins, R.J., 2009. Flow and dispersion in street intersections. *Atmos. Environ.* 43, 2981–2996.
52. Soulhac, L., Perkins, R.J., Salizzoni, P., 2008. Flow in a street canyon for any external wind direction. *Bound.-Layer Meteor.* 126, 365–388.
53. Soulhac, L., Puel, C., Duclaux, O., Perkins, R.J., 2003. Simulations of atmospheric pollution in Greater Lyon an example of the use of nested models. *Atmos. Environ.* 37, 5147–5156.
54. Soulhac, L., Salizzoni, P., Cierco, F.-X., Perkins, R., 2011. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmospheric Environment* 45, 7379–7395.
55. Soulhac, L., Salizzoni, P., Mejean, P., Didier, D., Rios, I., 2012. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmospheric Environment* 49, 320–337.
56. Soulhac, L., Salizzoni, P., Mejean, P., Perkins, R.J., 2013. Parametric laws to model urban pollutant dispersion with a street network approach. *Atmos. Environ.* 67, 229–241.
57. Strincone, M., Fino, A., Cattani, G., Catrambone, M., Pirrone, N., 2013. Emissions, air concentrations and atmospheric depositions of arsenic, cadmium, lead and nickel in Italy in the last two decades: A review of recent trends in relation to policy strategies adopted locally, regionally and globally. *E3S Web of Conferences* 1, 38003.
58. Sugita, K., Endo, O., Asada, S., Goto, S., Yajima, H., Ishii, T., 2004. Dioxin Concentration of Particulate and Gaseous Substances in the Ambient Air by Particle Size — Comparison between Specimens Collected in Summer and in Winter at Tokyo —. *Journal of Health Science* 50, 9–16.

59. Tchounwou, P.B., Yedjou, C.G., Patlolla, A.K., Sutton, D.J., 2012. Heavy Metals Toxicity and the Environment. *EXS* 101, 133–164.
60. Tung, J.W.T., Yu, J.Z., Lau, A.K.H., Louie, P.K.K., 2005. Abundance and sources of ambient dioxins in Hong Kong: A review of dioxin measurements from 1997 to 2001. *Chemosphere* 59, 1387–1398.
61. Umlauf, G., Christoph, E.H., Eisenreich, S.J., Mariani, G., Paradiz, B., Vives, I., 2010. Seasonality of PCDD/Fs in the ambient air of Malopolska Region, southern Poland. *Environ. Sci. Pollut. Res.* 17, 462–469.
62. U.S. Environmental Protection Agency (EPA). (2013) The National Dioxin Air Monitoring Network (NDAMN): Report of the results of atmospheric measurements of polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs), and dioxin-like polychlorinated biphenyls (PCBs) in rural and remote areas of the United States from June 1998 through November 2004. National Center for Environmental Assessment, Washington, DC; EPA/600/R-13/183F. Available at <http://epa.gov/ncea> (accessed 20 October 2017).
63. Zambon, P., Ricci, P., Bovo, E., Casula, A., Gattolin, M., Fiore, A.R., Chiosi, F., Guzzinati, S., 2007. Sarcoma risk and dioxin emissions from incinerators and industrial plants: a population-based case-control study (Italy). *Environmental Health* 6, 19.

### 3. Conclusion

Ce travail a été une étape préliminaire aux études épidémiologiques GEO3N et XENAIR qui étudient l'exposition atmosphérique aux dioxines et au cadmium des femmes de la cohorte E3N, pour évaluer le risque de cancer du sein associé à cette exposition.

A notre connaissance, il s'agit de la première étude qui évalue la variabilité spatio-temporelle des concentrations de dioxines et de cadmium sur près de deux décennies, en comparant les concentrations modélisées et mesurées, en prenant en compte une grande variété de types de sources (industrielle, diffuse, liée au trafic ou au chauffage) ainsi que les concentrations de fond. C'est également la première étude comparant les concentrations prédites par un modèle de dispersion aux concentrations de dioxines mesurées dans l'air ambiant.

Les résultats obtenus pour les estimations des émissions montrent que ceux-ci sont affectés par des différences non-négligeables par rapport aux mesures et se situent généralement entre  $\pm 100\%$  pour le cadmium et  $\pm 200\%$  pour la dioxine. Ces résultats restent en cohérence avec d'autres comparaisons observées dans la littérature (Breivik et al. 2006). Les comparaisons, sur deux ans (2007-2008), entre les concentrations modélisées et mesurées sont caractérisées par une sous-estimation des concentrations moyennes (résultats similaires observés par Liu et al (2015) pour le cadmium et Garcia et al (2013) pour d'autres métaux lourds) (Garcia et al. 2014; Liu et al. 2015b). Pour les dioxines, le modèle n'est pas capable de reproduire un certain nombre de pics de concentration, particulièrement en hiver. Cependant, il est à noter que les comparaisons avec des mesures réalisées en centre-ville montrent de meilleurs résultats que la comparaison avec les mesures ponctuelles réalisées en zone industrielles proche de sources de dioxines. En ce qui concerne le cadmium, le modèle capture mieux les variations hebdomadaires de concentrations, y compris les pics observés pendant l'hiver.

Malgré ces différences, les résultats remplissent les critères de validation pour les deux polluants, tels que définis par Chang et Hanna (2012) pour un modèle en milieu urbain (Hanna and Chang 2012). De plus, l'article a permis, en étudiant la variation du ratio dioxine/cadmium, d'illustrer l'impact important des événements locaux et imprédictibles (incendie, brûlages de câbles, etc.) sur les concentrations de dioxines.

La comparaison avec les résultats précédemment publiés montre que nos estimations des concentrations en dioxines et cadmium à la localisation des sujets de la cohorte E3N sont similaires aux valeurs fournies par d'autres études plus limitées dans l'espace et dans le temps (Cordier et al. 2010; Floret et al. 2006; Liu et al. 2015; Zambon et al. 2007).



Ce travail met en évidence que les sources industrielles peuvent être considérées, sur la décennie 1990-2000, comme les principales sources affectant l'exposition des sujets et les principales sources d'expositions différentielles. Au cours de la période suivante (2000-2008), nous avons identifié une baisse générale des niveaux d'exposition, mais surtout une disparition des zones de forte concentration et donc une réduction de la gamme des niveaux d'exposition (les différences d'exposition entre les sujets étaient beaucoup plus importantes dans la décennie 1990-2000 que dans la période 2000-2008). Pour le scénario le plus récent, à savoir 2008, nous observons que les principaux contributeurs aux concentrations sont dû à des sources diffuses diverses pour les dioxines et au trafic pour le cadmium. Ces observations soulignent la difficulté des études épidémiologiques à classer les sujets en fonction de leurs niveaux d'exposition, en raison des faibles concentrations observées actuellement. En raison de la classification des dioxines comme cancérigène sans seuil, des estimations de l'exposition avec une grande précision sont néanmoins nécessaires même si les niveaux d'exposition sont faibles au cours de la période 2000-2008.

L'une des principales limites de notre étude est l'évaluation de la contribution du brûlage de câbles. Les émissions sont estimées à partir de la quantité de câbles volés au niveau national et il est quasiment impossible de connaître où ces câbles sont brûlés. À défaut d'informations précises, les émissions sont positionnées en fonction de la densité de la population dans le cadastre d'ATMO-AURA. Sur notre zone d'étude, ces émissions, réparties sur des mailles de 1x1km, créent des zones d'émissions diffuses et homogènes centrées sur le centre-ville lyonnais. La contribution relative de ces émissions prend de l'importance au fil des années, ce qui induit une probable surestimation des expositions moyennes en centre-ville et donc un biais important dans l'estimation des expositions pour les années récentes.

Une autre limite est l'absence de mesures des émissions pour les sources diffuses ou le trafic. Ceci empêche de quantifier précisément les incertitudes liées à ce type d'émissions. Pour évaluer les implications de ces incertitudes sur la classification des sujets pour les années récentes, de nouvelles recherches doivent être menées. De plus, la part des émissions liées au trafic devra probablement être recalculée après mise à jour des facteurs d'émission suite au « scandale Volkswagen ».

En conclusion, les résultats de l'article confirment la nécessité d'utiliser des outils de simulation numériques pour estimer les expositions et la contribution des différentes sources à l'exposition (et leur variation dans le temps). L'utilisation de ce type de modèle est essentielle pour améliorer l'évaluation des expositions à long terme et réduire les biais des études épidémiologiques.





## **CHAPITRE IV. EVALUATION DE L'EXPOSITION ATMOSPHERIQUE CHRONIQUE AUX DIOXINES ET CADMIUM EN FRANCE (1990-2008)**

*Création d'un indicateur spatial de l'exposition chronique aux dioxines et au cadmium, validation en comparaison avec un modèle Gaussien et application dans le cadre du projet GEO3N*

## 1. Introduction

L'une des principales difficultés rencontrées par les études épidémiologiques portant sur les liens entre cancer et exposition atmosphérique, est de caractériser précisément l'exposition sur de grandes zones d'études et de longues périodes. Quantifier les expositions historiques aux polluants atmosphériques demande de connaître les antécédents résidentiels et professionnels des sujets et d'estimer les concentrations à une échelle spatiale et temporelle fine, en particulier dans les centres urbains, pour limiter au maximum les risques de biais de classement des sujets (Basagana et al. 2013; Cordioli et al. 2013). Alors que la couverture spatiale des réseaux de mesure de la qualité de l'air a régulièrement augmenté au cours des dernières années, différentes approches ont été développées pour représenter de manière adéquate la variation spatiale des polluants et reconstruire rétrospectivement les expositions. On peut citer notamment les systèmes d'information géographique (SIG), les modèles de Land Use Regression (LUR) (Eeftens et al. 2012; Gulliver et al. 2013; Hoek et al. 2008) ou encore les modèles de dispersions atmosphériques (de Hoogh et al. 2014).

Dans le cas des dioxines et du cadmium, le très faible nombre de mesures combiné à la forte diminution des émissions au cours des 30 dernières années (Serveau et al. 2013) limitent l'utilisation de modèles de dispersion ou de LUR pour évaluer les expositions historiques sur un large territoire. Pour ce type de polluants, l'utilisation d'un indicateur SIG semble donc plus pertinente. Un certain nombre d'études ont utilisé ces indicateurs SIG pour étudier les liens entre pollution atmosphérique et risque de cancer (Gulliver and Briggs 2011; Hoek et al. 2008; Pronk et al. 2013; White et al. 2009; Yu et al. 2006). Certaines études utilisent des indicateurs, se basant uniquement sur la distance sujet-source (Langlois et al. 2009; Pronk et al. 2013), d'autre emploient des indicateurs plus complexes incluant notamment des paramètres météorologiques ou topographiques (Gulliver and Briggs 2011; Han et al. 2013; Hoek et al. 2001a; Zou et al. 2009). Pour les dioxines, seuls Pronk et al (2013) ont étudié l'exposition chronique aux dioxines à l'aide d'un SIG, en utilisant un inventaire historique des émissions industrielles de dioxines (1987-2000). Cependant aucun paramètre météorologique n'était pris en compte dans cette étude.

Une autre problématique concerne l'exhaustivité des sources d'expositions. Pour les dioxines, comme les installations issues de divers secteurs industriels peuvent émettre ces polluants (incinérateur, cimenterie, fonderie, aciérie, sidérurgie) (Lohmann and Jones 1998), un certain nombre d'études, utilisant des SIG ou des modèles numériques, se concentrent uniquement sur les rejets des incinérateurs (Ashworth et al. 2013; Cordioli et al. 2013; Fabre et al. 2008). De plus, pour évaluer précisément les expositions chroniques aux dioxines, une caractérisation précise des technologies mises en place par les industriels au cours du temps est nécessaire. En effet, pour des installations semblables, l'absence ou la présence de système d'abattement spécifique peut entraîner des variations dans les émissions d'un facteur 20 (UNEP 2013b).

La plupart du temps, les émissions utilisées dans les études proviennent d'inventaires réalisés pour d'autres utilisations. Les auteurs n'ont donc pas la connaissance des paramètres techniques à l'échelle individuelle des sources (Pronk et al. 2013).

### *Objectifs et méthodes*

L'objectif de ce travail était de développer et de calibrer, par comparaison avec le modèle gaussien SIRANE, un indicateur de l'exposition atmosphérique des femmes de la cohorte E3N aux polluants émis par des sources industrielles, afin d'étudier l'association avec le risque de cancer du sein (dans le cadre des projets GEO3N et XENAIR). L'objectif primaire était de construire un indicateur viable pour estimer l'exposition industrielle aux dioxines entre 1990 et 2008, puis d'étudier son applicabilité pour évaluer les expositions industrielles de cadmium sur la même période et sur la même population.

L'indicateur avait comme paramètres de base : la localisation des sources industrielles sur l'ensemble du territoire français entre 1990 et 2008 (chapitre 2), les émissions annuelles de dioxines associées et la localisation des sujets E3N sur la même période. Nous avons testé six paramètres à inclure dans l'indicateur pour améliorer ses performances et le rapprocher d'un modèle gaussien (données météorologiques, vitesse de sortie des fumées, hauteur de cheminée, etc.). Ces paramètres ont été identifiés à partir d'une revue de la littérature (annexes 1 et 2) et des travaux sur la modélisation de la dispersion atmosphérique. Pour identifier la combinaison de paramètres la plus performante, nous avons reparti les sujets en quintiles d'expositions en utilisant plus de 80 combinaisons de paramètres. Puis nous avons identifié la combinaison permettant d'obtenir la classification des sujets la plus proche de celle fournie par le modèle SIRANE. Nous avons ainsi estimé les concordances entre les classifications en calculant les kappas pondérés ( $w_k$ ) et le coefficient de détermination ( $R^2$ ). Ce travail de calibration a été réalisé dans trois zones représentatives des lieux de vie des femmes E3N (zones urbaines, rurales et côtières) et sur trois années distinctes, soit neuf scénarios. Une fois fixée, la combinaison de paramètres permettant d'obtenir les meilleures concordances a été testée pour 18 autres scénarios permettant de tester la reproductibilité des résultats. Ces 18 scénarios étaient répartis en deux groupes : neuf avec une nouvelle localisation aléatoire de sujets, et neuf pour lesquels les valeurs d'émissions de dioxines avaient été remplacées par des valeurs d'émissions de cadmium. Ce dernier set de validation a donc également permis d'évaluer l'applicabilité de l'indicateur pour évaluer l'exposition de l'atmosphère au cadmium.

## 2. Article n°3

### **Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutant emissions from industrial sources**

*Thomas Coudon<sup>1,2</sup>, Aurélie Danjou<sup>1,2</sup>, Elodie Faure<sup>1</sup>, Delphine Praud<sup>1,3</sup>, Francesca Romana Mancini<sup>4</sup>, Gianluca Severi<sup>4</sup>, Pietro Salizzoni<sup>5</sup> and Beatrice Fervers<sup>1,2,3</sup>*

<sup>1</sup> *Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France*

<sup>2</sup> *Université Claude Bernard Lyon 1, Villeurbanne 69100, France*

<sup>3</sup> *INSERM 1052, CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, 69373, France.*

<sup>4</sup> *Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.*

<sup>5</sup> *Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully 69134, France*

#### *Reference:*

Coudon T, Danjou AMN, Faure E, Praud D, Mancini FR, Severi G, Salizzoni P, Fervers B. 2018. Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutant emissions from industrial sources. Accepted, Environment Health (2018).

# **Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutant emissions from industrial sources**

Thomas Coudon<sup>1,2</sup>, Aurélie Marcelle Nicole Danjou<sup>1,2</sup>, Elodie Faure<sup>1</sup>, Delphine Praud<sup>1,3</sup>, Gianluca Severi<sup>4</sup>, Francesca Romana Mancini<sup>4</sup>, Pietro Salizzoni<sup>5</sup>, Béatrice Fervers<sup>1,2,3</sup>

<sup>1</sup> Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France

<sup>2</sup> Université Claude Bernard Lyon 1, Villeurbanne 69100, France

<sup>3</sup> INSERM 1052, CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, 69373, France.

<sup>4</sup> Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.

<sup>5</sup> Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully 69134, France

## **E-mail addresses:**

Thomas.coudon@lyon.unicancer.fr, DanjouA@fellows.iarc.fr, Elodie.FAURE@lyon.unicancer.fr, delphine.praud@lyon.unicancer.fr, Gianluca.SEVERI@gustaveroussy.fr, francesca.mancini@gustaveroussy.fr, pietro.salizzoni@ec-lyon.fr, Beatrice.FERVERS@lyon.unicancer.fr

## **Corresponding author:**

Béatrice Fervers

Département Cancer & Environnement, Centre Léon Bérard, 69373 Lyon CEDEX 08, France

Email : beatrice.ferves@lyon.unicancer.fr Phone: +33 (0)4 78 78 28 01

Francesca R. Mancini

Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018),

Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.

Email : francesca.mancini@gustaveroussy.fr Phone: +33 (0) 1 4211 5864



# Abstract

## Background

Dioxins are environmental and persistent organic carcinogens with endocrine disrupting properties. A positive association with several cancers, including risk of breast cancer has been suggested.

## Objectives

This study aimed to develop and assess performances of an exposure metric based on a Geographic Information System (GIS) through comparison with a validated dispersion model to estimate historical industrial dioxin exposure for its use in a case-control study nested within a prospective cohort.

## Methods

Industrial dioxin sources were inventoried over the whole French territory ( $n > 2500$ ) and annual average releases were estimated between 1990 and 2008. In 3 selected areas (rural, urban and urban-costal), dioxin dispersion was modelled using SIRANE, an urban Gaussian model and exposure of the French E3N cohort participants was estimated. The GIS-based metric was constructed, calibrated and compared to SIRANE results using a set of parameters (local meteorological data, characteristics of industrial sources, e.g. emission intensity and stack height), by calculating weighted kappa statistics ( $wk$ ) and coefficient of determination ( $R^2$ ). Furthermore, as performance evaluation, the final GIS-based metric was tested to assess atmospheric exposure to cadmium.

## Results

The concordance between the GIS-based metric and the dispersion model for dioxin exposure estimate was strong ( $wk$  median=0.78 (1<sup>st</sup> quintile=0.72, 3<sup>rd</sup> quintile =0.82) and  $R^2$  median=0.82 (1<sup>st</sup> quintile=0.71, 3<sup>rd</sup> quintile=0.87)). We observed similar performance for cadmium.

## Conclusions

Our study demonstrated the ability of the GIS-based metric to reliably characterize long-term environmental dioxin and cadmium exposures as well as the pertinence of using dispersion modelling to construct and calibrate the GIS-based metric.

**Keywords:** dioxins, cadmium, exposure assessment, GIS, dispersion modelling, SIRANE, cancer

**Short title:** GIS-based metric to assess air pollutant exposure

## Background

Outdoor air pollution has been consistently linked to a range of adverse health effects, including cancer and has been estimated responsible for 3.1 million premature annual deaths worldwide [1]. Outdoor air pollution is a mixture of multiple pollutants originating from a large variety of sources, including various carcinogens classified as carcinogenic (Group 1) or probably carcinogenic (Group 2A) to humans by the International Agency for Research on Cancer (IARC) in 2013 [2]. Several recent epidemiological studies investigated the association between outdoor air pollution and breast cancer risk but results remain inconsistent. For traffic-related air pollutants (nitrogen oxides, particulate matter and polycyclic aromatic hydrocarbons), case-control studies highlighted positive associations [3–8] while prospective cohort studies did not report significant associations [6,9–12]. Only few studies investigated the effect of airborne exposure to dioxin and cadmium (Cd) on breast cancer risk and overall results are inconclusive and need to be further investigated [6,13,14]. Finally, inconsistency across results of studies on xenoestrogens exposure in ambient air and breast cancer risk from the literature could be explained by methodological limitations, including lack of historical measurements and insufficient statistical power [15–18]. The multiplicity of exposure sources and the latency between exposure and cancer occurrence represent major challenges and require to precisely characterize the spatial-temporal variability of exposures over large areas and long time-periods. In numerous studies, the lack of past residential history and historical air pollutant exposure assessment at a fine spatial and temporal scales may have resulted in exposure misclassification, hence likely to have contributed to imprecise risk estimates [13,15,19]. While numerous facilities, including metal industries and cement kilns are likely to emit dioxins [17,20,21], the majority of published studies restricted exposure assessment to incinerators [15,22,23]. Moreover, information on the evolution of the facility technologies and activity over time is needed to precisely assess long-term dioxin exposure [24].

Furthermore, to overcome the lack of measurement data, previous studies investigating the impact of dioxin exposure on linear distance or presence/absence of the source as a measure of exposure [15,17,25,26]. Yet, the use of these kinds of proxies has been shown to be subject to substantial misclassification [15]. Therefore dispersion modelling, is considered as more reliable to accurately assess exposure with a high spatial resolution [15].

While the spatial coverage of ambient air quality monitoring networks has steadily increased in recent years, different approaches have been developed to adequately represent the spatial variation of pollutants and reconstruct retrospective exposure for earlier periods, including atmospheric dispersion modeling [27], land use regression (LUR) models [28–30] and Geographical Information Systems (GIS). In the case of dioxins, the lack of monitoring data in France as input source and the

sharp decrease of emissions over the past 30 years [31] limit the use of LUR modelling to assess dioxin exposure over the French territory. Furthermore, the use of atmospheric dispersion models has to face several difficulties due the fact that the pollutant sources and receptors are distributed over wide area (the whole France). Performing simulations in domains of these size requires significant computational resources.

GIS are being increasingly used in environmental epidemiological studies and are based on the residential proximity to distinct types of environmental exposure sources (e.g. industrial facilities, traffic roads) considered as an exposure surrogate. Moreover, GIS allow integrating meteorological and topographical parameters influencing pollutant dispersion, into a GIS-based exposure metric [32–35]. The positional accuracy of subjects' residences is a key requisite to avoid exposure misclassification [36,37].

The objective of this study was to develop and calibrate, through comparison with a Gaussian dispersion model, a GIS-based metric assessing long-term airborne dioxin exposures of participants in a case-control study nested in the French E3N cohort (Etude Epidémiologique après de femmes de la Mutuelle Générale de l'Education Nationale) in order to investigate the association with breast cancer risk.

## **Methods**

### **Study population**

The E3N study is an ongoing prospective cohort involving 98,995 French female volunteers, born between 1925 and 1950, and members of a national teachers' health insurance plan, and aimed to identify female cancer risk factors [38]. Since 1990, participants completed self-administered questionnaires, mailed every 2-3 years, on health status, medical history and main cancer risk factors (hormonal, reproductive, dietary and lifestyle-related factors). The E3N cohort is the French component of the European Prospective Investigation into Cancer and nutrition (EPIC) study [39,40].

For this nested case-control study, 5,455 incident invasive breast cancer cases (confirmed by pathology report) and 5,455 matched controls were selected from the E3N cohort. Participants were included if they had filled in their home address at baseline, lived in the metropolitan French territory between 1990 and 2008, and had not reported cancer at baseline. According to an incidence density sampling, one control per case was randomly selected and matched to cases for age, department of residence, menopausal status, date of recruitment or blood draw and existence of a biological sample.

## **Preliminary steps to the GIS-based metric**

The GIS-based assessment of industrial airborne exposure to dioxins, was based on a national inventory of dioxin emitting sources and estimation of the annual dioxin emissions, the geocoding of dioxin sources and of the residential history of study participants over the study period.

### *Dioxin industrial sources: inventory and characterization*

A detailed retrospective inventory of industrial sources likely to emit or to have emitted dioxins between 1990 and 2008, over the whole France, was carried out, taking into account waste incineration, medical waste incineration, metal production, heat and power generation, production of mineral products, chemicals and consumer goods, and crematoria. Dioxin emitting sources were identified through institutional and national databases, namely GEREPA (annual emission reporting on pollutant and greenhouse gas releases), IREP (French National Registry of Pollutant Emissions) and ICPE (Inventory of French classified facilities). Industrial unions, nationally recognized associations and whistleblowers were contacted to identify additional facilities, in particular for earlier years and a structured questionnaire was sent to these identify facilities in order to collect additional information on technical characteristics. From 1990 to 2008, a total of 2,626 dioxin emitting sources were inventoried over the French territory.

Dioxin releases mostly depend on the combustion process and conditions, the type of material burned and the flue gas treatment. Along with the inventory, technical characteristics were collected, including geographic location of the facilities (geographic coordinates and addresses), operation periods and rates, stack height, process characteristics and flue gas cleaning technologies for the different periods. Using the technical characteristics of dioxin sources, intensity of dioxin emissions were estimated using the Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases developed by the United Nation Environmental Program [41]. The Toolkit allowed for a classification of dioxin emissions according to activity sectors and operating characteristics. The industrial sources inventoried were classified into Toolkit categories and a dioxin emission factor (in g-TEQ/t) was assigned. For each distinct operation periods, annual dioxin emission intensity (in g-TEQ/year) was estimated by multiplying the emission factor by the operation rate. A general decrease was observed in the dioxin source annual average emissions (Table 1) due to the improvement of gas cleaning technologies.

### *Geocoding of the residential history and emitting sources*

As this current work is a part of a GIS-based metric assessing long-term airborne dioxin exposures of the participants of the whole cohort-nested case-control study, location of facilities and geocoding of the participants' residential history was performed for the whole France, using ArcGIS

software (ArcGIS Locator, ESRI, Redlands, CA, USA), BD Adresse for ArcGIS and its reference street network database, BD Adresse® (National Geographic Institute, IGN, Saint Mandé, France) that includes 26 million addresses. The choice of the geocoding method was based on the accuracy of positioning to limit exposure misclassification and has been described in detail elsewhere [36]. Overall 28,511 residential addresses in metropolitan France were geocoded collected from the questionnaires sent between 1990 and 2005; 78.1% of the subject's residences were geocoded to the address, 26% required manual checking and among them 17.4% have been corrected.

Facilities identified through the inventory were located based on their geographic coordinates when available, or geocoded using the addresses collected. All automatically geocoded locations were manually checked and repositioned at the stack using current and historical aerial photography from IGN. Facilities for which the stack was not visible on current or past aerial images, were located as accurately as possible in descending order according to positional accuracy, *i.e.* at the centroid of the building; at the centroid of the parcel; or at the town hall of the municipality. Among the 2,626 sources inventoried in France between 1990 and 2008, 82% were positioned at the stack, 13% at the centroid of the building and 5% at the parcel.

## **Development of the GIS-based metric**

### *Identification of relevant parameters*

Based on the review of relevant publications in the literature [17,32,34,35,42–44] and previous work on dioxin and cadmium modelling [45], the following parameters were included in the GIS-based metric to characterize exposure and classify study subjects according to their airborne dioxin exposure from industrial sources (Table 2): subject's residence-to-source distance, wind direction and speed, exhaust smoke velocity and stack height. For all parameters, a setting sample was tested to identify the relevant combination of parameters (Table 2). The selected parameters were combined with the sources and subject locations, the sources' annual emission intensity (in g-TEQ/year) and the exposure duration (in years).

### *Integration of the selected parameters*

Proximity to dioxin emitting sources is a key parameter to assess individual dioxin exposure [15]. Based on the literature, three different buffer sizes were tested in the calibration step, corresponding to a circular buffer around each dioxin emitting source of respectively 3km, 5km and 10km [17,42–44]. A matrix of distance source-to-residence was calculated using the Point-to-Point function in ArcGIS software. Subjects residing outside the buffer were considered as non-exposed [44]. Inside the buffer, the decrease in dioxin concentrations was calculated testing different residence-to-source distance decline patterns (Table 2).

Pollutant atmospheric dispersion depends on meteorological conditions, in particular wind direction and speed. We included wind parameters in the GIS-based metric using data from the French national meteorological service, METEO France, based on 727 areas of homogeneous weather pattern (AHWP) in metropolitan France. To take into account continuity of meteorological monitoring over the study period (1990-2008), the 727 AHWP were grouped, in collaboration with METEO France, into 223 meteorological areas according to the proximity to the reference station, homogeneity of weather pattern and type of area (plain, mountain, hillside and valley). The measurement stations provided information on wind direction and speed between 1990 and 2008 on an hourly basis with an average completeness of data around 90% for the whole study period. Each inventoried industrial source was affected to a reference meteorological station representative of the local weather conditions. For a given year, if the completeness was below 75%, data from the nearest meteorological station were used.

To integrate wind directions into the GIS, a GIS data layer named contributing area for dioxin dispersion (CADD), was created for each dioxin emitting source, based on equal segments of the wind rose and proportion of annual wind blow and speed at each segment (Figure 1). In the calibration step, we tested CADDs of 10°, 30° and 90°. Furthermore, for CADDs of 10°, weighted contribution (50%, 25%) of adjacent segments was assessed [44]. To take into account the wind speed, we reported the average annual wind speed for each segment and adjusted the decrease in dioxin concentration according to this average wind speed. For a given subject, exposure was estimated based on the CADD in which the subject's residence was located. The process was managed with ArcMap 10.1.

## **GIS-based metric calibration**

### *Specific areas and periods selected for metric calibration*

We restricted the calibration of the GIS-based metric to three French geographical areas presenting typical topographical and meteorological patterns and distinct numbers and types of dioxin emitting sources (Figure 3), representative of the living environment of the majority of the E3N subjects: Lyon (a non-mountainous highly urban area), Le Bugey (a rural area) and Le Havre (a coastal medium-size urban area). We further selected three distinct years over the study period with a 6-years gap (1996, 2002 and 2008) presenting, for each area, different emission intensities and meteorological parameters.

In the Lyon area, addresses were collected from the E3N questionnaires sent in 1997, 2002 and 2005 (respectively for 1996, 2002 and 2008 scenarios). Due to residential mobility, loss to follow-up and, exclusion of case-control pairs following breast cancer diagnosis, the number of subjects decreased over the study period (312 subjects in 1996, 173 in 2002 and 68 in 2008). In this area, 80.8% of the study subjects were located at the address level and 16.9% at the street address. The

number of E3N subjects residing in Le Bugey and Le Havre was lower ( $n < 30$ ). In order to obtain a relevant number of subjects residences for the present calibration study, 150 simulated subjects' residences were randomly located in each of these two areas. The accuracy of the participant residential addresses geocoded was similar with the one observed in the whole France (see 2.2.).

Overall, 40 dioxin sources were identified and located in Lyon, nine in Le Havre and one in Le Bugey. All the 50 sources were located at the stack. In Lyon and Le Havre, the number of sources inventoried was sufficient to perform a comparison between the Gaussian model and the GIS-based metric (Table 1). As only a single industrial source was inventoried for the Bugey, we added three virtual sources with differing parameters (stack height, smoke exit velocity, emission intensity, etc.) and annually varying dioxin emissions. The choice of technical parameters and emissions of these three virtual sources were based on average values observed during the inventory step, for three emission domains: heat and power generation, metal production and crematoria.

Annual average dioxin emission estimates for each area and year are described in Table 1. Similarly to the national level, a decrease is observed, from 1996 to 2008, for each of the selected areas. These changes in emissions were due to new emission policies, applied at the end of the 90's [24]. The observed decrease is sharper for Le Havre (from 4.13 g/year in 2002 to 0.05g/year in 2008) than for the two other sites, and is mainly due to the closing of a major industrial sources after 2002.

#### *Modelling of airborne dioxin exposure*

Modelling of dioxin atmospheric dispersion was performed in Lyon, Le Bugey and Le Havre, for the three distinct periods (1996, 2002, and 2008) using the SIRANE atmospheric dispersion model. SIRANE is an urban dispersion model that integrates a specific module to simulate pollutant dispersion within a built environment [46,47] considering local meteorological conditions and geometry of the streets. The SIRANE model has been validated by means of wind tunnel experiments [48,49] and open field measurement data [46,50]. Note that a detailed validation of the model, based on NO<sub>2</sub> concentration levels, was performed over the whole Lyon urban agglomeration for the year 2008, i.e. one of the three domain considered in this study. Details on SIRANE dioxin modelling results are available in Additional file 1.

Average annual dioxin concentrations (in fg-TEQ/m<sup>3</sup>) were calculated at each E3N subject's residence location for each year and categorized into quintiles according to the average dioxin concentration. Modelled dioxin concentrations served as reference for the calibration and validation of the GIS-based metric. Modelled dioxin concentrations for 2008 in Lyon, were compared to weekly average concentrations provided by a monitoring station located in the city center of Lyon since 2007 [45].

## **Performance evaluation**

Once selected the parameters to be integrated in the GIS-based metric and best-performing parameter combination, two evaluations were completed to assess the performance of the GIS-based metric.

For the first performance evaluation, 150 new virtual subjects were randomly distributed in each of the three areas. The periods (1996, 2002 and 2008), source locations and emissions remained unchanged. Given the size of the Lyon area (34km x 30km), and the variation of population density over the area, the distribution was weighted according to population density in the Lyon area.

The second performance evaluation was made to assess the performance of the GIS-based metric to assess exposure to pollutants other than dioxins. For this purpose, we applied the GIS-based metric to assess cadmium exposure. Cadmium emitting industrial sources were inventoried through institutional and public databases, industrial unions and nationally recognized associations. Using emission factors provided by the OMINEA database (Organization and Methods of the National Inventories of the Atmospheric Releases in France), from the Inter-professional Technical Centre for studies of Air Pollution (CITEPA), and technical parameters collected through similar steps as for dioxins, annual cadmium emissions were estimated for each industrial source. A total of 2,686 cadmium emitting sources were inventoried over the French national territory from 1990 to 2008. At the national level, annual cadmium average emissions decreased from 24.4 kg/year to 9.5 kg/year between 1996 and 2008 (Table 1). More details are provided in a previous publication [45]. In the 3 selected areas, cadmium emissions estimates showed similar levels and trends to those observed at the national level (Table 1).

As for dioxins, the GIS-based metric was applied in Lyon, Le Bugey and Le Havre for 1996, 2002 and 2008 and modelling of cadmium atmospheric dispersion was performed using SIRANE for the same years and areas. Details on SIRANE cadmium modelling results are available in Additional file 2.

## **Statistical analyses**

The GIS-based metric was defined through comparison with annual dioxin concentrations estimated by SIRANE, for each scenario, which allowed selecting the most relevant parameters and their combination to be included in the GIS-based metric.

For the development of the GIS-based metric and its performance evaluation, we compared the categorical dioxin exposure classification (based on quintiles) of study subjects between the GIS-based metric and the SIRANE dispersion modelling for the three locations (Lyon, Le Bugey, Le Havre) and the three different years (1996, 2002 and 2008),



Agreement between quintiles of dioxin concentrations from modelling and quintiles of the GIS-based metric estimates was calculated using weighted kappa coefficients ( $w\kappa$ ) and their 95% confidence intervals (95% CI). Weighted kappa coefficients assign less importance to discrepancies between adjacent quintiles and higher weight to larger discrepancies [15]. The determination coefficient  $R^2$  was also computed for each scenario. Analyses were performed using SAS software version 9.4 (SAS Institute Inc, Cary,NC).

## Results

### *Calibration of the GIS-based metric*

We observed higher  $w\kappa$  for a buffer size of 10km ( $w\kappa$  ranging from 0.42 to 0.71 depending on the parameter combinations) compared to buffers of 3km and 5km ( $w\kappa$  ranging from 0.31 to 0.42 and from 0.34 to 0.60 for 3km and 5km respectively). Taking into account wind direction, using CADDs increased the metric performance, in particular for CADDs with 10° wind rose segments (Table 3 for Lyon area; see Additional file 3 for Le Havre area). The highest agreement was obtained for an inverse subject's residence-source square distance weighting. The addition of wind speed parameters reduced the agreement:  $w\kappa$  (95% CI) ranging from 0.71 (0.67, 0.76) to 0.81 (0.79, 0.88) without wind speed and from 0.58 (0.49, 0.66) to 0.71 (0.70, 0.82) after integration of wind speed in the Lyon scenario (see Additional file 4).

The integration of the source technical parameters (exhaust smoke velocity and stack height) did not further improve agreement between the categorical dioxin exposure classification by the two methods (see Additional file 5 and 6), except for Le Havre in 2008 where the integration of the stack height of the major industrial source (240 m) led to considerable improvement of the agreement between the GIS-based metric and the dispersion modelling ( $w\kappa$  (95%CI) from 0.64 (0.59, 0.71) and 0.78 (0.72, 0.84) without and with integration of stack height respectively). Given the absence of impact of the stack height on the weighted kappa coefficients in all other scenarios and the low completeness of data for stack height at the national level (36%), it was decided to integrate into the GIS-based metric only stack heights above 90 m, corresponding to 3 times the median stack height of the 2,626 sources over the whole France. Data were available for all sources with stack height above 90 m.

Based on the performance of the nearly 80 different parameter combinations in the nine calibration scenarios (3 areas over 3 years, see Figure 2 and 3), we retained the following formula for the GIS-based metric:

$$\text{GIS – based metric (g – TEQ.m}^{-2}\text{)} = \sum_{j=1}^J \sum_{i=1}^I EI_i \times t_i \times d_{ij}^{-2} \times F_{ij} \times \left( \frac{h_{median}}{h_{ij}} \right)^a \quad (i)$$

<sup>a</sup> if  $h_{ij}$  is greater than 90 m

where the indices  $j$  and  $i$  indicate the places of residence ( $j=1,\dots,J$ ) and the industrial sources ( $i=1,\dots,I$ ), respectively, and where  $EI_i$  is the annual dioxin emission intensity of each industrial source (in g-TEQ/year),  $t_i$  is its emission period duration (in years),  $d_{ij}$  is the residence-to-source distance (in m),  $F_{ij}$  is the factor taking into account the percentage of time during which the wind is blowing in a CADD (see “Integration of the selected parameters” subsection) as to induce a transport from the industrial source  $i$  to the location of the participant  $j$  (accounting for the weighted contribution of the sector of the buffer in which the participant is located in, *i.e.* 50% and that of the two adjacent sectors, *i.e.* 25% each),  $h_{ij}$  is the stack height (in m) and  $h_{median}$  is the median value of the other sources’ stack height (in m) in a 10km buffer. Atmospheric dioxin exposure was expressed in  $\mu\text{g-TEQ/m}^2$  and annually computed.

Using the formula (i), the nine calibration scenarios, yielded wk (95% CI) ranging from 0.71 (0.67, 0.76) to 0.84 (0.79, 0.88), corresponding to a “substantial” to “almost perfect” agreement between the categorical dioxin exposure classification into quintiles by the GIS-based metric, and the SIRANE dispersion modelling (Figure 3). The  $R^2$  ranged from 0.68 to 0.90 for the same scenarios (Table 4). The scatterplots (Figure 2) illustrates the ability of the GIS-based metric to provide robust estimates of the subject’s exposure in comparison to modelling results.

#### *Performance evaluation of the GIS-based metric*

Once established, the final GIS-based metric (i) was applied to the new samples of virtual subjects randomly located in Lyon, Le Bugey and Le Havre for 1996, 2002 and 2008. The calibration scenarios yielded wk (95% CI) ranging from 0.58 (0.49, 0.66) to 0.86 (0.81, 0.91) (Table 4). Weighted kappa were below 0.6 (0.58) for one scenario (Lyon, 2008). The determination coefficients ranged from 0.30 to 0.94 for these nine scenarios (Table 4) with one scenario under 0.5 (Havre, 1996;  $R^2$ : 0.3) and 4 scenarios above 0.85.

The GIS-based metric (i) was further applied to estimate airborne cadmium exposure (Table 4). We observed a “substantial” to “almost perfect” agreement for categorical dioxin exposure classification (quintiles) between the GIS-based metric and the SIRANE modelling with wk (95% CI) ranging from 0.69 (0.64, 0.73) to 0.86 (0.82, 0.91). The wk remained consistent across sites, periods, emission intensities and number of sources. The  $R^2$  for these nine scenarios ranged from 0.66 to 0.86.

## Discussion

GIS are being increasingly used in epidemiological studies to compute exposure surrogates based on distance between study population and exposure sources or using more advanced methods integrating meteorological and topographical data, residential history as well as characteristics of industrial sources [17,34,42–44]. We developed a GIS-based metric in this way, filling methodological gaps of the existing literature to improve accuracy of airborne dioxin exposure estimates. To our knowledge, this is the first study calibrating a GIS-based metric evaluating its performance to estimate dioxin or more largely, air pollutant exposure, from industrial sources, to be used in an epidemiological study, through comparison to the SIRANE model [50].

The combination of parameters demonstrated consistently reliable estimates for the two performance evaluations with differing number of sources, subjects, meteorological and topographical conditions. Weighted kappa coefficients indicated “substantial” to “almost perfect” [51] agreement with the modelled estimates, except for one scenario (Lyon, 2008; first performance evaluation set). The relative poor performance in this scenario ( $w_k$  (95% CI): 0.58 (0.49, 0.66)) may be explained by the situation for Lyon in 2008 involving a high density of sources ( $n=33$ ) with very low and homogenous source emissions increasing the difficulty to differentially classify study subjects into quintiles [52]. Note that in this scenario, we obtained a high determination coefficient ( $R^2$ : 0.88).

Likewise, the values of the coefficient of determination ( $R^2$ ) demonstrated reliable estimates by the GIS-based metric with median values of 0.80. Only one scenario (Le Havre, 1996, first performance evaluation set) showed a poor performance. This observation can be explained by two outliers (out of the 150 virtual subjects for this area) that were 10 times more exposed than all other subjects and underestimated by the GIS-based metric. The value of  $R^2$  increased from 0.30 to 0.95 after the exclusion of these two subjects, which were randomly located. It is worth noting that for this same situation, a substantial agreement was obtained for the categorical exposure classification ( $w_k$  (95% CI): 0.71(0.64, 0.78)). Regarding the large range of  $R^2$  obtained, it should be noted that the determination coefficient can be highly influenced by a few subjects with extreme values.

Pronk et al, studied dioxin exposure using a GIS but did not perform calibration or validation [17]. The performance observed in our study for parameter combinations such as used by Pronk et al. 2013 (inverse distances squared-weighted emission, winds not taken into account and buffer limiter to 5km), was much lower, with  $w_k$  ranging from 0.31 to 0.45.

Overall, the level of performance of our metric is comparable to other studies conducted in Europe on more frequently studied pollutants [15,32,34,42]. Unlike dioxins and cadmium, numerous measurement are available for NO<sub>2</sub> [42], PM<sub>10</sub> [15,32] and black smoke [34], and facilitates calibration and validation. Cordioli et al. (2013), using PM<sub>10</sub> as a tracer for incinerator pollutant

emissions, evaluated the agreement of categorical exposure classifications of subjects between PM10 concentration maps and different exposure methods, including a simple indicator based on distance between exact address location and incinerator. Using this simple indicator, the authors obtained a  $wk$  of 0.61. We observed similar results when our GIS-based metric was only based on source-subject distance ( $wk$  (95% CI) ranging from 0.61 (0.59, 0.96) to 0.78 (0.73, 0.84)). Three studies conducted a calibration using measurements [32,34,42] and two of them completed a performance evaluation. Gulliver and Briggs (2011) obtained a good agreement between ambient air measurements and metric estimates ( $R^2$ : 0.67-0.77) for annual PM10 concentrations in London [32]. These results were comparable to performance realised by a Gaussian model ( $R^2$ : 0.71-0.77). Vienneau et al. (2009), yielded a determination coefficient of 0.60, using a GIS-based moving window approach, in comparison with NO<sub>2</sub> measurements across Europe but provided an estimation with limited accuracy (1x1km<sup>2</sup>) [42].

The estimates observed for cadmium demonstrated the ability of the GIS-based metric to assess, as for dioxins, pollutant exposures from industrial sources with behaviours similar to dioxins, *i.e.* pollutants with particle size around 1 to 10  $\mu$ m and absence of chemical reaction in the atmosphere. Note that, the SIRANE model for both pollutants used a similar setting (pollutant modelled as a passive scalar with an average diameter of 1 $\mu$ m) but with different average densities (dioxins: 321.9 g/mol; cadmium: 112.4 g/mol) [45].

Strengths of our study included the use of a GIS and its application on a large area [34,42], over a long and retrospective time-period, at the individual subject's address and considering the residential history over the study period [13,15]. Moreover, they suggest that GIS-based metrics provide a robust alternative to LUR models in case studies with few measured data (limiting the use of LUR models) or wide domains with large number of sources (requiring high computational resources for the use of atmospheric dispersion models).

Our GIS-based metric required a retrospective inventory of industrial sources, the estimation of their emission intensity, the geocoding of the participants' residential history and of the industrial sources, and the computation of local meteorological data and source technical parameters in the GIS. The highest agreement of the parameters combination with the dispersion model was reported for the inversed square residence-to-source distance, as observed in three other studies [17,34,42]. The buffer size around sources was set at 10km, which is consistent with the literature, with buffer sizes ranging from 3 to 10km for industrial sources [17,42,44]. As the exposure due to traffic was not included in the GIS-based metric, smaller buffer sizes were not retained for the current study. Furthermore, while the parameters integrated in our GIS-based metric are consistent with several studies from the literature [15,32,35,43], the inclusion of the wind speed in the parameters' combination of the GIS-based metric, did not further improve the agreement statistics despite wind speed being known to impact pollutant dispersion, and this may constitute a possible source of error. While other studies

integrated wind direction [44] or both wind direction and wind speed in their metric [32,42,43], no other study evaluated the impact of wind speed on the metric performance. Similarly, we did not identify studies that evaluated the impact of stack height or other industrial sources technical parameters on the metric performance. Integration of additional parameters, such as pluviometry or outdoor temperature, as well as regional background concentrations [34], may further improve the performance of the GIS-based metric.

Our study was based on a multi-source approach, considering multiple emission sectors (waste incineration, metal production, cement industries, etc.) and the evolution over time of the facilities' technical characteristics. In the absence of dioxin monitoring data, emission intensity of the industrial sources was estimated using a standardized tool (<http://toolkit.pops.int/>). Pronk et al. also used a historical dioxin emission inventory (1987-2000) and a multi-source approach limited, however to some activity sectors [17]. The accuracy of our emission estimates were directly linked to the quality of the information collected from industrial facilities on technical characteristics of the sources. Previous studies often used emission inventories conducted for other purposes [17,42].

The accuracy of address location may have important implications on misclassification of individual exposure, depending on the spatial concentration gradient of the exposure. Although the residential addresses of the study subjects were not recorded initially to be geocoded and used for the assessment of environmental exposure, their accuracy can be considered precise enough to limit misclassification bias, in particular for urban subjects in the present study [15,33,36].

In the case of dioxin, domestic activities are known to poorly contribute to airborne dioxin exposure compared to industrial sources for earlier years [53]. Some non-industrial sources have however become non negligible for more recent periods and may lead to the underestimation of the exposure and to a non-differential misclassification bias. Other punctual and non-industrial sources can emit relatively high amounts of dioxins and cadmium [31,54], locally and in a short time scale, such as biomass fires, manufactured good burnings, cable burning, outdoor burning and illegal landfills in the early 1990s. These sources could not be considered in this GIS-based metric due to the difficulty of their retrospective inventory, their geolocalization and the estimation of their dioxin emissions. To reconstruct the subjects' historical dioxin exposure, additional methods should be employed to take into account these type of emissions as well as historical dietary exposure [55].

## **Conclusion**

In this study, a GIS-based metric has been developed and evaluated in order to estimate the retrospective airborne dioxin exposure of participants of a cohort-nested case-control study. The final metric combined residential distance to facilities, wind direction and proportion of the year blown and technical parameters of the facilities. This combination of parameters showed reliable estimates in comparison to an atmospheric dispersion model [50] across different scenarios. The GIS-based metric also provided reliable estimates for cadmium exposure from industrial sources and might be able to assess exposure to other air pollutants with similar properties and behaviour than dioxins and cadmium (i.e. heavy metals, PM10 etc.), in particular when monitoring data are lacking.

## **Declaration**

### **Ethics approval and consent to participate**

The E3N cohort study received ethical approval from the French National Commission for Computerized Data and Individual Freedom (Commission Nationale de l'Informatique et des Libertés, CNIL), and all women in the study provided signed informed consent for participation into the study. This approval covers all questionnaires designed within the framework of the cohort, and no new approval is needed for individual projects.

### **Consent for publication**

Not Applicable

### **Availability of data and materials**

The datasets generated or analysed during the current study are available from the corresponding author on reasonable request

### **Competing interests**

The authors declare that they have no competing interests.

### **Funding**

This work was supported by public funding from the French Environment and Energy Management Agency (ADEME, Grant N° 1306C0031), the Cancéropôle Lyon Auvergne Rhône-Alpes (CLARA), the Regional Committee of the French League against Cancer of the Saone et Loire Region, and was carried out in partnership with the ARC Foundations for Cancer Research. The E3N study is financially supported by the French League against Cancer, the Mutuelle Générale de l'Education Nationale, the Institut Gustave Roussy, the Institut National de la Santé et de la Recherche Médicale. Thomas Coudon and Aurélie Danjou were supported by a doctoral fellowship of University Lyon 1. Delphine Praud was supported by a post-doctoral fellowship of the French League against Cancer.

**Author's contribution**

TC and EF have made contribution to the data acquisition. TC performed the modeling of atmospheric exposure. EF and TC geocoded the addresses and performed the GIS analyses. AMND and DP performed the statistical analyses. TC drafted the manuscript. TC, EF, AMND, DP and BF have interpreted the results. EF, AMND, DP, BF, PS, GS and FRM have made contribution to the revision of the manuscript. All authors have read and approved the submitted manuscript.

**Acknowledgments**

We gratefully acknowledge Meteo France for providing meteorological data and the E3N cohort participants for providing data. We thank Camille Denis, Guillaume Harel and Hassan Hourani for their work on the dioxin and cadmium sources inventory and characterization. We also thank Charlotte Carretero and Maxime Guilou for geocoding.



## References

1. WHO. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report. 2013 [cited 2017 Aug 29]; Available from: <http://www.euro.who.int/en/health-topics/environment-and-health/air-quality/publications/2013/review-of-evidence-on-health-aspects-of-air-pollution-revihaap-project-final-technical-report>
2. Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al. The carcinogenicity of outdoor air pollution. *Lancet Oncol.* 2013;14:1262–3.
3. Bonner MR, Han DW, Nie L, Rogerson P, Vena JE, Muti P, et al. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer Epidemiol. Biomarkers Prev.* 2005;14:53–60.
4. Crouse DL, Goldberg MS, Ross NA, Chen H, Labrèche F. Postmenopausal breast cancer is associated with exposure to traffic-related air pollution in Montreal, Canada: a case-control study. *Environ. Health Perspect.* 2010;118:1578–83.
5. Hystad P, Villeneuve PJ, Goldberg MS, Crouse DL, Johnson K. Exposure to traffic-related air pollution and the risk of developing breast cancer among women in eight Canadian provinces: A case-control study. *Environ. Int.* 2015;74:240–8.
6. Liu R, Nelson DO, Hurley S, Hertz A, Reynolds P. Residential exposure to estrogen disrupting hazardous air pollutants and breast cancer risk: the California Teachers Study. *Epidemiol. Camb. Mass.* 2015;26:365–73.
7. Mordukhovich I, Beyea J, Herring AH, Hatch M, Stellman SD, Teitelbaum SL, et al. Vehicular Traffic-Related Polycyclic Aromatic Hydrocarbon Exposure and Breast Cancer Incidence: The Long Island Breast Cancer Study Project (LIBCSP). *Environ. Health Perspect.* 2016;124:30–8.
8. Nie J, Beyea J, Bonner MR, Han D, Vena JE, Rogerson P, et al. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer Causes Control.* 2007;18:947–55.
9. Andersen ZJ, Ravnskjaer L, Andersen KK, Loft S, Brandt J, Becker T, et al. Long-term Exposure to Fine Particulate Matter and Breast Cancer Incidence in the Danish Nurse Cohort Study. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 2017;26:428–30.
10. Hart JE, Bertrand KA, DuPre N, James P, Vieira VM, Tamimi RM, et al. Long-term Particulate Matter Exposures during Adulthood and Risk of Breast Cancer Incidence in the Nurses' Health Study II Prospective Cohort. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 2016;25:1274–6.
11. Raaschou-Nielsen O, Andersen ZJ, Hvidberg M, Jensen SS, Ketzel M, Sorensen M, et al. Air pollution from traffic and cancer incidence: a Danish cohort study. *Environ. Health.* 2011;10:67.
12. Reding KW, Young MT, Szpiro AA, Han CJ, DeRoo LA, Weinberg C, et al. Breast Cancer Risk in Relation to Ambient Air Pollution Exposure at Residences in the Sister Study Cohort. *Cancer Epidemiol. Biomark. Prev. Publ. Am. Assoc. Cancer Res. Cosponsored Am. Soc. Prev. Oncol.* 2015;24:1907–9.
13. Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. Environmental pollutants and breast cancer - Epidemiologic studies. *Cancer.* 2007;109:2667–711.

14. Xu J, Ye Y, Huang F, Chen H, Wu H, Huang J, et al. Association between dioxin and cancer incidence and mortality: a meta-analysis. *Sci. Rep.* [Internet]. 2016 [cited 2017 Aug 3];6. Available from: <http://www.nature.com/articles/srep38012>
15. Cordioli M, Ranzi A, De Leo GA, Lauriola P. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators. *J. Environ. Public Health* [Internet]. 2013 [cited 2017 Aug 29]; Available from: <https://www.hindawi.com/journals/jep/2013/129470/>
16. Huang YL, Batterman S. Residence location as a measure of environmental exposure: a review of air pollution epidemiology studies. *J. Expo. Anal. Environ. Epidemiol.* 2000;10:66–85.
17. Pronk A, Nuckols JR, De Roos AJ, Airola M, Colt JS, Cerhan JR, et al. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case-control study. *Environ. Health.* 2013;12:20.
18. Viel J-F, Clement M-C, Hagi M, Grandjean S, Challier B, Danzon A. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int. J. Health Geogr.* 2008;7:4.
19. Basagana X, Aguilera I, Rivera M, Agis D, Foraster M, Marrugat J, et al. Measurement Error in Epidemiologic Studies of Air Pollution Based on Land-Use Regression Models. *Am. J. Epidemiol.* 2013;178:1342–6.
20. Lohmann R, Jones KC. Dioxins and furans in air and deposition: a review of levels, behaviour and processes. *Sci. Total Environ.* 1998;219:53–81.
21. Pacyna EG, Pacyna JM, Pirrone N. European emissions of atmospheric mercury from anthropogenic sources in 1995. *Atmos. Environ.* 2001;35:2987–96.
22. Ashworth DC, Fuller GW, Toledano MB, Font A, Elliott P, Hansell AL, et al. Comparative assessment of particulate air pollution exposure from municipal solid waste incinerator emissions. *J. Environ. Public Health.* 2013;2013:560342.
23. Fabre P, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. *Synthèse St.-Maurice InVS* [Internet]. 2008 [cited 2016 Oct 18]; Available from: [http://www.pdfarchive.info/pdf/E/Et/Etude\\_d\\_incidence\\_des\\_cancers\\_a\\_proximite\\_des\\_usines\\_d\\_incineration\\_d\\_ordures\\_menageres.pdf](http://www.pdfarchive.info/pdf/E/Et/Etude_d_incidence_des_cancers_a_proximite_des_usines_d_incineration_d_ordures_menageres.pdf)
24. Nzihou A, Themelis NJ, Kemiha M, Benhamou Y. Dioxin emissions from municipal solid waste incinerators (MSWIs) in France. *Waste Manag.* 2012;32:2273–7.
25. Langlois PH, Brender JD, Suarez L, Zhan FB, Mistry JH, Scheuerle A, et al. Maternal residential proximity to waste sites and industrial facilities and conotruncal heart defects in offspring. *Paediatr. Perinat. Epidemiol.* 2009;23:321–31.
26. Zambon P, Ricci P, Bovo E, Casula A, Gattolin M, Fiore AR, et al. Sarcoma risk and dioxin emissions from incinerators and industrial plants: a population-based case-control study (Italy). *Environ. Health.* 2007;6:19.
27. de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, et al. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ. Int.* 2014;73:382–92.

28. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. Development of Land Use Regression Models for PM<sub>2.5</sub>, PM<sub>2.5</sub> Absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European Study Areas; Results of the ESCAPE Project. *Environ. Sci. Technol.* 2012;46:11195–205.
29. Gulliver J, de Hoogh K, Hansell A, Vienneau D. Development and Back-Extrapolation of NO<sub>2</sub> Land Use Regression Models for Historic Exposure Assessment in Great Britain. *Environ. Sci. Technol.* 2013;47:7804–11.
30. Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, et al. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos. Environ.* 2008;42:7561–78.
31. Serveau L, Boutang J, Allemand N, Fontelle J-P, Vincent J. source CITEPA / format SECTEN – INVENTAIRE DES EMISSIONS DE POLLUANTS ATMOSPHERIQUES ET DE GAZ A EFFET DE SERRE EN FRANCE SERIES SECTORIELLES ET ANALYSES ETENDUES [Internet]. Paris: CITEPA; 2013 Apr. Available from: <http://www.citepa.org/fr/activites/inventaires-des-emissions/secten>
32. Gulliver J, Briggs D. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Sci. Total Environ.* 2011;409:2419–29.
33. Han D, Bonner MR, Nie J, Freudenheim JL. Assessing bias associated with geocoding of historical residence in epidemiology research. *Geospatial Health.* 2013;7:369–74.
34. Hoek G, Fischer P, Van den Brandt P, Goldbohm S, Brunekreef B. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J. Expo. Anal. Environ. Epidemiol.* 2001;11:459–69.
35. Zou B, Wilson JG, Zhan FB, Zeng Y. Air pollution exposure assessment methods utilized in epidemiological studies. *J. Environ. Monit.* 2009;11:475–90.
36. Faure E, Danjou AMN, Clavel-Chapelon F, Boutron-Ruault M-C, Dossus L, Fervers B. Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environ. Health.* 2017;16:15.
37. Jacquemin B, Johanna L, Anne B, Valerie S. Environmental Health Perspectives – Impact of Geocoding Methods on Associations between Long-term Exposure to Urban Air Pollution and Lung Function. *Environ. Health Perspect.* 2013;1054–1060.
38. Clavel-Chapelon F. Cohort Profile: The French E3N Cohort Study. *Int. J. Epidemiol.* 2015;44:801–9.
39. Riboli E. Nutrition and Cancer - Background and Rationale of the European Prospective Investigation into Cancer and Nutrition (epic). *Ann. Oncol.* 1992;3:783–91.
40. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European prospective investigation into cancer and nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 2002;5:1113–24.
41. UNEP. United Nations Environment Programme | Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs [Internet]. 2013 [cited 2016 Aug 22]. Available from: [http://toolkit.pops.int/Publish/Main/01\\_Index.html](http://toolkit.pops.int/Publish/Main/01_Index.html)
42. Vienneau D, de Hoogh K, Briggs D. A GIS-based method for modelling air pollution exposures across Europe. *Sci. Total Environ.* 2009;408:255–66.

43. White N, teWaterNaude J, van der Walt A, Ravenscroft G, Roberts W, Ehrlich R. Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study. *Environ. Health.* 2009;8:45.
44. Yu C-L, Wang S-F, Pan P-C, Wu M-T, Ho C-K, Smith TJ, et al. Residential exposure to petrochemicals and the risk of leukemia: Using geographic information system tools to estimate individual-level residential exposure. *Am. J. Epidemiol.* 2006;164:200–7.
45. Coudon T, Hourani H, Fervers B, Salizzoni P. Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon Metropolitan Area (France). *Environ. Int.* 2017;
46. Soulhac L, Salizzoni P, Mejean P, Didier D, Rios I. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmos. Environ.* 2012;49:320–37.
47. Soulhac L, Salizzoni P, Cierco F-X, Perkins R. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmos. Environ.* 2011;45:7379–95.
48. Ben Salem N, Garbero V, Salizzoni P, Lamaison G, Soulhac L. Modelling Pollutant Dispersion in a Street Network. *Bound.-Layer Meteorol.* 2015;155:157–87.
49. Carpentieri M, Salizzoni P, Robins A, Soulhac L. Evaluation of a neighbourhood scale, street network dispersion model through comparison with wind tunnel data. *Environ. Model. Softw.* 2012;37:110–24.
50. Soulhac L, Nguyen CV, Volta P, Salizzoni P. The model SIRANE for atmospheric urban pollutant dispersion. PART III: Validation against NO<sub>2</sub> yearly concentration measurements in a large urban agglomeration. *Atmos. Environ.* 2017;167:377–88.
51. Viera AJ, Garrett JM, others. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 2005;37:360–363.
52. Gulliver J, de Hoogh K, Fecht D, Vienneau D, Briggs D. Comparative assessment of GIS-based methods and metrics for estimating long-term exposures to air pollution. *Atmos. Environ.* 2011;45:7072–80.
53. Quass U, Fermann M, Broker G. The European dioxin air emission inventory project - Final results. *Chemosphere.* 2004;54:1319–27.
54. Lee RGM, Green NJL, Lohmann R, Jones KC. Seasonal, Anthropogenic, Air Mass, and Meteorological Influences on the Atmospheric Concentrations of Polychlorinated Dibenzo-p-dioxins and Dibenzofurans (PCDD/Fs): Evidence for the Importance of Diffuse Combustion Sources. *Environ. Sci. Technol.* 1999;33:2864–71.
55. Danjou AM, Fervers B, Boutron-Ruault M-C, Philip T, Clavel-Chapelon F, Dossus L. Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. *Breast Cancer Res.* 2015;17:39.

## **Additional files**

Additional file 1 / FigureAF5.docx

Title: Boxplot of the average dioxin concentrations (fg/m<sup>3</sup>), modeled at the E3N location in Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008 with the SIRANE model

Additional file 2 / FigureAF6.docx

Title: Boxplot of the average cadmium concentrations (ng/m<sup>3</sup>), modeled at the E3N location in Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008 with the SIRANE model

Additional file 3 / Table AF1.docx

Title: Weighted kappa coefficients and CI95 % with different CADD and distance decline patterns in the Le Bugey scenarios.

Additional file 4 / Table AF4.docx

Title: Weighted kappa coefficients and CI95 % in Lyon and Le Bugey with and without taking into account wind speed.

Additional file 5 / Table AF2.docx

Title: Weighted kappa coefficients and CI95 % in Lyon, Le Bugey and Le Havre with different source technical parameter settings.

Additional file 6 / Table AF3.docx

Title: Weighted kappa coefficients and CI95 % in Lyon, Le Bugey and Le Havre with and without technical parameters.

## Tables

Table 1 – Average dioxin (g-TEQ/year) and cadmium (kg/year) release estimates of the sources inventoried for 1996, 2002 and 2008 in France, Lyon, Le Havre and Le Bugey

Dioxins	1996		2002		2008	
Areas	<i>n</i>	mean ± SD	<i>n</i>	mean ± SD	<i>n</i>	mean ± SD
France	918	2.76 ± 8.08	850	0.66 ± 2.64	854	0.08 ± 0.42
Lyon	23	3.08 ± 12.68	24	0.58 ± 1.91	28	0.02 ± 0.03
Le Havre	7	7.07 ± 11.93	8	4.13 ± 11.42	7	0.05 ± 0.05
Le Bugey <sup>a</sup>	4	1.63 ± 2.27	4	0.51 ± 0.38	4	0.31 ± 0.22
Cadmium	1996		2002		2008	
Areas	<i>n</i>	mean ± SD	<i>n</i>	mean ± SD	<i>n</i>	mean ± SD
France	963	24.4 ± 121.1	906	20.7 ± 121.2	909	9.5 ± 51.5
Lyon	23	11.6 ± 31.5	24	6.1 ± 5.7	28	1.1 ± 2.4
Le Havre	7	25.3 ± 25.6	8	11.1 ± 11.5	7	4.7 ± 4.9
Le Bugey <sup>a</sup>	4	27.8 ± 41.9	4	15.0 ± 23.4	4	3.5 ± 4.4

*n*: number of industrial sources; <sup>a</sup> For the need of the calibration three of the four sources in Le Bugey were virtual.

Table 2 – Parameters and modalities tested to calibrate the GIS-based metric

Type	Parameters	Modalities tested
Geographical	Impact zone of dioxin industrial emissions	Circular buffer with 3, 5 and 10 km radius
	Residence to source distance decline	$\frac{1}{d}$ ; $\frac{1}{d^{1.5}}$ ; $\frac{1}{d^2}$ ; $\frac{1}{d^3}$ ; $e^d$ ; $\frac{e^d}{d}$ ; $e^{-d} \cdot d$
Meteorological	Wind frequency	Equal wind rose segment sizes (CADDs) of 90°, 30°, and 10° ± Weighted contribution of adjacent segments
	Wind speed	Variation of the residence to source distance decline pattern according to the hourly wind speed
Technical	Exhaust smoke velocity	$\frac{1}{v}$ ; $\frac{1}{v^2}$
	Stack height	$\frac{1}{h}$ ; $\frac{1}{h^{0.5}}$ ; $\frac{1}{h^2}$

*d*: residence to source distance; *v*: exhaust smoke velocity; *h*: facility's stack height

Table 3 – Weighted kappa coefficients ( $\kappa$ ) and corresponding 95% confidence intervals (95% CI) for the calibration in the Lyon area (1996, 2002 and 2008)

Integration of wind direction into the GIS-based metric ( $\kappa$ (95% CI))						
Distance decline	Year	Without integration of wind direction	90°; 4 CADD	30°; 12 CADD	10°; 36 CADD	10° with adjacent segments; weighted at 50%
1/d	1996	0.64 (0.59, 0.69)	0.65 (0.59, 0.70)	0.57 (0.51, 0.63)	0.64 (0.59, 0.69)	0.56 (0.50, 0.63)
	2002	0.61 (0.54, 0.68)	0.63 (0.55, 0.70)	0.61 (0.54, 0.68)	0.60 (0.52, 0.67)	0.60 (0.52, 0.67)
	2008	0.71 (0.61, 0.82)	0.48 (0.34, 0.63)	0.56 (0.42, 0.70)	0.69 (0.58, 0.80)	0.67 (0.56, 0.78)
1/d <sup>2</sup>	1996	0.61 (0.55, 0.66)	0.70 (0.66, 0.75)	0.76 (0.72, 0.81)	0.75 (0.71, 0.79)	0.71 (0.67, 0.76)
	2002	0.60 (0.53, 0.68)	0.71 (0.65, 0.78)	0.80 (0.75, 0.85)	0.83 (0.76, 0.88)	0.82 (0.78, 0.87)
	2008	0.64 (0.52, 0.75)	0.62 (0.49, 0.74)	0.66 (0.54, 0.77)	0.83 (0.75, 0.90)	0.79 (0.70, 0.88)
e <sup>d</sup>	1996	0.52 (0.46, 0.59)	0.63 (0.56, 0.67)	0.63 (0.57, 0.68)	0.39 (0.32, 0.46)	0.65 (0.58, 0.70)
	2002	0.52 (0.43, 0.60)	0.65 (0.57, 0.72)	0.68 (0.60, 0.75)	0.32 (0.23, 0.42)	0.65 (0.58, 0.73)
	2008	0.58 (0.45, 0.71)	0.60 (0.47, 0.73)	0.62 (0.49, 0.74)	0.43 (0.27, 0.58)	0.71 (0.61, 0.81)
e <sup>d</sup> /d	1996	0.48 (0.41, 0.55)	0.57 (0.51, 0.63)	0.59 (0.53, 0.65)	0.64 (0.59, 0.69)	0.59 (0.54, 0.65)
	2002	0.52 (0.42, 0.61)	0.63 (0.54, 0.71)	0.65 (0.57, 0.73)	0.60 (0.52, 0.67)	0.65 (0.57, 0.73)
	2008	0.58 (0.46, 0.70)	0.54 (0.40, 0.68)	0.60 (0.47, 0.73)	0.69 (0.58, 0.80)	0.62 (0.49, 0.74)
1/d <sup>1.5</sup>	1996	0.63 (0.57, 0.68)	0.71 (0.66, 0.75)	0.70 (0.65, 0.75)	0.72 (0.68, 0.76)	0.66 (0.61, 0.71)
	2002	0.61 (0.54, 0.68)	0.73 (0.67, 0.79)	0.74 (0.68, 0.80)	0.75 (0.69, 0.81)	0.73 (0.67, 0.79)
	2008	0.66 (0.54, 0.77)	0.58 (0.46, 0.70)	0.62 (0.49, 0.75)	0.85 (0.78, 0.92)	0.79 (0.70, 0.88)
1/d <sup>3</sup>	1996	0.52 (0.46, 0.58)	0.64 (0.58, 0.69)	0.67 (0.62, 0.72)	0.69 (0.65, 0.74)	0.67 (0.62, 0.72)
	2002	0.55 (0.47, 0.64)	0.67 (0.60, 0.74)	0.77 (0.70, 0.83)	0.77 (0.70, 0.83)	0.75 (0.69, 0.81)
	2008	0.58 (0.45, 0.71)	0.60 (0.46, 0.74)	0.62 (0.48, 0.75)	0.66 (0.54, 0.77)	0.69 (0.58, 0.81)
e <sup>d</sup> *d	1996	0.55 (0.49, 0.61)	0.64 (0.59, 0.69)	0.67 (0.63, 0.72)	0.24 (0.16, 0.32)	0.69 (0.65, 0.74)
	2002	0.55 (0.47, 0.63)	0.68 (0.62, 0.75)	0.74 (0.68, 0.81)	0.16 (0.06, 0.27)	0.74 (0.68, 0.80)
	2008	0.60 (0.47, 0.72)	0.62 (0.50, 0.74)	0.69 (0.58, 0.81)	0.31 (0.15, 0.47)	0.69 (0.60, 0.79)

d: residence to source distance; CADD: contributing area for dioxin dispersion

Table 4 - Weighted kappa coefficients ( $w\kappa$ ), corresponding 95% confidence intervals (95% CI) and  $R^2$  obtained with the final the final metric for Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008.

Years	Calibration					
	Areas	Lyon	Le Bugey	Le Havre		
1996	$w\kappa$ (95% CI) $R^2$	0.71 (0.67, 0.76) 0.78	0.79 (0.73, 0.85) 0.83	0.74 (0.68, 0.80) 0.68		
2002	$w\kappa$ (95% CI) $R^2$	0.84 (0.79, 0.88) 0.83	0.82 (0.76, 0.87) 0.90	0.77 (0.72, 0.83) 0.69		
2008	$w\kappa$ (95% CI) $R^2$	0.81 (0.72, 0.89) 0.87	0.73 (0.66, 0.79) 0.77	0.78 (0.72, 0.84) 0.84		
Years	Performance evaluation 1: new virtual subjects randomly distributed					
	Areas	Lyon <sup>a</sup>	Le Bugey	Le Havre		
1996	$w\kappa$ (95% CI) $R^2$	0.71 (0.64, 0.78) 0.95	0.86 (0.81, 0.90) 0.86	0.71 (0.64, 0.78) 0.30		
2002	$w\kappa$ (95% CI) $R^2$	0.67 (0.60, 0.74) 0.94	0.86 (0.81, 0.91) 0.55	0.74 (0.68, 0.80) 0.80		
2008	$w\kappa$ (95% CI) $R^2$	0.58 (0.49, 0.66) 0.88	0.84 (0.79, 0.89) 0.75	0.79 (0.73, 0.84) 0.62		
Years	Performance evaluation 2: cadmium emissions					
	Areas	Lyon	Le Bugey	Le Havre		
1996	$w\kappa$ (95% CI) $R^2$	0.69 (0.64, 0.73) 0.74	0.83 (0.78, 0.87) 0.71	0.72 (0.67, 0.78) 0.65		
2002	$w\kappa$ (95% CI) $R^2$	0.86 (0.82, 0.91) 0.83	0.82 (0.77, 0.87) 0.86	0.77 (0.71, 0.82) 0.66		
2008	$w\kappa$ (95% CI) $R^2$	0.73 (0.64, 0.83) 0.82	0.69 (0.63, 0.76) 0.83	0.74 (0.68, 0.80) 0.80		

<sup>a</sup> Adjusted on population density



# Figures

Figure 1 – Integration of wind direction parameters into the GIS-based metric through creating a data layer of contributing area for dioxin dispersion (CADD)

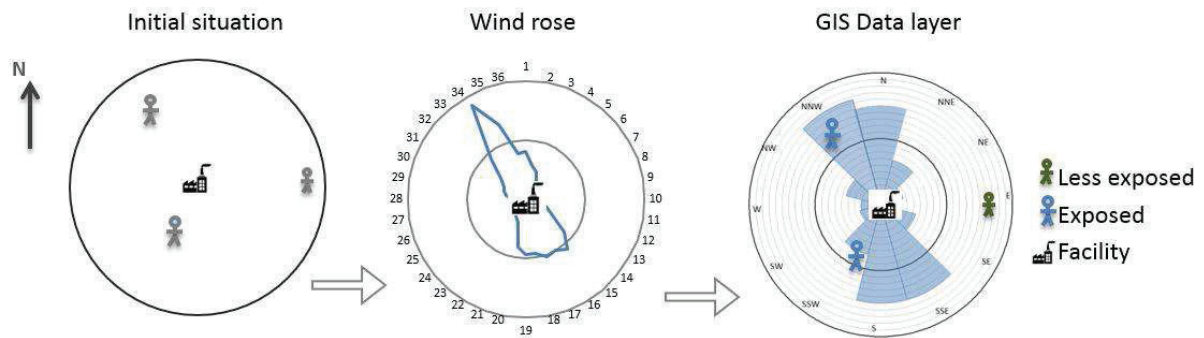


Fig 1.a: Creation of GIS data layer: Division of the wind rose into equal segments and assignment of wind direction and proportion of the year blown to each segment

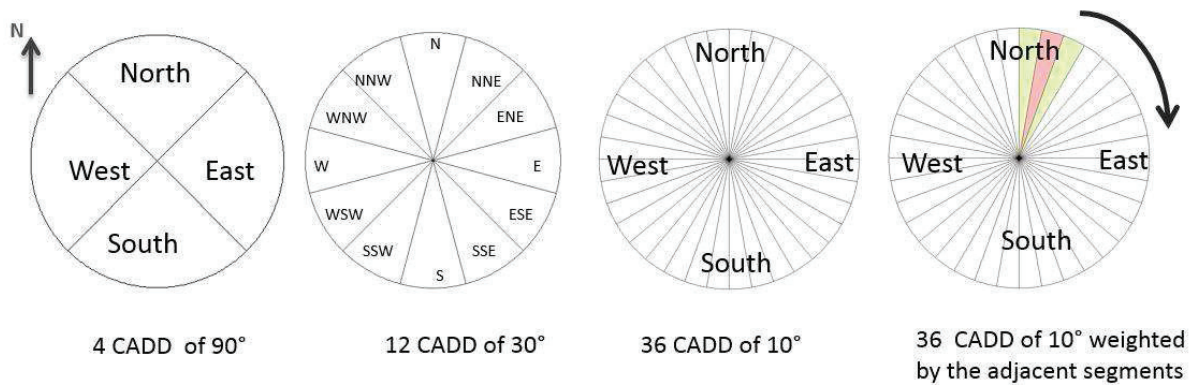


Fig 1.b: Degrees of CADD tested to consider wind directions in the GIS metric

Figure 2 – Scatterplots of the calibration of the subjects exposure estimates using the GIS-based metric and the SIRANE model in Lyon, Le Havre and Le Bugey for the years 1996, 2002 and 2008

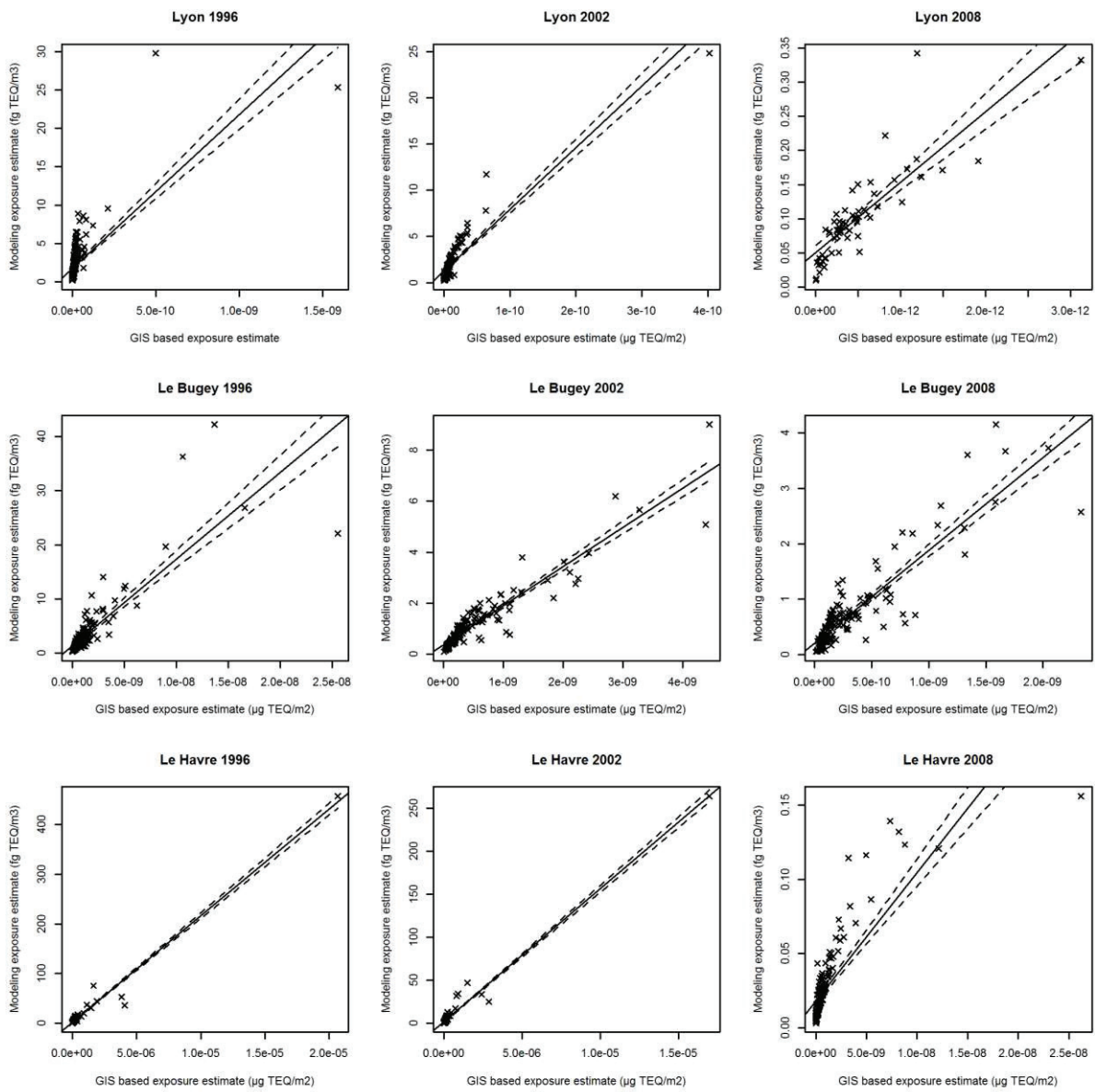


Figure 3 – Spatial representation of the subject classification into quintiles of exposure, using the GIS-based metric (A) and the SIRANE model (B) in Lyon, Le Bugey and Le Havre in 2008 – Calibration set

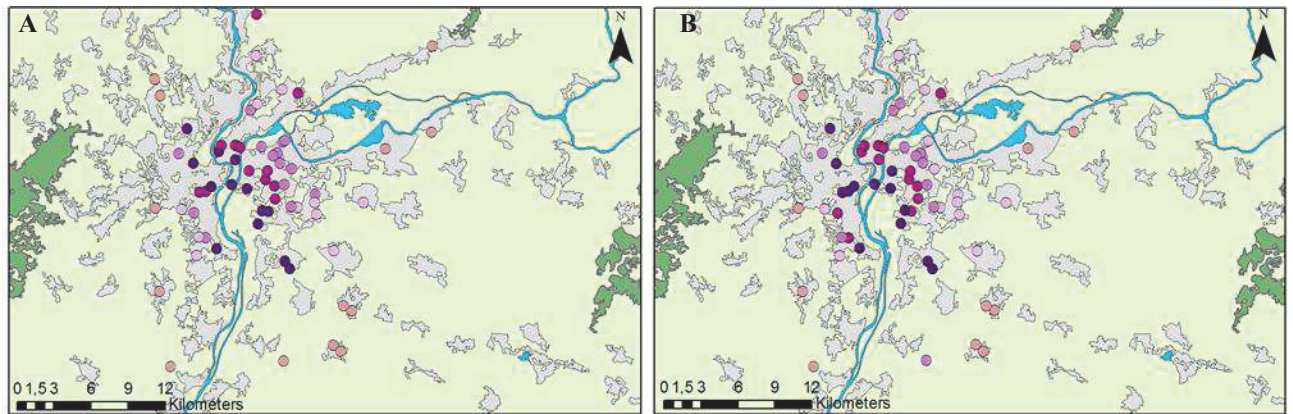


Figure 3.a – Lyon area, subject classification into quintiles of exposure, calibration set, 2008

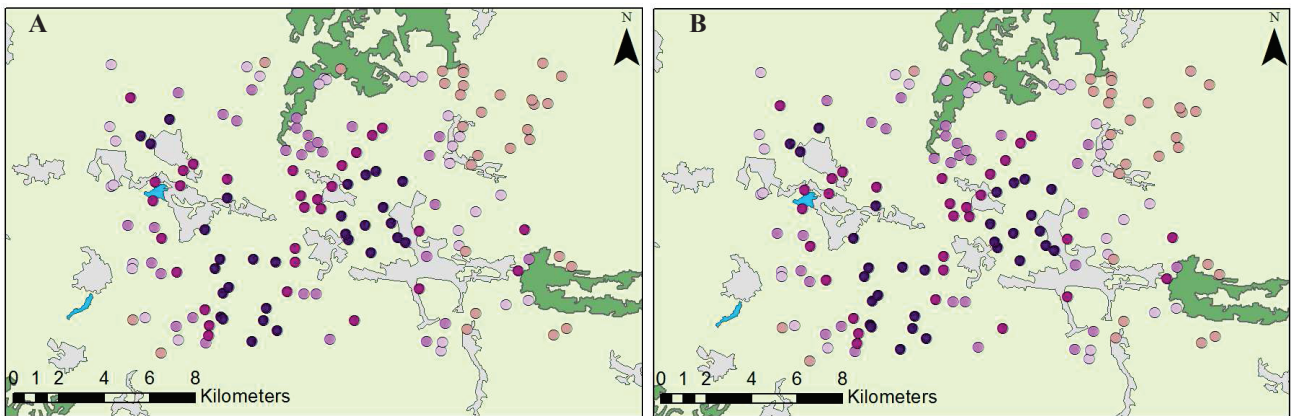


Figure 3.b – Le Bugey area, subject classification into quintiles of exposure, calibration set, 2008

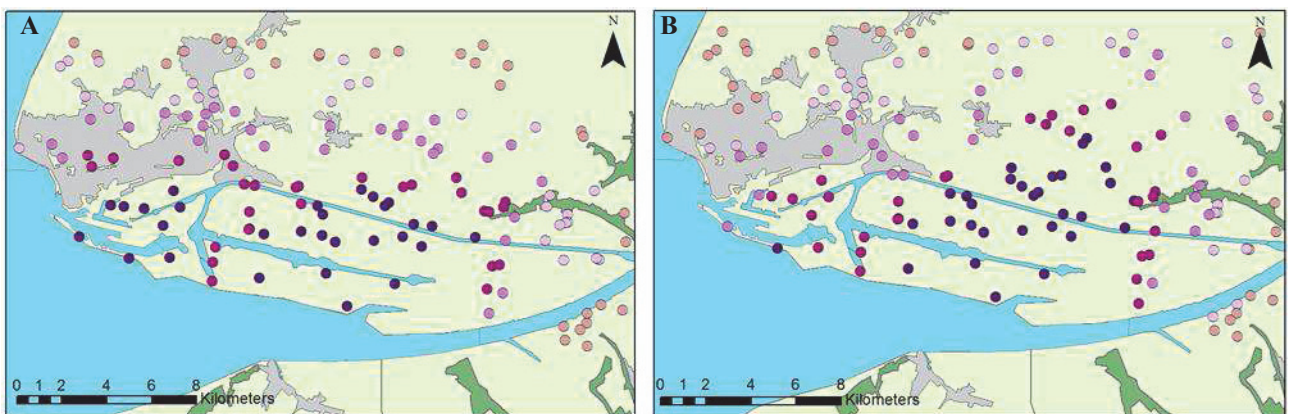
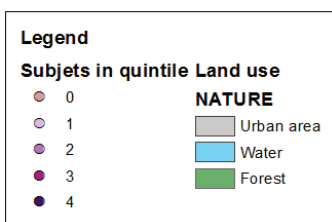


Figure 3.c – Le Havre area, subject classification into quintiles of exposure, calibration set, 2008



# Supplemental Material

## Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutants from industrial sources

*T. Coudon<sup>1,2</sup>, A. Danjou<sup>1,2</sup>, E. Faure<sup>1</sup>, D. Praud<sup>1</sup>, FR. Mancini<sup>4</sup>, P. Salizzoni<sup>3</sup>, B. Fervers<sup>1,2,5</sup>*

*<sup>1</sup> Département Cancer & Environnement, Centre Léon Bérard, Lyon, 69008, France*

*<sup>2</sup> Université Claude Bernard Lyon 1, Villeurbanne 69100, France*

*<sup>3</sup> Laboratory of Fluid Mechanics and Acoustics, Ecole Centrale de Lyon, Ecully 69134, France*

*<sup>4</sup> Centre de Recherche en Épidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS, UVSQ, Gustave Roussy, Villejuif, France.*

*<sup>5</sup> INSERM 1052, CNRS 5286, Centre de Recherche en Cancérologie de Lyon, Lyon, 69373, France.*

### Table of Contents

Table S1 - Weighted kappa coefficients and CI95 % with different CADD and distance decline patterns in the Le Bugey scenarios

Table S2 - Weighted kappa coefficients and CI95 % in Lyon, Le Bugey and Le Havre with different source technical parameter settings (smock exhaust velocity and stack height)

Table S3 - Weighted kappa coefficients and CI95 % in Lyon, Le Bugey and Le Havre with and without technical parameters (smock exhaust velocity and stack height)

Table S4- Weighted kappa coefficients and CI95 % in Lyon and Le Bugey with and without taking into account wind speed.

Figure S1 – Boxplot of the average dioxin concentrations (fg/m<sup>3</sup>), modeled at the E3N location in Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008 with the SIRANE model

Figure S2 - Boxplot of the average cadmium concentrations (ng/m<sup>3</sup>), modeled at the E3N location in Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008 with the SIRANE model

Table S1 - Weighted kappa coefficients and CI95 % with different CADD and distance decline patterns in le Bugey scenarios

Distance decline	Year	Without integration of wind direction	90°; 4 CADD	30°; 12 CADD	10°; 36 CADD	10° with adjacent segments; weighted at 25%	10° with adjacent segments; weighted at 50%
1/d	1996	0.64 (0.56-0.72)	0.58 (0.50-0.67)	0.63 (0.55-0.70)	0.65 (0.58-0.72)	0.66 (0.59-0.73)	0.63 (0.55-0.70)
	2002	0.78 (0.73-0.84)	0.58 (0.51-0.66)	0.59 (0.51-0.70)	0.58 (0.50-0.67)	0.58 (0.49-0.66)	0.58 (0.50-0.65)
	2008	0.68 (0.60-0.75)	0.51 (0.42-0.60)	0.53 (0.44-0.63)	0.53 (0.45-0.62)	0.53 (0.44-0.61)	0.52 (0.43-0.61)
1/d <sup>2</sup>	1996	0.59 (0.51-0.68)	0.67 (0.59-0.74)	0.73 (0.67-0.80)	0.75 (0.69-0.81)	0.76 (0.71-0.84)	0.79 (0.73-0.85)
	2002	0.74 (0.68-0.80)	0.78 (0.72-0.83)	0.84 (0.79-0.89)	0.83 (0.77-0.88)	0.83 (0.78-0.88)	0.82 (0.76-0.87)
	2008	0.64 (0.57-0.72)	0.67 (0.60-0.74)	0.72 (0.65-0.78)	0.73 (0.67-0.80)	0.74 (0.68-0.80)	0.73 (0.66-0.79)
e <sup>-d</sup>	1996	0.50 (0.40-0.60)	0.56 (0.47-0.65)	0.63 (0.54-0.71)	0.64 (0.56-0.72)	0.65 (0.57-0.73)	0.65 (0.57-0.73)
	2002	0.65 (0.58-0.72)	0.73 (0.67-0.80)	0.77 (0.71-0.83)	0.78 (0.72-0.84)	0.78 (0.72-0.84)	0.78 (0.72-0.84)
	2008	0.58 (0.50-0.66)	0.67 (0.60-0.74)	0.66 (0.59-0.73)	0.66 (0.59-0.73)	0.67 (0.60-0.74)	0.67 (0.60-0.74)
e <sup>-d</sup> /d	1996	0.45 (0.35-0.55)	0.49 (0.39-0.59)	0.51 (0.42-0.60)	0.54 (0.45-0.63)	0.53 (0.44-0.61)	0.53 (0.44-0.61)
	2002	0.63 (0.56-0.71)	0.66 (0.58-0.73)	0.73 (0.66-0.79)	0.74 (0.67-0.81)	0.73 (0.66-0.80)	0.73 (0.67-0.80)
	2008	0.57 (0.48-0.65)	0.60 (0.52-0.68)	0.59 (0.51-0.67)	0.62 (0.54-0.69)	0.62 (0.54-0.69)	0.62 (0.54-0.69)
1/d <sup>1.5</sup>	1996	0.61 (0.53-0.69)	0.66 (0.58-0.74)	0.73 (0.67-0.80)	0.72 (0.65-0.78)	0.73 (0.67-0.80)	0.73 (0.67-0.80)
	2002	0.78 (0.72-0.83)	0.69 (0.63-0.75)	0.78 (0.73-0.84)	0.74 (0.68-0.80)	0.76 (0.70-0.82)	0.74 (0.68-0.80)
	2008	0.68 (0.61-0.76)	0.62 (0.54-0.69)	0.64 (0.56-0.72)	0.66 (0.58-0.73)	0.66 (0.58-0.73)	0.65 (0.58-0.72)
1/d <sup>3</sup>	1996	0.51 (0.42-0.60)	0.58 (0.49-0.66)	0.69 (0.62-0.77)	0.70 (0.63-0.77)	0.72 (0.64-0.79)	0.72 (0.65-0.79)
	2002	0.68 (0.61-0.76)	0.78 (0.71-0.84)	0.83 (0.77-0.88)	0.83 (0.78-0.89)	0.85 (0.80-0.90)	0.85 (0.80-0.90)
	2008	0.62 (0.54-0.69)	0.68 (0.61-0.75)	0.68 (0.61-0.74)	0.67 (0.60-0.73)	0.69 (0.62-0.76)	0.69 (0.62-0.76)
e <sup>-d</sup> *d	1996	0.53 (0.44-0.63)	0.66 (0.60-0.75)	0.70 (0.63-0.77)	0.75 (0.68-0.82)	0.75 (0.68-0.82)	0.75 (0.68-0.82)
	2002	0.65 (0.58-0.72)	0.74 (0.68-0.80)	0.78 (0.73-0.84)	0.80 (0.75-0.85)	0.82 (0.77-0.87)	0.81 (0.76-0.86)
	2008	0.60 (0.52-0.68)	0.64 (0.57-0.71)	0.68 (0.62-0.75)	0.68 (0.62-0.75)	0.68 (0.62-0.75)	0.68 (0.62-0.75)

Table S2 - Weighted kappa coefficients and CI95 % in Lyon, Le Bugey and Le Havre with different source technical parameter settings (smock exhaust velocity and stack height)

Parameters	Years	Lyon	Le Bugey	Havre
$\frac{1}{h}$	1996	0.64 (0.59-0.69)	0.72 (0.66-0.79)	0.70 (0.63-0.77)
	2002	0.77 (0.70-0.83)	0.77 (0.71-0.82)	0.78 (0.72-0.84)
	2008	0.75 (0.64-0.86)	0.77 (0.71-0.83)	0.73 (0.66-0.79)
$\frac{1}{\sqrt{h}}$	1996	0.68 (0.64-0.73)	0.72 (0.66-0.79)	0.73 (0.66-0.79)
	2002	0.81 (0.77-0.86)	0.77 (0.72-0.83)	0.80 (0.74-0.86)
	2008	0.79 (0.70-0.88)	0.73 (0.66-0.79)	0.76 (0.70-0.82)
$\frac{1}{h^2}$	1996	0.57 (0.51-0.64)	0.71 (0.65-0.78)	0.55 (0.46-0.64)
	2002	0.70 (0.62-0.78)	0.77 (0.71-0.83)	0.64 (0.56-0.72)
	2008	0.56 (0.42-0.70)	0.69 (0.63-0.76)	0.61 (0.53-0.69)
$\frac{1}{h * v}$	1996	0.67 (0.62-0.71)	0.71 (0.65-0.78)	0.71 (0.64-0.78)
	2002	0.78 (0.72-0.84)	0.77 (0.71-0.83)	0.81 (0.75-0.86)
	2008	0.81 (0.72-0.90)	0.79 (0.73-0.85)	0.73 (0.66-0.79)
$\frac{1}{\sqrt{h} * v}$	1996	0.70 (0.66-0.75)	0.71 (0.65-0.78)	0.77 (0.71-0.83)
	2002	0.82 (0.77-0.87)	0.76 (0.70-0.82)	0.81 (0.75-0.86)
	2008	0.81 (0.73-0.89)	0.77 (0.71-0.83)	0.73 (0.67-0.80)
$\frac{1}{h^2 * v}$	1996	0.60 (0.54-0.66)	0.70 (0.64-0.77)	0.57 (0.48-0.65)
	2002	0.71 (0.63-0.78)	0.77 (0.71-0.83)	0.70 (0.63-0.77)
	2008	0.66 (0.53-0.78)	0.69 (0.62-0.75)	0.63 (0.56-0.71)

Table S3 - Weighted kappa coefficients and CI95 % in Lyon, Le Bugey and Le Havre with and without technical parameter (smock exhaust velocity and stack height)

Areas	Years	GIS metric ( $g - TEQ.m^{-2}$ )	GIS metric ( $g - TEQ.m^{-2}$ )
		$= \sum_j \sum_i t_i \times \frac{1}{d_{ij}^2} \times EI_i \times F_{ij}$	$= \sum_j \sum_i t_i \times \frac{1}{d_{ij}^2} \times EI_i \times F_{ij} \times \frac{1}{\sqrt{h_i * v_i}}$
Lyon	1996	0.71 (0.67-0.76)	0,70 (0,66-0,75)
	2002	0.84 (0.79-0.88)	0,82 (0,77-0,87)
	2008	0.81 (0.72-0.89)	0,81 (0,73-0,89)
Le Bugey	1996	0.79 (0.73-0.85)	0,71 (0,65-0,78)
	2002	0.82 (0.76-0.87)	0,76 (0,70-0,82)
	2008	0.73 (0.66-0.79)	0,77 (0,71-0,83)
Le Havre	1996	0,74 (0,68-0,80)	0,77 (0,71-0,83)
	2002	0,77 (0,72-0,83)	0,81 (0,75-0,86)
	2008	0,64 (0,57-0,72)	0,73 (0,67-0,80)

We observed a clear improvement of the results for the scenario Le Havre 2008. In this specific scenario, the major source had a stack height of 240m that impacted dioxin dispersion and decrease dioxin exposure. As explained in 3.1, given the absence of impact of the stack height on the kappa coefficients in all other scenarios and the low completeness of data for stack height at the national level (36%), it was decided to integrate into the GIS metric only stacks above 90m, corresponding to 3 times the median stack height of the 2,626 sources at the national level. This led us to the final GIS metric:

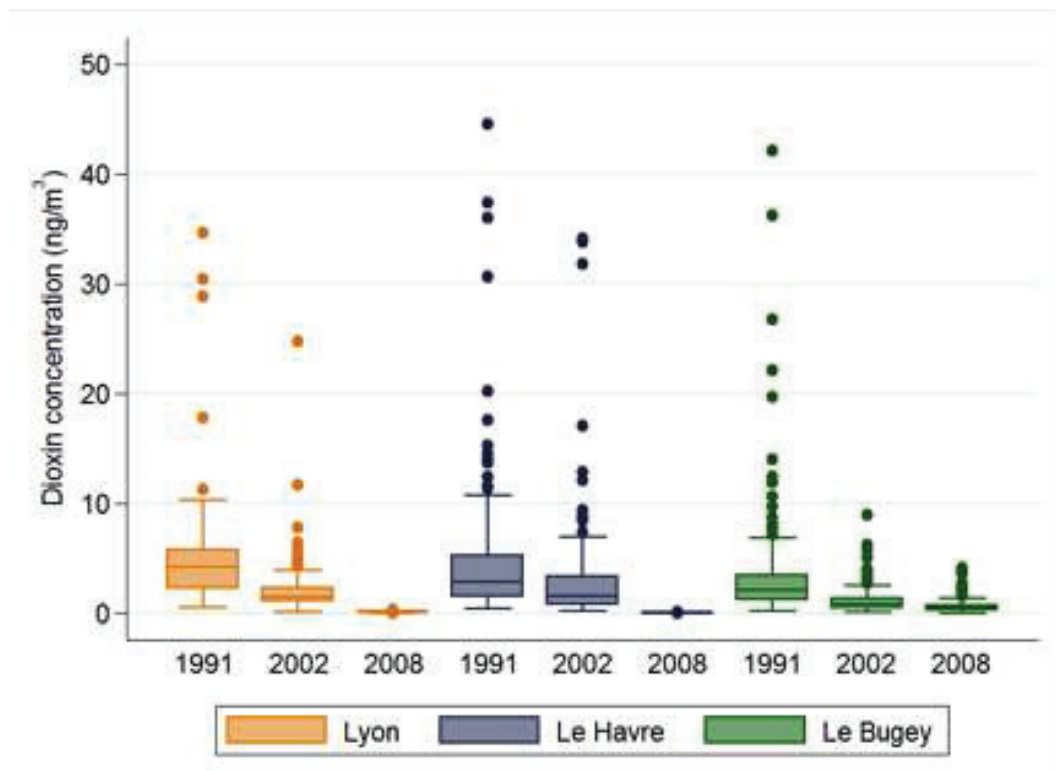
$$\text{GIS - based metric } (g - TEQ.m^{-2}) = \sum_j \sum_i t_i \times \frac{1}{d_{ij}^2} \times EI_i \times F_{ij} \times \left( \frac{h_{\text{median}}}{h_{ij}} \right)^* \quad (i)$$

Table S4- Weighted kappa coefficients and CI95 % in Lyon and Le Bugey with and without taking into account wind speed.

Area	Year	Without wind speed	With wind speed
Le Bugey	1996	0.79 (0.73-0.85)	0.76 (0.70-0.82)
	2002	0.82 (0.76-0.87)	0.58 (0.49-0.66)
	2008	0.73 (0.66-0.79)	0.59 (0.51-0.67)
Lyon	1996	0.71 (0.67-0.76)	0.68 (0.63-0.73)
	2002	0.84 (0.79-0.88)	0.59 (0.52-0.66)
	2008	0.81 (0.72-0.89)	0.71 (0.61-0.81)

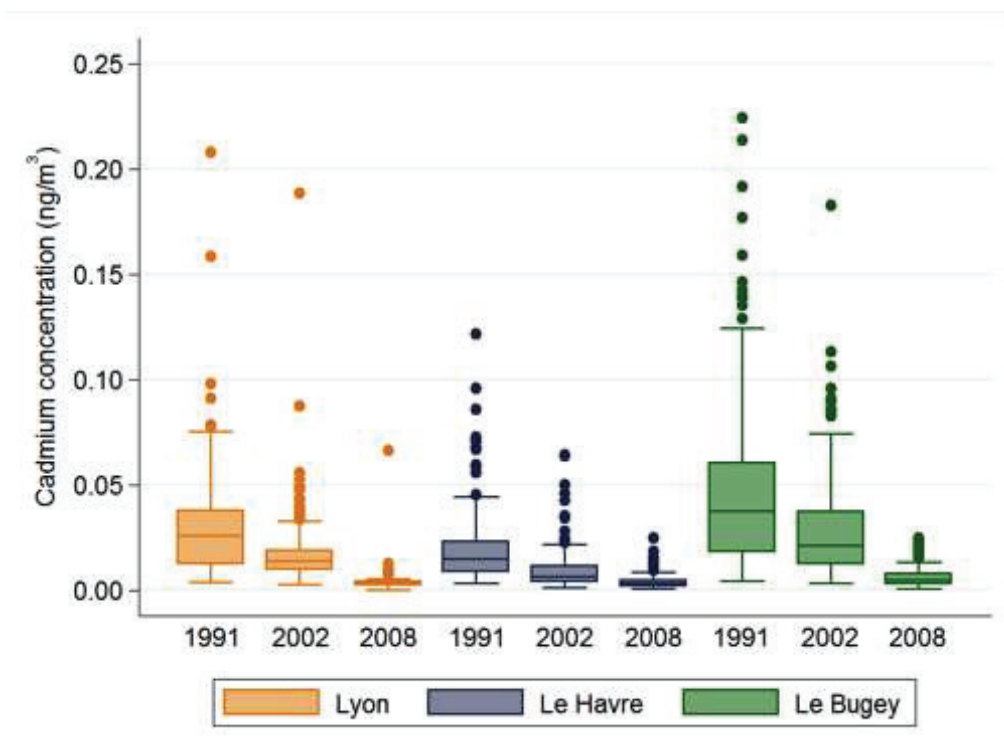


Figure S1 – Boxplot of the average dioxin concentration (fg/m<sup>3</sup>), modeled at the E3N location in Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008 with the SIRANE model



The highest median dioxin concentrations (fg-TEQ/m<sup>3</sup>) were observed in 1996 in the three areas (median (min-max): Lyon 4.8 (0.6-34.7); Le Havre 2.87 (0.46- 5.7); Le Bugey 2.18 (0.25-42.2)). From 1996 to 2008, annual average dioxin concentrations decreased by 98%, 99% and 80% for Lyon, Le Bugey and Le Havre, respectively. Over the same period, we observed a relevant change in the spatial concentration patterns with the virtual extinction of the local concentration peaks due to industrial sources. As a result, the concentration field appears more homogeneous in 2008 compared to the previous years. In 2008, dioxin concentrations were lower than in 1996 and 2002 with median concentrations less than 1 fg-TEQ/m<sup>3</sup> (median (min-max): Lyon 0.08 (0.005-0.34); Le Havre 0.02 (0.003- 0.15); Le Bugey 0.51 (0.51-4.19)).

Figure S2 - Boxplot of the average cadmium concentration ( $\text{ng}/\text{m}^3$ ), modeled at the E3N location in Lyon, Le Havre and Le Bugey for 1996, 2002 and 2008 with the SIRANE model



The highest median concentrations were observed in 1996 (median (min-max) in  $\text{ng}/\text{m}^3$ : Lyon 0.032 (0.001-0.378); Havre 0.015 (0.003- 0.649); Bugey 0.039 (0.004-0.744)). We also observed a decrease in cadmium concentration over the 1996-2008 periods (by 89%, 86% and 79% for Lyon, Le Bugey and Le Havre, respectively). In 2008, cadmium concentrations were less than 0.05  $\text{fg-TEQ}/\text{m}^3$  (median (min-max): Lyon 0.003 (0.0003-0.008); Havre 0.003 (0.001- 0.025); Bugey 0.005 (0.001-0.082)). This change in cadmium concentration is due to the decrease in emissions. Cadmium concentration field was more homogeneous and subjects were exposed to lower cadmium concentrations in 1996 than in 2002 and 2008.

### 3. Conclusion

L'indicateur SIG développé au cours de ce travail a permis de caractériser de manière fiable les expositions environnementales aux dioxines et au cadmium sur une longue durée. Ces résultats soulignent l'intérêt de se baser sur des outils de modélisation de la dispersion pour calibrer et évaluer ce type de métrique SIG.

Les résultats obtenus pour le cadmium démontrent la possibilité d'utiliser cette métrique pour estimer les expositions à d'autres polluants provenant de sources industrielles, avec comme condition que ceux-ci aient des propriétés proches de celles des dioxines (polluants sous forme particulaire avec une taille comprise entre [10-1]  $\mu\text{m}$  et inerte chimiquement dans l'atmosphère).

Parmi les forces de cet indicateur, nous pouvons citer l'approche multi-sources, prenant en compte de multiples secteurs d'émissions (incinération des déchets, production de métaux et industries du ciment, etc.) ainsi que la connaissance de l'évolution des caractéristiques techniques des installations dans le temps. Pronk et al (2013) a également utilisé un inventaire historique des émissions de dioxines (1987-2000) et une approche multi-sources (mais limitée à dix secteurs d'activité). Toutes ces sources ont été géolocalisées, avec une précision moyenne élevée (46% à la cheminée, 70% au minimum à la parcelle). Le niveau de précision de la localisation des sources peut avoir des implications importantes sur le classement des individus en différentes catégories d'exposition, dû à la forte variation spatiale des concentrations autour des sources. De manière similaire, alors que les adresses résidentielles des sujets E3N n'ont pas été renseignées lors des questionnaires dans l'objectif d'être géolocalisées, leur précision peut être considérée comme assez bonne pour limiter les biais de mauvaise classification, en particulier pour les sujets urbains (Cordioli et al. 2013; Faure et al. 2017; Han et al. 2013).

Il est à noter que les valeurs des paramètres retenus dans l'indicateur sont en accord avec d'autres études utilisant des indicateurs spatialisés : l'exposition pondérée par le carré de la distance sujet-source (Hoek et al. 2001a; Pronk et al. 2013; Vienneau et al. 2009), ou encore la taille de la zone d'impact de 10 km (Vienneau et al. 2009). Certaines études utilisent des buffers de taille plus réduite, mais elles portent sur l'exposition aux rejets liés au trafic routier ayant une décroissance rapide avec la distance à la source.

Cet indicateur comporte néanmoins plusieurs limites. La précision de nos estimations des émissions est directement liée à la qualité des informations recueillies auprès des industriels (article n°1). Mais alors que les autres études ont souvent utilisé des inventaires d'émissions construits pour d'autres objectifs et sans avoir connaissance des incertitudes liées aux émissions (Pronk et al. 2013; Vienneau et al. 2009), nous avons pu estimer les incertitudes des émissions estimées.

Si les paramètres intégrés dans l'indicateur sont en accord avec plusieurs études dans la littérature (Cordioli et al. 2013c; Gulliver and Briggs 2011; White et al. 2009; Zou et al. 2009), certains paramètres n'ont pas été conservés durant la phase de calibration. Ainsi, la prise en compte de la vitesse de sortie des fumées améliorerait très légèrement les performances de l'indicateur (amélioration du kappa pondéré moyen de 0.02) mais ce paramètre était disponible pour moins de 15% dans sources et n'as pas été retenu dans l'équation finale. L'inclusion de la vitesse du vent dans la combinaison de paramètres n'a pas amélioré les résultats alors qu'il est admis que ce paramètre influence la dispersion des polluants dans l'atmosphère. D'autres études intègrent la direction du vent (Yu et al. 2006) ou la direction du vent et la vitesse du vent dans leur indicateur (Gulliver and Briggs 2011; Vienneau et al. 2009; White et al. 2009) mais aucune étude n'a évalué l'impact de la vitesse du vent sur la performance de ces indicateurs. De même, nous n'avons pas identifié d'études ayant évalué l'impact de la hauteur de cheminée ou d'autres paramètres techniques sur la performance des indicateurs. Enfin, certains paramètres pris en compte dans les modèles de dispersion atmosphérique n'ont pas été testés durant la phase de calibration comme la pluviométrie ou la température extérieure.

Durant cette étude, un indicateur SIG a été développé et évalué pour estimer l'exposition atmosphérique aux dioxines dans l'air. Cette métrique combine la distance sujet-source, la direction du vent, l'intensité d'émissions des sources et la hauteur de cheminée. La classification des sujets estimés à partir de cet indicateur montre une concordance forte avec la classification issue d'un modèle de dispersion. Plus important, la qualité de la concordance reste quasi-constante quel que soit le nombre de sources, le nombre de sujets, l'intensité des sources, la position des sources et des sujets ou les conditions météorologiques et topographiques. Des résultats similaires sont observés pour le cadmium. Ces résultats laissent à penser que l'indicateur pourrait également permettre d'estimer l'exposition à d'autres polluants ayant des propriétés physico-chimiques similaires lorsque les données de surveillance seront indisponibles. Les résultats obtenus ici suggèrent également que des mesures simples basées principalement sur la distance sujet-source et utilisées dans plusieurs autres études (Cordioli et al. 2013; Hoek et al. 2001; Pronk et al. 2013), montrent des degrés de concordance moindre avec les résultats de la modélisation, et peuvent en conséquence générer un biais de classement. Ces résultats suggèrent que les outils SIG offrent une alternative robuste aux modèles LUR dans les situations où les données de surveillance sont manquantes et où l'utilisation des modèles de dispersion atmosphérique est limitée en raison de la taille du domaine d'étude.

## **Application de l'indicateur dans l'étude cas-témoins nichée GEO3N**

L'indicateur SIG développé dans cette thèse a été utilisée dans l'étude GEO3N pour étudier l'hypothèse d'une association entre l'exposition aux dioxines et le risque de cancer du sein.

L'application de l'indicateur nécessite, en plus de la localisation des sources et de la quantification de leurs émissions, une localisation précise des femmes de la cohorte (Faure 2017) et un recueil de données météorologiques.

Les antécédents résidentiels des participants ont été extraits des questionnaires E3N. Malgré le fait que ces questionnaires n'avaient pas été créés pour extraire des informations permettant le géocodage des sujets, les résultats obtenus sont bon avec une majorité des adresses géocodées à l'adresse exacte ou à la rue (Faure et al. 2017). Ce qui permet de limiter les erreurs de classification.

Pour récupérer les données météorologiques, nous nous sommes basé sur les zones symposiums (zone de météo homogène) créée par Météo France, et nous avons cherché à identifier une station permettant de fournir 5 paramètres (vitesse et direction du vent, pluviométrie, température et nébulosité) sur toute la période d'étude, avec un pas de temps horaire, pour chaque zone incluant une source recensée. Nous avons finalement récupéré les données de 223 zones symposiums qui couvrent l'ensemble des sources recensées (Figure 11). Plusieurs sources isolées ne possédaient pas de stations remplissant les critères dans leurs zones symposium. En collaboration avec Météo France, nous leur avons attribué les données issues de stations météorologiques dans des zones symposium similaire et proche. La base finale contient les 5 paramètres horaires, disponibles pour chaque source et sera exploitable pour d'autres travaux.

L'inventaire des sources de dioxines et de cadmium a été mené en premier lieu dans la région Rhône-Alpes qui avait servi de zone test lors de l'étude pilote du projet GEO3N (financement Oncostarter 2014). De plus, l'inventaire dans cette région a bénéficié de l'accès à certaines bases de données spécifiques (notamment l'étude menée par Cordier et al (2011) et des documents provenant de l'association robin des bois) permettant de terminer l'inventaire plus rapidement. Dans le cadre de la thèse d'Aurélie Danjou, une première version de l'indicateur a été appliquée en Rhône-Alpes pour estimer l'exposition atmosphérique aux dioxines de 429 cas de cancer du sein et 716 témoins (annexe 3 : Danjou et al (2018), *en cours de finalisation*). Ce première indicateur ne prenait pas en compte les incinérateurs hospitaliers (car non recensées à l'époque) et n'incluait pas le paramètre hauteur de cheminée (car considéré comme non-impactant au vu des premières résultats sur les zones de Lyon et du Bugey). Les résultats n'ont pas montré d'augmentations du risque du cancer du sein lié à une exposition atmosphérique aux dioxines (annexe 3).

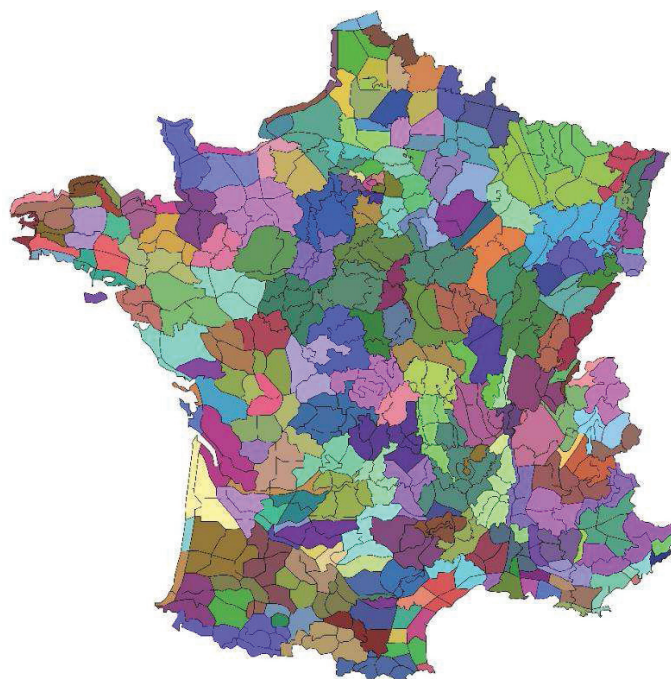


Figure 11 - Carte des zones symposiums définis pour le projet GEO3N (n=223)

L'étude de l'association entre l'exposition atmosphérique et le risque de cancer du sein dans l'ensemble de la population de l'étude GEO3N (5 500 cas de cancer du sein et 5 500 témoins appariés) a été réalisée en 2017 en utilisant l'indicateur SIG construit pendant cette thèse. Il a permis d'attribuer pour chaque femme de la cohorte une valeur représentant l'exposition cumulative sur 18 ans (Figure 12, 13 et 14). L'analyse statistique des résultats est en cours et fait l'objet d'un article (Praud, D, *en préparation*). L'analyse des résultats issus de l'application de l'indicateur ouvre une série de problématiques développées dans la partie « perspectives de recherches » de cette thèse.

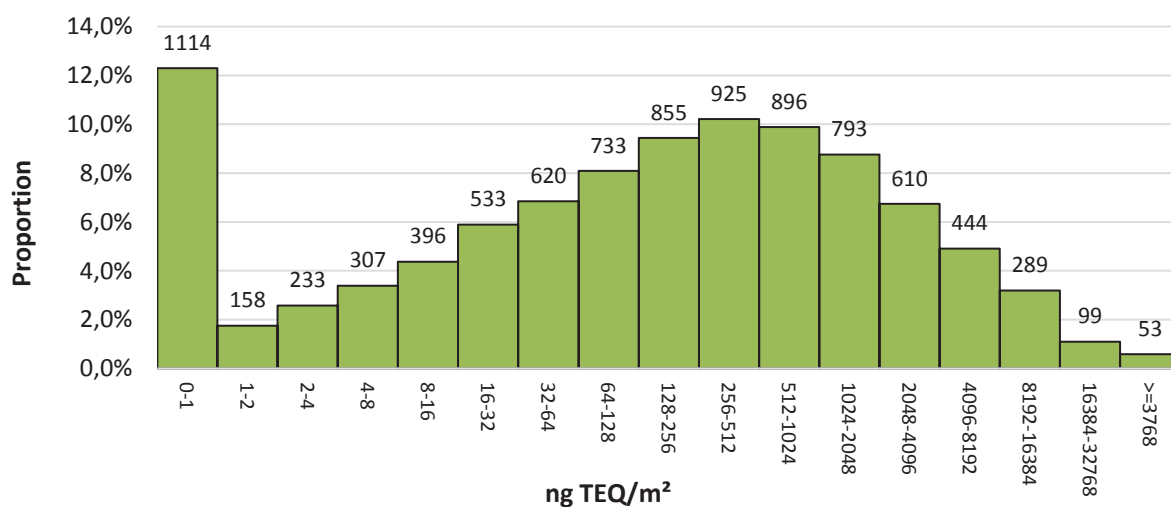


Figure 12 - Application de l'indicateur SIG dans l'étude GEO3N. Distribution des sujets (pourcentage sur l'axe verticale et effectifs sur les barres) selon l'exposition environnementale aux dioxines estimé par l'indicateur spatialisé (exposition cumulée) dans l'étude GEO3N (N=9058)

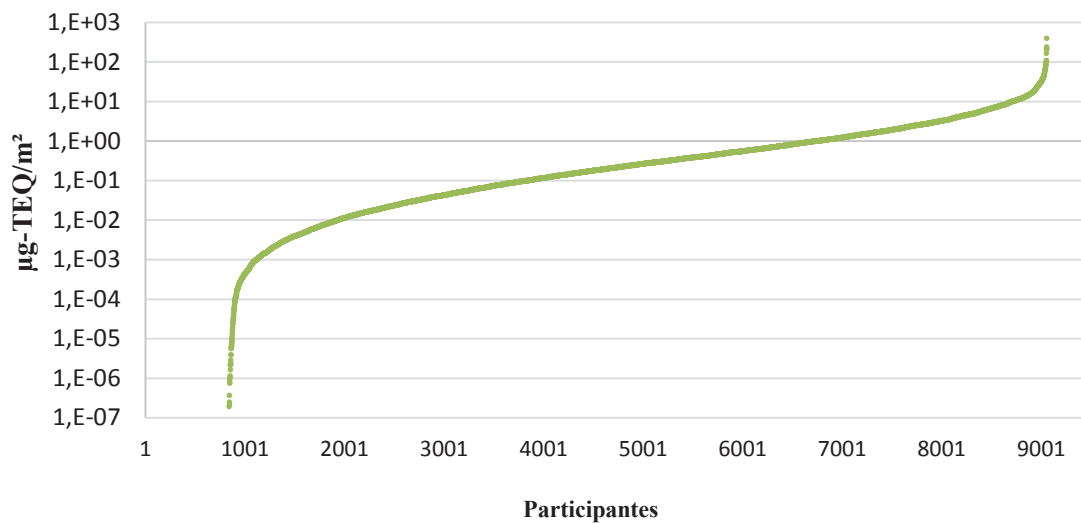


Figure 13 - Application de l'indicateur SIG dans l'étude GEO3N. Répartition des sujets selon leur l'exposition cumulative (1990-2008) aux dioxines (en  $\mu\text{g-TEQ}/\text{m}^2$ ) de la moins exposées à la plus exposées. Environ 900 sujets sont considéré comme non-exposées car en dehors de toutes zones d'exposition.

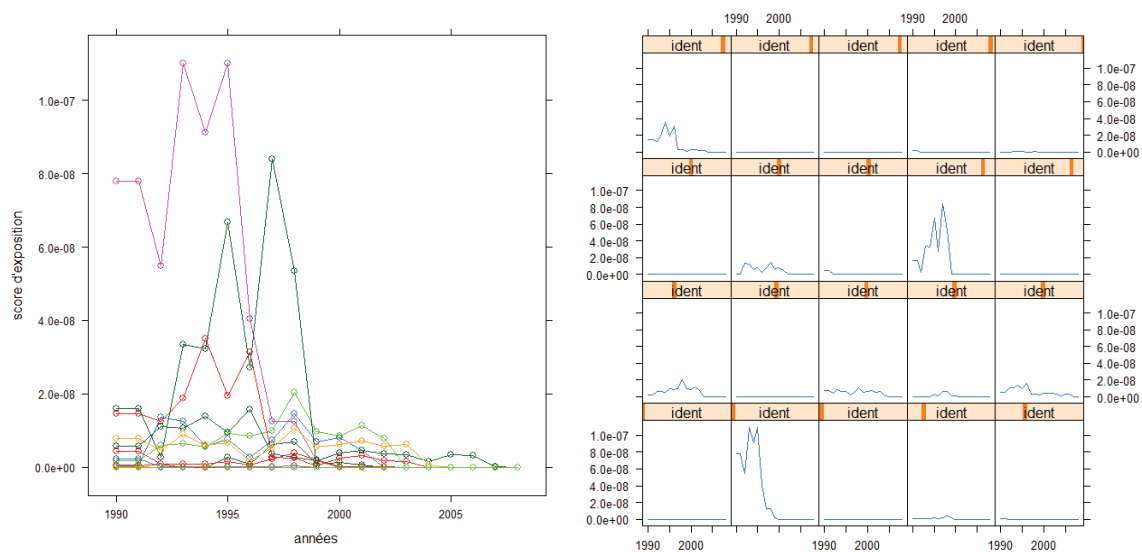


Figure 14 - Exemple de variations des expositions annuelles de 20 sujets, choisit aléatoirement, entre 1990 et 2008, calculée avec l'indicateur SIG (en  $\mu\text{g}/\text{m}^3$ ).





## **CHAPITRE V. DISCUSSION ET PERSPECTIVES**

## 1. Discussion

Mon travail de thèse visait à étudier les expositions atmosphériques aux dioxines et au cadmium, dans le cadre d'études épidémiologiques conduites pour estimer le risque de cancer du sein associés à ces expositions, en fournissant de nouvelles connaissances sur les expositions historiques, chroniques et multi-sources de la population générale et des sujets issus de la cohorte E3N et en essayant de combler certaines lacunes méthodologiques de la littérature concernant l'évaluation des expositions atmosphériques. Le premier objectif de cette thèse était d'inventorier les principales sources de dioxines et cadmium en France, entre 1990 et 2008, et d'évaluer les émissions associées (article n°1). Le deuxième objectif a été d'estimer les expositions chroniques aux dioxines et au cadmium sur le territoire de la Métropole de Lyon en utilisant un modèle de dispersion atmosphérique et d'étudier la contribution à l'exposition des différentes sources existantes (Coudon et al. 2018). Le troisième et principal objectif était de développer un indicateur SIG de l'exposition atmosphérique aux dioxines et au cadmium pour estimer les expositions à l'échelle de la France, sur 20 ans, et capable de fournir des estimations robustes, validées par confrontation avec un modèle de dispersion gaussien (Coudon et al, soumis).

Nous avons créé la première base de données contenant les émissions annuelles de dioxines et de cadmium entre 1990 et 2008 à l'échelle de l'installation pour toute la France. Les émissions estimées ont été confrontées à des mesures pour certaines sources ainsi qu'aux données d'inventaires disponibles à l'échelle départementale et nationale. En utilisant les données techniques recueillies, cette base de données pourra être étendue à d'autres polluants.

Grâce au modèle SIRANE, nous avons réalisé la première étude visant à évaluer la variabilité spatio-temporelle des concentrations de dioxines et de cadmium sur près de 20 ans, en comparant les concentrations modélisées et mesurées en air ambiant, et en tenant compte d'une grande variété de sources (industrie, trafic routier, chauffage urbain ou sources diffuses, pollution de fond). De plus, il s'agit de la première étude comparant les concentrations estimées par un modèle, aux concentrations de dioxines dans l'air ambiant.

Par rapport aux études précédentes, nous avons évalué l'exposition atmosphérique aux dioxines et au cadmium via une approche multi-sources et historique. Afin de surmonter les limites posées par l'étude de l'exposition sur un territoire vaste avec un faible nombre de mesures disponibles, nous avons développé un indicateur implémenté dans un SIG. Cet indicateur prend en compte des paramètres identifiés à partir d'une revue de la littérature et des travaux sur la modélisation de la dispersion atmosphérique. Pour la première fois, ce type d'indicateur SIG de l'exposition atmosphérique aux dioxines et au cadmium a été calibré et évalué.

Bien que cette thèse ait proposé des améliorations de la méthodologie relative à l'étude de l'exposition atmosphérique aux dioxines et au cadmium, nous avons rencontré plusieurs difficultés induisant un certain nombre de limites.

#### *Estimations des émissions de dioxines et de cadmium*

L'inventaire et la caractérisation des sources émettrices de dioxines a permis de recenser 2620 sources de dioxines et 2700 sources de cadmium (pour une majorité d'entre elles localisée à la cheminée ou à l'adresse postale) et d'estimer leurs émissions. Mais cette approche exhaustive s'est révélée fastidieuse, particulièrement la collecte des données techniques. De plus, nous avons eu peu de retour des questionnaires envoyé aux industriels (taux de retour de 32%). Ceci implique de fortes incertitudes pour les émissions de certains secteurs, notamment les incinérateurs de déchets hospitaliers (qui représentent près de la moitié de sources inventoriées et pour lesquels nous n'avons pas pu obtenir d'information auprès des exploitants) et les petits incinérateurs de déchets municipaux existant au début des années 90 (qui sont parmi les sources les plus émettrices).

Les émissions ont été estimées avec des jeux de facteurs d'émission reconnus au niveau européen (UNEP) ou français (CITEPA). Le niveau de détail des informations collectées a permis de connaître les technologies mises en place à l'échelle individuelle et d'identifier les « hotspots », en particulier dans le cas des dioxines où la présence de filtres spécifiques peut fortement diminuer les émissions (UNEP 2013). Contrairement à un certain nombre d'études sur les expositions historiques, les incertitudes ont été estimées (en se basant sur la qualité des informations recueillies). Néanmoins, l'incertitude propre à chaque facteur d'émission n'était pas précisée dans les jeux de données UNEP et CITEPA et n'a pas pu être prise en compte.

L'utilisation d'une méthodologie se basant sur des facteurs d'émission a permis de s'affranchir de l'utilisation de données de mesures et de fournir une qualité d'évaluation des émissions homogène sur la période d'étude. Cependant, la comparaison de nos estimations avec un faible échantillon de mesures de l'émission à la source, comprenant uniquement les émissions déclarées depuis 2003 au-dessus des limites officielles (soit environ 2% des sources), a montré une forte dispersion des résultats. Ces écarts peuvent s'expliquer par plusieurs facteurs : incertitudes sur les facteurs d'émission, incertitude sur les données recueillies, mesures effectuées une fois par an dans le meilleur des cas et uniquement auprès des sources importantes. Ces mesures se concentrent sur la période où les émissions sont les plus faibles (2003-2008), ainsi les différences observées ne peuvent pas être extrapolées pour toute la période d'étude. Pour toutes ces raisons, si les mesures d'émissions sont les valeurs les plus proches des émissions réelles, l'interprétation des divergences observées lors de la comparaison avec les estimations doit être faite avec prudence. Les études précédentes comparant les estimations de différents inventaires ont également observées des discordances pouvant aller jusqu'à plusieurs ordres de grandeur (Breivik et al. 2006).

Certains types d'émissions liées aux activités domestiques, au trafic et aux activités illégales n'ont pas pu être pris en compte dans l'inventaire, ce qui pourrait entraîner une sous-estimation des émissions et donc des expositions. Le projet GEO3N avait pour objectif d'étudier l'exposition aux émissions atmosphériques de dioxines et a choisi de se concentrer sur les principales sources sur la période d'étude (1990-2008): les sources industrielles ou affiliées (crématorium, incinérateur hospitalier). Néanmoins, au niveau national, le CITEPA estime qu'en 2008, 38% des émissions atmosphériques de dioxines étaient dues au secteur résidentiel et 18% au transport routier. Parmi les émissions du secteur résidentiel, la majorité (75%) est estimée provenir du brûlage de câbles. Le brûlage de câbles produit de fortes émissions de dioxines dues à la présence autour du cuivre de gaines protectrices contenant des produits chlorés. Cependant, il n'existe pas de registre recensant cette activité illégale. Les facteurs d'émission issus de la littérature vont de quelques  $\mu\text{g-TEQ/kg}$  (LUA 2000) à plusieurs milliers de  $\mu\text{g-TEQ/kg}$  (UNEP 2013). Réaliser l'inventaire des activités de brûlage de câbles serait un travail long, très complexe, et associé à une grande incertitude, c'est pourquoi nous n'avons pas estimé ces émissions. La même problématique s'applique pour les différents types d'évènements ponctuels comme les feux de décharge ou les incendies. Le cadmium est moins touché par cette problématique : d'après le CITEPA, en 2008, 78% des émissions provenaient des secteurs de l'incinération et de la production de métaux.

#### *Modélisation de la dispersion des polluants*

Pour approfondir les différentes contributions à la pollution atmosphérique, j'ai étudié l'exposition aux dioxines et au cadmium des sujets E3N dans une zone urbaine (Métropole de Lyon) et estimé la variation de la contribution de différents types de sources à la concentration au cours de la période d'étude. Ce travail ne permet pas de prendre en compte précisément les sources ponctuelles et illégales, mais en combinant les données d'émissions issues de l'inventaire des sources de dioxines et de cadmium avec les données issues du cadastre d'ATMO AURA (Atmo Auvergne-Rhône-Alpes 2016), nous avons modélisé la dispersion des émissions des sources industrielles, des sources diffuses, des sources liées au chauffage urbain ainsi que celles liées au trafic routier.

Les simulations des concentrations de polluants reconstitués au cours des 20 dernières années montrent une diminution des concentrations de dioxines et de cadmium dans l'air, qui s'explique principalement par la réduction des rejets industriels, et se traduit par une tendance générale à l'homogénéisation de l'exposition des populations. En 2008, les principales contributions étaient les sources diffuses dans le cas des dioxines et les émissions liées au trafic pour le cadmium. Il est important de noter que la concentration de fond représente une part majoritaire des expositions sur l'ensemble de la période d'étude. Ce travail illustre que, depuis le début des années 2000, en milieu urbain, les sources industrielles ne sont plus les contributeurs majeurs de l'exposition aux dioxines et au cadmium. Ces résultats sont extrapolables à d'autres villes similaires à Lyon dans la mesure où ces villes ne sont pas fortement industrialisées.

Dans le cas des dioxines, la limite majeure de ces travaux de modélisation de la dispersion atmosphérique, est l'évaluation de la contribution du brûlage de câbles qui fait partie des émissions « résidentielles » selon les classifications d'ATMO AURA et du CITEPA. Outre l'incertitude sur les émissions, estimées à partir de la quantité de câbles volés en France, elles sont positionnées par le cadastre en fonction de la densité de la population. Il en résulte des émissions diffuse et homogène centrés sur le centre-ville lyonnais, dont la contribution prend de l'importance au fil des années. Le positionnement de ces émissions en centre-ville n'est pas réaliste car elles sont plus probablement localisées dans des zones faiblement peuplé (campagne, friches industrielles, ...). Ceci peut induire un biais important dans l'estimation des expositions pour les années récentes.

Cette limite se retrouve lors de la comparaison entre les mesures réalisées par ATMO AURA et les concentrations modélisées avec le modèle SIRANE. La comparaison avec les mesures hebdomadaires réalisées dans le centre-ville lyonnais, en 2007 et 2008, a montré que les concentrations moyennes modélisées de dioxines étaient deux fois inférieures aux concentrations mesurées. Le modèle reproduit de manière correcte les concentrations en été et au printemps, mais sous-estime fortement les concentrations en hiver et lors de certains pics ponctuels pouvant atteindre 10 fois la valeur modélisée. Ces pics sont probablement dus à des émissions de dioxines non contrôlées et imprévisibles, qui ne peuvent pas être prises en compte dans nos scénarios d'émission. Cette hypothèse est renforcée par l'étude de la variation du ratio des concentrations de dioxines et de cadmium qui varie fortement lorsque ces pics sont observés (ce qui sous-entend de fortes émissions spécifiques de dioxines).

Pour le cadmium, la comparaison montre une meilleure concordance, avec des concentrations annuelles mesurées proches des concentrations annuelles modélisées. Le modèle montre également une meilleure capacité à prédire des pics de concentrations de cadmium ainsi que la variation saisonnière des concentrations.

Malgré ces différences dans les séries temporelles des concentrations, les résultats remplissent les critères de validation applicable pour un modèle gaussien en milieu urbain, comme définit par Hanna and Chang (2012). Il est à noter que, dans la littérature, les études ont montré des difficultés à prédire des concentrations mesurées de métaux lourds ou de HAP, affichant parfois même des désaccords d'un facteur 10 (e.g. EPA 2002; Garcia et al. 2014). Les données de l'EPA montrent, dans le cas du cadmium, une sous-estimation des concentrations annuelles d'un facteur 5.

En tirant parti de la quantification des incertitudes réalisée lors de l'inventaire, nous avons évalué la propagation des incertitudes sur les émissions des sources industrielles dans la chaîne de modélisation. Due à la faible contribution des sources industrielles en 2008, ces incertitudes représentaient des variations des concentrations moyennes de seulement  $\pm 1\%$  (pour les dioxines) et  $\pm 4\%$  (pour le cadmium) à la station de mesure continue. En 1990, ces incertitudes ont généré des variations des concentrations estimées à l'adresse des sujets E3N plus élevées :  $\pm 10,6\%$  pour les dioxines et  $\pm 5,9\%$  pour le cadmium. En revanche, il n'était pas possible d'estimer les incertitudes liées aux émissions du trafic routier, du chauffage et des sources diffuses dû au manque d'information sur les données d'entrée du cadastre d'ATMO AURA.

De manière générale, nous pouvons observer que les points forts de l'inventaire des sources (exhaustivité, quantification des incertitudes, grande quantité de données techniques sur les sources) se sont répercutés sur le travail de modélisation. De manière similaire, les points faibles et limites de l'inventaire (manque de données sur les autres types de sources, non prise en compte des émissions ponctuelles, manque de mesure) impactent également le travail plus en aval. Cependant, malgré le manque de données, nous avons pu comparer à la fois les émissions et les concentrations estimées avec des mesures. Il est à noter que si nos estimations des émissions des sources industrielles sont supérieures aux autres inventaires, elle n'entraîne pas de surestimations des concentrations estimées avec le modèle SIRANE (au contraire on observe une sous-estimation des concentrations moyennes). L'inventaire des sources et le travail de modélisation ont servi de base à la création d'un indicateur implémenté dans un SIG pour évaluer l'exposition aux dioxines sur la France métropolitaine entre 1990 et 2008.

#### *Création d'un indicateur de l'exposition atmosphérique*

Par le passé, les études épidémiologiques utilisaient des indicateurs se basant principalement sur la distance sujet-source, pour estimer les expositions. Cependant, d'après Cordioli et al. (2013), ces indicateurs simples augmentent fortement les risques de biais de classement des expositions (en comparaison avec un modèle gaussien). A partir d'une revue de la littérature et des travaux de modélisations nous avons identifié des paramètres météorologiques et technologiques pouvant influencer l'exposition atmosphérique aux dioxines (Gulliver and Briggs 2011; Hoek et al. 2001a; Pronk et al. 2013; Vienneau et al. 2009; White et al. 2009; Yu et al. 2006). Sur un ensemble de 9 scénarios, nous avons identifié une combinaison de paramètres fournissant des classifications des expositions en accord "substantiel" (Viera and Garrett 2005) avec les classifications fournies par le modèle SIRANE. L'application de cet indicateur sur un nouveau jeu de données, incluant des émissions de cadmium a montré des performances élevées et constantes à travers les différents scénarios.

Ces résultats suggèrent que les outils SIG offrent une alternative robuste aux modèles LUR dans les situations où les données de surveillance sont manquantes et où l'utilisation des modèles de dispersion atmosphérique est limitée en raison de la taille du domaine d'étude.

Bien que nous ayons créé un indicateur produisant de meilleures performances que les indicateurs simples historiquement utilisés dans la littérature scientifique, d'autres paramètres connus pour impacter la dispersion des polluants n'ont pas été pris en compte dans l'indicateur final : vitesse et température d'éjections des fumées, température ambiante, vitesse du vent et pluviométrie. Ces paramètres auraient pu potentiellement améliorer davantage l'évaluation de l'exposition en produisant des estimations plus précises de l'exposition. Parmi les paramètres non-retenus, certains diminuaient les performances de l'indicateur (comme la vitesse du vent), d'autres avait peu d'impact sur les résultats ou n'était disponible pour l'ensemble de sources au niveau national (vitesse et température d'éjection des fumées). Enfin, d'autres doivent encore faire l'objet d'étude dans le futur afin d'évaluer leur apport dans la performance de l'indicateur (pluviométrie). L'application de l'indicateur au niveau régionale et nationale ouvre une série de problématique et suggère un certain nombre de perspectives d'amélioration.

## 2. Perspectives de recherches

### *Augmentation du degré de validité de l'indicateur*

Malgré de bonnes performances en termes d'étude des variations spatio-temporelle des expositions, les indicateur SIG ou les modèles gaussiens restent considérés comme des indicateurs de l'exposition moins fiables que des mesures biologiques. La mesure des concentrations de dioxines ou de cadmium dans les tissus adipeux permet de refléter l'exposition cumulative et historique. Pour renforcer le degré de validation de notre indicateur, il serait pertinent de prélever des échantillons de tissus adipeux chez un certain nombre de sujets pour ensuite comparer la concordance entre la classification de l'indicateur et la classification issus des mesures de concentrations dans les tissus adipeux. Mais ces mesures sont coûteuses et le prélèvement de tissu adipeux est une procédure invasive et difficile à mettre en œuvre.

### *Amélioration de l'indicateur*

L'indicateur a attribué aux sujets une valeur d'exposition cumulative pouvant varier d'un facteur  $1 \times 10^9$  d'un sujet à l'autre (Figures 12 et 13). Il paraît donc nécessaire de définir un seuil en dessous duquel il n'est plus possible d'affirmer avec certitude que la contribution des sources industrielles est supérieure à la contribution des autres types de sources, ou d'identifier le seuil correspondant à une exposition inférieure au background (définis à  $10 \text{ fg/m}^3$  dans l'article n°2). Pour cela, un travail sera mené pour établir une correspondance entre l'unité de l'indicateur en  $\mu\text{g/m}^2$  et les valeurs de concentrations modélisées avec le modèle SIRANE en  $\text{fg/m}^3$ .

De plus, au vu des résultats obtenus avec le modèle numérique sur le Grand Lyon (article n°2), il apparaît que les sources industrielles ne sont pas les seuls déterminants de l'exposition en milieu urbain et particulièrement sur la fin de la période d'étude.

Dans le cas des dioxines, les émissions dominantes sur la fin de la période d'étude viennent des sources diffuses et plus particulièrement des émissions liées aux brûlages de câbles. Pour les différentes raisons détaillées dans les chapitres précédents (activité illégale, absence de recensement, difficulté de localisation et des quantifications des émissions), il est difficile, voire impossible, de prendre en compte de manière correcte cette exposition.

Pour le cadmium, l'étude de l'exposition dans la métropole lyonnaise (article n°2) a mis en avant la forte contribution du trafic routier dans les expositions. Dans la littérature scientifique, un certain nombre d'études ont déjà développé des indicateurs pour estimer l'exposition liées au trafic (e.g. Gulliver and Briggs 2011; Visser et al. 2004). L'indicateur développé pourrait être complété par un indicateur basé sur la proximité avec les différents tronçons routiers, la « densité » de routes autour du sujet, les flux des véhicules ou encore la proportion de poids lourds. Ceci permettrait d'avoir une seule métrique pour évaluer l'exposition atmosphérique aux sources industrielles et au trafic routier. Ce nouvel indicateur pourra être calibré et validé en comparant ces performances avec le modèle SIRANE.

Enfin l'inventaire des sources de cadmium et de dioxines et l'indicateur SIG seront utilisés dans le projet XENAIR pour l'évaluation de l'exposition atmosphérique au cadmium et aux dioxines. Dans le cadre de ce projet, nous disposerons des résultats du modèle Chimère, qui fournira une estimation de la concentration annuelle de dioxines et de cadmium sur le territoire français entre 1990 et 2010 avec une résolution de 7x7km. Ceci permettra de prendre en compte la pollution de fond « régionale » et une réflexion devra être menée pour l'intégrer à l'indicateur.

#### *Perspectives d'application futures de l'indicateur*

Cet outil de quantification des expositions atmosphériques s'inscrit dans la mouvance globale d'amélioration des méthodologies d'évaluation des expositions. Plusieurs études récentes ont montré la nécessité d'étudier les expositions avec précision sur des périodes courtes et des fenêtres d'expositions spécifiques (Martens et al. 2017). Ren et al. (2017) ont ainsi récemment montré que l'exposition aux particules fines 1 mois avant et après le début de la grossesse augmentait le risque d'anomalies congénitales.

Grâce à la finesse des informations météorologiques recueillies, l'indicateur SIG est capable d'estimer une exposition journalière sur l'ensemble du territoire français entre 1990 et 2008. La possibilité de prendre en compte une exposition chronique est également pertinente pour l'étude d'un grand nombre de pathologies et de mécanismes biologiques (Ying et al. 2014).



De nouvelles approches statistiques ont été développées permettant d'évaluer les mobilités résidentielles au niveau de l'IRIS (Wahida et al. 2016). Ces approches, combinées à l'utilisation de l'indicateur que nous avons développé, permettraient d'estimer l'exposition de sujets dont l'histoire résidentielle n'est pas disponible et ainsi de minimiser les potentiels biais de classification.

Dans le cas de l'exposition aux dioxines et le risque de cancer du sein, les effets peuvent dépendre du timing d'exposition, en particulier lors de périodes critiques du développement mammaire (in utero, puberté, grossesse et allaitement) (Teitelbaum et al. 2015). La figure 14 illustre la capacité de l'indicateur à identifier des périodes spécifiques de fortes expositions pour chaque sujet. Les femmes E3N étaient âgées de 45 à 60 ans lors du recrutement, nous n'avons pas été en mesure d'évaluer l'exposition atmosphérique pour ces fenêtres critiques d'exposition situées avant 1990. La troncature à gauche des données d'exposition peut avoir un impact sur l'estimation du risque, et ce d'autant plus que les émissions de dioxines ont été plus important avant 1990 (Kjeller et al. 1996; Pacyna et al. 2001).

#### *Exploitation des données E3N*

Les données recueillies par la cohorte E3N sont un des points fort des projets GEO3N et XENAIR et ont permis de répondre à certaines limites des précédentes études, notamment grâce aux données sur l'histoire résidentielle des femmes depuis 1990, ainsi que les données sur l'état de santé des femmes E3N, leur mode de vie et leurs facteurs reproductifs.

Grace au nouveau questionnaire d'histoire résidentielle envoyé en 2017 aux femmes incluses dans l'étude GEO3N, nous aurons la connaissance des lieux d'habitation des sujets E3N sur toute leur vie et nous permettant de définir le statut urbain/rural de chaque lieu de résidence. La connaissance des trajectoires résidentielles permettra d'explorer les effets d'une exposition à long terme et à différents moments de la vie sur le risque de cancer du sein en se basant sur l'inventaire des sources pour la période 1990-2008 et en utilisant la résidence urbaine comme substitut d'exposition pour les périodes antérieures (Lambrechtsen et al. 2012).

#### *Evaluation des expositions via différentes voies*

Nous avons été les premiers, dans le cadre du projet GEO3N, à prendre en compte les deux principales voies d'exposition aux dioxines dans le cadre de l'étude du risque de cancer du sein. Le projet GEO3N a permis d'évaluer les expositions par voies alimentaire (Danjou et al, 2015) et atmosphérique (Danjou et al, 2018 *en cours de finalisation*, Praud et al, *en préparation*). Il serait pertinent d'évaluer l'exposition aux dioxines en tenant compte du risque associé aux expositions par voie alimentaire et aérienne dans un modèle étudiant l'effet conjoint de ces deux voies d'exposition (Petit et al. 2012).

Il serait également envisageable de développer un indicateur pour quantifier les dépôts de dioxines à l'adresse des femmes (grâce aux données de pluviométries recueillies) et de combiner ces données d'exposition avec les données issues des questionnaires alimentaires en ciblant les femmes pratiquant l'autoconsommation.

#### *Evaluation des multi-expositions*

Les installations industrielles émettent non seulement des dioxines, mais aussi d'autres polluants atmosphériques comme le cadmium (Ranzi et al. 2011). Le projet GEO3N portait uniquement sur l'exposition aux dioxines et peut avoir manqué l'effet de l'exposition à d'autres polluants sur le risque de cancer du sein. Des essais expérimentaux sur les effets combinés d'un mélange de xénoestrogènes ont montré qu'ils pouvaient agir ensemble et produire des effets significatifs même lorsque les composants individuels étaient présents à de faibles concentrations, ce qui montre l'importance de considérer l'action conjointe des produits chimiques oestrogéniques pour éviter une sous-estimation potentielle des estimations de risque (Silva et al. 2002). Bien qu'aucun consensus clair ne se soit encore dégagé sur la façon dont une approche multi-polluante peut être abordée dans les études épidémiologiques, les recherches futures, comme le projet XENAIR, sont orientées vers cet objectif.

### **Le projet XENAIR**

Le projet de recherche XENAIR vise à étudier les effets de l'exposition chronique à faibles doses à 8 polluants atmosphériques aux propriétés xénoestrogéniques (PM<sub>10</sub>, PM<sub>2.5</sub>, dioxyde d'azote, ozone, B[a]P, dioxines, PCB et cadmium) sur le risque de cancer du sein, dans une étude cas-témoin nichée dans la cohorte prospective française E3N.

Le projet XENAIR a pour objectif spécifique d'évaluer l'exposition atmosphérique à ces polluants sur une période de 20 ans. La méthodologie repose sur des modèles intervenant à diverses échelles spatiales et temporelles. Ils permettront d'étudier l'impact de l'exposition atmosphérique chronique sur le risque de cancer du sein, en estimant le poids relatif de l'exposition en fonction du temps d'exposition et de l'exposition cumulée. De plus, les données recueillies lors des questionnaires permettront, une quantification des expositions à la pollution atmosphérique intérieure (chauffage domestique, etc.) sera effectuée.

Grâce à la bourse de mobilité du CLARA, j'ai pu me rendre à l'Impérial College dans l'équipe de John Gulliver, pour étudier le fonctionnement des modèles LUR et travailler sur le développement d'un modèle LUR national pour estimer les expositions aux NO<sub>2</sub>, PM<sub>10</sub>, PM<sub>2.5</sub> et O<sub>3</sub>.

A l'échelle nationale, le modèle LUR, utilisé pour modéliser les concentrations de PM<sub>10</sub>, PM<sub>2,5</sub>, NO<sub>2</sub>, O<sub>3</sub> pour les années les plus récentes de l'étude, suivra les étapes de développement suivantes :

1. Élaboration de modèles LUR à l'échelle nationale pour les années récentes du projet (2009-2010) et de modèles "hybrides" combinant les concentrations de fond issu du modèle CHIMERE et les variables expliquant la variabilité locale des concentrations atmosphérique. Pour développer et évaluer les modèles LUR, des données issues de sites de mesures seront utilisées. Ces sites seront séparés en deux groupes : pour calibrer les modèles et pour les évaluer.
2. En utilisant les mesures historiques de concentrations en polluants atmosphériques (base GAZEL AIR), des méthodes de backextrapolation des modèles LUR seront utilisées pour prédire les expositions dans le passé (jusqu'en 1990). Les sites de surveillance de la pollution atmosphérique placée en zones rurales (sites de concentration de fond) serviront de base pour établir des tendances de fond et ainsi tester différents types de backextrapolation.
3. Application pour prédire les concentrations à l'adresse résidentielle des participantes. Ces sorties seront utilisées comme estimation de l'exposition dans les analyses épidémiologiques du projet.

Pour évaluer l'évolution de la concentration de fond entre 1990 et 2010 sur l'ensemble du territoire, le modèle de chimie-transport CHIMERE sera utilisé pour fournir des valeurs horaires de concentrations de fond (résolution spatiale de 7 km<sup>2</sup>) pour tous les polluants. Devant le manque de données concernant le BaP et le PCB 153 pour la France, les concentrations de fonds issus de CHIMERE seront directement utilisées comme référence pour estimer l'exposition à ces deux composées.

Des analyses de sensibilités seront réalisées en comparant les simulations numériques du modèle SIRANE sur la métropole lyonnaise et les résultats issus des modèles LUR locaux et nationaux. Les simulations menées avec le modèle SIRANE permettront également d'étudier la relation entre les concentrations de PM<sub>10</sub>, PM<sub>2,5</sub>, NO<sub>2</sub>, O<sub>3</sub> prédites par un LUR local et celles calculées par le modèle SIRANE. Ce qui permettra d'explorer l'effet du modèle choisi sur la qualité des estimations (de Hoogh et al. 2014). Les corrélations éventuelles entre l'évolution spatiale et temporelle des concentrations de PM<sub>10</sub>, PM<sub>2,5</sub>, NO<sub>2</sub> et O<sub>3</sub> et celles de PCB 153, BaP, TCDD et cadmium seront étudiées, pour potentiellement utiliser les concentrations d'un polluant comme indicateur de l'exposition pour d'autres polluants.

En plus des perspectives innovantes en termes d'évaluation des expositions, le projet XENAIR vise également à étudier, si une association statistiquement positive est observée :

- les expositions à de multiples polluants et les expositions combinées atmosphériques et alimentaires (pour les dioxines, le cadmium, les PCB et le B[a]P).
- les interactions entre l'exposition à long terme aux polluants de l'air ambiant et les polymorphismes génétiques impliqués dans le métabolisme des xénoestrogènes (testé parmi 2500 cas de cancer du sein et 2500 témoins appariés).
- le rôle de la méthylation de l'ADN en tant que marqueur intermédiaire du risque associé à l'exposition aux polluants atmosphérique et au risque de cancer du sein dans une étude cas-témoin nichée incluant 200 cas de cancer du sein chez des femmes non-fumeuses.

### 3. Conclusion

Le rôle des facteurs environnementaux dans la survenue du cancer devient de plus en plus évident et constitue une préoccupation majeure de santé publique. Parmi ces facteurs environnementaux, les polluants xénoestrogènes et les perturbateurs endocriniens sont plus particulièrement suspectés d'être associés à une augmentation du risque des cancers hormonodépendants, tels que le cancer du sein.

Un certain nombre d'études se sont penchées sur le lien entre cancer du sein et exposition à la pollution atmosphérique et notamment aux dioxines et au cadmium. Toutefois, les résultats de ces études ne sont pas concluants en raison des différences méthodologiques et de plusieurs limites. Ma thèse s'est concentrée sur l'amélioration des méthodes d'évaluation des expositions aux dioxines et au cadmium dans le cadre de deux études épidémiologiques.

Un indicateur SIG capable d'évaluer l'exposition atmosphérique chronique multi-source à l'échelle individuelle sur le territoire métropolitain a été développée. Cet indicateur appliqué dans le cadre du projet GEO3N contribue à clarifier la relation entre l'exposition à ces polluants et le risque de cancer du sein en considérant l'exposition combinée des individus par voies multiples, en apportant de nouvelles connaissances par le biais d'une recherche interdisciplinaire et en comblant les lacunes méthodologiques de la littérature scientifique existante. Cet indicateur pourra être adapté à l'évaluation de l'exposition à d'autres polluants environnementaux liés au cancer du sein ou utilisé dans l'étude d'autres pathologies.

Les recherches menées durant cette thèse ont également permis d'approfondir les connaissances sur les déterminants de l'exposition atmosphérique chronique aux dioxines et au cadmium en milieu urbain, via l'utilisation d'un modèle de dispersion atmosphérique.

De manière plus globale, ces travaux de recherche apportent des éléments de réponse aux questionnements des différents plans nationaux en santé environnement (PNSE 2004, 2009, 2015).



## Références

1. Andersen ZJ, Ravnskjaer L, Andersen KK, Loft S, Brandt J, Becker T, et al. 2017. Long-term Exposure to Fine Particulate Matter and Breast Cancer Incidence in the Danish Nurse Cohort Study. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 26:428–430; doi:10.1158/1055-9965.EPI-16-0578.
2. Anzivino-Viricel L, Falette N, Carretier J, Montestrucq L, Guye O, Philip T, et al. 2012. Gestion des déchets ménagers et assimilés : bilan des connaissances et évaluation des effets sanitaires en population générale et au travail. *Environ Risques Santé* 11:360–377; doi:10.1684/ers.2012.0559.
3. Arisawa K, Matsumura T, Tohyama C, Saito H, Satoh H, Nagai M, et al. 2003. Fish intake, plasma  $\omega$ -3 polyunsaturated fatty acids, and polychlorinated dibenzo-p-dioxins/polychlorinated dibenzo-furans and co-planar polychlorinated biphenyls in the blood of the Japanese population. *Int Arch Occup Environ Health* 76:205–215; doi:10.1007/s00420-002-0400-y.
4. Aronson KJ, Miller AB, Woolcott CG, Sterns EE, McCready DR, Lickley LA, et al. 2000. Breast adipose tissue concentrations of polychlorinated biphenyls and other organochlorines and breast cancer risk. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 9: 55–63.
5. Ashworth DC, Fuller GW, Toledano MB, Font A, Elliott P, Hansell AL, et al. 2013. Comparative assessment of particulate air pollution exposure from municipal solid waste incinerator emissions. *J Environ Public Health* 2013:560342; doi:10.1155/2013/560342.
6. Atmo Auvergne-Rhône-Alpes. 2016. Méthode d'élaboration de l'inventaire des émissions atmosphériques en Rhône-Alpes.
7. Baan R, Grosse Y, Straif K, Secretan B, El Ghissassi F, Bouvard V, et al. 2009. A review of human carcinogens—Part F: Chemical agents and related occupations. *Lancet Oncol* 10:1143–1144; doi:10.1016/S1470-2045(09)70358-4.
8. Basagana X, Aguilera I, Rivera M, Agis D, Foraster M, Marrugat J, et al. 2013. Measurement Error in Epidemiologic Studies of Air Pollution Based on Land-Use Regression Models. *Am J Epidemiol* 178:1342–1346; doi:10.1093/aje/kwt127.
9. Ben Salem N, Garbero V, Salizzoni P, Lamaison G, Soulhac L. 2015. Modelling Pollutant Dispersion in a Street Network. *Bound-Layer Meteorol* 155:157–187; doi:10.1007/s10546-014-9990-7.
10. Bertazzi A, Pesatori AC, Consonni D, Tironi A, Landi MT, Zocchetti C. 1993. Cancer incidence in a population accidentally exposed to 2,3,7,8-tetrachlorodibenzo-para-dioxin. *Epidemiol Camb Mass* 4: 398–406.
11. Bessagnet B, Seigneur C, Menut L. 2010. Impact of dry deposition of semi-volatile organic compounds on secondary organic aerosols. *Atmos Environ* 44:1781–1787; doi:10.1016/j.atmosenv.2010.01.027.
12. Bidgoli SA, Ahmadi R, Zavarhei MD. 2010. Role of hormonal and environmental factors on early incidence of breast cancer in Iran. *Sci Total Environ* 408:4056–4061; doi:10.1016/j.scitotenv.2010.05.018.
13. Binachon B, Dossus L, Danjou AMN, Clavel-Chapelon F, Fervers B. 2014. Life in urban areas and breast cancer risk in the French E3N cohort. *Eur J Epidemiol* 29:743–751; doi:10.1007/s10654-014-9942-z.
14. Bonner MR, Han DW, Nie L, Rogerson P, Vena JE, Muti P, et al. 2005. Breast cancer risk and exposure in early life to polycyclic aromatic hydrocarbons using total suspended particulates as a proxy measure. *Cancer Epidemiol Biomarkers Prev* 14: 53–60.

15. Brama M, Gnessi L, Basciani S, Cerulli N, Politi L, Spera G, et al. 2007. Cadmium induces mitogenic signaling in breast cancer cell by an ERalpha-dependent mechanism. *Mol Cell Endocrinol* 264:102–108; doi:10.1016/j.mce.2006.10.013.
16. Breivik K, Vestreng V, Rozovskaya O, Pacyna JM. 2006. Atmospheric emissions of some POPs in Europe: a discussion of existing inventories and data needs. *Environ Sci Policy* 9:663–674; doi:10.1016/j.envsci.2006.09.001.
17. Briggs DJ, de Hoogh C, Guiliver J, Wills J, Elliott P, Kingham S, et al. 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments. *Sci Total Environ* 253:151–167; doi:10.1016/S0048-9697(00)00429-0.
18. Brody JG, Moysich KB, Humblet O, Attfield KR, Beehler GP, Rudel RA. 2007. Environmental pollutants and breast cancer - Epidemiologic studies. *Cancer* 109:2667–2711; doi:10.1002/cncr.22655.
19. Carpentieri M, Salizzoni P, Robins A, Soulhac L. 2012. Evaluation of a neighbourhood scale, street network dispersion model through comparison with wind tunnel data. *Environ Model Softw* 37:110–124; doi:10.1016/j.envsoft.2012.03.009.
20. Chen F-P, Chien M-H. 2014. Lower concentrations of phthalates induce proliferation in human breast cancer cells. *Climacteric J Int Menopause Soc* 17:377–384; doi:10.3109/13697137.2013.865720.
21. Cheruiyot NK, Lee W-J, Yan P, Mwangi JK, Wang L-C, Gao X, et al. 2016. An Overview of PCDD/F Inventories and Emission Factors from Stationary and Mobile Sources: What We Know and What is Missing. *Aerosol Air Qual Res* 16:2965–2988; doi:10.4209/aaqr.2016.10.0447.
22. CITEPA. 2017. CITEPA, 2017. Rapport OMINEA – 14ème édition.
23. Clavel-Chapelon F. 2015. Cohort Profile: The French E3N Cohort Study. *Int J Epidemiol* 44:801–809; doi:10.1093/ije/dyu184.
24. Clavel-Chapelon F, van Liere MJ, Giubout C, Niravong MY, Goulard H, Le Corre C, et al. 1997. E3N, a French cohort study on cancer risk factors. E3N Group. Etude Epidémiologique auprès de femmes de l'Education Nationale. *Eur J Cancer Prev Off J Eur Cancer Prev Organ ECP* 6: 473–478.
25. Consonni D, Pesatori AC, Zocchetti C, Sindaco R, D'Oro LC, Rubagotti M, et al. 2008. Mortality in a population exposed to dioxin after the Seveso, Italy, accident in 1976: 25 years of follow-up. *Am J Epidemiol* 167:847–858; doi:10.1093/aje/kwm371.
26. Consonni D, Sindaco R, Bertazzi PA. 2012. Blood levels of dioxins, furans, dioxin-like PCBs, and TEQs in general populations: a review, 1989-2010. *Environ Int* 44:151–162; doi:10.1016/j.envint.2012.01.004.
27. Cordier S, Lehébel A, Amar E, Anzivino-Viricel L, Hours M, Monfort C, et al. 2010. Maternal residence near municipal waste incinerators and the risk of urinary tract birth defects. *Occup Environ Med* 67:493–499; doi:10.1136/oem.2009.052456.
28. Cordioli M, Ranzi A, De Leo GA, Lauriola P. 2013a. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators. *J Environ Public Health*; doi:10.1155/2013/129470.
29. Coudon T, Hourani H, Nguyen C, Faure E, Mancini FR, Fervers B, et al. 2018. Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon metropolitan area (France). *Environ Int* 111:177–190; doi:10.1016/j.envint.2017.11.027.
30. Dai D, Oyana TJ. 2008a. Spatial variations in the incidence of breast cancer and potential risks associated with soil dioxin contamination in Midland, Saginaw, and Bay Counties, Michigan, USA. *Environ Health* 7:49; doi:10.1186/1476-069X-7-49.



31. de Hoogh K, Korek M, Vienneau D, Keuken M, Kukkonen J, Nieuwenhuijsen MJ, et al. 2014. Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environ Int* 73:382–392; doi:10.1016/j.envint.2014.08.011.
32. Demers A, Ayotte P, Brisson J, Dodin S, Robert J, Dewailly É. 2002. Plasma Concentrations of Polychlorinated Biphenyls and the Risk of Breast Cancer: A Congener-specific Analysis. *Am J Epidemiol* 155:629–635; doi:10.1093/aje/155.7.629.
33. Den Hond E, Roels HA, Hoppenbrouwers K, Nawrot T, Thijs L, Vandermeulen C, et al. 2002. Sexual maturation in relation to polychlorinated aromatic hydrocarbons: Sharpe and Skakkebaek's hypothesis revisited. *Environ Health Perspect* 110: 771–776.
34. Diamanti-Kandarakis E, Bourguignon J-P, Giudice LC, Hauser R, Prins GS, Soto AM, et al. 2009. Endocrine-disrupting chemicals: an Endocrine Society scientific statement. *Endocr Rev* 30:293–342; doi:10.1210/er.2009-0002.
35. Domenico A di, Cerlesi S, Ratti S. 1990. A two-exponential model to describe the vanishing trend of 2,3,7,8-tetrachlorodibenzodioxin (TCDD) in the soil at Seveso, northern Italy. *Chemosphere* 20: 1559–1566.
36. Eeftens M, Beelen R, de Hoogh K, Bellander T, Cesaroni G, Cirach M, et al. 2012a. Development of Land Use Regression Models for PM<sub>2.5</sub>, PM<sub>2.5</sub> Absorbance, PM<sub>10</sub> and PM<sub>coarse</sub> in 20 European Study Areas; Results of the ESCAPE Project. *Environ Sci Technol* 46:11195–11205; doi:10.1021/es301948k.
37. El Ghissassi F, Baan R, Straif K, Grosse Y, Secretan B, Bouvard V, et al. 2009. A review of human carcinogens—Part D: radiation. *Lancet Oncol* 10:751–752; doi:10.1016/S1470-2045(09)70213-X.
38. EPA. 2002. Comparison of 2002 model-predicted concentration to monitored data.
39. European Commission. 2000. Ambient air pollution by AS, CD and NI compounds. 27–28.
40. Fabre P, Daniau C, Gorla S, de Crouy-Chanel P, Empereur-Bissonnet P. 2008. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. Synthèse St-Maurice InVS.
41. Faure E, Danjou AMN, Clavel-Chapelon F, Boutron-Ruault M-C, Dossus L, Fervers B. 2017. Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environ Health* 16:15; doi:10.1186/s12940-017-0217-5.
42. Fenton SE, Birnbaum LS. 2015. Timing of Environmental Exposures as a Critical Element in Breast Cancer Risk. *J Clin Endocrinol Metab* 100:3245–3250; doi:10.1210/jc.2015-2848.
43. Flaounas E, Coll I, Armengaud A, Schmechtig C. 2009. The representation of dust transport and missing urban sources as major issues for the simulation of PM episodes in a Mediterranean area. *Atmospheric Chem Phys* 9: 8091–8101.
44. Floret N, Mauny F, Challier B, Arveux P, Cahn J-Y, Viel J-F. 2003. Dioxin Emissions from a Solid Waste Incinerator and Risk of Non-Hodgkin Lymphoma. *Epidemiology* 14:392; doi:10.1097/01.ede.0000072107.90304.01.
45. Floret N, Viel J-F, Lucot E, Dudermeil P-M, Cahn J-Y, Badot P-M, et al. 2006. Dispersion Modeling as a Dioxin Exposure Indicator in the Vicinity of a Municipal Solid Waste Incinerator: A Validation Study. *Environ Sci Technol* 40:2149–2155; doi:10.1021/es052309u.
46. Gammon MD, Santella RM, Neugut AI, Eng SM, Teitelbaum SL, Paykin A, et al. 2002. Environmental toxins and breast cancer on Long Island. I. Polycyclic aromatic hydrocarbon DNA adducts. *Cancer Epidemiol Biomarkers Prev* 11: 677–685.

47. Garcia E, Hurley S, Nelson DO, Gunier RB, Hertz A, Reynolds P. 2014. Evaluation of the agreement between modeled and monitored ambient hazardous air pollutants in California. *Int J Environ Health Res* 24:363–377; doi:10.1080/09603123.2013.835031.
48. Garcia E, Hurley S, Nelson DO, Hertz A, Reynolds P. 2015. Hazardous air pollutants and breast cancer risk in California teachers: a cohort study. *Environ Health* 14:UNSP 14; doi:10.1186/1476-069X-14-14.
49. Gulliver J, Briggs D. 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Sci Total Environ* 409:2419–2429; doi:10.1016/j.scitotenv.2011.03.004.
50. Gulliver J, de Hoogh K, Hansell A, Vienneau D. 2013. Development and Back-Extrapolation of NO<sub>2</sub> Land Use Regression Models for Historic Exposure Assessment in Great Britain. *Environ Sci Technol* 47:7804–7811; doi:10.1021/es4008849.
51. Gusev A, Aas W, Modlik M, Srnensky R. 2016. Persistent Organic Pollutants: assessment of transboundary pollution on regional and global scales.
52. Han D, Bonner MR, Nie J, Freudenheim JL. 2013. Assessing bias associated with geocoding of historical residence in epidemiology research. *Geospatial Health* 7:369–374; doi:10.4081/gh.2013.93.
53. Hanna S, Chang J. 2012. Acceptance criteria for urban dispersion model evaluation. *Meteorol Atmospheric Phys* 116:133–146; doi:10.1007/s00703-011-0177-1.
54. Hardell L. 1993. Phenoxy Herbicides, Chlorophenols, Soft-Tissue Sarcoma (sts) and Malignant-Lymphoma. *Br J Cancer* 67:1154–1155; doi:10.1038/bjc.1993.212.
55. Hart JE, Bertrand KA, DuPre N, James P, Vieira VM, Tamimi RM, et al. 2016. Long-term Particulate Matter Exposures during Adulthood and Risk of Breast Cancer Incidence in the Nurses' Health Study II Prospective Cohort. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 25:1274–1276; doi:10.1158/1055-9965.EPI-16-0246.
56. Hodzic A, Jimenez JL, Madronich S, Canagaratna MR, DeCarlo PF, Kleinman L, et al. 2010. Modeling organic aerosols in a megacity: potential contribution of semi-volatile and intermediate volatility primary organic compounds to secondary organic aerosol formation. *Atmospheric Chem Phys* 10:5491–5514; doi:10.5194/acp-10-5491-2010.
57. Hodzic A, Vautard R, Bessagnet B, Lattuati M, Moreto F. 2005. Long-term urban aerosol simulation versus routine particulate matter observations. *Atmos Environ* 39:5851–5864; doi:10.1016/j.atmosenv.2005.06.032.
58. Hoek G, Beelen R, de Hoogh K, Vienneau D, Gulliver J, Fischer P, et al. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmos Environ* 42:7561–7578; doi:10.1016/j.atmosenv.2008.05.057.
59. Hoek G, Fischer P, Van den Brandt P, Goldbohm S, Brunekreef B. 2001a. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J Expo Anal Environ Epidemiol* 11:459–469; doi:10.1038/sj.jea.7500189.
60. Huang YL, Batterman S. 2000. Residence location as a measure of environmental exposure: a review of air pollution epidemiology studies. *J Expo Anal Environ Epidemiol* 10: 66–85.
61. Hystad P, Villeneuve PJ, Goldberg MS, Crouse DL, Johnson K. 2015. Exposure to traffic-related air pollution and the risk of developing breast cancer among women in eight Canadian provinces: A case–control study. *Environ Int* 74:240–248; doi:10.1016/j.envint.2014.09.004.
62. IARC. 1997. IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS (1997) Polychlorinated Dibenzo-para-dioxins and Polychlorinated Dibenzofurans.

63. IARC, ed. 2012. *IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 F, chemical agents and related occupations: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 20 - 27 October 2009*. IARC:Lyon.
64. IARC IWG on the E of CR to. 2012d. *CADMIUM AND CADMIUM COMPOUNDS*. International Agency for Research on Cancer.
65. Inserm, Afsset. 2008. *Cancer et environnement - Expertise collective*. Les éditions Inserm.
66. Jacquemin B, Johanna L, Anne B, Valerie S. 2013. Environmental Health Perspectives – Impact of Geocoding Methods on Associations between Long-term Exposure to Urban Air Pollution and Lung Function. *Environ Health Perspect* 1054–1060; doi:10.1289/ehp.1206016.
67. Järup L, Akesson A. 2009. Current status of cadmium as an environmental health problem. *Toxicol Appl Pharmacol* 238:201–208; doi:10.1016/j.taap.2009.04.020.
68. Johnson MD, Kenney N, Stoica A, Hilakivi-Clarke L, Singh B, Chepko G, et al. 2003. Cadmium mimics the in vivo effects of estrogen in the uterus and mammary gland. *Nat Med* 9:1081–1084; doi:10.1038/nm902.
69. Josyula S, Lin J, Xue X, Rothman N, Lan Q, Rohan TE, et al. 2015. Household air pollution and cancers other than lung: a meta-analysis. *Environ Health Glob Access Sci Source* 14:24; doi:10.1186/s12940-015-0001-3.
70. Julin B, Wolk A, Bergkvist L, Bottai M, Akesson A. 2012. Dietary cadmium exposure and risk of postmenopausal breast cancer: a population-based prospective cohort study. *Cancer Res* 72:1459–1466; doi:10.1158/0008-5472.CAN-11-0735.
71. Kavlock RJ, Daston GP, DeRosa C, Fenner-Crisp P, Gray LE, Kaattari S, et al. 1996. Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. EPA-sponsored workshop. *Environ Health Perspect* 104 Suppl 4: 715–740.
72. Kazantzis G, Lam TH, Sullivan KR. 1988. Mortality of cadmium-exposed workers. A five-year update. *Scand J Work Environ Health* 14: 220–223.
73. Kiesewetter G, Borcken-Kleefeld J, Schoepp W, Heyes C, Thunis P, Bessagnet B, et al. 2014. Modelling NO<sub>2</sub> concentrations at the street level in the GAINS integrated assessment model: projections under current legislation. *Atmospheric Chem Phys* 14:813–829; doi:10.5194/acp-14-813-2014.
74. Kjeller L-O, Jones KC, Johnston AE, Rappe C. 1996. Evidence for a Decline in Atmospheric Emissions of PCDD/Fs in the U.K. *Environ Sci Technol* 30:1398–1403; doi:10.1021/es950708r.
75. Kogevinas M. 2001. Human health effects of dioxins: cancer, reproductive and endocrine system effects. *Hum Reprod Update* 7: 331–339.
76. Kok H, Pulles T, Quass U. 2005. Dioxin emission estimates and reduction options for Central and Eastern Europe. In: *Environmental Health Risk III* (C.A. Brebbia, V. Popov, and D. Fayzieva, eds). Vol. 9 of. Wit Press:Southampton. 189–197.
77. Lambrechtsen J, Gerke O, Egstrup K, Sand NP, Nørgaard BL, Petersen H, et al. 2012. The relation between coronary artery calcification in asymptomatic subjects and both traditional risk factors and living in the city centre: a DanRisk substudy. *J Intern Med* 271:444–450; doi:10.1111/j.1365-2796.2011.02486.x.
78. Langlois PH, Brender JD, Suarez L, Zhan FB, Mistry JH, Scheuerle A, et al. 2009. Maternal residential proximity to waste sites and industrial facilities and conotruncal heart defects in offspring. *Paediatr Perinat Epidemiol* 23:321–331; doi:10.1111/j.1365-3016.2009.01045.x.

79. Lauby-Secretan B, Loomis D, Grosse Y, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al. 2013. Carcinogenicity of polychlorinated biphenyls and polybrominated biphenyls. *Lancet Oncol* 14:287–288; doi:10.1016/S1470-2045(13)70104-9.
80. Lauwerys RR, Bernard AM, Roels HA, Buchet JP. 1994. Cadmium: exposure markers as predictors of nephrotoxic effects. *Clin Chem* 40: 1391–1394.
81. Leijds MM, Koppe JG, Olie K, Van WA, Voogt P, Vulsma T, et al. 2008. Delayed initiation of breast development in girls with higher prenatal dioxin exposure; a longitudinal cohort study. *Chemosphere* 73:999–1004; doi:10.1016/j.chemosphere.2008.05.053.
82. Leijds MM, Koppe JG, Vulsma T, Olie K, van Aalderen WMC, de Voogt P, et al. 2017. Alterations in the programming of energy metabolism in adolescents with background exposure to dioxins, dl-PCBs and PBDEs. *PLoS One* 12:e0184006; doi:10.1371/journal.pone.0184006.
83. Lin J, Zhang F, Lei Y. 2016. Dietary intake and urinary level of cadmium and breast cancer risk: A meta-analysis. *Cancer Epidemiol* 42:101–107; doi:10.1016/j.canep.2016.04.002.
84. Liu R, Nelson DO, Hurley S, Hertz A, Reynolds P. 2015a. Residential Exposure to Estrogen Disrupting Hazardous Air Pollutants and Breast Cancer Risk The California Teachers Study. *Epidemiology* 26:365–373; doi:10.1097/EDE.0000000000000277.
85. Liu R, Nelson DO, Hurley S, Hertz A, Reynolds P. 2015b. Residential exposure to estrogen disrupting hazardous air pollutants and breast cancer risk: the California Teachers Study. *Epidemiol Camb Mass* 26:365–373; doi:10.1097/EDE.0000000000000277.
86. Lohmann R, Jones KC. 1998. Dioxins and furans in air and deposition: a review of levels, behaviour and processes. *Sci Total Environ* 219: 53–81.
87. Loomis D, Grosse Y, Lauby-Secretan B, El Ghissassi F, Bouvard V, Benbrahim-Tallaa L, et al. 2013. The carcinogenicity of outdoor air pollution. *Lancet Oncol* 14:1262–1263; doi:10.1016/S1470-2045(13)70487-X.
88. López-Carrillo L, Hernández-Ramírez RU, Calafat AM, Torres-Sánchez L, Galván-Portillo M, Needham LL, et al. 2009. Exposure to Phthalates and Breast Cancer Risk in Northern Mexico. *Environ Health Perspect* 118:539–544; doi:10.1289/ehp.0901091.
89. LUA. 2000. The European Dioxin Emission Inventory. Stage II.
90. Mandal PK. 2005. Dioxin: a review of its environmental effects and its aryl hydrocarbon receptor biology. *J Comp Physiol [B]* 175:221–230; doi:10.1007/s00360-005-0483-3.
91. Manuwald U, Garrido MV, Berger J, Manz A, Baur X. 2012. Mortality study of chemical workers exposed to dioxins: follow-up 23 years after chemical plant closure. *Occup Environ Med* 69: 636–642.
92. Martens DS, Gouveia S, Madhloum N, Janssen BG, Plusquin M, Vanpoucke C, et al. 2017. Neonatal Cord Blood Oxylipins and Exposure to Particulate Matter in the Early-Life Environment: An ENVIRONAGE Birth Cohort Study. *Environ Health Perspect* 125:691–698; doi:10.1289/EHP291.
93. Millikan R, DeVoto E, Duell EJ, Tse CK, Savitz DA, Beach J, et al. 2000. Dichlorodiphenyldichloroethene, polychlorinated biphenyls, and breast cancer among African-American and white women in North Carolina. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 9: 1233–1240.
94. Momeniha F, Faridi S, Amini H, Shamsipour M, Naddafi K, Yunesian M, et al. 2017. Estimating national dioxins and furans emissions, major sources, intake doses, and temporal trends in Iran from 1990–2010. *J Environ Health Sci Eng* 15:20; doi:10.1186/s40201-017-0283-1.

95. Mordukhovich I, Beyea J, Herring AH, Hatch M, Stellman SD, Teitelbaum SL, et al. 2016. Vehicular Traffic-Related Polycyclic Aromatic Hydrocarbon Exposure and Breast Cancer Incidence: The Long Island Breast Cancer Study Project (LIBCSP). *Environ Health Perspect* 124:30–38; doi:10.1289/ehp.1307736.
96. Moysich KB, Ambrosone CB, Vena JE, Shields PG, Mendola P, Kostyniak P, et al. 1998. Environmental organochlorine exposure and postmenopausal breast cancer risk. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 7: 181–188.
97. Moysich KB, Shields PG, Freudenheim JL, Schisterman EF, Vena JE, Kostyniak P, et al. 1999. Polychlorinated biphenyls, cytochrome P4501A1 polymorphism, and postmenopausal breast cancer risk. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 8: 41–44.
98. Nagata C, Nagao Y, Nakamura K, Wada K, Tamai Y, Tsuji M, et al. 2013. Cadmium exposure and the risk of breast cancer in Japanese women. *Breast Cancer Res Treat* 138:235–239; doi:10.1007/s10549-013-2414-4.
99. Nawrot T, Plusquin M, Hogervorst J, Roels HA, Celis H, Thijs L, et al. 2006. Environmental exposure to cadmium and risk of cancer: a prospective population-based study. *Lancet Oncol* 7:119–126; doi:10.1016/S1470-2045(06)70545-9.
100. Nesatyy VJ, Ammann AA, Rutishauser BV, Suter MJF. 2006. Effect of cadmium on the interaction of 17beta-estradiol with the rainbow trout estrogen receptor. *Environ Sci Technol* 40: 1358–1363.
101. Nie J, Beyea J, Bonner MR, Han D, Vena JE, Rogerson P, et al. 2007. Exposure to traffic emissions throughout life and risk of breast cancer: the Western New York Exposures and Breast Cancer (WEB) study. *Cancer Causes Control* 18:947–955; doi:10.1007/s10552-007-9036-2.
102. Nuckols JR, Ward MH, Jarup L. 2004. Using Geographic Information Systems for Exposure Assessment in Environmental Epidemiology Studies. *Environ Health Perspect* 112:1007–1015; doi:10.1289/ehp.6738.
103. Nzihou A, Themelis NJ, Kemiha M, Benhamou Y. 2012. Dioxin emissions from municipal solid waste incinerators (MSWIs) in France. *Waste Manag* 32:2273–2277; doi:10.1016/j.wasman.2012.06.016.
104. Olendrzyński K, Anderberg S, Stigliani W, Bartnicki J, Pacyna J. 1996. Atmospheric emissions and depositions of cadmium, lead, and zinc in Europe during the period 1955–1987. *Environ Rev* 4:300–320; doi:10.1139/a96-016.
105. Pacyna EG, Pacyna JM, Pirrone N. 2001. European emissions of atmospheric mercury from anthropogenic sources in 1995. *Atmos Environ* 35:2987–2996; doi:10.1016/S1352-2310(01)00102-9.
106. Pay MT, Jimenez-Guerrero P, Baldasano JM. 2012. Assessing sensitivity regimes of secondary inorganic aerosol formation in Europe with the CALIOPE-EU modeling system. *Atmos Environ* 51:146–164; doi:10.1016/j.atmosenv.2012.01.027.
107. Pedersen M, Giorgis-Allemand L, Bernard C, Aguilera I, Andersen A-MN, Ballester F, et al. 2013. Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med* 1:695–704; doi:10.1016/S2213-2600(13)70192-9.
108. Pesatori A, Consonni D, Rubagotti M, Grillo P, Bertazzi P. 2009. Cancer incidence in the population exposed to dioxin after the “Seveso accident”: twenty years of follow-up. *Environ Health* 8:39; doi:10.1186/1476-069X-8-39.
109. PNSE. 2004. 1er Plan National Santé Environnement (2004-2008).
110. PNSE. 2009. 2eme Plan National Santé Environnement (2009-2014).

111. PNSE. 2015. 3eme Plan National Santé Environnement (2015-2019).
112. Pronk A, Nuckols JR, De Roos AJ, Airola M, Colt JS, Cerhan JR, et al. 2013. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case-control study. *Environ Health* 12:20; doi:10.1186/1476-069X-12-20.
113. Pulles T, Kok H, Quass U. 2006. Application of the emission inventory model TEAM: Uncertainties in dioxin emission estimates for central Europe. *Atmos Environ* 40:2321–2332; doi:10.1016/j.atmosenv.2005.12.015.
114. Quass U, Fermann M, Broker G. 2004. The European dioxin air emission inventory project - Final results. *Chemosphere* 54:1319–1327; doi:10.1016/S0045-6535(03)00251-0.
115. Quina MJ, Pedro RS, Gando-Ferreira LM, Quinta-Ferreira RM. 2011. A national inventory to estimate release of polychlorinated dibenzo-p-dioxins and dibenzofurans in Portugal. *Chemosphere* 85:1749–1758; doi:10.1016/j.chemosphere.2011.09.028.
116. Raaschou-Nielsen O, Andersen ZJ, Hvidberg M, Jensen SS, Ketznel M, Sorensen M, et al. 2011. Air pollution from traffic and cancer incidence: a Danish cohort study. *Environ Health* 10:67; doi:10.1186/1476-069X-10-67.
117. Raaschou-Nielsen O, Pavuk M, Leblanc A, Dumas P, Philippe Weber J, Olsen A, et al. 2005. Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 14: 67–74.
118. Ranzi A, Fano V, Erspamer L, Lauriola P, Perucci CA, Forastiere F. 2011. Mortality and morbidity among people living close to incinerators: a cohort study based on dispersion modeling for exposure assessment. *Environ Health* 10:22; doi:10.1186/1476-069X-10-22.
119. Ranzi A, Porta D, Badaloni C, Cesaroni G, Lauriola P, Davoli M, et al. 2014. Exposure to air pollution and respiratory symptoms during the first 7 years of life in an Italian birth cohort. *Occup Environ Med* 71:430–436; doi:10.1136/oemed-2013-101867.
120. Read D, Wright C, Weinstein P, Borman B. 2007. Cancer incidence and mortality in a New Zealand community potentially exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin from 2,4,5-trichlorophenoxyacetic acid manufacture. *Aust N Z J Public Health* 31:13–18; doi:10.1111/j.1753-6405.2007.00003.x.
121. Reding KW, Young MT, Szpiro AA, Han CJ, DeRoo LA, Weinberg C, et al. 2015. Breast Cancer Risk in Relation to Ambient Air Pollution Exposure at Residences in the Sister Study Cohort. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol* 24:1907–1909; doi:10.1158/1055-9965.EPI-15-0787.
122. Ren S, Haynes E, Hall E, Hossain M, Chen A, Muglia L, et al. 2017. Periconception Exposure to Air Pollution and Risk of Congenital Malformations. *J Pediatr*; doi:10.1016/j.jpeds.2017.09.076.
123. Revich B, Aksel E, Ushakova T, Ivanova I, Zhuchenko N, Klyuev N, et al. 2001. Dioxin exposure and public health in Chapaevsk, Russia. *Chemosphere* 43: 951–966.
124. Ries L, Harkins D, Krapcho M, Mariotto A, Miller B, Feuer E, et al. 2006. SEER Cancer Statistics Review, 1975-2003. *Public Health Fac Publ*.
125. Salehi F, Turner MC, Phillips KP, Wigle DT, Krewski D, Aronson KJ. 2008. Review of the etiology of breast cancer with special attention to organochlorines as potential endocrine disruptors. *J Toxicol Environ Health B Crit Rev* 11:276–300; doi:10.1080/10937400701875923.
126. Saral A, Gunes G, Demir S. 2014. Atmospheric PCDD/PCDF emission inventory for Turkey. *Atmospheric Pollut Res* 5:24–33; doi:10.5094/APR.2014.004.

127. Serveau L, Boutang J, Allemand N, Fontelle J-P, Vincent J. 2013. source CITEPA / format SECTEN – INVENTAIRE DES EMISSIONS DE POLLUANTS ATMOSPHERIQUES ET DE GAZ A EFFET DE SERRE EN FRANCE SERIES SECTORIELLES ET ANALYSES ETENDUES.
128. Silva E, Rajapakse N, Kortenkamp A. 2002. Something from “Nothing” – Eight Weak Estrogenic Chemicals Combined at Concentrations below NOECs Produce Significant Mixture Effects. *Environ Sci Technol* 36:1751–1756; doi:10.1021/es0101227.
129. Skinner MK, Manikkam M, Guerrero-Bosagna C. 2011. Epigenetic transgenerational actions of endocrine disruptors. *Reprod Toxicol Elmsford N* 31:337–343; doi:10.1016/j.reprotox.2010.10.012.
130. Slama R, Vernet C, Nassan FL, Hauser R, Philippat C. 2017. Characterizing the effect of endocrine disruptors on human health: The role of epidemiological cohorts. *C R Biol* 340:421–431; doi:10.1016/j.crvi.2017.07.008.
131. Sonawane BR. 1995. Chemical contaminants in human milk: an overview. *Environ Health Perspect* 103: 197–205.
132. Soulhac L, Nguyen CV, Volta P, Salizzoni P. 2017. The model SIRANE for atmospheric urban pollutant dispersion. PART III: Validation against NO<sub>2</sub> yearly concentration measurements in a large urban agglomeration. *Atmos Environ* 167:377–388; doi:10.1016/j.atmosenv.2017.08.034.
133. Soulhac L, Puel C, Duclaux O, Perkins RJ. 2003. Simulations of atmospheric pollution in Greater Lyon an example of the use of nested models. *Atmos Environ* 37:5147–5156; doi:10.1016/j.atmosenv.2003.03.002.
134. Soulhac L, Salizzoni P, Cierco F-X, Perkins R. 2011. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmos Environ* 45:7379–7395; doi:10.1016/j.atmosenv.2011.07.008.
135. Soulhac L, Salizzoni P, Mejean P, Didier D, Rios I. 2012. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmos Environ* 49:320–337; doi:10.1016/j.atmosenv.2011.11.031.
136. Steenland K, Whelan E, Deddens J, Stayner L, Ward E. 2003. Ethylene oxide and breast cancer incidence in a cohort study of 7576 women (United States). *Cancer Causes Control CCC* 14: 531–539.
137. Stoica A, Katzenellenbogen BS, Martin MB. 2000. Activation of estrogen receptor-alpha by the heavy metal cadmium. *Mol Endocrinol Baltim Md* 14:545–553; doi:10.1210/mend.14.4.0441.
138. Straif K, Benbrahim-Tallaa L, Baan R, Grosse Y, Secretan B, El Ghissassi F, et al. 2009. A review of human carcinogens--Part C: metals, arsenic, dusts, and fibres. *Lancet Oncol* 10: 453–454.
139. Takiguchi M, Yoshihara S. 2006. New aspects of cadmium as endocrine disruptor. *Environ Sci* 13: 107–116.
140. Teitelbaum SL, Belpoggi F, Reinlib L. 2015. Advancing Research on Endocrine Disrupting Chemicals in Breast Cancer: Expert Panel Recommendations. *Reprod Toxicol Elmsford N* 54:141–147; doi:10.1016/j.reprotox.2014.12.015.
141. Terrenoire E, Bessagnet B, Rouil L, Tognet F, Pirovano G, Letinois L, et al. 2015. High-resolution air quality simulation over Europe with the chemistry transport model CHIMERE. *Geosci Model Dev* 8:21–42; doi:10.5194/gmd-8-21-2015.
142. Tombette M, Sportisse B. 2007. Aerosol modeling at a regional scale: Model-to-data comparison and sensitivity analysis over Greater Paris. *Atmos Environ* 41:6941–6950; doi:10.1016/j.atmosenv.2006.10.037.

143. UNEP. 2013. United Nations Environment Programme | Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs. Available: [http://toolkit.pops.int/Publish/Main/01\\_Index.html](http://toolkit.pops.int/Publish/Main/01_Index.html) [accessed 22 August 2016].
144. US EPA. 2010. Recommended Toxicity Equivalence Factors (TEFs) for Human Health Risk Assessments of 2,3,7,8-Tetrachlorodibenzo-p-dioxin and Dioxin-Like Compounds.
145. Vandenberg LN, Colborn T, Hayes TB, Heindel JJ, Jacobs DR, Lee D-H, et al. 2012. Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses. *Endocr Rev* 33:378–455; doi:10.1210/er.2011-1050.
146. Viel J-F, Clement M-C, Hagi M, Grandjean S, Challier B, Danzon A. 2008a. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int J Health Geogr* 7:4; doi:10.1186/1476-072X-7-4.
147. Vienneau D, de Hoogh K, Bechle MJ, Beelen R, van Donkelaar A, Martin RV, et al. 2013. Western European Land Use Regression Incorporating Satellite- and Ground-Based Measurements of NO<sub>2</sub> and PM<sub>10</sub>. *Environ Sci Technol* 47:13555–13564; doi:10.1021/es403089q.
148. Vienneau D, de Hoogh K, Briggs D. 2009. A GIS-based method for modelling air pollution exposures across Europe. *Sci Total Environ* 408:255–266; doi:10.1016/j.scitotenv.2009.09.048.
149. Viera AJ, Garrett JM. 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med* 37: 360–363.
150. Visser O, van Wijnen JH, van Leeuwen FE. 2004. Residential Traffic Density and Cancer Incidence in Amsterdam, 1989–1997. *Cancer Causes Control* 15:331–339; doi:10.1023/B:CACO.0000027480.32494.a3.
151. Wahida K-T, Padilla CM, Denis Z-N, Olivier B, Géraldine LN, Philippe Q, et al. 2016. A Conceptual Framework for the Assessment of Cumulative Exposure to Air Pollution at a Fine Spatial Scale. *Int J Environ Res Public Health* 13; doi:10.3390/ijerph13030319.
152. Warner M, Mocarelli P, Samuels S, Needham L, Brambilla P, Eskenazi B. 2011. Dioxin exposure and cancer risk in the Seveso Women’s Health Study. *Environ Health Perspect* 119:1700–1705; doi:10.1289/ehp.1103720.
153. White N, teWaterNaude J, van der Walt A, Ravenscroft G, Roberts W, Ehrlich R. 2009. Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study. *Environ Health* 8:45; doi:10.1186/1476-069X-8-45.
154. WHO. 1992. International Program on Chemical Safety.
155. WHO. 2013. Review of evidence on health aspects of air pollution – REVIHAAP project: final technical report.
156. WHO. 2000. WHO, Air quality guidelines for Europe, 2nd Edition, 2000.
157. Wilton D, Szpiro A, Gould T, Larson T. 2010. Improving spatial concentration estimates for nitrogen oxides using a hybrid meteorological dispersion/land use regression model in Los Angeles, CA and Seattle, WA. *Sci Total Environ* 408:1120–1130; doi:10.1016/j.scitotenv.2009.11.033.
158. Xu J, Ye Y, Huang F, Chen H, Wu H, Huang J, et al. 2016. Association between dioxin and cancer incidence and mortality: a meta-analysis. *Sci Rep* 6:38012; doi:10.1038/srep38012.
159. Yang C-Y, Chiou S-L, Wang J-D, Guo YL. 2015. Health related quality of life and polychlorinated biphenyls and dibenzofurans exposure: 30 years follow-up of Yucheng cohort. *Environ Res* 137:59–64; doi:10.1016/j.envres.2014.11.020.



160. Ying Z, Xu X, Bai Y, Zhong J, Chen M, Liang Y, et al. 2014. Long-term exposure to concentrated ambient PM<sub>2.5</sub> increases mouse blood pressure through abnormal activation of the sympathetic nervous system: a role for hypothalamic inflammation. *Environ Health Perspect* 122:79–86; doi:10.1289/ehp.1307151.
161. Yoshioka H, Hiromori Y, Aoki A, Kimura T, Fujii-Kuriyama Y, Nagase H, et al. 2012. Possible aryl hydrocarbon receptor-independent pathway of 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced antiproliferative response in human breast cancer cells. *Toxicol Lett* 211:257–265; doi:10.1016/j.toxlet.2012.04.005.
162. Yu C-L, Wang S-F, Pan P-C, Wu M-T, Ho C-K, Smith TJ, et al. 2006. Residential exposure to petrochemicals and the risk of leukemia: Using geographic information system tools to estimate individual-level residential exposure. *Am J Epidemiol* 164:200–207; doi:10.1093/aje/kwj182.
163. Zambon P, Ricci P, Bovo E, Casula A, Gattolin M, Fiore AR, et al. 2007. Sarcoma risk and dioxin emissions from incinerators and industrial plants: a population-based case-control study (Italy). *Environ Health* 6:19; doi:10.1186/1476-069X-6-19.
164. Zhang S, Lei P, Liu X, Li X, Walker K, Kotha L, et al. 2009. The aryl hydrocarbon receptor as a target for estrogen receptor-negative breast cancer chemotherapy. *Endocr Relat Cancer* 16:835–844; doi:10.1677/ERC-09-0054.
165. Zhang T, Fiedler H, Yu G, Solorzano Ochoa G, Carroll WF, Gullett BK, et al. 2011. Emissions of unintentional persistent organic pollutants from open burning of municipal solid waste from developing countries. *Chemosphere* 84:994–1001; doi:10.1016/j.chemosphere.2011.04.070.
166. Zou B, Wilson JG, Zhan FB, Zeng Y. 2009. Air pollution exposure assessment methods utilized in epidemiological studies. *J Environ Monit* 11:475–490; doi:10.1039/b813889c.

## Annexes

### Annexe 1. Revue de la littérature : liste des équations de recherche bibliographique utilisées pour identifier les paramètres d'intérêts dans la construction de l'indicateur SIG.

- Dioxines:

(dioxins[mesh terms] OR dioxin\*[tiab] OR tetrachlorodibenzodioxin[mesh terms] OR tetrachlorodibenzo p dioxin[tiab]OR TCDD[tiab] OR polychlorodibenzo-4-dioxin[tiab] OR PCDD[tiab] OR PCDF[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (geographic information systems[mesh terms] OR geographic information system[tiab] OR geograph\*[tiab] OR metric[tiab] OR (spatial\*[tiab] and (analys\*[tiab] OR analyt\*[tiab])) OR GIS\*[tiab] OR spatial analysis[mesh terms]) AND (plume\*[tiab] OR buffer\*[tiab] OR distance\*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor\*[tiab] OR meteorolog\*[tiab] OR ((residen\*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

- HAP, dioxin-like et POPs:

(polycyclic hydrocarbons,aromatic[mesh terms] OR polycyclic aromatic hydrocarbon[tiab] OR polychlorinated biphenyls[mesh terms] OR polychlorinated biphenyls[tiab] OR PCB\*[tiab] OR PAH\*[tiab] OR dioxin-like[tiab] OR persistent organic pollutant\*[tiab] OR POP[tiab] OR POPs[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (geographic information systems[mesh terms] OR geographic information system[tiab] OR geograph\*[tiab] OR metric[tiab] OR (spatial\*[tiab] and (analys\*[tiab] OR analyt\*[tiab])) OR GIS\*[tiab] OR spatial analysis[mesh terms]) AND (plume\*[tiab] OR buffer\*[tiab] OR distance\*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor\*[tiab] OR meteorolog\*[tiab] OR ((residen\*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

- Pollution de l'air :

((air pollut\*[mesh terms] OR air pollut\*[tiab]) OR (environmental pollut\*[mesh terms] OR environmental pollut\*[tiab]) OR (carbon dioxide[mesh terms] OR carbon dioxide[tiab] OR CO2[tiab]) OR (nitrogen oxides[mesh terms] OR nitrogen oxide\*[tiab]) OR (particulate matter[mesh terms] OR particulate matter\*[tiab] OR PM[tiab]) OR industrial waste[mesh terms] OR emission\*[tiab] OR hazardous waste[mesh terms] OR hazardous waste[tiab] OR industry[mesh terms] OR indust\*[tiab] OR incineration[mesh terms] OR incinerat\*[tiab] OR combustion[tiab] OR facilit\*[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (geographic information systems[mesh terms] OR geographic information system[tiab] OR geograph\*[tiab] OR metric[tiab] OR (spatial\*[tiab] and (analys\*[tiab] OR analyt\*[tiab])) OR GIS\*[tiab] OR spatial analysis[mesh terms]) AND (plume\*[tiab] OR buffer\*[tiab] OR distance\*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor\*[tiab] OR meteorolog\*[tiab] OR ((residen\*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

- Algorithmes utilisés pour identifier les études ayant modélisé la dispersion atmosphérique :

(dioxins[mesh terms] OR dioxin\*[tiab] OR tetrachlorodibenzodioxin[mesh terms] OR tetrachlorodibenzo p dioxin[tiab]OR TCDD[tiab] OR polychlorodibenzo-4-dioxin[tiab] OR PCDD[tiab] OR PCDF[tiab]) AND (environmental exposure[mesh terms] OR environmental exposure[tiab] OR air exposure[tiab] OR exposure assessment[tiab] OR exposure[tiab]) AND (((dispersion[tiab] AND model\*[tiab]) OR dispersion model\*[tiab]) OR (normal distribution[mesh terms] OR gaussian[tiab] OR gaussian model\*[tiab]) OR (ADMS[tiab] OR atmospheric dispersion modeling system[tiab]) OR (lagrangian[tiab] OR lagrangian model\*[tiab]) OR air movements[mesh terms]) AND (plume\*[tiab] OR buffer\*[tiab] OR distance\*[tiab] OR wind[mesh terms] OR wind[tiab] OR wind direction[tiab] OR topography[tiab] OR stack[tiab] OR climatic factor\*[tiab] OR meteorolog\*[tiab] OR ((residen\*[tiab] AND proximity[tiab]) OR vicinity[tiab] OR surrounding[tiab] OR living area[tiab]))

## Annexe 2. Présentation des 7 articles sélectionnés dans la revue de littérature

Tableau 5 - Présentation des 7 articles sélectionnés dans le cadre de la revue de la littérature réalisée pour définir la métrique SIG

Auteur et année de publication	Titre	Objectifs	Méthode et paramètres utilisés
Pronk, 2013	Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma (NHL): a case-control study	Estimer le risque de lymphome non Hodgkinien liée à la proximité résidentielle avec des installations de combustion	Variation de l'exposition en fonction de la distance sujet-source : $1/d^2$ ; histoire résidentielles (connu sur 15 ans); Pas de données météorologique
Gulliver and Briggs, 2011	STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment	Décrire les méthodes utilisées pour élaborer le modèle STEMS-Air Présenter les résultats d'une étude de validation à l'aide des données de surveillance Faire des observations sur les améliorations potentielles du modèle.	Modèle de dispersion gaussien simplifié : prise en compte de la vitesse et de la direction du vent (moyenne par $45^\circ$ ); considération d'une exposition uniquement pour les sujets sous le vent (fonction FOCALSUM dans ArcGIS); utilisation de données techniques (hauteur de cheminée)
Vienneau, 2009	A GIS-based method for modeling air pollution exposures across Europe	Décrire et tester une technique de « fenêtre mobile » pour établir un modèle pondéré entre des émissions et des concentrations mesurées Appliquer cette technique pour estimer les concentrations dans des sites de l'UE.	Direction du vent : angle de $45^\circ$ ; Variation de l'exposition en fonction de la distance sujet-source : $1/d^2$ puis calibrer avec des mesures; distance maximum d'exposition : 11 km; Vitesse du vent en 3 catégories (<1 m/s; 1-2.5 m/s; >2.5 m/s)
White, 2009	Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study	Déterminer s'il y avait une association entre la prévalence des symptômes de l'asthme et l'estimation de l'exposition à des émissions industrielles.	Direction et vitesse du vent (données de la station météorologique la plus proche de la source); géocodage manuel des adresses de résidence; SIG pour calculer la distance entre l'adresse du domicile et la source.
Yu, 2006	Residential exposure to petrochemicals and the risk of leukemia: using geographic information system tools to estimate individual-level residential exposure	Étudier le lien entre l'exposition aux produits pétrochimiques résidentiels et le risque de leucémie chez les sujets de 29 ans et moins.	Données sur l'histoire résidentielle; géocodage des adresses et des sources; zone d'exposition maximum: 3 km de rayon; Prise en compte de la direction du vent (zone de $45^\circ$ degrés; moyenne mensuelle et uniquement pour les zones sous le vent); Variation de l'exposition en fonction de la distance : $1/d$

Auteur et année de publication	Titre	Objectifs	Méthode et paramètres utilisés
Zou, 2009	Air pollution exposure assessment methods utilized in epidemiological studies (review)	<p>Présenter les méthodes, les concepts et les critères utilisés pour l'évaluation de l'exposition à la pollution atmosphérique;</p> <p>Examiner les différents modèles d'exposition (développements, applications, mérites et faiblesses)</p> <p>Offrir des perspectives sur le développement futur des méthodes d'évaluation de l'exposition</p>	<p>Modèles de proximité (distance par rapport à la source; densité du trafic pondérée par la distance; système d'information géographique, utilisation des données météorologiques et topographiques);</p> <p>Modèles de dispersion atmosphérique ;</p> <p>Modèles hybrides (mesures à l'échelle individuelle et régionale ; mesures individuelles + modélisation de dispersion atmosphérique; maillage modèle régional avec modèles dispersion échelle locale);</p> <p>Modèles d'inhalation</p> <p>Biomarqueurs</p>
Hoek, 2001	Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality	<p>Étudier la relation entre la pollution atmosphérique liée au trafic et la mortalité</p> <p>Élaborer des méthodes d'évaluation de l'exposition et évaluer le contraste de l'exposition à la pollution atmosphérique dans la cohorte au sein d'une sous-cohorte aléatoire de 5 000 sujets.</p>	<p>Exposition définit comme la somme des contributions régionales, urbaines et locales</p> <p>Régionale: test de plusieurs fonctions décroissantes pour distance source-sujet avec calibration sur des mesures proches des sujets;</p> <p>Urbaine: densité d'adresse utilisée comme indicateur de l'intensité des activités humaines;</p> <p>Locale: exposition fonction de la distance aux différents types de routes (&lt;50 m d'une route urbaine = exposée; &gt;50 m = non exposée)</p>

**Annexe 3. Danjou et al. 2018. *En cours de finalisation*. 2018**

# Author's draft

## **Airborne Dioxin Exposure and Breast Cancer Risk in a case-control study nested within the French E3N Prospective Cohort**

**Aurélie Marcelle Nicole Danjou** <sup>a,b</sup>, Thomas Coudon <sup>a,b</sup>, Delphine Praud <sup>a,c,d</sup>, Emilie Lévêque <sup>e</sup>, Elodie Faure <sup>a</sup>, Pietro Salizzoni <sup>f</sup>, Muriel Le Romancer <sup>b,c,d</sup>, Gianluca Severi <sup>g</sup>, Francesca Romana Mancini <sup>g\*</sup>, Karen Leffondré <sup>e</sup>, Laure Dossus <sup>a,g,h</sup> and Béatrice Fervers <sup>a,b,c,d\*</sup>

### **Authors' affiliations**

<sup>a</sup> Département Cancer et Environnement, Centre Léon Bérard, Lyon, France

<sup>b</sup> Université de Lyon, Université Claude Bernard Lyon 1, Villeurbanne, France

<sup>c</sup> Inserm U1052, Centre de Recherche en Cancérologie de Lyon, Lyon, France

<sup>d</sup> CNRS UMR5286, Centre de Recherche en Cancérologie de Lyon, Lyon, France

<sup>e</sup> Université de Bordeaux, Institut de Santé Publique, d'Épidémiologie et de Développement, Centre Inserm U1219 Epidemiology and Biostatistics, Bordeaux, France

<sup>f</sup> Laboratoire de Mécanique des Fluides et d'Acoustique, UMR CNRS 5509, Université de Lyon, Ecole Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon 1, Ecully, France

<sup>g</sup> Centre de Recherche en Epidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS UVSQ, Gustave Roussy, Villejuif, France

<sup>h</sup> Nutrition and Metabolism Section, International Agency for Research on Cancer, Lyon, France

\* Address correspondence to:

Béatrice Fervers, Département Cancer et Environnement, Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon Cedex 08, France. Phone: +33(0)478782801. E-mail: [beatrice.fervers@lyon.unicancer.fr](mailto:beatrice.fervers@lyon.unicancer.fr)

Francesca R Mancini, Centre de Recherche en Epidémiologie et Santé des Populations (CESP, Inserm U1018), Facultés de Médecine, Université Paris-Saclay, UPS UVSQ, Gustave Roussy, Villejuif, France. Phone: +33(0)142115864. E-mail: [francesca.mancini@gustaveroussy.fr](mailto:francesca.mancini@gustaveroussy.fr)

### **Short title**

Airborne dioxin exposure and breast cancer risk

# Author's draft

## **Acknowledgments**

We gratefully acknowledge the participants for providing data and the practitioners for providing pathology reports. We thank Guillaume Harel for his work on the inventory and characterization of the industrial sources located in the Rhône-Alpes region. We thank the scientific committee of the project for its advice on the exposure assessment.

This project was supported by funding from the Oncostarter Program of the *Cancéropôle Lyon Auvergne Rhône-Alpes*, the French Environment and Energy Management Agency (Grant No 1306C0031), the Regional Committee of the French League against Cancer of the Saône et Loire Region (mrt-za262) and the Rhône-Alpes Region, and was carried out in partnership with the ARC Foundations for Cancer Research. The E3N cohort is financially supported by the French League Against Cancer, the *Mutuelle Générale de l'Éducation Nationale*, the *Institut Gustave Roussy*, the *Institut National de la Santé et de la Recherche Médicale*. Aurélie Danjou and Thomas Coudon were supported by an academic doctoral fellowship from the *Ecole Doctorale Interdisciplinaire Sciences-Santé, Université de Lyon*. Aurélie Danjou received a mobility grant from the *Cancéropôle Lyon Auvergne Rhône-Alpes*. Delphine Praud is supported by a post-doctoral fellowship from the National French Cancer League.

## **Competing financial interest declaration**

The authors declare that they have no competing interests.

**Manuscript in preparation, to be submitted to Environmental Health Perspectives**

# Author's draft

## Abstract

### Background:

Dioxins are emitted by industrial chlorinated combustion processes and its most potent congener, TCDD (2,3,7,8-tetrachlorodibenzo-p-dioxin), has been classified as a Group 1 carcinogen by the International Agency for Research on Cancer. TCDD is suspected to be involved in breast cancer etiology and may influence hormone-mediated pathways.

### Objectives:

We aimed to estimate breast cancer risk associated with airborne dioxin exposure, using geographic information system (GIS) methods and historical exposure data.

### Methods:

We designed a case-control study nested within the E3N (Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale) cohort and restricted to the Rhône-Alpes region, France. Between 1990 and 2008, 429 breast cancer cases were diagnosed and matched to 716 controls on relevant factors. Airborne dioxin exposure was assessed based on a detailed inventory of dioxin sources and residential history of the participants, using a GIS-based metric that included proximity to dioxin sources, their technical characteristics, exposure duration and wind direction. We first estimated odds ratios (OR) and 95% confidence intervals (CI) associated with quintiles of cumulative exposure using logistic regression. We then estimated effects of annual airborne dioxin exposure on breast cancer risk according to time prior to diagnosis using a flexible time-dependent weight function.

### Results:

We observed no increased risk of breast cancer for higher dioxin exposure levels (OR for Q5 *versus* Q1: 1.12, 95% CI: 0.69-1.82) and no linear trend across quintiles of exposure ( $P$  for trend=0.81). We however observed a statistically significant OR for Q2 *versus* Q1 (1.61, 95% CI: 1.04-2.49). ORs associated with an increase of 0.1  $\mu\text{g-TEQ}/\text{m}^2$  in the annual intensity of exposure in each year before diagnosis were all not statistically different from one, whatever the timing of exposure. No association was observed for breast cancer risk according to status of estrogen and progesterone-receptors.

### Conclusions:

No association was observed between airborne dioxin exposure and overall breast cancer risk in our study population, as well as for hormone-receptor breast cancers. Confirmation of our findings is required in larger populations. Our study may provide a new tool for the assessment of long-term exposure with GIS-based metric as well as new ways of considering temporal dimensions for environmental exposures in disease risk assessment.



## 1. Introduction

Dioxins are a mixture of related chemicals, namely polychlorinated dibenzo-*para*-dioxins (PCDD) and polychlorinated dibenzofurans (PCDF), generated from combustion processes involving chlorine-bearing materials. The main sources of dioxin release into the environment include chemical manufacturing of chlorinated products such as herbicides and insecticides and industrial activities from metallurgy, steel and municipal solid waste incineration (MSWI) (Anzivino-Viricel et al., 2012; ATSDR, 1998; Travis and Hattemer-Frey, 1989; Zook and Rappe, 1994). Emissions of dioxins have greatly decreased since the 1990s, because of changes in regulation and technological processes (Anzivino-Viricel et al., 2012; Coudon et al., 2017; Nzihou et al., 2012). In France, emissions have been reduced by a factor 1000 between 1990 and 2008 (CITEPA, 2015). As dioxins are persistent in the environment and bioaccumulate in the liver and adipose tissues, most of human exposure to dioxins occurs from consumption of contaminated fat-rich food (IARC, 2012).

TCDD (2,3,7,8-tetrachlorodibenzo-*para*-dioxin) is the most potent dioxin congener and has been classified as carcinogenic to humans (Group 1) by the International Agency for Research on Cancer (IARC), with sufficient evidence for all cancers combined (Baan et al., 2009). Experimental evidence showed that dioxins have no genotoxic activity, but rather act as tumor promoter through activation of cellular replication, interruption of apoptosis and increase in oxidative stress (IARC, 2012; Knerr and Schrenk, 2006; Mandal, 2005). As a persistent endocrine disruptor, TCDD may contribute to breast cancer by binding to the aryl hydrocarbon receptor (AhR), a nuclear transcription factor widely present in humans and animals that influences gene expression, enzyme metabolism and hormone signaling pathway, including estrogen- and progesterone-mediated pathways (Birnbaum and Fenton, 2002; Boverhof, 2006; Jenkins et al., 2012; Matthews and Gustafsson, 2006; Tsuchiya et al., 2005).

An increased risk of breast cancer associated with dioxin exposure has been suggested, but overall, epidemiological studies did not provide a consistent link between dioxin exposure and risk of breast cancer in women (Rodgers et al., 2018; Xu et al., 2016). Following the 1976 industrial accident in Seveso, Italy, results from the cohort of local residents in the contaminated zones showed no increase in breast cancer risk (Bertazzi et al., 1993; Pesatori et al., 2009). In addition, in the Seveso Women's Health Study (SWHS), no association was observed between TCDD serum levels measured in blood samples provided after the accident and breast cancer risk among women who were 0-40 years old at the time of the accident and lived in the most contaminated zones (Warner et al., 2011). Positive associations were found in cohorts of herbicide and insecticide workers (Manuwald et al., 2012), among Russian women living near a chemical plant, compared with women from a non-contaminated area (Revich et al., 2001), and among women from a French ecological study of residents exposed to dioxin emissions from several MSWIs (Fabre et al., 2008). On the contrary, a statistically significant decrease in breast cancer risk was found among women over 60 years of age

# Author's draft

and living in the most exposed zone around a French MSWI (Viel et al., 2008). The most recent meta-analysis, involving ten studies conducted on the association between external TCDD exposure and cancer incidence, reported no statistically significant increased risk for breast cancer (Xu et al., 2016).

The inconsistency in published results may be explained by differences in study populations and methodology (*e.g.* study design and exposure assessment) as well as some specific study limitations including the lack of adjustment for confounding by established risk factors, inadequate exposure assessment and small number of cases (Cordioli et al., 2013; Rodgers et al., 2018; Xu et al., 2016). The long-term nature of airborne dioxin exposure may imply variations in exposure intensities over time. Given their tumor-promoter properties, studying the association between airborne dioxin exposure and breast cancer risk should consider the impact of time-varying and exposures close to the time of diagnosis as well as hormone receptor subtypes of breast tumors (Rodgers et al., 2018).

Recently, we investigated the relationship between estimated dietary dioxin exposure and breast cancer risk in the French E3N (Étude Épidémiologique auprès des femmes de la Mutuelle Générale de l'Éducation Nationale) prospective cohort. No association was observed overall, except for a statistically significant decrease in postmenopausal hormone receptor-negative breast cancer risk among women most exposed (Danjou et al., 2015). Here, we estimated breast cancer risk associated with airborne dioxin exposure in a case-control study nested within the E3N cohort and restricted to the Rhône-Alpes region, France, using geographic information system (GIS) methods and historical exposure data for the assessment of airborne dioxin exposure and considering temporal dimensions of exposure in risk estimates.

## 2. Material and methods

### 2.1. The E3N cohort study

The E3N study is an ongoing prospective cohort, initiated to identify female cancer risk factors, such as dietary, hormonal and reproductive factors (Clavel-Chapelon, 2015). The cohort includes 98,995 French female volunteers, born between 1925 and 1950, subscribers of the national health insurance covering mostly teachers. The study protocol was approved by the French National Commission for Data Protection and Privacy. The study was initiated in 1989-1991 when participants returned the first self-administered questionnaire and provided written informed consent for data collection. Since then, participants completed self-administered questionnaires, mailed every 2 to 3 years, on health status, medical history and a large number of lifestyle factors. Overall, nine questionnaires were sent between 1990 and 2008. Epidemiological data were enriched with 25,000 blood samples and 47,000 saliva samples. Participants' place of birth (zip code and commune) was obtained from the first questionnaire and an urban/rural status was assigned; definition of the urban/rural status has been previously described (Binachon et al., 2014). The E3N cohort is the French

# Author's draft

component of the European Prospective Investigation into Cancer and nutrition (EPIC) study (Riboli, 1992; Riboli et al., 2002).

## 2.2. Study population

For this nested case-control study, we selected participants who filled in their home address at baseline, lived in the Rhône-Alpes region, France between 1990 and 2008 and had not reported cancer at baseline. The study population was selected from the Rhône-Alpes region because of its high population density (147.8 inhabitants/km<sup>2</sup> in 2014 (Insee, 2017)), its large area (43,700 km<sup>2</sup>), and its dense industrial network (2<sup>nd</sup> French industrial region, (Insee, 2012)). Data on health status and medical history of participants, including cancer occurrence, were collected and updated throughout the follow-up. Between 1990 and 2008, 429 primary invasive breast cancer cases were reported, of which 93% have been validated with pathology reports; not confirmed cases were considered as breast cancer cases, as the proportion of false-positive self-reported breast cancer cases is less than 5% in the E3N cohort. Information on estrogen receptor (ER) and progesterone receptor (PR) status was obtained from the pathology reports: n=244 tumors were ER positive (ER+); n=80 were ER-negative (ER-); n=215 were PR positive (PR+); and n=108 were PR-negative (PR-).

Using incidence density sampling, up to two controls per case were randomly selected and matched for age ( $\pm$  1 year), department of residence, date ( $\pm$  3 months) and menopausal status at blood collection for the participants who donated blood, or at baseline for the participants without a blood sample who were then matched for existence or not of a saliva sample. A total of 716 controls were selected.

## 2.3. Assessment of airborne dioxin exposure

The methodology for the assessment of airborne dioxin exposure has been previously described in detail (Coudon et al., 2017, *Submitted*); a brief description is given below.

### 2.3.1. Inventory and characterization of industrial dioxin sources

A detailed retrospective inventory of industrial sources likely to emit or to have emitted dioxins between 1990 and 2008 in the Rhône-Alpes region was performed. The industrial sources were identified through national databases, institutional information sources and structured questionnaires. Information on technical and process characteristics (*e.g.* stack height, exhaust flow rate and flue gas cleaning technologies), operating periods and rates, input materials and geographic location of the facilities were collected (Coudon et al., *Submitted*). We also used data from an existing

# Author's draft

inventory of MSWIs that operated in the Rhône-Alpes region between 2001 and 2003, and for which technical characteristics and dioxin emissions have been collected (Cordier et al., 2004).

A total of 222 industrial dioxin sources operated in the Rhône-Alpes region from 1990 to 2008, corresponding to 286 distinct operation periods according to the evolution of technical characteristics. The predominant sector was waste incineration (N=119 distinct operation periods-sources), followed by production of mineral (N=87), heat and power generation (N=40), metal production (N=26), crematoria (N=13) and chemicals and consumer goods (N=1).

Using the 2013 Standardized Toolkit for Identification and Quantification of Dioxin and Furan Releases of the United Nation Environment Program (UNEP, 2013), the industrial sources were classified according to their main activity sector and technical characteristics to determine a dioxin emission factor (g-toxic equivalents (TEQ)/t). For each distinct operation period, the annual dioxin emission intensity (g-TEQ/year) was estimated by multiplying the emission factor by the operating rate (Coudon et al., *Submitted*).

## 2.3.2. Geocoding of the residential history and industrial sources

From 1990 to 2008, the residential history of the participants was extracted from the E3N follow-up questionnaires and geocoded (X and Y coordinates, addresses) using the ArcGIS Software (ArcGIS Locator version 10.0, Environmental System Research Institute – ESRI, Redlands, CA, USA) and the address database, BD Adresse®, from the National Geographic Institute (IGN) (Faure et al., 2017). Geocoding was performed by a trained technician blinded to the case-control status of the participants. In addition, each industrial source inventoried was located in the GIS, based on collected geographic coordinates when available or addresses, and manually checked or repositioned at the location of the flue-gas stack (Coudon et al., *Submitted*).

## 2.3.3. Exposure assessment

Airborne dioxin exposure was estimated for each participant at the individual address level using a GIS-based metric. Relevant parameters to be integrated in the GIS-based metric were identified from the literature (Gulliver and Briggs, 2011; Hoek et al., 2001; Pronk et al., 2013; Vienneau et al., 2009; White et al., 2009; Yu, 2006) and from a previous work on dioxin dispersion modeling (Coudon et al., 2017).

We considered that each industrial source had an impact limited to a buffer zone, *i.e.* a circle of radius  $d$ . The buffer zone was subsequently divided azimuthally into sectors of equal size, each of them characterized by a percentage of wind direction frequency (evaluated by means of hourly data registered at the nearest *Météo France*® station, over the 1990-2008 period). The exposure of a participant to a given industrial source depended then on its positioning, within the buffer zone, in one

# Author's draft

of these sectors. In order to take into account the effects of atmospheric turbulence on the dispersion of dioxin emission over a wide area downwind the industrial source, we included two weighting factors: a higher one for the sector in which the participant is located in, and a lower one for the two adjacent sectors.

In order to verify the reliability of the GIS-based metric, we compared its exposure classifications, obtained with different values of the above mentioned parameters, with those evaluated by means of the atmospheric dispersion model, SIRANE (Soulhac et al., 2011, 2012, 2017). These comparisons were performed for three distinct years (1996, 2002 and 2008) and in three areas with typical weather and topographical conditions, time-varying dioxin emission intensities and distinct number of participants and types of industrial sources: Lyon, a highly urban area; Le Bugey, a rural area; and Le Havre, a coastal medium size urban area (Coudon et al., *Submitted*).

We tested the values of the GIS-metric parameters: a radius  $d$  equal to 3, 5 or 10 km for the buffer zone, a residence-to-source distance decline pattern of different forms ( $d^{-1}$ ;  $d^{-1.5}$ ;  $d^{-2}$ ;  $d^{-3}$ ;  $e^{-d}$ ;  $e^{-d/d}$  and  $e^{-d*d}$ ), different azimuthal amplitudes of the sectors of the buffer zone ( $90^\circ$ ;  $30^\circ$  and  $10^\circ$ ) and different weighting factors for the adjacent sectors. The tested parameters were combined with the Toolkit-based annual dioxin emission intensity and the exposure duration.

The agreement between classifications into quintiles of the dispersion modeling results and the GIS-based metric was quantified through weighted kappa coefficients and their 95% confidence intervals (CI). The highest weighted kappa coefficients were obtained imposing a buffer zone of  $d$  equal to 10 km radius, an inverse of the squared residence-to-source distance decline, *i.e.*  $d^{-2}$ , a wind rose divided into  $10^\circ$  sectors and attributing a weighting factor of 50% to the sector in which the participant is located in and a weighting factor of 25% to the two adjacent sectors. In the final GIS-based metric, the airborne dioxin exposure of each participant was computed as:

$$\text{GIS – based metric} = \sum_{j=1}^J \sum_{i=1}^I (EI_i \times t_i \times d_{ij}^{-2} \times F_{ij})$$

where the indices  $j$  and  $i$  indicate the places of residence ( $j=1, \dots, J$ ) and the industrial sources ( $i=1, \dots, I$ ), respectively, and where  $EI_i$  is the annual dioxin emission intensity of each industrial source (in g-TEQ/year),  $t_i$  is its emission period duration (in years),  $d_{ij}$  is the residence-to-source distance (in m) and  $F_{ij}$  is the factor taking into account the percentage of time during which the wind is blowing in a direction so as to induce a transport from the industrial source  $i$  to the location of the participant  $j$  (accounting for the weighted contribution of the sector of the buffer in which the participant is located in, *i.e.* 50% and that of the two adjacent sectors, *i.e.* 25% each). Airborne dioxin exposure was expressed in  $\mu\text{g-TEQ}/\text{m}^2$  and computed for each calendar year 1990 to 2008. Weighted kappa coefficients ranged from 0.71 (0.67-0.76) to 0.84 (0.79-0.88) and were consistent across periods and areas (rural/ urban/ coastal).

## 2.4. Statistical analyses

Airborne dioxin exposure was summarized into a cumulative exposure metric from baseline to the index date (date of diagnosis of the case in the case-control pair) by summing up the annual metric estimates. The cumulative metric of airborne dioxin exposure was *a priori* categorized according to quintiles of the control distribution.

Baseline characteristics of the participants were described according to quintiles of the cumulative metric of airborne dioxin exposure and their case-control status using mean and SD for continuous covariates and frequency and percentage for categorical covariates. Baseline characteristics were compared between cases and controls with univariate conditional logistic regression models, and across quintiles of cumulative airborne dioxin exposure with Chi-square statistical test and analysis of variance.

Odds ratios (OR) and corresponding 95% CIs for invasive breast cancer were estimated for quintiles of the cumulative exposure metric with the first quintile as the reference group, using multivariate conditional logistic regression models. Tests for linear trend across quintiles were performed and derived from the Wald test of the models with the continuous variable. All models were conditioned for matching factors (age, date, department of residence and menopausal status at blood collection or at baseline and existence of a biological sample (blood, saliva, none)). Models included terms for individual breast cancer risk factors selected through backward stepwise selection; covariates were retained if they modified the risk estimate by more than 10%: recreational physical activity at baseline (quartiles: <25.3, 25.3-37.2, 37.3-57.5,  $\geq$ 57.6 metabolic equivalent task-hour per week (MET-h/w)); previous family history of breast cancer (no, yes); age at first full-term pregnancy ( $\leq$ 24, >24 years old; median value among controls); breastfeeding (never, ever) and status of birthplace (rural, urban). For covariates with less than 5% missing data, the latter were replaced by their modal or median value in the control population. Age at first full-term pregnancy had more than 5% of missing data, thus a category "missing data" was generated.

We estimated invasive breast cancer risk according to hormone-receptor status (ER and PR) of the breast tumors, excluding cases for whom the status was missing and their matched controls from the corresponding analyses (N=105 cases for ER; N=106 cases for PR). Heterogeneity in subgroups (ER+/ER-; PR+/PR-) was assessed using the Q-statistic. In addition, models were further adjusted for estimated dietary dioxin intake, which was assessed in the E3N cohort in a previous study (Danjou et al., 2015). To investigate whether the effect of cumulative metric of airborne dioxin exposure was homogeneous across strata of selected covariates, we conducted analyses stratified by menopausal status at baseline (pre-menopause, post-menopause), body mass index (BMI) at baseline ( $\leq$ 21.9, >21.9 kg/m<sup>2</sup>; median value among controls), smoking status at baseline (never, ever), breastfeeding (never, ever), age at first full-term pregnancy ( $\leq$ 24, >24 years old; median value among controls), parity (<3,  $\geq$ 3 children), age at menarche (<13,  $\geq$ 13 years old; modal value among controls) and status of the birthplace (rural, urban). Association between invasive breast cancer risk and quintiles of the

# Author's draft

cumulative metric of airborne dioxin exposure was examined in the above-defined subgroups using unconditional logistic regression models adjusted for matching factors in order to retain all subjects in the models. Heterogeneity across strata was based on likelihood ratio tests comparing the models including and excluding interaction terms. In addition, the interaction between BMI and menopausal status at baseline was tested. We performed a first sensitivity analysis by separating the participants that have been living more than 10 km away from any source between 1990 and 2008 (*i.e.* for which cumulative airborne dioxin exposure was estimated to be equal to 0 over the study period by the cumulative metric and presented differences for several baseline characteristics) and categorizing the other participants into quintiles of cumulative airborne dioxin exposure (first quintile as reference). In a second sensitivity analysis, we excluded cases diagnosed within the first five years after inclusion in the cohort (from 1990 to 1995, N=382, 33%) in order to exclude the impact of short-term airborne dioxin exposure and take into account left truncation of the cumulative exposure on ORs for breast cancer (Hazelbag et al., 2015).

Furthermore, we investigated the possible time-dependent impact of the annual intensity of airborne dioxin exposure on breast cancer risk, using a weighted cumulative index of exposure in logistic regression. For each subject  $i$ , the weighted cumulative index was equal to  $\sum_t w(t)x_i(t)$ , where  $x_i(t)$  was the annual intensity of exposure received for the subject  $i$  in the year  $t$  before the index date ( $t=1, \dots, 19$  maximum). The weights were estimated from the data using a cubic B-spline function with one to three interior knots placed at equal distance (Hauptmann et al., 2000; Lacourt et al., 2017). We chose the number of knots that minimize the Akaike's information criteria (AIC). The estimated weight function  $w(t)$  and the estimated regression coefficient associated with the cumulative weighted index of exposure were estimated simultaneously using the SAS Proc NLP (SAS). These estimates allowed us to calculate the OR associated with the annual intensity of exposure at each year of exposure before the index date. Nonparametric bootstrap sampling was performed to obtain 95% confidence interval for the estimated weight function. All unconditional logistic regression models included matching factors and covariates. In a sensitivity analysis, models were also run in a sub-population excluding cases and matched controls with airborne dioxin exposure that lasted less than 5 years from inclusion.

All  $P$ -values were two-sided and the nominal level of statistical significance was set at 0.05. We used the SAS statistical software version 9.4 (SAS Institute Inc., Cary, NC, USA) and the R software version 3.2.0 (R Foundation for Statistical Computing, Vienna, Austria) for data analysis.

## 3. Results

The baseline characteristics of the 429 cases of invasive breast cancer and the 716 matched controls are shown in Table 1. Cases reported less recreational physical activity than controls; they were more educated, younger at age at menarche and older at age at first full-term pregnancy. Performing a mammogram, family history of breast cancer and personal history of benign breast disease were more frequent among cases than controls. There was no difference in estimated dietary dioxin intake between breast cancer cases and matched controls ( $p=0.07$ ; Table 1). The average cumulative metric of airborne dioxin exposure was estimated at  $0.14 \pm 1.20$   $\mu\text{g-TEQ}/\text{m}^2$  (range: 0 to 24.43) for breast cancer cases and  $0.12 \pm 0.61$   $\mu\text{g-TEQ}/\text{m}^2$  (range: 0 to 10.35) for controls. The average exposure duration between recruitment and diagnosis among cases was  $8.7 \pm 4.9$  years (range: 1 month to 18 years) and the mean age at breast cancer diagnosis was  $58.0 \pm 7.8$  years (range: 41.9 to 78.2). A total of 105 participants (9.2%) had been living more than 10 km away from any dioxin source between 1990 and 2008, having therefore a cumulative airborne dioxin exposure of 0 according to our formula (Table 2). Compared with participants highly exposed to airborne dioxins (in the fifth quintile), those non-exposed participants were more frequently never smoker at baseline, born and living at baseline in a rural area, and less educated.

*Table 1 and Table 2 should appear here.*

Table 3 shows the distribution of breast cancer cases and matched controls and the odds ratios according to quintiles of the cumulative metric of airborne dioxin exposure, overall and for hormone-receptor status. We observed no increased risk of overall breast cancer for higher cumulative airborne dioxin exposure levels (OR for Q5 *versus* Q1: 1.12, 95% CI: 0.69-1.82) and no linear trend across quintiles of exposure ( $P$  for trend=0.81). We however observed a modest, statistically significant, increase in breast cancer risk for Q2 *versus* Q1 (OR: 1.61, 95% CI: 1.04-2.49). We found no evidence for heterogeneity in the associations for ER-positive *versus* ER-negative and PR-positive *versus* PR-negative tumors ( $P_{\text{het}}=0.42$  and  $P_{\text{het}}=0.41$ , respectively). None of the estimated ORs for specific breast cancer subtypes defined according to ER and PR status were statistically significant. Further adjustment for estimated dietary dioxin intake did not materially modify the odds ratios (data not shown).

*Table 3 should appear here.*

Odds ratios for breast cancer according to quintiles of cumulative metric of airborne dioxin exposure in strata of selected covariates are presented in Table 4. Although no statistically significant effect modification by age at first full-term pregnancy was found ( $P$  for interaction=0.09), we observed a statistically significant decrease in breast cancer risk for the highest *versus* the lowest quintile of cumulative metric of airborne dioxin exposure with an OR of 0.40 (95% CI: 0.20-0.81;  $P$  for trend=0.03) among women aged 24 years old or less at first full-term pregnancy. No association was found between cumulative metric of airborne dioxin exposure and breast cancer across strata of



# Author's draft

menopausal status, breastfeeding, BMI, age at menarche, status of birthplace, smoking status and parity. There was no effect modification by menopausal status and BMI on the association between cumulative metric of airborne dioxin exposure and breast cancer risk ( $P$  for interaction=0.39). When separating participants with a null estimated cumulative airborne dioxin exposure from the other participants categorized into quintiles, no association with overall breast cancer risk and no linear trend across categories were found ( $P$  for trend=0.61), and the association previously observed for Q2 *versus* Q1 did not remain significant (OR: 0.86, 95% CI: 0.54-1.37) (data not shown). When excluding cases with breast cancer diagnosis between 1990 and 1995 and matched controls, OR for overall breast cancer remained statistically significant for Q2 *versus* Q1 (OR: 2.06, 95% CI: 1.14-3.71;  $P$  for trend=0.51), however, no association and no linear trend were further observed in the subgroup of participants aged 24 years or less at first full-term pregnancy (OR for Q5 *versus* Q1: 0.44, 95% CI: 0.18-1.06;  $P$  for trend=0.15) (data not shown).

*Table 4 should appear here.*

The best model including the weighted cumulative index of exposure was obtained with one interior knot for the weight function  $w(t)$  (AIC=1,511.5). Table 5 shows the resulting estimated ORs associated with an increase of 0.1  $\mu\text{g-TEQ}/\text{m}^2$  (which corresponds to one SD of the annual doses) in the intensity of airborne dioxin exposure in different years before the index date (5, 10, 15, 18 years), overall and in the subgroup excluding case-controls pairs with diagnosis between 1990-1995. When using all subjects, all ORs were close to one, suggesting no effect of exposure intensity, whatever the timing of exposure. In the subgroup excluding short exposure duration, recent doses of airborne dioxin exposure received before diagnosis had a statistically significant impact on breast cancer risk. For example, an increase of 0.1  $\mu\text{g-TEQ}/\text{m}^2$  in the dose of airborne dioxin exposure received in the 5th year preceding diagnosis was associated with an OR of 1.06 (95% CI: 1.01-1.16), adjusted for all exposure doses estimated for the other years included in the study. The impact of the doses of airborne dioxin exposure on breast cancer risk seemed to decrease according to time prior to diagnosis: compared to the OR at year 5, the OR at year 10 was 1.03 (95% CI: 1.00-1.07) and the OR at year 15 was 1.01 (1.00-1.05).

*Table 5 should appear here.*

## 4. Discussion

Among women from the French E3N cohort, no increased risk of breast cancer for higher dioxin exposure levels was observed, as well as considering hormone-receptor status of breast cancers separately. Our results suggested an increased risk of breast cancer for the second *versus* the first quintile of exposure. A suggestive decrease in breast cancer risk associated with higher dioxin exposure levels was observed among women younger than 24 at first full-term pregnancy. An increase in annual airborne dioxin exposure was not associated with the risk of overall breast cancer at each year prior to diagnosis.

Several studies have reported no association between dioxin exposure and female breast cancer risk (Bertazzi et al., 1993; Warner et al., 2011; Xu et al., 2016). In a cohort involving the population exposed to dioxins following the Seveso industrial accident (1976, Italy), no association with breast cancer occurrence was found after a 10-year follow-up (1977-1986) comparing women living in one of the two contaminated zones with women living in a non-contaminated surrounding area (Bertazzi et al., 1993), and breast cancer risks remained non-statistically significant after a 20-year follow-up (Pesatori et al., 2009). In a specific retrospective cohort of women aged 0 to 40 years at the time of the accident and living in the two contaminated areas, an increase in TCDD serum concentration was not associated with increased breast cancer risk (Warner et al., 2011). The most recent meta-analysis, conducted on the association between external TCDD exposure and cancer incidence, reported no statistically significant risks for breast cancer (Xu et al., 2016). However, weak increases in breast cancer risk have been observed in studies conducted in populations accidentally exposed (5 cases, RR: 2.6, 95% CI: 1.1-6.2; (Pesatori et al., 2009)) and living in the vicinity of industrial facilities (18,824 cases, RR: 1.1, 95% CI: 1.0-1.2; (Fabre et al., 2008)). On the contrary, Viel et al. reported a decreased breast cancer risk among women aged 60 years and more and living in a highly exposed zone around a MSWI (434 cases, OR: 0.31, 95% CI: 0.1-0.9; (Viel et al., 2008)). Results from the literature are difficult to compare with our findings because of differences in methods for assessing dioxin exposure, exposure levels and presence of study limitations. Dioxin exposure levels were higher in accidentally exposed populations than in our study population. Monitoring of soil contamination by dioxins after the Seveso accident led to mean soil levels of TCDD ranging from 15.5 to 580  $\mu\text{g}/\text{m}^2$  in the most contaminated zone (Bertazzi et al., 1998) whereas in our study, average cumulative airborne dioxin exposure was estimated at 0.14 and 0.12  $\mu\text{g}\text{-TEQ}/\text{m}^2$  among cases and controls respectively. Fabre et al. used an atmospheric dispersion model and estimated an average cumulative annual dioxin deposit of  $7.9 \times 10^{-3}$   $\mu\text{g}/\text{m}^2/\text{year}$  from 1972 to 1990 over four French departments (Fabre et al., 2008). In the study by Viel et al., dioxin exposure consisted in exposure zones (very low, low, intermediate and high exposure) based on predicted ground-level air concentrations and measurements in soil samples (Viel et al., 2008). The results of these studies have to be interpreted with caution, due to the lack of adjustment for individual breast cancer risk factors

# Author's draft

(Fabre et al., 2008; Viel et al., 2008). Moreover, most of the above studies did not assess dioxin exposure at the individual address level (Fabre et al., 2008; Pesatori et al., 2009; Viel et al., 2008) and none considered dietary dioxin exposure in the statistical analysis. Our findings are in agreement with the results from our previous study conducted in the E3N cohort, in which we observed no association between estimated dietary dioxin exposure and overall breast cancer risk (Danjou et al., 2015).

Our results suggested a modest increase in breast cancer risk when comparing participants of the second *versus* the first quintile of exposure. After performing stratified analyses, the increased risk of breast cancer was observed only for ER-positive tumors and remained statistically significant when excluding short-term exposure. These results could be explained by the potential low-dose and non-monotonic dose-response effect of dioxins on cancer risk, as it has been suggested for other endocrine disruptors on human health outcomes (Vandenberg et al., 2012). Non-monotonic dose-response curves, including U-shaped and J-shaped curves, have been defined as non-linear relationships between dose and effect, where the slope of the curve changes sign within the range of doses examined, and may imply the use of other methods of risk assessment. Although the implications of the non-monotonic dose-response of endocrine disruptor continue to be debated (Lee and Jacobs, 2015; Solecki et al., 2017; Vandenberg et al., 2012), future research should investigate potential non-monotonic curves in risk assessment in relation to exposure to endocrine disruptors.

Statistically significant decreased breast cancer risks were found associated with higher airborne dioxin exposure among women aged 24 years old or less at first birth – although there was no effect modification by age at first full-term pregnancy ( $P=0.09$ ). This observation may be consistent with the protective effect of early maternal age on breast cancer risk, which has been shown in other studies (Kobayashi et al., 2012; Sun et al., 2017) and the inverse association between organochlorine compounds and breast cancer risk suggested in some studies (Gammon et al., 2002; Itoh et al., 2009; Iwasaki et al., 2008; Raaschou-Nielsen et al., 2005; Rusiecki et al., 2004). However, our subgroups analyses were based on small numbers of participants, possibly yielding unstable measures of association. Given also the high number of tests performed, we cannot exclude that some of the statistically significant findings may have occurred by chance.

We estimated the time-dependent effect of the intensity of airborne dioxin exposure received over the study period on the risk of breast cancer, by using a flexible time-dependent weight function. Overall, no association was observed with breast cancer risk. However, when excluding participants with short exposure duration (less than five years from inclusion), an increase in the annual intensity of airborne dioxin exposure the five years before diagnosis had a statistically significant impact on breast cancer risk. These results are in favor of the tumor-promoting activities of dioxins on carcinogenesis (Rodgers et al., 2018) and confirm the importance of considering dioxin exposures close to the time of diagnosis in the estimation of breast cancer risk. However, it is important to note that the weighted cumulative index of exposure in our study was not estimated over lifetime. The estimated impact of the annual dose in a given year was not adjusted for all other years over lifetime.

# Author's draft

Indeed, in our study, prevalent dioxin exposure was assessed from 1990 to 2008 as the participants were included in the cohort in 1990 and the residential history of the E3N participants before inclusion was not recorded (except for place of residence at birth). Moreover, information on past emissions of dioxins from industrial sources was not available as far back in time, participants being born between 1925 and 1950, leading to the left truncation of the exposure estimates. Although we intended to encompass this limitation by adjusting the models for the status urban or rural of the birthplace, participants might not be comparable regarding the dioxin exposure levels before 1990. Future studies should consider the impact of left-truncation on exposure in the association with breast cancer risk (Hazelbag et al., 2015; Vandenbroucke and Pearce, 2015). Finally, the weighted cumulative index of exposure assumed a linear effect of annual intensity. Indeed, the effect of intensity was allowed to flexibly vary from year to year of exposure, but in each year, we assumed the effect was linear. This might be a too strong assumption, as suggested by our results obtained with the quintiles of the standard cumulative index of exposure. Further studies relaxing this linearity assumption should be performed to further investigate the impact of the timing of exposure.

A number of observations have suggested that breast cancer may originate in early life and that woman may experience multiple time-variable windows of susceptibility, including the prenatal period, puberty and pregnancy, when mammary cells rapidly proliferate and differentiate, in which dioxins could affect hormonal pathways and induce the development of breast cancer later in life (Cohn, 2011; Rodgers et al., 2018; Teitelbaum et al., 2015). Therefore, assessing dioxin exposure during these critical periods of breast development may be important. In our study, we were not able to estimate the risk of breast cancer at these specific time windows, because we could not assess airborne dioxin exposure over the lifetime. Future studies should investigate the impact of dioxin exposure during the windows of breast susceptibility on breast cancer later in life.

In our study, dioxin exposure was not directly measured in blood samples or adipose tissues. Although the E3N women provided blood samples in 1995-1997, the volume in storage was insufficient to measure dioxin levels, and a new blood sampling would not be relevant regarding the etiology of breast cancer (Rodgers et al., 2018). GIS-based methods have been previously used as an alternative to assess exposure to air pollutants in absence of measured data and their use in epidemiological studies have grown in the last years (Cordioli et al., 2013; Rodgers et al., 2018). We assessed airborne dioxin exposure through a GIS-based metric that allowed the estimation of long-term exposure at the individual address level (Coudon et al., *Submitted*). The GIS included parameters known to influence individual airborne dioxin exposure according to the literature: residence-to-source proximity, dioxin emission intensity of industrial sources, exposure duration and wind direction (Gulliver and Briggs, 2011; Hoek et al., 2001; Pronk et al., 2013; Vienneau et al., 2009; White et al., 2009; Yu, 2006). Exposure classification into quintiles using the GIS-based metric showed “substantial” to “almost perfect” agreement with dioxin dispersion modeling (SIRANE, (Soulhac et al., 2011, 2012, 2017)) across different settings (Coudon et al., *Submitted*; Viera et al., 2005). We were

# Author's draft

able to consider the residential history of the participants over the study period and, in absence of dioxin monitoring data, we implemented a standardized method for the estimation of dioxin emissions using the UNEP Toolkit (UNEP, 2013). Geocoding of the participants' residential addresses and industrial facilities was achieved through an automatic method whose accuracy was assessed in a previous study (Faure et al., 2017).

Our study was based on a multi-source approach, considering major industrial dioxin sources (waste incineration, metal production, cement industry, etc.) and the evolution over time of the facilities' technical characteristics. Traffic-related exposure was not considered as its contribution to average dioxin concentration was estimated to be negligible in a previous study on dioxin modeling (<3%; (Coudon et al., 2017)) and stable over the study period (CITEPA, 2015). Emissions from domestic activities, including heating, chimney fire, cooking methods (*e.g.* stove, wood stove and barbecue) as well as backyard burning of domestic and green waste, may contribute to the airborne dioxin exposure and have been associated with breast cancer risk (White et al., 2014). Dioxin release from illegal cable burning (*i.e.* the process in which copper and lead are recovered by burning the insulating material from electricity and electronics) may be a critical source of dioxin emissions and potentially relevant exposure sources nowadays (Stockholm Convention Clearing House). However, because of the lack of data over the French territory (geographical and monitoring data), these emissions could not be considered in the exposure assessment. These punctual and non-industrial sources, in addition to the lack of past residential history and historical dioxin exposure estimates before 1990, may have resulted in exposure misclassification likely to contribute to imprecise measures of association drawn toward unity (Basagaña et al., 2013). Information on domestic activities and lifetime residential history should be collected in future studies and additional methods should be employed to take into account these types of dioxin sources in the exposure assessment (Coudon et al., *Submitted*).

Strengths of our study included its prospective design and the quality of the information collected and regularly updated. The statistical models were adjusted for known individual breast cancer risk factors, minimizing residual confounding. Moreover, factors that may influence individual dioxin exposure were considered in the estimation on breast cancer risk, such as age, breastfeeding and pregnancy. We estimated breast cancer risk according to hormone-receptor subtypes and were able to consider the potential impact of airborne dioxin exposure close to diagnosis on breast cancer risk. Although diet is quantitatively the main pathway for non-occupationally dioxin exposure, adjustment for dioxin intake in our models did not change our results (Danjou et al., 2015). Moreover, most of the E3N women are teachers and from affiliated occupations, thus their occupational dioxin exposure was assumed to be negligible and homogenous among the study participants.

Focusing our analysis on the E3N women that had permanently lived in the Rhône-Alpes region over the study period lowered the number of participants and thus the statistical power of our study. Confirmation of our findings is required in further studies and larger populations. In particular,

# Author's draft

further enlargement to breast cancer cases and their matched controls of the entire E3N cohort and identification of dioxin sources at the national level would be relevant. Moreover, future studies should investigate of the joint effect of dietary and airborne dioxin exposure in statistical models. Finally, the GIS-based metric may be adapted for the exposure assessment of other environmental pollutants in relation to breast cancer risk or used in the investigation of other pathologies.

## **5. Conclusions**

Among women from the E3N cohort, no association was observed between long-term and cumulative airborne dioxin exposure and breast cancer risk, as well as for hormone-receptor defined breast cancer. These results need confirmation in larger populations. The GIS-based metric developed to assess long-term and low-dose airborne dioxin exposure at the individual address level may provide an alternative in absence of measurement of ambient dioxin concentrations and may be used to assess exposure to other pollutants acting similarly as dioxins. Finally, our study may provide new ways of considering temporal dimensions for environmental exposures in disease risk assessment.

## References

1. Anzivino-Viricel, L., Falette, N., Carretier, J., Montestrucq, L., Guye, O., Philip, T., Fervers, B., 2012. Gestion des déchets ménagers et assimilés: bilan des connaissances et évaluation des effets sanitaires en population générale et au travail. *Environ. Risques Santé* 11, 360–377.
2. ATSDR, 1998. Toxicological profile for Chlorinated Dibenzop-Dioxins. US Department of Health and Human Services, Agency for Toxic Substances and Disease Registry, Atlanta, GA.
3. Baan, R., Grosse, Y., Straif, K., Secretan, B., El Ghissassi, F., Bouvard, V., Benbrahim-Tallaa, L., Guha, N., Freeman, C., Galichet, L., Coglianò, V., WHO International Agency for Research on Cancer Monograph Working Group, 2009. A review of human carcinogens--Part F: chemical agents and related occupations. *Lancet Oncol.* 10, 1143–1144.
4. Basagaña, X., Aguilera, I., Rivera, M., Agis, D., Foraster, M., Marrugat, J., Elosua, R., Künzli, N., 2013. Measurement Error in Epidemiologic Studies of Air Pollution Based on Land-Use Regression Models. *Am. J. Epidemiol.* 178, 1342–1346. <https://doi.org/10.1093/aje/kwt127>
5. Bertazzi, A., Pesatori, A.C., Consonni, D., Tironi, A., Landi, M.T., Zocchetti, C., 1993. Cancer incidence in a population accidentally exposed to 2,3,7,8-tetrachlorodibenzo-para-dioxin. *Epidemiol. Camb. Mass* 4, 398–406.
6. Bertazzi, P.A., Bernucci, I., Brambilla, G., Consonni, D., Pesatori, A.C., 1998. The Seveso studies on early and long-term effects of dioxin exposure: a review. *Environ. Health Perspect.* 106, 625.
7. Binachon, B., Dossus, L., Danjou, A.M.N., Clavel-Chapelon, F., Fervers, B., 2014. Life in urban areas and breast cancer risk in the French E3N cohort. *Eur. J. Epidemiol.* 29, 743–751. <https://doi.org/10.1007/s10654-014-9942-z>
8. Birnbaum, L.S., Fenton, S.E., 2002. Cancer and Developmental Exposure to Endocrine Disruptors. *Environ. Health Perspect.* 111, 389–394. <https://doi.org/10.1289/ehp.5686>
9. Boverhof, D.R., 2006. Dioxin Induces an Estrogen-Like, Estrogen Receptor-Dependent Gene Expression Response in the Murine Uterus. *Mol. Pharmacol.* 69, 1599–1606. <https://doi.org/10.1124/mol.105.019638>
10. Brody, J.G., Moysich, K.B., Humblet, O., Attfield, K.R., Beehler, G.P., Rudel, R.A., 2007. Environmental pollutants and breast cancer: Epidemiologic studies. *Cancer* 109, 2667–2711. <https://doi.org/10.1002/cncr.22655>
11. CITEPA, 2015. Dioxines et Furanes - PCDD-F [WWW Document]. URL <http://www.citepa.org/fr/air-et-climat/polluants/polluant-organiques-persistants/dioxines-et-furanes> (accessed 9.1.16).
12. Clavel-Chapelon, F., 2015. Cohort Profile: The French E3N Cohort Study. *Int. J. Epidemiol.* 44, 801–809. <https://doi.org/10.1093/ije/dyu184>
13. Cohn, B.A., 2011. Developmental and Environmental Origins of Breast Cancer DDT as a Case Study. *Reprod. Toxicol. Elmsford N* 31, 302–311. <https://doi.org/10.1016/j.reprotox.2010.10.004>
14. Cordier, S., Chevrier, C., Robert-Gnansia, E., Lorente, C., Brula, P., Hours, M., 2004. Risk of congenital anomalies in the vicinity of municipal solid waste incinerators. *Occup. Environ. Med.* 61, 8–15.
15. Cordioli, M., Ranzi, A., De Leo, G.A., Lauriola, P., 2013. A Review of Exposure Assessment Methods in Epidemiological Studies on Incinerators. *J. Environ. Public Health* 2013. <https://doi.org/10.1155/2013/129470>
16. Coudon, T., Danjou, A.M.N., Faure, E., Praud, D., Severi, G., Mancini, F.R., Salizzoni, P., Fervers, B., n.d. Development and performance evaluation of a GIS-based metric to assess exposure to airborne pollutant emissions from industrial sources.
17. Coudon, T., Hourani, H., Nguyen, C., Faure, E., Mancini, F.R., Fervers, B., Salizzoni, P., 2017. Assessment of long-term exposure to airborne dioxin and cadmium concentrations in the Lyon metropolitan area (France). *Environ. Int.* 111, 177–190. <https://doi.org/10.1016/j.envint.2017.11.027>
18. Coudon, T., Salizzoni, P., Praud, D., Danjou, A., Dossus, L., Faure, E., Fervers, B., n.d. National inventory of historical dioxin air emission sources in France.
19. Danjou, A., Fervers, B., Boutron-Ruault, M.-C., Philip, T., Clavel-Chapelon, F., Dossus, L., 2015. Estimated dietary dioxin exposure and breast cancer risk among women from the French E3N prospective cohort. *Breast Cancer Res.* 17, 39. <https://doi.org/10.1186/s13058-015-0536-9>

# Author's draft

20. Fabre, P., Daniau, C., Gorla, S., de Crouy-Chanel, P., Empereur-Bissonnet, P., 2008. Étude d'incidence des cancers à proximité des usines d'incinération d'ordures ménagères. Synthèse St.-Maurice InVS.
21. Faure, E., Danjou, A.M.N., Clavel-Chapelon, F., Boutron-Ruault, M.-C., Dossus, L., Fervers, B., 2017. Accuracy of two geocoding methods for geographic information system-based exposure assessment in epidemiological studies. *Environ. Health Glob. Access Sci. Source* 16, 15. <https://doi.org/10.1186/s12940-017-0217-5>
22. Gammon, M.D., Wolff, M.S., Neugut, A.I., Eng, S.M., Teitelbaum, S.L., Britton, J.A., Terry, M.B., Levin, B., Stellman, S.D., Kabat, G.C., others, 2002. Environmental toxins and breast cancer on Long Island. II. Organochlorine compound levels in blood. *Cancer Epidemiol. Biomarkers Prev.* 11, 686–697.
23. Gulliver, J., Briggs, D., 2011. STEMS-Air: A simple GIS-based air pollution dispersion model for city-wide exposure assessment. *Sci. Total Environ.* 409, 2419–2429. <https://doi.org/10.1016/j.scitotenv.2011.03.004>
24. Hauptmann, M., Wellmann, J., Lubin, J.H., Rosenberg, P.S., Kreienbrock, L., 2000. Analysis of Exposure-Time-Response Relationships Using a Spline Weight Function. *Biometrics* 56, 1105–1108.
25. Hazelbag, C.M., Klungel, O.H., van Staa, T.P., de Boer, A., Groenwold, R.H.H., 2015. Left truncation results in substantial bias of the relation between time-dependent exposures and adverse events. *Ann. Epidemiol.* 25, 590–596. <https://doi.org/10.1016/j.annepidem.2015.03.019>
26. Hoek, G., Fischer, P., Van Den Brandt, P., Goldbohm, S., Brunekreef, B., 2001. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. *J. Expo. Anal. Environ. Epidemiol.* 11.
27. IARC (Ed.), 2012. IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 F, chemical agents and related occupations: this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 20 - 27 October 2009. IARC, Lyon.
28. Insee, 2017. Insee - Statistiques locales - Indicateurs : cartes, données et graphiques [WWW Document]. URL <https://statistiques-locales.insee.fr/#bbox=-1052393,6661347,2596908,1619645&c=indicator&i=bdcom.dens&s=2014&view=map2> (accessed 3.14.18).
29. Insee, 2012. Rhône-Alpes : une région riche et industrielle que la crise n'a pas épargnée - La Lettre Analyses | Insee [WWW Document]. URL <https://www.insee.fr/fr/statistiques/1290761> (accessed 9.26.17).
30. Itoh, H., Iwasaki, M., Hanaoka, T., Kasuga, Y., Yokoyama, S., Onuma, H., Nishimura, H., Kusama, R., Tsugane, S., 2009. Serum organochlorines and breast cancer risk in Japanese women: a case-control study. *Cancer Causes Control CCC* 20, 567–580. <https://doi.org/10.1007/s10552-008-9265-z>
31. Iwasaki, M., Inoue, M., Sasazuki, S., Kurahashi, N., Itoh, H., Usuda, M., Tsugane, S., 2008. Plasma organochlorine levels and subsequent risk of breast cancer among Japanese women: A nested case-control study. *Sci. Total Environ.* 402, 176–183. <https://doi.org/10.1016/j.scitotenv.2008.05.009>
32. Jenkins, S., Betancourt, A.M., Wang, J., Lamartiniere, C.A., 2012. Endocrine-active chemicals in mammary cancer causation and prevention. *J. Steroid Biochem. Mol. Biol.* 129, 191–200. <https://doi.org/10.1016/j.jsbmb.2011.06.003>
33. Knerr, S., Schrenk, D., 2006. Carcinogenicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in experimental models. *Mol. Nutr. Food Res.* 50, 897–907. <https://doi.org/10.1002/mnfr.200600006>
34. Kobayashi, S., Sugiura, H., Ando, Y., Shiraki, N., Yanagi, T., Yamashita, H., Toyama, T., 2012. Reproductive history and breast cancer risk. *Breast Cancer Tokyo Jpn.* 19, 302–308. <https://doi.org/10.1007/s12282-012-0384-8>
35. Lacourt, A., Lévêque, E., Guichard, E., Ilg, A.G.S., Sylvestre, M.-P., Leffondré, K., 2017. Dose-time-response association between occupational asbestos exposure and pleural mesothelioma. *Occup. Environ. Med.* 74, 691–697. <https://doi.org/10.1136/oemed-2016-104133>
36. Lee, D.-H., Jacobs, D.R., 2015. Methodological issues in human studies of endocrine disrupting chemicals. *Rev. Endocr. Metab. Disord.* 16, 289–297. <https://doi.org/10.1007/s11154-016-9340-9>



37. Mandal, P.K., 2005. Dioxin: a review of its environmental effects and its aryl hydrocarbon receptor biology. *J. Comp. Physiol. B* 175, 221–230. <https://doi.org/10.1007/s00360-005-0483-3>
38. Manuwald, U., Garrido, M.V., Berger, J., Manz, A., Baur, X., 2012. Mortality study of chemical workers exposed to dioxins: follow-up 23 years after chemical plant closure. *Occup. Environ. Med.* 69, 636–642.
39. Matthews, J., Gustafsson, J.-Å., 2006. Estrogen receptor and aryl hydrocarbon receptor signaling pathways. *Nucl. Recept. Signal.* 4. <https://doi.org/10.1621/nrs.04016>
40. Nzihou, A., Themelis, N.J., Kemiha, M., Benhamou, Y., 2012. Dioxin emissions from municipal solid waste incinerators (MSWIs) in France. *Waste Manag.* 32, 2273–2277. <https://doi.org/10.1016/j.wasman.2012.06.016>
41. Pesatori, A., Consonni, D., Rubagotti, M., Grillo, P., Bertazzi, P., 2009. Cancer incidence in the population exposed to dioxin after the “Seveso accident”: twenty years of follow-up. *Environ. Health* 8, 39. <https://doi.org/10.1186/1476-069X-8-39>
42. Pronk, A., Nuckols, J.R., De Roos, A.J., Airola, M., Colt, J.S., Cerhan, J.R., Morton, L., Cozen, W., Severson, R., Blair, A., others, 2013. Residential proximity to industrial combustion facilities and risk of non-Hodgkin lymphoma: a case–control study. *Environ. Health* 12, 1.
43. Raaschou-Nielsen, O., Pavuk, M., LeBlanc, A., Dumas, P., Weber, J.P., Olsen, A., Tjønneland, A., Overvad, K., Olsen, J.H., 2005. Adipose organochlorine concentrations and risk of breast cancer among postmenopausal Danish women. *Cancer Epidemiol. Biomarkers Prev.* 14, 67–74.
44. Revich, B., Aksel, E., Ushakova, T., Ivanova, I., Zhuchenko, N., Klyuev, N., Brodsky, B., Sotnikov, Y., 2001. Dioxin exposure and public health in Chapaevsk, Russia. *Chemosphere* 43, 951–966.
45. Riboli, E., 1992. Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann. Oncol.* 3, 783–791.
46. Riboli, E., Hunt, K., Slimani, N., Ferrari, P., Norat, T., Fahey, M., Charrondière, U., Hémon, B., Casagrande, C., Vignat, J., Overvad, K., Tjønneland, A., Clavel-Chapelon, F., Thiébaud, A., Wahrendorf, J., Boeing, H., Trichopoulos, D., Trichopoulou, A., Vineis, P., Palli, D., Bueno-de-Mesquita, H., Peeters, P., Lund, E., Engeset, D., González, C., Barricarte, A., Berglund, G., Hallmans, G., Day, N., Key, T., Kaaks, R., Saracci, R., 2002. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 5, 1113. <https://doi.org/10.1079/PHN2002394>
47. Rodgers, K.M., Udesky, J.O., Rudel, R.A., Brody, J.G., 2018. Environmental chemicals and breast cancer: An updated review of epidemiological literature informed by biological mechanisms. *Environ. Res.* 160, 152–182. <https://doi.org/10.1016/j.envres.2017.08.045>
48. Rusiecki, J.A., Holford, T.R., Zahm, S.H., Zheng, T., 2004. Polychlorinated biphenyls and breast cancer risk by combined estrogen and progesterone receptor status. *Eur. J. Epidemiol.* 19, 793–801.
49. SAS, n.d. Chapter 4. The NLP procedure, SAS/OR® 9.22 user's Guide: constraint programming. ed. SAS Institute Inc.
50. Solecki, R., Kortenkamp, A., Bergman, Å., Chahoud, I., Degen, G.H., Dietrich, D., Greim, H., Håkansson, H., Hass, U., Husoy, T., Jacobs, M., Jobling, S., Mantovani, A., Marx-Stoelting, P., Piersma, A., Ritz, V., Slama, R., Stahlmann, R., van den Berg, M., Zoeller, R.T., Boobis, A.R., 2017. Scientific principles for the identification of endocrine-disrupting chemicals: a consensus statement. *Arch. Toxicol.* 91, 1001–1006. <https://doi.org/10.1007/s00204-016-1866-9>
51. Soulhac, L., Nguyen, C.V., Volta, P., Salizzoni, P., 2017. The model SIRANE for atmospheric urban pollutant dispersion. PART III: Validation against NO<sub>2</sub> yearly concentration measurements in a large urban agglomeration. *Atmos. Environ.* 167, 377–388. <https://doi.org/10.1016/j.atmosenv.2017.08.034>
52. Soulhac, L., Salizzoni, P., Cierco, F.-X., Perkins, R., 2011. The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model. *Atmos. Environ.* 45, 7379–7395. <https://doi.org/10.1016/j.atmosenv.2011.07.008>
53. Soulhac, L., Salizzoni, P., Mejean, P., Didier, D., Rios, I., 2012. The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study. *Atmos. Environ.* 49, 320–337. <https://doi.org/10.1016/j.atmosenv.2011.11.031>

# Author's draft

54. Stockholm Convention Clearing House, n.d. Section VI. Guidance/guidelines by source category: Part III of Annex C. VI.L. Smouldering of copper cables [WWW Document]. URL <http://chm.pops.int/Implementation/BATandBEP/BATBEPGuidelinesArticle5/tabid/187/Default.aspx> (accessed 10.7.16).
55. Sun, Y.-S., Zhao, Z., Yang, Z.-N., Xu, F., Lu, H.-J., Zhu, Z.-Y., Shi, W., Jiang, J., Yao, P.-P., Zhu, H.-P., 2017. Risk Factors and Preventions of Breast Cancer. *Int. J. Biol. Sci.* 13, 1387–1397. <https://doi.org/10.7150/ijbs.21635>
56. Teitelbaum, S.L., Belpoggi, F., Reinlib, L., 2015. Advancing research on endocrine disrupting chemicals in breast cancer: Expert panel recommendations. *Reprod. Toxicol.* 54, 141–147. <https://doi.org/10.1016/j.reprotox.2014.12.015>
57. Travis, C.C., Hattemer-Frey, H.A., 1989. A perspective on dioxin emissions from municipal solid waste incinerators. *Risk Anal. Off. Publ. Soc. Risk Anal.* 9, 91–97.
58. Tsuchiya, Y., Nakajima, M., Yokoi, T., 2005. Cytochrome P450-mediated metabolism of estrogens and its regulation in human. *Cancer Lett.* 227, 115–124. <https://doi.org/10.1016/j.canlet.2004.10.007>
59. UNEP, 2013. United Nations Environment Programme | Toolkit for Identification and Quantification of Releases of Dioxins, Furans and Other Unintentional POPs [WWW Document]. URL [http://toolkit.pops.int/Publish/Main/01\\_Index.html](http://toolkit.pops.int/Publish/Main/01_Index.html) (accessed 8.22.16).
60. Vandenberg, L.N., Colborn, T., Hayes, T.B., Heindel, J.J., Jacobs, D.R., Lee, D.-H., Shioda, T., Soto, A.M., vom Saal, F.S., Welshons, W.V., Zoeller, R.T., Myers, J.P., 2012. Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocr. Rev.* 33, 378–455. <https://doi.org/10.1210/er.2011-1050>
61. Vandembroucke, J., Pearce, N., 2015. Point: incident exposures, prevalent exposures, and causal inference: does limiting studies to persons who are followed from first exposure onward damage epidemiology? *Am. J. Epidemiol.* 182, 826–833. <https://doi.org/10.1093/aje/kwv225>
62. Viel, J.-F., Clément, M.-C., Hägi, M., Grandjean, S., Challier, B., Danzon, A., 2008. Dioxin emissions from a municipal solid waste incinerator and risk of invasive breast cancer: a population-based case-control study with GIS-derived exposure. *Int. J. Health Geogr.* 7, 4. <https://doi.org/10.1186/1476-072X-7-4>
63. Vienneau, D., de Hoogh, K., Briggs, D., 2009. A GIS-based method for modelling air pollution exposures across Europe. *Sci. Total Environ.* 408, 255–266. <https://doi.org/10.1016/j.scitotenv.2009.09.048>
64. Viera, A.J., Garrett, J.M., others, 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med* 37, 360–363.
65. Warner, M., Mocarelli, P., Samuels, S., Needham, L., Brambilla, P., Eskenazi, B., 2011. Dioxin exposure and cancer risk in the Seveso Women's Health Study. *Environ. Health Perspect.* 119, 1700–1705. <https://doi.org/10.1289/ehp.1103720>
66. White, A.J., Teitelbaum, S.L., Stellman, S.D., Beyea, J., Steck, S.E., Mordukhovich, I., McCarty, K.M., Ahn, J., Rossner, P., Santella, R.M., others, 2014. Indoor air pollution exposure from use of indoor stoves and fireplaces in association with breast cancer: a case-control study. *Environ. Health* 13, 1.
67. White, N., teWaterNaude, J., van der Walt, A., Ravenscroft, G., Roberts, W., Ehrlich, R., 2009. Meteorologically estimated exposure but not distance predicts asthma symptoms in schoolchildren in the environs of a petrochemical refinery: a cross-sectional study. *Environ. Health* 8. <https://doi.org/10.1186/1476-069X-8-45>
68. Xu, J., Ye, Y., Huang, F., Chen, H., Wu, H., Huang, J., Hu, J., Xia, D., Wu, Y., 2016. Association between dioxin and cancer incidence and mortality: a meta-analysis. *Sci. Rep.* 6. <https://doi.org/10.1038/srep38012>
69. Yu, C.-L., 2006. Residential Exposure to Petrochemicals and the Risk of Leukemia: Using Geographic Information System Tools to Estimate Individual-Level Residential Exposure. *Am. J. Epidemiol.* 164, 200–207. <https://doi.org/10.1093/aje/kwj182>
70. Zook, D., Rappe, C., 1994. Environmental sources, distribution, and fate of polychlorinated dibenzodioxins, dibenzofurans, and related organochlorines., in: Schecter A. Plenum Press, New York, pp. 80–113.

**Tables**

Table 1. Baseline characteristics of 429 breast cancer cases and 716 controls, E3N study, Rhône-Alpes region, France, 1990-2008.

Characteristics	Cases		Controls		P <sup>a</sup>
<b>Age (years), mean (SD)</b>	49.4	6.1	49.5	6.2	0.81
<b>Body Mass Index (kg/m<sup>2</sup>), median value among controls, n (%)</b>					0.56
<=21.9	229	53.4	366	51.1	
>21.9	200	46.6	350	48.9	
<b>Alcohol drinking, n (%)</b>					0.11
Never drinker	37	8.6	85	11.9	
Drinker <5.9 g/day	144	33.6	229	32.0	
Drinker >= 5.9 g/day	217	50.6	309	43.2	
<i>Missing data</i>	31	7.2	93	13.0	
<b>Smoking status, n (%)</b>					0.30
Never smoker	235	54.8	416	58.1	
Current smoker	55	12.8	98	13.7	
Former smoker	139	32.4	202	28.2	
<b>Status of birthplace, n (%)</b>					0.07
Rural	89	20.7	180	25.1	
Urban	310	72.3	496	69.3	
<i>Missing data</i>	30	7.0	40	5.6	
<b>Status of residence at baseline</b>					0.12
Rural	120	28.0	237	33.1	
Urban	279	65.0	439	61.3	
<i>Missing data</i>	30	7.0	40	5.6	
<b>Recreational physical activity (METs-h/week), n (%)</b>					0.02
<25.3	105	24.5	141	19.7	
25.3-37.3	117	27.3	211	29.5	
37.4-56.9	132	30.8	191	26.7	
>=57.0	75	17.5	173	24.2	
<b>Education, n (%)</b>					0.02
Undergraduate	44	10.3	88	12.3	
Post-graduate with a 1- to 2-year university degree	198	46.2	373	52.1	
Post-graduate with a >= 3 year university degree	187	43.6	255	35.6	
<b>Age at menarche (years), modal value among controls, n (%)</b>					0.01
<13	222	51.7	311	43.4	
>=13	207	48.3	405	56.6	
<b>Previous use of oral contraceptives, n (%)</b>	264	61.5	426	59.5	0.60
<b>Previous use of progestin before menopause, n (%)</b>	200	46.6	300	41.9	0.21
<b>Menopausal status, n (%)</b>					0.72
Premenopausal women	261	60.8	426	59.5	
Postmenopausal women	168	39.2	290	40.5	
<b>Use of menopausal hormone treatment, n (%)</b>	62	14.5	94	13.1	0.70
<b>Mammography during the previous follow-up period, n (%)</b>	342	79.7	527	73.6	0.01
<b>Age at first full-term pregnancy (years), median value among controls, n (%)</b>					0.02
<=24	161	37.5	314	43.9	
>24	208	48.5	313	43.7	
<i>Missing data</i>	60	14.0	89	12.4	

# Author's draft

Characteristics	Cases		Controls		P <sup>a</sup>
<b>Parity, n (%)</b>					0.38
<=2	260	60.6	413	57.7	
>=3	109	25.4	213	29.7	
<i>Missing data</i>	60	14.0	90	12.6	
<b>Breastfeeding, n (%)</b>	243	56.6	433	60.5	0.20
<b>Previous family history of breast cancer, n (%)</b>	76	17.7	67	9.4	<0.001
<b>Previous history of personal benign breast disease, n (%)</b>	166	38.7	206	28.8	<0.001
<b>Estimated dietary dioxin intake (pg TEQ/kg body weight/day), mean (SD)</b>	1.3	0.4	1.2	0.4	0.07
<i>Missing data</i>	119	27.7	135	18.9	

<sup>a</sup> P-values from univariate conditional logistic regression models – except for age, menopausal status and use of menopausal treatment which were matching factors

Table 2. Baseline characteristics of study participants (N=1,145) according to status of living more than 10 km from any source between 1990 and 2008 and quintiles<sup>a</sup> of cumulative metric of airborne dioxin exposure, E3N study, Rhône-Alpes region, France, 1990-2008

Case/Control status	Cumulative metric of airborne dioxin exposure										P <sup>b</sup>	
	Living > 10 km from any source		I		II		III		IV			V
	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)
Cases	32	(30.5)	81	(38.8)	67	(34.2)	93	(41.9)	83	(39.3)	73	(36.1)
Controls	73	(69.5)	128	(61.2)	129	(65.8)	129	(58.1)	128	(60.7)	129	(63.9)
Age (years), mean (SD)	50.2	(5.9)	50.0	(6.1)	48.7	(6.2)	49.6	(6.0)	48.9	(6.1)	49.4	(6.7)
Body Mass Index (kg/m <sup>2</sup> ), median value among controls												
<=21.9	52	(49.5)	109	(52.2)	99	(50.5)	112	(50.5)	116	(55.0)	107	(53.0)
>21.9	53	(50.5)	100	(47.8)	97	(49.5)	110	(49.5)	95	(45.0)	95	(47.0)
Alcohol drinking												
Never drinker	9	(8.6)	28	(13.4)	25	(12.8)	21	(9.5)	14	(6.6)	25	(12.4)
Drinker <5.9 g/day	40	(38.1)	62	(29.7)	63	(32.1)	72	(32.4)	76	(36.0)	60	(29.7)
Drinker >=5.9 g/day	37	(35.2)	89	(42.6)	83	(42.3)	95	(42.8)	108	(51.2)	114	(56.4)
Missing data	19	(18.1)	30	(14.4)	25	(12.8)	34	(15.3)	13	(6.2)	3	(1.5)
Smoking status												
Never smoker	69	(65.7)	130	(62.2)	121	(61.7)	116	(52.3)	108	(51.2)	107	(53.0)
Current smoker	12	(11.4)	18	(8.6)	19	(9.7)	36	(16.2)	43	(20.4)	25	(12.4)
Former smoker	24	(22.9)	61	(29.2)	56	(28.6)	70	(31.5)	60	(28.4)	70	(34.7)
Status of birthplace												
Rural	36	(34.3)	53	(25.4)	51	(26.0)	44	(19.8)	32	(15.2)	53	(26.2)
Urban	65	(61.9)	150	(71.8)	136	(69.4)	159	(71.6)	165	(78.2)	131	(64.9)
Missing data	4	(3.8)	6	(2.9)	9	(4.6)	19	(8.6)	14	(6.6)	18	(8.9)
Status of residence at baseline												
Rural	79	(75.2)	70	(33.5)	75	(38.3)	63	(28.4)	40	(19.0)	30	(14.9)

<0.0001

Urban	22	(21.0)	133	(63.6)	112	(57.1)	140	(63.1)	157	(74.4)	154	(76.2)
<i>Missing data</i>	4	(3.8)	6	(2.9)	9	(4.6)	19	(8.6)	14	(6.6)	18	(8.9)
Recreational physical activity (METs-h/week)												0.09
<25.3	11	(10.5)	55	(26.3)	44	(22.4)	40	(18.0)	47	(22.3)	49	(24.3)
25.3-37.3	25	(23.8)	57	(27.3)	63	(32.1)	66	(29.7)	60	(28.4)	57	(28.2)
37.4-56.9	34	(32.4)	56	(26.8)	53	(27.0)	64	(28.8)	57	(27.0)	59	(29.2)
>=57.0	35	(33.3)	41	(19.6)	36	(18.4)	52	(23.4)	47	(22.3)	37	(18.3)
Education												<0.0001
Undergraduate	16	(15.2)	28	(13.4)	18	(9.2)	21	(9.5)	26	(12.3)	23	(11.4)
Post-graduate with a 1- to 2-year university degree	60	(57.1)	112	(53.6)	125	(63.8)	105	(47.3)	97	(46.0)	72	(35.6)
Post-graduate with a >= 3 year university degree	29	(27.6)	69	(33.0)	53	(27.0)	96	(43.2)	88	(41.7)	107	(53.0)
Age at menarche (years), modal value among controls												0.66
<13	46	(43.8)	88	(42.1)	91	(46.4)	111	(50.0)	100	(47.4)	97	(48.0)
>=13	59	(56.2)	121	(57.9)	105	(53.6)	111	(50.0)	111	(52.6)	105	(52.0)
Previous use of oral contraceptives	57	(54.3)	122	(58.4)	136	(69.4)	135	(60.8)	126	(59.7)	114	(56.4)
Previous use of progestin before menopause	51	(48.6)	81	(38.8)	77	(39.3)	112	(50.5)	95	(45.0)	84	(41.6)
Menopausal status at baseline												0.15
Premenopausal women	58	(55.2)	115	(55.0)	131	(66.8)	133	(59.9)	133	(63.0)	117	(57.9)
Postmenopausal women	47	(44.8)	94	(45.0)	65	(33.2)	89	(40.1)	78	(37.0)	85	(42.1)
Use of menopausal hormone treatment at baseline	16	(15.2)	27	(12.9)	21	(10.7)	30	(13.5)	31	(14.7)	31	(15.3)
Mammography during the previous follow-up period	82	(78.1)	157	(75.1)	144	(73.5)	173	(77.9)	164	(77.7)	149	(73.8)
Age at first full-term pregnancy (years), median value among controls												0.80
<=24	38	(36.2)	93	(44.5)	90	(45.9)	93	(41.9)	95	(45.0)	66	(32.7)
>24	54	(51.4)	97	(46.4)	90	(45.9)	100	(45.0)	87	(41.2)	93	(46.0)
<i>Missing data</i>	13	(12.4)	19	(9.1)	16	(8.2)	29	(13.1)	29	(13.7)	43	(21.3)
Parity												0.26
<=2	63	(60.0)	134	(64.1)	124	(63.3)	119	(53.6)	118	(55.9)	115	(56.9)
>=3	29	(27.6)	55	(26.3)	56	(28.6)	74	(33.3)	64	(30.3)	44	(21.8)

<i>Missing data</i>	13	(12.4)	20	(9.6)	16	(8.2)	29	(13.1)	29	(13.7)	43	(21.3)
Breastfeeding	54	(51.4)	122	(58.4)	122	(62.2)	126	(56.8)	127	(60.2)	125	(61.9)
Previous history of personal benign breast disease	28	(26.7)	80	(38.3)	67	(34.2)	70	(31.5)	67	(31.8)	60	(29.7)
Previous family history of breast cancer	9	(8.6)	25	(12.0)	24	(12.2)	27	(12.2)	29	(13.7)	29	(14.4)
Estimated dietary dioxin intake, mean (SD)	1.3	(0.5)	1.2	(0.4)	1.3	(0.4)	1.2	(0.4)	1.3	(0.4)	1.2	(0.4)
<i>Missing data from the E3N diet history questionnaire of 1993</i>	32	(30.5)	59	(28.2)	50	(25.5)	60	(27.0)	34	(16.1)	19	(9.4)

<sup>a</sup> quintiles' cut offs: 2.6E-4, 4.7E-3, 2.9E-2, 1.1E-1 µg-TEQ/m<sup>2</sup>

<sup>b</sup> *P* from Chi-square statistical test and analysis of variance

# Author's draft

Table 3. Odds ratios (OR) <sup>a</sup> and 95% confidence intervals (CI) for the association between invasive breast cancer and quintiles <sup>b</sup> of cumulative metric of airborne dioxin exposure overall and according to hormone receptor status (N=1,145), E3N study, 1990-2008.

Populations		Cases, N (%)		Controls, N (%)		OR <sup>a</sup>	Lower 95%CI	Upper 95%CI	P-trend	P-het <sup>e</sup>
Overall									0.81	
	I	65	15.2	142	19.8	1.00				
	II	101	23.5	143	20.0	1.61	1.04	2.49		
	III	87	20.3	145	20.3	1.40	0.89	2.21		
	IV	100	23.3	143	20.0	1.51	0.95	2.38		
	V	76	17.7	143	20.0	1.12	0.69	1.82		
ER positive <sup>c</sup>									0.86	0.42
	I	34	13.9	81	20.0	1.00				
	II	60	24.6	75	18.5	1.95	1.08	3.51		
	III	48	19.7	83	20.4	1.32	0.72	2.44		
	IV	61	25.0	83	20.4	1.66	0.89	3.08		
	V	41	16.8	84	20.7	1.04	0.53	2.02		
ER negative <sup>c</sup>									0.42	0.42
	I	17	21.3	21	16.2	1.00				
	II	10	12.5	18	13.8	0.66	0.21	2.04		
	III	17	21.3	26	20.0	0.87	0.27	2.80		
	IV	20	25.0	35	26.9	0.63	0.20	1.93		
	V	16	20.0	30	23.1	0.61	0.20	1.86		
PR positive <sup>d</sup>									0.14	0.41
	I	34	15.8	64	18.0	1.00				
	II	48	22.3	61	17.1	1.36	0.74	2.50		
	III	41	19.1	75	21.1	0.83	0.44	1.58		
	IV	56	26.0	81	22.8	0.99	0.53	1.86		
	V	36	16.7	75	21.1	0.62	0.31	1.26		
PR negative <sup>d</sup>									0.48	0.41
	I	17	15.7	39	21.8	1.00				
	II	22	20.4	30	16.8	1.91	0.70	5.24		
	III	23	21.3	33	18.4	2.11	0.74	6.02		
	IV	24	22.2	37	20.7	2.07	0.69	6.22		
	V	22	20.4	40	22.3	1.71	0.63	4.63		

<sup>a</sup> adjusted for family history of breast cancer, age at first full-term pregnancy, recreational physical activity, status of birthplace and breastfeeding

<sup>b</sup> quintiles' cut offs: 1.7E-5, 2.2E-3, 1.9E-2 and 9.6E-2  $\mu\text{g-TEQ}/\text{m}^2$

<sup>c</sup> n=105 invasive breast cancer cases with missing estrogen receptor status, and n=205 matched controls

<sup>d</sup> n=106 invasive breast cancer cases with missing progesterone receptor status, and n=251 matched controls

<sup>e</sup> P-heterogeneity comparing ER+ versus ER- and PR+ versus PR- breast tumors.



# Author's draft

Table 4. Odds ratios (OR)<sup>a</sup> and 95% confidence intervals (CI) for the association between invasive breast cancer and quintiles<sup>b</sup> of cumulative metric of airborne dioxin exposure, stratified analyses (N=1,145), E3N study, 1990-2008

Populations		Cases, N (%)	Controls, N (%)	OR <sup>a</sup>	Lower 95%CI	Upper 95%CI	P-trend	P for interaction
Menopausal status at baseline								0.92
Premenopause								0.53
	I	37 14.5	74	17.8	1.00			
	II	63 24.6	90	21.7	1.17	0.76	1.81	
	III	50 19.5	86	20.7	1.04	0.65	1.67	
	IV	65 25.4	85	20.5	1.16	0.73	1.84	
	V	41 16.0	80	19.3	0.86	0.51	1.44	
Postmenopause								0.86
	I	26 15.8	68	24.0	1.00			
	II	37 22.4	51	18.0	1.50	0.87	2.58	
	III	37 22.4	58	20.5	1.40	0.80	2.44	
	IV	32 19.4	50	17.7	1.31	0.73	2.35	
	V	33 20.0	56	19.8	1.17	0.64	2.14	
Age at first full-term pregnancy <sup>c</sup>								0.09
<=24 years								0.03
	I	32 19.9	53	16.9	1.00			
	II	38 23.6	72	22.9	0.89	0.53	1.47	
	III	37 23.0	59	18.8	1.04	0.61	1.76	
	IV	40 24.8	70	22.3	0.79	0.47	1.34	
	V	14 8.7	60	19.1	0.40	0.20	0.81	
>24 years								0.96
	I	27 13.0	72	23.0	1.00			
	II	56 26.9	61	19.5	1.63	1.00	2.65	
	III	39 18.8	68	21.7	1.29	0.76	2.18	
	IV	48 23.1	49	15.7	1.58	0.92	2.71	
	V	38 18.3	63	20.1	1.16	0.66	2.04	
Breastfeeding								0.45
Ever								0.09
	I	38 15.6	73	16.9	1.00			
	II	61 25.1	93	21.5	1.08	0.70	1.68	
	III	47 19.3	83	19.2	1.02	0.64	1.62	
	IV	59 24.3	86	19.9	1.04	0.66	1.64	
	V	38 15.6	98	22.6	0.64	0.38	1.06	
Never								0.60
	I	27 14.5	69	24.4	1.00			
	II	40 21.5	50	17.7	1.54	0.92	2.59	
	III	40 21.5	62	21.9	1.34	0.78	2.32	
	IV	41 22.0	57	20.1	1.33	0.76	2.34	
	V	38 20.4	45	15.9	1.36	0.75	2.45	
Body mass index at baseline <sup>d</sup>								0.49
<=21.9 kg/m <sup>2</sup>								0.10
	I	38 16.6	66	18.0	1.00			
	II	55 24.0	74	20.2	1.10	0.71	1.71	
	III	47 20.5	66	18.0	0.98	0.61	1.58	
	IV	54 23.6	80	21.9	0.90	0.56	1.45	
	V	35 15.3	80	21.9	0.69	0.40	1.17	
>21.9 kg/m <sup>2</sup>								0.49

# Author's draft

Populations		Cases, N (%)		Controls, N (%)		OR <sup>a</sup>	Lower 95%CI	Upper 95%CI	P-trend	P for interaction
	I	27	13.5	76	21.7	1.00				
	II	46	23.0	69	19.7	1.57	0.94	2.61		
	III	40	20.0	79	22.6	1.48	0.86	2.53		
	IV	46	23.0	63	18.0	1.58	0.92	2.72		
	V	41	20.5	63	18.0	1.32	0.75	2.33		
Age at menarche <sup>c</sup>										0.96
<13 years									0.64	
	I	30	13.5	59	19.0	1.00				
	II	53	23.9	62	19.9	1.46	0.89	2.40		
	III	44	19.8	65	20.9	1.30	0.77	2.18		
	IV	56	25.2	60	19.3	1.59	0.94	2.69		
	V	39	17.6	65	20.9	1.20	0.68	2.12		
>=13 years									0.18	
	I	35	16.9	83	20.5	1.00				
	II	48	23.2	81	20.0	1.09	0.68	1.73		
	III	43	20.8	80	19.8	0.97	0.60	1.59		
	IV	44	21.3	83	20.5	0.91	0.55	1.50		
	V	37	17.9	78	19.3	0.73	0.43	1.25		
Status of birthplace										0.57
Rural									0.19	
	I	20	22.5	43	23.9	1.00				
	II	26	29.2	35	19.4	1.59	0.83	3.05		
	III	20	22.5	35	19.4	1.39	0.69	2.80		
	IV	9	10.1	25	13.9	0.93	0.38	2.23		
	V	14	15.7	42	23.3	0.69	0.31	1.53		
Urban									0.73	
	I	44	14.2	93	18.8	1.00				
	II	70	22.6	104	21.0	1.19	0.80	1.77		
	III	59	19.0	102	20.6	1.14	0.74	1.74		
	IV	83	26.8	107	21.6	1.27	0.84	1.92		
	V	54	17.4	90	18.1	1.07	0.68	1.69		
Smoking status at baseline										0.99
Ever									0.22	
	I	24	12.4	50	16.7	1.00				
	II	41	21.1	47	15.7	1.41	0.82	2.43		
	III	44	22.7	70	23.3	1.24	0.72	2.12		
	IV	48	24.7	66	22.0	1.17	0.68	2.03		
	V	37	19.1	67	22.3	0.79	0.43	1.44		
Never									0.88	
	I	41	17.4	92	22.1	1.00				
	II	60	25.5	96	23.1	1.19	0.78	1.82		
	III	43	18.3	75	18.0	1.08	0.67	1.73		
	IV	52	22.1	77	18.5	1.17	0.73	1.88		
	V	39	16.6	76	18.3	0.98	0.58	1.65		
Parity										0.92
<3 children									0.46	
	I	42	16.2	84	20.3	1.00				
	II	70	26.9	91	22.0	1.21	0.81	1.82		
	III	48	18.5	80	19.4	1.08	0.69	1.70		
	IV	60	23.1	71	17.2	1.27	0.81	1.99		

# Author's draft

Populations		Cases, N (%)		Controls, N (%)		OR <sup>a</sup>	Lower 95%CI	Upper 95%CI	<i>P</i> -trend	<i>P</i> for interaction
>=3 children	V	40	15.4	87	21.1	0.81	0.49	1.33	0.26	
	I	17	15.6	40	18.8	1.00				
	II	24	22.0	42	19.7	1.43	0.72	2.82		
	III	28	25.7	47	22.1	1.34	0.68	2.63		
	IV	28	25.7	48	22.5	1.08	0.55	2.14		
	V	12	11.0	36	16.9	0.64	0.28	1.47		

<sup>a</sup> adjusted for age, date, department of residence and menopausal status at blood collection or at baseline, existence of a biological sample (blood, saliva, none), family history of breast cancer, age at first full-term pregnancy, recreational physical activity, status of birthplace and breastfeeding

<sup>b</sup> quintiles' cut offs: 1.7E-5, 2.2E-3, 1.9E-2 and 9.6E-2 µg-TEQ/m<sup>2</sup>

<sup>c</sup> median age at first full-term pregnancy based on the distribution among controls

<sup>d</sup> median value of body mass index based on the distribution among controls

<sup>e</sup> modal value of age at menarche based on the distribution among controls

## Author's draft

Table 5. Odds ratios (OR) and 95% confidence intervals (CI) for the association between invasive breast cancer and time-weighted cumulative airborne dioxin exposure (for a 0.1 µg-TEQ/m<sup>2</sup> annual increase) according to time before diagnosis, 1990-2008.

Time <i>t</i> prior to breast cancer diagnosis (years)	Overall, N=1,145		Sensitivity analysis <sup>b</sup> , N=763	
	N cases / N controls	OR (95%CI) <sup>a</sup>	N cases / N controls	OR (95%CI) <sup>a</sup>
5	330 / 524	1.00 (0.94-1.01)	219 / 342	1.06 (1.01-1.15)
10	183 / 306	1.01 (0.92-1.02)	128 / 205	1.03 (1.00-1.07)
15	57 / 113	1.00 (0.93-1.01)	46 / 72	1.01 (1.00-1.05)
18	9 / 20	1.00 (0.96-1.00)	6 / 18	1.01 (1.00-1.02)

<sup>a</sup> adjusted for age, date, department of residence and menopausal status at blood collection or at baseline, existence of a biological sample (blood, saliva, none), family history of breast cancer, age at first full-term pregnancy, recreational physical activity, status of birthplace and breastfeeding

<sup>b</sup> excluding cases and controls for whom the index date is situated within the five first years from the inclusion in the cohort (from 1990 to 1995; N=382, 33%)