



HAL
open science

**Systèmes d'Information Géographique et Lien
Environnement – Santé (SIGLES) : contribution au
développement d'outils cartographiques d'aide à la
décision face aux risques sanitaires liés à l'environnement**

Florent Occelli

► **To cite this version:**

Florent Occelli. Systèmes d'Information Géographique et Lien Environnement – Santé (SIGLES) : contribution au développement d'outils cartographiques d'aide à la décision face aux risques sanitaires liés à l'environnement. Médecine humaine et pathologie. Université du Droit et de la Santé - Lille II, 2014. Français. NNT : 2014LIL2S043 . tel-01221796

HAL Id: tel-01221796

<https://theses.hal.science/tel-01221796v1>

Submitted on 28 Oct 2015

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.

Ecole doctorale 446 Biologie Santé – Lille Nord de France
Année universitaire 2013 – 2014

THESE DE DOCTORAT

Systemes d'Information Géographique et Lien Environnement – Santé SIGLES

**Contribution au développement d'outils cartographiques d'aide à la décision face aux
risques sanitaires liés à l'environnement**

Par Florent OCCELLI

en vue de l'obtention du grade de

DOCTEUR DE L'UNIVERSITE DROIT ET SANTE DE LILLE 2

DISCIPLINE : RECHERCHE CLINIQUE, INNOVATION TECHNOLOGIQUE, SANTE PUBLIQUE

SPECIALITE : SANTE PUBLIQUE

Présentée publiquement le 24 Septembre 2014

Constitution du Jury :

Directeurs de thèse :	Damien CUNY	Professeur Université Lille 2
	Annabelle DERAM	Professeur Université Lille 2
Rapporteurs :	Daniel JOLY	Directeur de Recherches CNRS
	Michel MENCH	Directeur de Recherches INRA
Examineurs :	Alain DUHAMEL	Professeur Université Lille 2 Praticien Hospitalier CHRU Lille
	Cécile COUCHOUD	Coordinatrice Réseau REIN – Agence de Biomédecine
	François GLOWACKI	Praticien Hospitalier CHRU Lille
	Corinne GOWER	Praticien Hospitalier CHRU Lille

REMERCIEMENTS

Ces travaux n'auraient pu être menés à bien sans l'aide de différents financeurs : l'Université Lille 2, l'École Doctorale Biologie Santé de Lille (ED446), le Conseil Régional du Nord – Pas de Calais et l'EA4483 de l'Université Lille 2. Je tiens donc en tout premier lieu à remercier les personnes qui, au travers de leur soutien matériel, ont manifesté un intérêt particulier à mes recherches et m'ont fait confiance.

Je tiens ensuite à exprimer mes plus vifs remerciements aux Professeurs Annabelle Deram et Damien Cuny, pour leur encadrement tout au long de cette thèse et même bien avant. Confier cette vaste et nouvelle thématique de recherche à un jeune étudiant était un pari osé. Pour cela je vous suis reconnaissant. Je mesure également la chance que j'ai d'avoir été guidé, encouragé, conseillé, poussé (parfois menacé ?), mais néanmoins avec une grande liberté par des personnes passionnées de la recherche. Votre dévouement à m'intégrer dans divers projets de recherche et à tout mettre en œuvre pour que je puisse m'épanouir dans cette équipe sur le long terme témoigne la confiance que vous m'accordez. Pour tout cela merci.

J'aimerais remercier M. Daniel Joly, Directeur de Recherches au CNRS et M. Michel Mench, Directeur de Recherches à l'INRA, pour me faire l'honneur de relire ce manuscrit et d'en être rapporteurs.

J'adresse aussi mes remerciements au Pr. Alain Duhamel et aux Dr. Corinne Gower et François-Xavier Glowacki pour leur présence à la fois dans mon jury de comité de suivi de thèse et dans ce jury de soutenance. Je les remercie également pour les collaborations scientifiques que nous avons débutées et qui, je l'espère, vont se poursuivre.

A ce titre, je tiens aussi à remercier le Dr. Cécile Couchoud, qui me fait l'honneur de participer au jury de soutenance.

J'en profite également pour remercier l'ensemble des personnes ayant participé à mes comités de suivi de thèse et que je n'ai pas encore cité. Pr. Denis Bard, Dr. Sarah Goria, Dr. Karine Ligier, Dr. Valérie Mitchell, Dr. Catherine Nisse, Dr. Corinne Schadkowski, Dr. Julien Caudeville, vos critiques et suggestions ont toujours été avisées et m'ont permis de mieux cerner la portée de mon sujet.

La particularité de l'étude des liens santé – environnement et plus particulièrement à l'aide des Systèmes d'Information Géographique, est certainement sa pluridisciplinarité. Ce sujet de recherche m'a donc offert l'opportunité de rencontrer un grand nombre de personnes de divers horizons (toxicologues, écotoxicologues, médecins, épidémiologistes, biostatisticiens, mathématiciens, sigistes, géostatisticiens, gestionnaires de collectivités territoriales), qui m'ont toutes apporté de nouveaux savoirs et un grand nombre de perspectives. Je tiens donc à remercier toutes ces personnes qui ont enrichi mes travaux.

Je vais porter une attention particulière au Dr. Michaël Génin pour m'avoir apporté son expertise dans les statistiques spatiales de santé, mais aussi pour toutes nos longues discussions que je nommerai « de parvis » sur la thèse et l'après-thèse, et maintenant la maison ! Ma deuxième attention se porte vers Rachel Bavdek, avec qui j'ai récemment commencé à travailler et rédiger. Merci pour ta réactivité ces derniers temps.

Je tiens maintenant à mentionner le plaisir que j'ai eu (et que j'aurai !!) à travailler au sein du LSVF. Je tiens à tous vous remercier, pour votre aide et votre soutien, mais aussi pour vos moqueries sur le fait que « je joue aux cartes toute la journée » et que je supporte les Sang et Or. Merci à Audrey, Fabien, Franck, Inès, Julie (tic tac tic tac... ton tour viendra...), Ludivine, Pierre-Arthur, Régis, Stéphane et Sylvain.

Je remercie également les stagiaires avec qui j'ai eu la chance de travailler pendant ces trois années, Céline, Matthieu, Antoine et Théo.

Y paraît qu'ch'est lô qu'in peut s'lâcher dins eune thèse... Alors te croyos que j't'avos oublié hein ! J'voulos faire un tchio qué qu'cosse pour ti, qui conno ch'patos par cœur à c't'heure. Un grind merci à ti d'avoir caché après ches fautes dins c'que j'avos bavé. T'as paché grindmint d'timps à raviser mes bafouilles et même chi t'avos mau à t'tiète, te t'es nin laiché abatte. Ch'us fin bénache eud cha ! Vu que j'vas cor ouvrir ichi un momint, j'vas pouvoir t'vire braire avec tes zis tous fraique eul midi et faire la goulaffe avec ches chuques.

Mes remerciements vont aussi à la famille, la belle famille (merci pour tout Véro...), mes amis footix Kaisers du FC Douliou et les collègues de l'ILIS, notamment les ex lofteurs du 26 (si je fais un double, je fais un Doctorat !)... J'ai à cette heure une grosse pensée pour ma binôme Mely ! On aura été au bout ! YEPEE YA YA YEPEE YEPEE YEAH !

C'est difficile de trouver les mots pour te dire à quel point je suis fier de ce que nous sommes en train de construire (au sens propre, comme au figuré) et heureux que tu m'aies suivi et soutenu dans tous mes choix, encouragé pendant ces trois dernières années et supporté ces dernières semaines. Alors juste, « parce que c'est toi... »

RESUME

La question des inégalités environnementales et sociales de santé (IESS) à l'échelle des territoires est rapportée à deux dimensions cumulatives : un niveau d'exposition des populations à un environnement de vie dégradé et un niveau de vulnérabilité face à ces facteurs de risques, capables d'affecter la santé. Ces travaux de recherche concernent les Systèmes d'Information Géographique (SIG) appliqués au domaine de la santé environnementale. Ils ont pour objectifs généraux la caractérisation de la qualité des milieux environnementaux et l'évaluation des IESS.

L'atteinte de ces objectifs passe préalablement par la récolte, la genèse et la mise en forme de bases de données environnementales spatialisées. Ces bases issues de la surveillance physico-chimique et de la biosurveillance des milieux sont ensuite cartographiées à l'aide des outils SIG, notamment les méthodes géostatistiques d'interpolation spatiale. En parallèle, les disparités spatiales d'incidence de pathologies sont étudiées grâce à des méthodes de cartographie des maladies (Ratios d'Incidence Standardisés : SIR) et de détection de clusters atypiques d'évènements (statistique de scan) mises en œuvre sur la base de registres de maladies. Finalement des études écologiques spatialisées sont élaborées afin de rapprocher les cartographies environnementales générées aux indicateurs sanitaires et socio-économiques. Ces travaux s'attachent à répondre à la question « les populations présentant un état de santé dégradé vivent-elles dans un environnement dégradé ? » Ce questionnement a été étudié à travers trois axes de recherche.

Le premier concerne la caractérisation de l'imprégnation de l'environnement par les éléments traces métalliques (ETM) et l'évaluation des IESS à fine résolution sur trois bassins de vie de la région Nord-Pas de Calais (NPdC). Cette recherche est menée à partir de mesures d'imprégnations biologiques réalisées à la fois chez des lichens épiphytes et chez l'Homme et d'un indicateur de défaveur localisé. Les ETM mesurés ont été considérés de manière individuelle, mais également de manière globale à travers le développement d'un indice intégré multimétallique, afin d'exprimer le niveau d'imprégnation général par plusieurs métaux. Des inégalités environnementales ont été mises en évidence à l'échelle des quartiers du bassin de Dunkerque. Nos résultats supposent également l'implication de l'environnement dans l'imprégnation des populations pour certains ETM.

Dans le deuxième axe, nous avons mis en évidence des disparités spatiales d'incidence de l'insuffisance rénale chronique terminale entre les cantons du NPdC et souligné l'implication du statut socio-économique dans l'apparition de ces disparités, contrairement à d'autres facteurs (diabète, maladies cardiovasculaires, pratiques médicales). Seule une partie des disparités de cette pathologie est aujourd'hui expliquée. Il est donc nécessaire de poursuivre sur le rôle des contaminants de l'environnement.

Enfin, le troisième axe s'intéresse à l'analyse spatiale et spatio-temporelle de contaminations des eaux souterraines par les ETM, en vue d'identifier de potentiels facteurs de risques environnementaux des maladies inflammatoires chroniques de l'intestin.

La réalisation de ces travaux s'appuie sur plusieurs collaborations établies avec le Réseau REIN, le Registre EPIMAD et des équipes de recherche (EA4483 et EA2694 Université Lille 2, TVES EA4477 ULCO).

A la suite de ces travaux de thèse, les perspectives de recherche sont de poursuivre le développement d'indicateurs intégrés d'exposition aux contaminations multiples des milieux environnementaux. Les résultats montrent également un manque de renseignement des bases de données environnementales par rapport aux registres de santé. Un travail est donc nécessaire afin de définir le contenu de ces bases, nécessaire pour caractériser la qualité des milieux et faciliter l'étude de l'interaction de l'Homme avec son environnement.

Mots clés : biosurveillance, bases de données, cartographie, géostatistiques, indice d'imprégnation, santé environnementale, statistique de scan, systèmes d'information géographique

ABSTRACT

GEOGRAPHIC INFORMATION SYSTEMS AND ENVIRONMENTAL HEALTH

Environmental and social inequalities in health (ESIH) over territories are related to two cumulative dimensions: populations exposed to their living poor quality environment and the vulnerability of these populations to the environmental risk factors, which can affect health. This research deals with the Geographic Information Systems (GIS) applied to the field of environmental health. General purposes are the characterization of environmental media quality and the assessment of ESIH.

Achieving these objectives requires a first step of harvest, genesis and formatting spatialized environmental databases. Such data are resulting from physico-chemical monitoring and biomonitoring. They were then mapped using GIS tools, including geostatistical spatial interpolation methods. On the other hand, spatial variability in the incidence of diseases were investigated using disease mapping methods (Standardized Incidence Ratios: SIR) and the detection of atypical clusters of events (scan statistics), which are based on disease registries. Finally, geographical ecological studies are developed to associate the environmental maps generated to health and socio-economic status. Thus, this work aims to answer the question "do people with poor state of health live in a poor quality environment?" This question has been studied through three main researches.

The first relates the characterization of trace elements burdens in the environment and the assessment of ESIH on neighborhood scale, over three territories in the Nord-Pas de Calais (NPdC) region. This research is conducted from measurements of biological burdens performed both in epiphytic lichens and humans and from a localized index of deprivation. The measured metals were considered individually, but also holistically by developing an integrated multimetallic index, in order to describe the general status of environmental pollution by metals. Environmental inequalities were observed on neighborhood scale in Dunkerque. Our results assume that trace elements burdens in populations are affected by environmental burdens.

In our second research we revealed spatial disparities in the incidence of end stage renal disease (ESRD) on small area in the NPdC. Unlike other factors (diabetes, cardiovascular disease, medical practices), we highlighted the role of socio-economic status in the occurrence of such disparities. Only a part of the ESRD variability is currently explained. It is therefore necessary to focus on the environmental hypothesis.

The third research focuses on the spatial and spatio-temporal analysis of groundwater contamination by trace elements, in order to identify potential environmental risk factors in the incidence of chronic inflammatory bowel disease.

This work is based on several collaborations with the REIN network, the EPIMAD registry, and several research teams (EA4483 and EA2694 Université Lille 2, TVES EA4477 ULCO).

Following this thesis, the research prospects are to pursue the development of integrated indicators to assess population exposure to the multiple environmental media contamination. The results also indicate a lack of information in environmental databases compared to health registries. A work is thus needed to define the content of such databases. These are necessary to characterize the environmental quality and to help the assessment of interaction between the populations and their living environment.

Key words: biomonitoring, database, environmental health, geographic information systems, geostatistics, impregnation index, mapping, scan statistics

VALORISATION DES TRAVAUX DE THESE

PUBLICATIONS

Article publié

Florent Ocelli, Marie-Amélie Cuny, Inès Devred, Annabelle Deram, Stéphanie Quarré et Damien Cuny « Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques », *Pollution atmosphérique* [En ligne], N° 220, mis à jour le : 13/01/2014, URL : <http://lodel.irevues.inist.fr/pollution-atmospherique/index.php?id=2497>.

Article en révision

Ocelli F, Deram A, Génin M, Cuny D, Glowacki F-X and Néphronor Network. Mapping end-stage renal disease (ESRD): spatial variations on small area level in France, and association with deprivation, earlier renal replacement care and related diseases. *Plos One* (soumis le 13.02.2014).

Articles à soumettre

Ocelli F, Bavdek R, Deram A, Hellequin A-P, Cuny M-A, Cuny D and Zwarterook I. Using lichen biomonitoring to assess environmental justice over neighborhoods in an industrial area of Northern France. *Environmental Health Perspectives* (A soumettre).

Ocelli F, Génin M, Deram A, Glowacki F-X, Cuny D and Néphronor Network. End-Stage Renal Disease in Northern France: isotonic cluster detection. *Nephrology Dialysis Transplantation* (A soumettre).

COMMUNICATIONS ORALES

Ocelli F, Deram A, Génin M, Glowacki F-X, Cuny D et Réseau Néphronor. Disparités spatiales d'incidence de l'insuffisance rénale chronique terminale (IRCT) en Nord – Pas de Calais en lien avec la défaveur sociale. 16ème réunion commune Société Francophone de Dialyse & Société de Néphrologie. Du 30 septembre au 3 octobre 2014 (Communication orale).

Ocelli F, Bavdek R, Deram A, , Cuny M-A, Dallery A, Leroyer A, Nisse C, Hellequin A-P, Cuny D. Atmospheric trace element deposition in epiphytic lichens, deprivation and human biomonitoring: an assessment of environmental and social health inequalities in the Nord-Pas-de-Calais region, France. Workshop International "Biosurveillance végétale et fongique de la Qualité de l'Air". 13-14 octobre 2014 (Communication orale).

Ocelli F, Bavdek R, Deram A, Cuny M-A, Dallery T, Hellequin A-P, Cuny D. Analyse spatiale de l'imprégnation des lichens épiphytes par les éléments traces métalliques en région Nord – Pas de Calais et relation à la défaveur sociale à fine échelle. Journées Interdisciplinaires de la Qualité de l'Air. 10-11 février 2014 (Communication orale).

Ocelli F, Cuny M-A, Deram A, Cuny D. Analyse spatiale de l'imprégnation des lichens épiphytes par les éléments traces métalliques en région Nord – Pas de Calais. SIG 2013 la conférence francophone ESRI, 2-3 octobre 2013, Versailles (Communication orale).

Ocelli F, Deram A, Cuny D. Maladies Inflammatoires Chroniques de l'Intestin & Facteurs de risques environnementaux – Caractérisation de la qualité des milieux environnementaux. Journée commune MICI et environnement. 26 septembre 2013 (Communication orale).

Ocelli F, Deram A, Glowacki F-X, Cuny D. Mapping chronic kidney disease: spatial variations and cluster detection at a regional scale. 13ème Journée André Verbert EDBS Lille, 18 septembre 2013, Lille (Communication orale).

Ocelli F, Cuny M-A, Deram A, Cuny D. Système d'Information Géographique et Lien Environnement – Santé (SIGLES) – Application à l'évaluation de l'imprégnation en éléments traces métalliques à l'aide des lichens épiphytes. Journée de restitution IRENI Vulnérabilités et perceptions des populations à la pollution atmosphérique, 24 mai 2013, Dunkerque (Communication orale).

Ocelli F, Deram A, Glowacki F-X, Cuny D. Analyse cartographique de l'incidence de néphropathie dans le Nord – Pas de Calais : le Registre REIN. Conférence du Groupement des Néphrologues Francophones de Belgique, 7 mars, 2013, Tournai, Belgique (Communication orale).

COMMUNICATION AFFICHEE

Ocelli F, Cuny M-A, Deram A, Cuny D. Croisement de données environnementales et sanitaires grâce à un Système d'Information Géographique en vue d'évaluer les liens entre la pollution atmosphérique et la santé humaine à l'échelle de la région Nord – Pas de Calais. Journées Interdisciplinaires de la Qualité de l'Air. 2-3 février 2012 (Poster).

SOMMAIRE

Remerciements	3
Résumé.....	5
Abstract	6
Valorisation des travaux de thèse	7
Sommaire	9
Table des figures.....	11
Liste des tableaux.....	14
Glossaire.....	16
Liste des abréviations.....	17
Contexte et problématique.....	19
Introduction générale.....	24
Partie 1. Les Systèmes d'Information Géographique (SIG).....	24
Partie 2. L'interpolation spatiale.....	27
Partie 3. L'épidémiologie spatiale.....	35
Partie 4. La biosurveillance lichénique de la qualité de l'air.....	39
Partie 5. Les indices intégrés de qualité des milieux	47
Partie 6. Les études de corrélation écologique	52
Matériel et méthodes.....	64
Partie 1. La région Nord – Pas de Calais.....	64
Partie 2. Ressources.....	71
Partie 3. Matériels et méthodes du chapitre I.....	76
Partie 4. Matériels et méthodes du chapitre II.....	80
Partie 5. Matériels et méthodes du chapitre III.....	83
Chapitre I. Inégalités environnementales et sociales de santé à fine échelle en région Nord – Pas de Calais.....	88
Partie 1. Introduction.....	88
Partie 2. Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques	90
Partie 3. Inégalités environnementales	113
Partie 4. Inégalités environnementales de santé	133
Partie 5. Conclusion et perspectives du chapitre I	145
Chapitre II. Disparités spatiales d'incidence de l'insuffisance rénale chronique terminale en Nord – Pas de Calais	150
Partie 1. Introduction.....	150
Partie 2. Disparités spatiales d'incidence de l'IRCT en Nord – Pas de Calais en lien avec la défaveur sociale	151

Partie 3.	Détection de clusters atypiques d'IRCT en Nord – Pas de Calais	166
Partie 4.	Conclusion et perspectives du chapitre II	180
Chapitre III.	Facteurs de risques environnementaux des maladies inflammatoires chroniques de l'intestin (MICI)	184
Partie 1.	Introduction.....	184
Partie 2.	Optimisation d'une base de données environnementale en vue de construire un outil cartographique d'aide à la décision – cas de la contamination des eaux souterraines.....	187
Partie 3.	Analyse spatiale et spatio-temporelle de la contamination en aluminium dans les eaux souterraines.....	199
Partie 4.	ADES comme outil d'aide à la décision en santé – environnement dans le contexte des MICI	210
Partie 5.	Discussion générale et perspectives du chapitre III	214
Discussion générale		222
Références bibliographiques		236
Annexes		I

TABLE DES FIGURES

Figure 1. Couches d'informations spatialisées au sein d'un SIG (modifié à partir d'ESRI, 2014).....	25
Figure 2. Principales règles de sémiologie graphique (Rican, 2007)	27
Figure 3. Campagne de mesure du NO ₂ atmosphérique (en µg/m ³) sur la communauté urbaine de Lille en été 2003 par le réseau ATMO Nord – Pas de Calais (modifié à partir de Declercq, 2007)	28
Figure 4. Schéma simplifié de l'interpolation spatiale	28
Figure 5. Estimation de la concentration en dioxyde d'azote en Alsace (A) par plus proche voisin, (B) par moyenne mobile, (C) par inverse des distances et (D) par inverse des distances au carré (modifié à partir de Lemarchand & Jeannée, 2009)	29
Figure 6. Exemple de variogramme expérimental (logiciel ArcGIS®) reprenant les valeurs des paires de points (rouge), leur moyenne (croix bleue) et le variogramme modélisé (ligne bleue)	31
Figure 7. Résultats du krigeage (logiciel ArcGIS®). La cartographie de l'estimation de la variable aléatoire (en haut) et la cartographie de l'erreur d'estimation (en bas)	32
Figure 8. Spectre de variogrammes modélisés par krigeage empirique bayésien (en bleu) – les lignes rouges représentent la médiane (ligne continue), les 25 ^e et 75 ^e percentiles (pointillés) de la distribution des variogrammes (logiciel ArcGIS®)	33
Figure 9. Résultat du krigeage empirique bayésien pour les concentrations en radiocésium dans les sols de la région de Fukushima Daiichi au Japon (à gauche : Krivoruchko, 2012) et « La corbeille de fruits » de Giuseppe Arcimboldo (à droite : ca. 1527–1593)	34
Figure 10. Instabilité du SIR en fonction du nombre de cas attendus (logiciel R).....	37
Figure 11. Un petit échantillon des nombreuses fenêtres de scan utilisées par SaTScan™ (modifié à partir de Kulldorff, 1999).....	38
Figure 12. Echelle d'estimation de la qualité de l'air en Angleterre et au Pays de Galles à partir des lichens observés sur des écorces non eutrophisées et cartographie réalisée à partir de cette échelle (Hawksworth et Rose, 1970) (Gavériaux, 2005).....	42
Figure 13. Evolution de la qualité de l'air entre 1973 et 1992 dans le Nord – Pas de Calais (Van Haluwyn, 1993).....	43
Figure 14. Distribution des concentrations en cadmium, plomb et arsenic chez le lichen <i>Xanthoria parietina</i> sur la région de Dunkerque (a) en 1995 et (b) en 2002 (modifié à partir de Cuny <i>et al.</i> , 2004)	44
Figure 15. Cartographies des indices de diversité lichénique (à gauche) et de la mortalité par cancers pulmonaires chez les hommes de moins de 55 ans (à droite) en Vénétie, Nord de l'Italie (Cislaghi & Nimis, 1997).....	45
Figure 16. Corrélation entre les concentrations de cérium et les infections nasales (à gauche) et les concentrations de sélénium et le cancer de la prostate (à droite). Concernant le sélénium, des relations similaires ont été décrites avec les pathologies du système circulatoire et digestif (Wolterbeck & Verburg, 2004)	46
Figure 17. Densité de population par commune en Nord – Pas de Calais (logiciel ArcGIS®).....	65
Figure 18. Occupation du sol en Nord – Pas de Calais (logiciel ArcGIS®)	65
Figure 19. Localisation des sols potentiellement pollués inventoriés (DREAL, 2014a).....	67
Figure 20. État chimique des eaux de surface en 2011 (DREAL, 2014a).....	67

Figure 21. Nombre de jours de dépassements de la valeur limite journalière de 50 µg/m ³ pour les PM10 en 2011 (ATMO Nord – Pas de Calais)	69
Figure 22. Indice de défaveur sociale de Townsend par commune en 2006 (ORS Nord – Pas de Calais, 2011)	69
Figure 23. Mortalité prématurée (avant 65 ans) par zone d’emploi, deux sexes confondus – un ICM supérieur à 100 indique une surmortalité par rapport à la population de référence – un ICM de 130 implique une surmortalité de 30% (logiciel ArcGIS®).....	70
Figure 24. Diffusion cartographique du RMQS (INDIQUASOL, 2014)	74
Figure 25. Variabilité des RIM en fonction du nombre d’éléments traces métalliques intégrés au calcul, sous forme de box plot – les indices sont calculés à partir des 15, 10 et 5 éléments les plus corrélés (a) ou les moins corrélés (b) au RIM18 – les différences entre les RIM sont évaluées à l’aide du test de Dunn (p<0.001)	111
Figure 26. Corrélacion entre le RIM et le PLI pour les trois bassins (échelle log10).....	112
Figure 27. Exemple de distribution statistique des concentrations (en µg/L) dans les urines pour l’aluminium (Alu).....	135
Figure 28. Risque relatif lissé de la MC (à gauche) et la RCH (à droite) par canton dans le nord de la France entre 1990 et 2003 (Declercq <i>et al.</i> , 2010).....	185
Figure 29. Risque relatif de la maladie de Crohn dans le nord de la France entre 1990 et 2006 – clusters spatiaux constants pendant la période (à gauche) et clusters spatio-temporels non constants pendant la période (à droite) (Génin <i>et al.</i> , 2013)	185
Figure 30. La zone d’étude, composée des départements du Nord, du Pas de Calais, de la Somme et de la Seine-Maritime (Chouraki <i>et al.</i> , 2011)	188
Figure 31. Variabilité des LDD ou LDQ (en µg/L d’aluminium) en fonction du laboratoire d’analyse	191
Figure 32. Variabilité des LDD et LDQ (en µg/L d’aluminium) en fonction du temps.....	192
Figure 33. Distribution des résultats d’analyse en aluminium (µg/L) dans les eaux souterraines (base B3)	194
Figure 34. Distribution des résultats d’analyse en aluminium (µg/L) dans les eaux souterraines (base B4 validée).....	195
Figure 35. Répartition spatiale des points d’eau faisant l’objet d’un suivi en aluminium.....	196
Figure 36. Nombre d’analyses effectuées par année en fonction du territoire étudié (source des données : ADES Eau France, 2014).....	197
Figure 37. Concentration moyenne en aluminium par point d’eau entre 1990 et 2012 (µg/L) par cartographie ponctuelle – les classes sont déterminées à l’aide des seuils de Jenks – les points A à H sont supérieurs à 200 µg/L – les points N et S ont des pollutions ponctuelles sur les départements du Nord, de Pas de Calais et de la Somme	200
Figure 38. Distribution des concentrations moyennes en aluminium par point d’eau entre 1990 et 2012 (µg/L)	201
Figure 39. Concentration moyenne en aluminium par point d’eau entre 1990 et 2012 (µg/L) par cartographie géostatistique (logiciel ArcGIS®)	203
Figure 40. Erreur standard de prédiction de la concentration moyenne en aluminium entre 1990 et 2012 (µg/L) (logiciel ArcGIS®).....	203
Figure 41. Suivi temporel des concentrations en aluminium dans les eaux souterraines (ADES Eau France, 2014).....	206

Figure 42. Variabilité spatiale et temporelle de l'échantillonnage et des concentrations en aluminium dans les eaux souterraines (suite).....	209
Figure 43. Concentration moyenne en aluminium par point d'eau (µg/L) entre 1990 et 2012 (A), clusters de MC entre 1990 et 2006 (B : repris à partir de Génin <i>et al.</i> , 2013) et superposition des deux cartes (C)	212
Figure 44. RIM par point d'eau entre 1990 et 2012 (A), clusters de MC entre 1990 et 2006 (B : repris à partir de Génin <i>et al.</i> , 2013) et superposition des deux cartes (C)	213
Figure 45. Arbre de décision pour l'utilisation et l'interprétation des données de qualitomètres de la banque ADES dans le cadre d'une analyse spatiale et spatio-temporelle	216
Figure 46. Réseaux de contrôle opérationnel et de surveillance des masses d'eau souterraines (www.artois-picardie.eaufrance.fr).....	217
Figure 47. Principaux transferts d'eau potable sur le bassin Artois – Picardie (ADES Eau France, 2014)	219

LISTE DES TABLEAUX

Tableau 1. Méthodes de discrétisation de variables continues en vue d'une analyse spatiale (Goria <i>et al.</i> , 2011)	26
Tableau 2. Diversité des indices intégrés de qualité des milieux environnementaux – avec C_i (concentration mesurée du paramètre i), B_i (valeur de fond, norme, etc.), T_i (facteur de toxicité ou écotoxicité), P_i (indice de pollution du paramètre i), W_i (poids attribué au paramètre i), $MaxP_i$ (P_i maximal observé sur l'ensemble des paramètres intégrés), \bar{P}_i (moyenne des P_i) et n (nombre de paramètres intégrés).....	48
Tableau 3. Indice ATMO synonyme d'un état de la qualité de l'air en France (Fédération ATMO France, 2014).....	49
Tableau 4. Ressources de données sanitaires récoltées lors de la thèse.....	71
Tableau 5. Ressources de données environnementales existantes sur le territoire de la région Nord – Pas de Calais	73
Tableau 6. Valeurs référence (en mg/kg poids sec) des concentrations en ETM dans les lichens ...	79
Tableau 7. Normes environnementales utilisées pour le calcul du RIM eaux souterraines (Arrêté du 11/01/07)	84
Tableau 8. Matrice de corrélation de Spearman entre les 18 métaux et le RIM pour les trois bassins – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 5%. Les valeurs en vert sont ≥ 0.75	108
Tableau 9. Classement des éléments en fonction de leur degré de corrélation (coefficient de Spearman) avec le RIM, concentration moyenne de chaque élément dans les lichens (en mg/kg) et ratio d'imprégnation monométallique.....	108
Tableau 10. Classement des éléments en fonction de leur niveau de corrélation (coefficient de Spearman) avec le RIM – analyse par bassin – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 5%.....	110
Tableau 11. Valeurs de référence utilisées pour le calcul des indices multimétalliques d'imprégnation	134
Tableau 12. Résultats du test Mann et Whitney de comparaison de moyennes pour les ETM sanguins (s) et urinaires (u) entre les fumeurs et les non-fumeurs (NS : non-significatif ; *** : significatif) – les valeurs en gras sont significativement supérieures, à un risque alpha de 0.05	136
Tableau 13. Résultats du test Mann et Whitney de comparaison de moyennes pour les ETM sanguins (s) et urinaires (u) entre les moins de 30 ans et les plus de 50 ans (NS : non-significatif ; *** : significatif) – les valeurs en gras sont significativement supérieures, à un risque alpha de 0.05.....	138
Tableau 14. Influence de l'âge et du tabac sur les concentrations sanguines et urinaires en métaux (NS : non significatif ; + : augmentation significative des concentrations ; - : diminution significative des concentrations).....	138
Tableau 15. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) et urinaires (u) pour l'ensemble des individus – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.a	139

Tableau 16. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour l'ensemble des individus – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.b	140
Tableau 17. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus non-fumeurs – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.c	140
Tableau 18. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus de moins de 40 ans (A) et de 40 ans et plus (B) – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.d	141
Tableau 19. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus en fonction des quatre groupes de risque (A : non-fumeur et moins de 40 ans et B : non-fumeur et plus de 40 ans) – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.e	141
Tableau 20. Synthèse des résultats d'analyse de corrélation entre les dosages sanguins (s) ou urinaires (u) et les dosages dans les lichens en fonction des différents groupes de population – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05	142
Tableau 21. Sélection des points d'eau en fonction du mode d'utilisation	190
Tableau 22. Couples de valeurs statut/qualification autorisés dans le schéma Sandre et pourcentage de valeurs correspondantes dans la base aluminium	193
Tableau 23. Synthèse des traitements de la base aluminium	194
Tableau 24. Statistiques descriptives de la distribution des résultats d'analyse en aluminium dans les eaux souterraines ($\mu\text{g/L}$)	196
Tableau 25. Résultats du test multiple de Dunn pour la comparaison des concentrations en aluminium entre les départements (niveau de signification : $p < 0.0083$)	201
Tableau 26. Statistiques descriptives de la distribution du nombre d'analyses en aluminium effectuées au niveau des points d'eau	204
Tableau 27. Normes environnementales utilisées pour le calcul du RIM eaux souterraines (Arrêté du 11/01/07)	211

GLOSSAIRE

Agrégation spatiale : transformation de données pour obtenir une valeur par unité spatiale

Autocorrélation spatiale : phénomène qui implique que les entités géographiquement proches se ressemblent

Biométrie : mesure dans les liquides biologiques

Choroplethe (carte) : représentation cartographique de quantités relatives à des espaces

Cluster : concentration d'évènements anormalement faible ou élevée par rapport à celle attendue sur la zone

Discrétisation : division de la série statistique en classes ou intervalles

Géocodage : attribution de coordonnées spatiales

Géomatique : science des données à référence spatiale

Géoréférencement : attribution d'une position à la surface de la Terre

Géostatistiques (données) : informations quantitatives mesurées au niveau de points d'échantillonnage géoréférencés

Géostatistique (méthode) : méthodes empiriques d'interpolation spatiale, comme le krigeage

Hétérogénéité spatiale : variabilité dans l'espace

Inégalité environnementale : populations défavorisées et exposées à un environnement dégradé

Inégalité environnementale de santé : populations exposées à un environnement dégradé et présentant un état de santé dégradé

Inégalité environnementale et sociale de santé : populations défavorisées, exposées à un environnement dégradé et présentant un état de santé dégradé

Interpolation spatiale : estimation de la valeur d'une variable en tous points de l'espace à partir d'un échantillon de points mesurés

Krigeage : méthode géostatistique d'interpolation spatiale

Lissage (modèle de) : modèle mathématique permettant de prendre en compte le phénomène d'autocorrélation spatiale lors de la cartographie des maladies

Raster : image géoréférencée

Sémiologie graphique : équilibre des formes, tailles, couleurs et grains pour la représentation cartographique

Shape : format utilisé pour la diffusion de fichiers SIG (.shp)

Variogramme : fonction mathématique utilisée en géostatistique, en particulier pour le krigeage

Vecteur : forme ou objet géoréférencé(e)

LISTE DES ABREVIATIONS

AASQA : Associations Agréées pour la Surveillance de la Qualité de l’Air
ADEME : Agence de l’Environnement et de la Maîtrise de l’Energie
ADES : Accès aux Données sur les Eaux Souterraines
Al : aluminium
API : Air Pollution Index
APPA-NPC : Association pour la Prévention de la Pollution Atmosphérique Nord – Pas de Calais
AQI : Air Quality Index
As : arsenic
BASIAS : Base des Anciens Sites Industriels et Activités de Service
BASOL : Base des Sites et sols pollués
Be : beryllium
BRGM : Bureau de Recherches Géologiques et Minières
BSS : Banque du Sous-Sol
Cd : cadmium
CERESTE : Centre de Recherches en Santé - Travail - Environnement
CGDD : Commissariat Général au Développement Durable
CITEPA : Centre Interprofessionnel Technique d’Etudes de la Pollution Atmosphérique
CMU : Couverture Maladie Universelle de base
CNIL : Commission Nationale de l’Informatique et des Libertés
Co : cobalt
CO : monoxyde de carbone
Cr : chrome
Cu : cuivre
DC : Degree of Contamination
DREAL : Direction Régionale de l’Environnement de l’Aménagement et du Logement
EBK : Krigeage Empirique Bayésien
EPIMAD : Registre des Maladies Inflammatoires Chroniques du Tube Digestif
ERS : Evaluation des Risques Sanitaires
ETM : Elément Trace Métallique
HAP : Hydrocarbures Aromatiques Polycycliques
Hg : mercure
IARC : International Agency for Research on Cancer
IC : Intervalle de Confiance
ICMM : Indicateur de Charge Multi-Métallique
ICPE : Installations Classées pour la Protection de l’Environnement
IDL : Indice de Défavor Localisé
IGN : Institut National de l’Information Géographique et Forestière
IMEPOGE : Imprégnation par les MEtaux de la POPulation GENérale
INERIS : Institut National de l’EnviRonnement Industriel et des Risques
INSEE : Institut National de Statistique et d’Etudes Economiques
InVS : Institut de Veille Sanitaire
Ipg : Indice de pollution global
IPI : Integrated Pollution Index
IRCT : Insuffisance Rénale Chronique Terminale
IRENI : Institut de Recherche en Environnement Industriel
IRIS : Îlots Regroupés pour l’Information Statistique
LDD : Limite De Détection
LDQ : Limite De Quantification
LMCU : Lille Métropole Communauté Urbaine

LSVF : Laboratoire des Sciences Végétales et Fongiques
MC : Maladie de Crohn
MICI : Maladies Inflammatoires Chroniques de l'Intestin
Mn : manganèse
MONICA : Multinational MONItoring of trends and determinants in Cardiovascular disease
MPI : Metal Pollution Index
MSPI : Marine Sediment Pollution Index
Ni : nickel
NO₂ : dioxyde d'azote
NPC : Nord – Pas de Calais
O₃ : ozone
OIP : Overall Index of Pollution
OMS : Organisation Mondiale de la Santé
Pb : plomb
PCB : PolyChloroBiphényles
Pd : palladium
PIN : New Pollution Index
PLI : Pollution Load Index
PM_{2,5} : particules fines inférieures à 2,5 µm
PM₁₀ : particules fines inférieures à 10 µm
PNSE : Plan National Santé – Environnement
PPA : Plan de Protection de l'Atmosphère
PRSE : Plan Régional Santé – Environnement
Pt : platine
RCH : Rectocolite Hémostatique
RECORD : REcherche COopérative sur les Déchets et l'Environnement
REIN : Réseau Epidémiologie et Information en Néphrologie
Rh : rhodium
RI : Ratio d'Imprégnation
RIM : Ratio d'Imprégnation Moyen
RMQS : Réseau de Mesure de la Qualité des Sols
RPM : Ratio de Pollution Moyen
RR : Risque Relatif
Sb : antimoine
SDAGE : Schéma Directeur d'Aménagement et de Gestion des Eaux
SEQ : Système d'Evaluation de la Qualité des cours d'eau
SGBD : Système de Gestion de Bases de Données
SGMAS : Secrétariat Général des Ministères Chargés des Affaires Sociales
SIG : Systèmes d'Information Géographique
SIGLES : Systèmes d'Information Géographique et Lien Environnement - Santé
SIR : Ratio d'Incidence Standardisé
SO₂ : dioxyde d'azote
Ti : titane
Tis : Taux comparatif d'Incidence
TVES : Territoires Villes Environnement et Société
V : vanadium
WPI : Water Pollution Index
WQI : Water Quality Index
Zn : zinc

CONTEXTE ET PROBLEMATIQUE

L'état de santé d'une population peut être affecté par un ensemble complexe de déterminants individuels et de facteurs sociaux et environnementaux. Dans un rapport du Secrétariat Général des Ministères chargés des Affaires Sociales (SGMAS) et du Commissariat Général au Développement Durable (CGDD), la question des inégalités environnementales et sociales de santé, à l'échelle des territoires, est rapportée à deux dimensions cumulatives : un niveau d'exposition des populations à un environnement de vie dégradé et un niveau de vulnérabilité face à ces facteurs de risques, capables d'affecter la santé (Besse *et al.*, 2014). Ainsi, les populations faisant face à une précarité plus prononcée seraient à la fois plus exposées à des nuisances environnementales et plus sensibles aux effets sanitaires résultants (OMS, 2012). En France, le second Plan National Santé Environnement (PNSE2, 2009) fait de la lutte contre ces disparités territoriales en termes d'inégalités environnementales et sociales de santé une priorité.

La santé environnementale, telle qu'elle est définie dans le premier PNSE (PNSE1, 2004) constitue « l'ensemble des interactions entre l'Homme et son environnement et les effets sur la santé liés aux conditions de vie (expositions liées à la vie privée et/ou professionnelle...) et à la contamination des différents milieux (eau, air, sol...) ». La notion d'environnement renvoie alors aux différents milieux de vie des populations et écarte ce qui relève de l'exposition volontaire et du comportement individuel comme le tabagisme, la consommation d'alcool ou les habitudes alimentaires. Le milieu de vie peut être constitué de l'environnement domestique, extérieur, mais aussi professionnel. Dans le cadre de ce travail, nous nous intéressons à la population générale telle que le définit le code de santé publique. Par conséquent, **nous considérons l'environnement essentiellement en tant que milieu de vie extérieur (air, eau, sols)**. Celui-ci peut également être dénommé comme environnement général ou naturel et relève du Code de l'environnement. Nous écartons ainsi l'exposition professionnelle, régie par la prévention des risques au travail (Code du travail), et les expositions à l'intérieur des locaux.

L'impact que peuvent avoir les différentes composantes de l'environnement, c'est-à-dire la qualité de l'air, la qualité de l'eau et la qualité des sols, sur la santé est reconnu depuis de nombreuses années. De manière générale, l'Organisation Mondiale de la Santé (OMS) estimait en 2006 à 17% la part de décès prématurés imputables à l'environnement au sein des pays développés (Prüss-Üstün & Corvalán, 2006). Concernant la pollution atmosphérique, l'OMS indique dans son dernier rapport que près de 7 millions de personnes sont décédées prématurément dans le monde en 2012 – soit 6,5 % des décès – du fait de l'exposition à la pollution de l'air intérieur et/ou extérieur (OMS, 2014). Suite aux décisions récentes de classer les émissions diesel et la pollution de l'air elle-même comme cancérigènes certains, la part de ces décès prématurés attribuée à l'environnement atmosphérique est d'ailleurs deux fois plus élevée que celle estimée en 2008 (OMS, 2011). Lorsqu'on s'intéresse aux décès prématurés attribuables uniquement la pollution de l'air extérieur, parmi les 3,7 millions, 455 000 décès sont observés dans les régions au niveau de revenu élevé. Cela confirme également l'implication de facteurs socio-économiques sur les effets de santé. Ce rapport assume finalement un lien particulier entre la pollution de l'air et les maladies cardio-vasculaires, respiratoires, ainsi que les cancers. A ce titre, le projet européen APHEKOM a récemment conclu que le dépassement des recommandations de l'OMS sur les particules fines atmosphériques (PM_{2,5}) génèrerait 19 000 décès supplémentaires par année dans les grandes villes européennes, dont la majorité serait causée par des maladies cardio-vasculaires (Pascal *et al.*, 2013).

Pourtant, ce lien reste difficile à caractériser, car les maladies mises en lien avec l'environnement se caractérisent par une étiologie multifactorielle et une faible spécificité des symptômes. De plus, l'Homme est exposé à son environnement d'une manière globale, et donc à l'ensemble des contaminants environnementaux qu'il contient. La multiplicité des voies d'exposition rend encore

plus complexe l'identification et la détermination de facteurs de risques environnementaux. Afin de mieux caractériser la qualité de l'environnement de vie des populations, un cadre réglementaire national et international de plus en plus présent (la Loi sur l'air et l'utilisation rationnelle de l'énergie en 1996 qui fait partie intégrante du Code de l'Environnement, la Directive Cadre Eau en 2000 traduite au niveau local en schémas directeurs d'aménagement et de gestion des eaux (SDAGE), la Directive Cadre Sols actuellement en cours d'écriture, etc.) fixe la mise en place d'importants programmes de surveillance des milieux par les instances agréées (associations agréées pour la surveillance de la qualité de l'air (AASQA), agences de l'eau, groupement d'intérêt scientifique sol). Malgré cette surveillance continue, la connaissance de l'état qualitatif des milieux environnementaux reste en retard par rapport à celle de l'état de santé des populations, essentiellement pour des raisons de coût et d'accès aux données existantes.

L'ensemble de ces constats explique qu'aujourd'hui, en dehors de quelques pathologies particulières telles que le saturnisme (principalement lié à l'ingestion de plomb) ou le mésothéliome de la plèvre suite à une exposition à l'amiante, peu de maladies peuvent être indubitablement associées à la qualité de l'environnement. Certaines pathologies bénéficient également d'une littérature conséquente concernant les facteurs de risques environnementaux. La pollution atmosphérique et notamment les particules fines a par exemple récemment été classée cancérigène certain par le Centre International de Recherche sur le Cancer (IARC, 2013) et a fait l'objet de nombreuses associations aux événements respiratoires et cardio-vasculaires. Néanmoins, pour d'autres pathologies telles que l'insuffisance rénale chronique ou les maladies inflammatoires chroniques de l'intestin (MICI), les déterminants environnementaux sont encore peu ciblés. Le cadmium pourrait par exemple accélérer le processus de passage au stade terminal de l'insuffisance rénale (Lauwerys *et al.*, 1993 ; Hellström *et al.*, 2001 ; Järup *et al.*, 2002). Une augmentation rapide et de grande ampleur de l'incidence de la maladie de Crohn depuis 1950 dans les pays développés, ainsi que son apparition récente dans les pays émergents, témoignent par ailleurs que des modifications environnementales jouant un rôle dans la pathogénèse (Jantchou *et al.*, 2006).

Pour pallier ce manque de caractérisation et déterminer un lien de causalité entre une pathologie et l'environnement, il existe les méthodes d'évaluation des risques sanitaires (ERS), basées sur le continuum source – vecteur – cible (National Research Council, 1983), ou les approches épidémiologiques, réalisées sur des cohortes de patients ou en population générale. Les ERS sont spécifiques au site étudié, c'est-à-dire que les résultats ne peuvent être extrapolés à la population générale. Les études épidémiologiques sont quant à elles difficiles à mettre en œuvre et présentent notamment un coût élevé.

Dans ce contexte, les études écologiques géographiques (ou temporelles) se présentent comme d'autant plus intéressantes, afin d'explorer les hypothèses étiologiques en amont d'études plus interventionnelles. Les études écologiques visent à détecter les variations de l'occurrence d'une maladie dans l'espace ou dans le temps, et à relier ces variations à des facteurs environnementaux ou sociaux. Ces études utilisent des données agrégées (groupes d'individus) et non des données individuelles. Elles constituent un véritable outil de diagnostic et d'aide à la décision en santé publique, puisqu'elles offrent une réponse à un coût limité, dans un temps relativement bref et à une échelle spatiale modulable (Goria *et al.*, 2011). Ces études produisent néanmoins des résultats globaux, liés à l'existence d'un certain nombre de biais (Wakefield, 2008), comme le biais écologique. Ce biais écologique se traduit par exemple par des facteurs de risques individuels différents au sein des groupes de populations étudiés. Il est ainsi difficile de se focaliser sur une substance précise, puisque les facteurs environnementaux étudiés ne peuvent être considérés en tant qu'exposition réelle des groupes de population. Dans une étude écologique, le lien santé – environnement ne repose donc pas sur l'estimation d'un risque, mais sur la superposition d'informations de différentes natures et l'étude de leur évolution conjointe ou non. Celles-ci sont ainsi susceptibles de traduire un lien, caractérisé par des populations ayant un état de santé dégradé dans un environnement dégradé.

Pour territorialiser et étudier de manière conjointe des paramètres de l'état de l'environnement, des statistiques démographiques et socio-économiques, ainsi que la morbi-mortalité des populations, les Systèmes d'Information Géographiques (SIG) alimentés par des bases de données spatialisées représentent un outil indispensable en vue de détecter les points noirs environnementaux, sanitaires et socio-économiques (Vine *et al.*, 1997). Renforcés par des approches statistiques spatiales (méthodes géostatistiques d'interpolation, cartographie des maladies, détection de clusters spatiaux) appliquées à ces indicateurs géoréférencés, ils constituent une base conséquente de puissants outils d'analyse spatialisée.

La concrétisation d'une étude spatialisée en santé environnementale dépend non seulement de la disponibilité des données, sur un territoire identique et une période de temps adéquate (notamment en fonction du temps de latence du facteur de santé étudié), mais aussi de leur format pour permettre leur mise en relation statistique. Or les données environnementales, sociales et sanitaires sont de natures et de types très divers. En effet pour des raisons institutionnelles d'éthique et de confidentialité, les indicateurs de santé (issus de registres de pathologies par exemple) et les indicateurs socio-économiques des populations (issus des recensements de l'Institut National de Statistique et d'Etudes Economiques (Insee) par exemple) sont calculés et communiqués de manière agrégée. C'est-à-dire que l'information (incidence de maladie, indicateur de précarité, etc.) caractérise un groupe de population, en général une unité administrative (canton, commune, quartier, etc.). Pour l'environnement, la majorité des données renseignant la qualité de milieux sont de nature géostatistique. Celles-ci sont constituées d'un échantillon de points géoréférencés auxquels est attribué un indicateur mesuré (une concentration en polluant ou un indice biologique). Les différents formats de ces données nécessitent la mise en œuvre de traitements spatiaux variés et peuvent engendrer des difficultés lors de leur mise en relation. Par ailleurs, les données environnementales issues des modèles de surveillance ou de biosurveillance sont caractérisées par des valeurs continues (une mesure de concentration), alors que les données brutes de santé représentent le plus fréquemment des valeurs discrètes (un nombre de cas de maladie par unité géographique) ou binaires (malade versus non-malade).

Ce travail a pour objectifs de caractériser la qualité des milieux de vie de différentes populations grâce aux bases de données environnementales disponibles et aux outils SIG, puis de rapprocher les cartographies générées aux indicateurs sanitaires et socio-économiques par des études écologiques. Ainsi, nous chercherons à évaluer si les populations ayant un état de santé dégradé vivent dans un environnement dégradé, tout en intégrant les paramètres sociodémographiques. Ce questionnement sera étudié à travers trois axes de recherche.

Dans le premier axe de recherche (Chapitre I), nous allons étudier comment pallier la différence de format des données caractérisant l'environnement et les populations. Cet axe s'appuiera sur la caractérisation de l'imprégnation de l'environnement atmosphérique par les éléments traces métalliques (ETM) et l'évaluation des inégalités environnementales et sociales de santé à fine échelle sur trois bassins de vie de la région Nord – Pas de Calais. Nous aurons pour celui-ci l'occasion de travailler à partir de mesures d'imprégnation biologique réalisées à la fois chez l'Homme¹ (sang et urines) et dans l'environnement (lichens épiphytes), ainsi que d'un indicateur de défaveur sociale². Les mesures d'imprégnation représentant le résultat d'une exposition globale à plusieurs métaux,

¹ Données obtenues dans le cadre d'un partenariat avec le Centre de Recherches en Santé - Travail - Environnement (CERESTE) du CHRU de Lille (Dr. Catherine Nisse et Dr. Ariane Leroyer, EA4483, Université Lille 2).

² Données obtenues dans le cadre d'un partenariat avec le laboratoire Territoires Villes Environnement et Société (TVES) de l'Université Littoral Côte d'Opale (Dr. Anne Peggy Hellequin et Rachel Bavdek, EA4477, ULCO).

nous avons considéré ces éléments de manière individuelle, mais également de manière concomitante à travers le développement d'un indicateur d'imprégnation multimétallique.

Dans le deuxième axe de recherche (Chapitre II), nous traiterons de données sanitaires de format différent. Il ne s'agira plus de mesures d'imprégnation mais de données d'incidence de l'insuffisance rénale chronique terminale (IRCT). Après avoir évalué les disparités d'incidence de cette pathologie à l'échelle de la région Nord – Pas de Calais³ à l'aide des outils de cartographie des maladies et de détection de clusters, nous appliquerons une étude écologique spatialisée, afin de confirmer certains facteurs de risques connus, tels que le statut socio-économique et les maladies à risques (diabète, maladies cardiovasculaires).

Enfin, l'objectif de notre troisième axe de recherche (Chapitre III) est d'étudier la contribution des études spatialisées dans l'identification de facteurs de risques environnementaux, ici suspectés dans le cas des MICI⁴, afin de suggérer des hypothèses étiologiques. Pour cela, nous nous intéresserons à l'analyse spatiale et spatio-temporelle d'une base de données publique de contamination des eaux souterraines.

En amont de ces trois chapitres, une introduction générale est tout d'abord consacrée à la présentation des outils et méthodes permettant de caractériser la qualité des milieux environnementaux et la santé des populations, ainsi que l'analyse spatiale d'inégalités environnementales et sociales de santé. Cette introduction générale est suivie d'un chapitre détaillant les ressources et les méthodes utilisées dans le cadre de nos recherches. Une discussion et une conclusion générale analysent et critiquent finalement les résultats obtenus dans le cadre de cette thèse, puis offre une vue des perspectives de recherche.

³ Données obtenues dans le cadre d'un partenariat avec le réseau REIN Néphronor de l'Agence de Biomédecine (Dr. François-Xavier Glowacki, EA4483, CHRU de Lille).

⁴ Données obtenues dans le cadre d'un partenariat avec le registre des maladies inflammatoires chroniques du tube digestif (EPIMAD) du CHRU de Lille (Dr. Corinne Gower-Rousseau, EA2694, Université Lille 2) et le Centre d'Etudes et de Recherche en Informatique Médicale (CERIM) du CHRU de Lille (Pr. Alain Duhamel et Dr. Michaël Génin, EA2694, Université Lille 2)

REFERENCES BIBLIOGRAPHIQUES

- Besse G, Caudeville J, Michelot F, Prisse N, Trugeon A. 2014. Inégalités territoriales, environnementales et sociales de santé. Regards croisés en régions : de l'observation à l'action. SGMAS & CGDD, 68p.
- Directive Cadre sur l'Eau. 2000. Directive 2000/60/CE du Parlement européen et du Conseil établissant un cadre pour une politique communautaire dans le domaine de l'eau.
- Goria S, Stempfelet M, de Crouy-Chanel P. 2011. Introduction aux méthodes statistiques et aux systèmes d'information géographique en santé environnement – Application aux études écologiques. Saint-Maurice: Institut de veille sanitaire. 65 p. Disponible à partir de l'URL : <http://www.invs.sante.fr>.
- Hellström L, Elinder C-G, Dahlberg B, Lundberg M, Järup L, Persson B, Axelson O. 2001. Cadmium exposure and end-stage renal disease. *Am J Kidney Dis*, 38: 1001-8.
- IARC. 2013. Outdoor air pollution a leading environmental cause of cancer deaths. Press release N°221. 4p.
- Jantchou P, Monnet E, Carbonnel F. 2006. Les facteurs d'environnement dans la maladie de Crohn et la rectocolite hémorragique (tabac et appendicectomie exclus). *Gastroenterol Clin Biol*, 30 : 859-867.
- Järup L. 2002. Cadmium overload and toxicity. *Nephrol Dial Transplant*, 17: 35–9.
- Lauwerys RR, Bernard AM, Buchet J-P, Roels H. 1993. Assessment of the health impact of environmental exposure to cadmium: contribution of the epidemiologic studies carried out in Belgium. *Environ Res*, 62: 200-6.
- Loi sur l'air et l'utilisation rationnelle de l'énergie. 1996. Loi N° 96-1236 du 30 décembre 1996 sur l'air et l'utilisation rationnelle de l'énergie.
- National Research Council, Committee on the institutional means for assessment of risks to public health. 1983. Risk assessment in the Federal government: managing the process. National Academy Press Washington, D.C. 191p.
- OMS. 2011. Burden of disease attributable to outdoor air pollution. Geneva
- OMS. 2012. Environmental health inequalities in Europe. Copenhagen, OMS Europe. 190p.
- OMS. 2014. Burden of disease from the joint effects of Household and Ambient Air Pollution for 2012. Geneva.
- Pascal M, Corso M, Chanel O, Declecq C, Badaloni C, Cesaroni G, Henschel S, Maister K, Haluza D, Martin-Olmedo P, Medina S. 2013. Assessing the public health impact of urban air pollution in 25 European cities: results of the Aphekom project. *Science of the Total Environment*, 449: 390-400.
- PNSE1. 2004. Plan National Santé Environnement 2004-2008. 88p.
- PNSE2. 2009. Deuxième Plan National Santé-Environnement 2009-2013. 67p.
- Prüss-Üstün A & Corvalán C. 2006. Preventing disease through healthy environments: Towards an estimate of the environmental burden of disease. WHO Library Cataloguing-in-Publication Data. 104p.
- Vine M, Degnan D, Hanchette C. 1997. Geographic Information Systems: Their Use in Environmental Epidemiologic Research. *Environmental Health Perspectives*, Volume 105, Number 6: 598-605.
- Wakefield J. 2008. Ecologic studies revisited. *Annual Review of Public Health*, 29: 75–90.

INTRODUCTION GENERALE

Ce chapitre d'introduction générale est consacré à la présentation des méthodes et outils mis en œuvre pour évaluer les inégalités environnementales et sociales de santé, à travers des analyses spatiales et spatio-temporelles. Les systèmes d'information géographique (1^{ère} partie) permettent d'implémenter des données spatialisées issues de sources diverses et de les représenter sous forme de cartographies. Au sein des SIG, de puissants outils d'analyse permettent d'évaluer les comportements de ces données dans l'espace et le temps. Les méthodes géostatistiques d'interpolation spatiale (2^e partie) sont utilisées pour cartographier la qualité de milieux environnementaux à partir de campagnes de mesures, alors que les outils d'épidémiologie spatiale (3^e partie), telles que la cartographie des taux d'incidence et la détection de clusters atypiques des maladies, sont utilisés afin d'évaluer les disparités spatiales et spatio-temporelles des pathologies. Dans le cadre de cette thèse, des données issues de la biosurveillance lichénique de la qualité de l'air (4^e partie) ont été cartographiées à l'aide des outils géostatistiques et intégrées dans une étude épidémiologique. Celles-ci traitent de mesures d'imprégnation des lichens épiphytes par différents ETM. A ce titre, un indice intégré de contamination multimétallique (5^e partie) a été développé afin d'évaluer de manière synthétique la qualité globale du milieu et d'agrémenter l'étude du lien environnement – santé. Finalement, des études écologiques (6^e partie) ont été menées pour évaluer les liens statistiques entre les différents indicateurs sanitaires, environnementaux et socioéconomiques.

Partie 1. Les Systèmes d'Information Géographique (SIG)

Le contexte de l'information géographique s'est largement développé depuis la fin du XX^e siècle grâce à l'évolution des outils informatiques. L'apparition conjointe des technologies de l'information et de la communication et la circulation accrue des données, notamment suite à l'avènement de l'internet, font ainsi de la géomatique (science des données à référence spatiale : Pornon, 2007) une discipline nouvelle et indispensable, utilisée pour répondre à des problématiques très variées à travers une approche multidisciplinaire et multithématique (Bergeron, 1992 ; Clarke, 2001 ; Gadal, 2008). Elle fait appel à un vocabulaire spécifique et emploie des technologies singulières : les systèmes d'information géographique.

Un consensus de définitions montre que les SIG ne peuvent être considérés comme de simples logiciels de traitement de données spatialisées. Parmi ces définitions, celle de Françoise De Blomac illustre bien ce concept : un SIG est « un ensemble organisé de matériels informatiques, de logiciels, de données géographiques et de personnels capables de saisir, stocker, mettre à jour, manipuler, analyser et présenter toutes formes d'informations géographiquement référencées » (De Blomac *et al.*, 1994).

Plusieurs fonctions peuvent donc être attribuées à ces systèmes. La première est la création et la collecte de données géoréférencées, issues de sources multiples et de formats très divers (par exemple la numérisation de fonds de cartes ou de photographies aériennes, la digitalisation de données de terrain, l'harmonisation spatiale, temporelle et structurelle des données). La deuxième fonction est le stockage, l'organisation et la spatialisation de ces informations grâce à un système de gestion de bases de données (SGBD). De manière bien plus spécifique, les fonctions d'affichage et de superposition de différentes couches d'information géographique sous forme de cartographies permettent une visualisation spatialisée des phénomènes terrestres. Finalement, les fonctions d'interrogation et d'analyse spatiale et/ou statistique évaluent ensuite les tendances et interactions géographiques entre ces différents phénomènes. De ce fait, les SIG sont considérés comme des outils complets de connaissance, d'aide à la décision et de communication (Goria *et al.*, 2011).

Quels types de données peuvent être intégrés dans un SIG ? Fernandez-Falcon *et al.* (1993) estiment à plus de 80 % les informations possédant une référence spatiale. Techniquement, ces données peuvent être de deux natures : les images et les vecteurs. Les images, également appelées rasters, sont constituées d'une matrice de pixels géoréférencés auxquels sont attribuées des valeurs radiométriques de couleur. Les photographies aériennes et les cartes topographiques sont des exemples de rasters couramment utilisés. Les données vecteurs sont composées d'objets géographiques de forme élémentaire. Les points sont des coordonnées XY, les lignes sont une succession de points de coordonnées XY et les polygones sont une succession de points de coordonnées XY délimitant une surface fermée (figure 1).

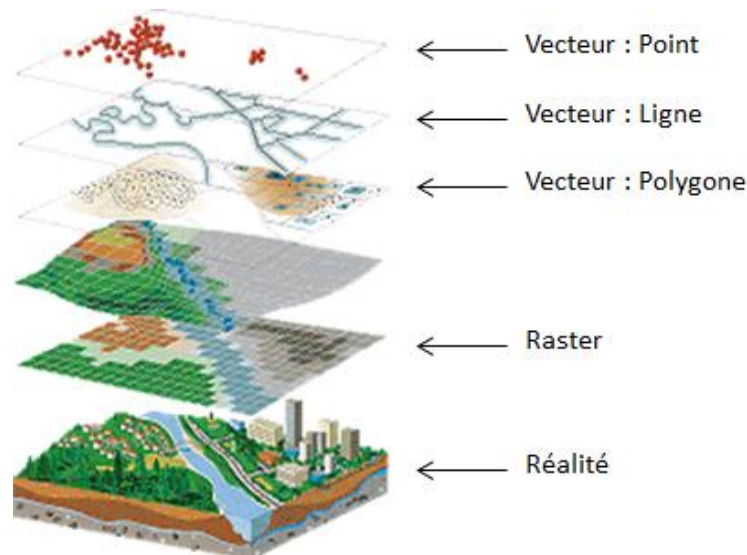


Figure 1. Couches d'informations spatialisées au sein d'un SIG (modifié à partir d'ESRI, 2014)

Grâce aux SIG, nous disposons d'un grand nombre d'outils d'analyse spatiale pour évaluer les structures et processus géographiques de plusieurs jeux de données. Ces outils se basent notamment sur plusieurs concepts comme la distance, l'interaction spatiale ou la centralité. Dans le champ de la santé environnementale, de puissants outils d'analyse spatiale peuvent être employés. Le géocodage est par exemple mis en œuvre pour localiser les patients enregistrés dans les registres de maladies, en transformant leur adresse postale en coordonnées spatiales (Vine *et al.*, 1997). Les données d'occupation du sol sont régulièrement utilisées pour définir des plans et stratégies d'échantillonnage des milieux environnementaux (AFNOR NF X43-903, 2008). Suite à cet échantillonnage, l'interpolation spatiale estime les valeurs de paramètres mesurés (concentrations en polluants, indicateurs, etc.) au niveau de points non échantillonnés, afin de construire par exemple des proxys d'exposition des populations (Nuckols *et al.*, 2004). La cartographie des maladies (lissage des taux d'incidence) et la détection de clusters atypiques évaluent la distribution spatiale des événements de santé et leur tendance à se regrouper (Clayton & Kaldor, 1987 ; Kulldorff, 1997). Grâce à ces outils SIG, il est finalement possible de rapprocher statistiquement les cartographies d'indicateurs environnementaux, sanitaires et sociodémographiques. Certaines de ces fonctions seront plus détaillées dans les paragraphes suivants.

La cartographie peut finalement représenter le livrable final d'une étude SIG. La carte est un outil adapté de synthèse et de communication de résultats issus d'analyses spatiales. La carte choroplèthe, qui représente des quantités relatives à des espaces (une densité de population par commune par exemple) et la carte géostatistique, qui représente la variabilité continue d'une information quantitative à travers un espace défini, sont les formats les plus utilisés en santé environnementale. La production de tels supports implique de considérer certaines règles de

sémiologie graphique et de discrétisation des données. Dans un premier temps, il est important de s'attarder sur le choix d'une méthode de discrétisation, c'est-à-dire de division de la série statistique que l'on veut cartographier en classes ou intervalles. Afin d'éviter des erreurs de représentation, certaines méthodes de discrétisation sont privilégiées en fonction de l'objectif de la carte et de la forme de la distribution statistique des données (Tableau 1). Pour optimiser la lisibilité de la carte, le nombre de classes doit préférentiellement être au maximum de 6.

Tableau 1. Méthodes de discrétisation de variables continues en vue d'une analyse spatiale (Goria *et al.*, 2011)

Nature de l'information		Méthode appropriée	Méthode inappropriée
Forme de la distribution	Normale	Moyenne écart-type ou classes équiprobables	Classes d'amplitude égale ou progression géométrique (à droite) ou moyenne écart-type ou classes équiprobables
	Dissymétrique	Progression géométrique (à gauche) ou méthode de Jenks	Amplitude égale/moyenne-écart-type
	Uniforme	Classes d'amplitude égale	Progression géométrique
	Plurimodale	Seuils naturels – Jenks	Amplitudes égales
Objectif recherché	Faire ressortir les valeurs extrêmes	Seuils naturels – Jenks ou comparaison à une valeur standard/de référence	Amplitudes égales
	Comparaison	Moyenne écart-type ou classes équiprobables (si distribution normale) ou discrétisation par quantiles	Arbitraire/seuils naturels/amplitude égale/Jenks
	Mettre en évidence une configuration spatiale	Jenks	Arbitraire

La réalisation d'une carte repose ensuite sur la sémiologie graphique. Le choix sémiologique repose sur trois paramètres : la forme des objets (ou la texture des surface), la taille des objets et la couleur. La figure 2 reprend les principales règles de sémiologie cartographique.









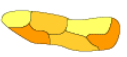






Objets géographiques	Points	Lignes ou Flux	Aires
Information thématique	 <i>Ex : Villes / Structures de soins / Stations de mesure etc.</i>	 <i>Ex : infrastructures routières / réseau hydrographique</i> <i>Ex : échanges / recours aux soins</i>	 <i>Ex : entités administratives / bassins d'emploi etc.</i>
nominale <i>ex : type d'objet, classification</i>	cartes en symboles (variation de formes, d'orientations ou de couleurs)  <i>(Cartes d'inventaire : Ex : Villes selon leurs fonctions, structures de soins suivant leur nature)</i>	symboles linéaires (variation selon la forme, ou la couleur)  <i>(Ex : Réseaux, types d'accès aux soins)</i>	Plages variant selon la couleur ou le motif  <i>(Ex : Typologie)</i>
Qualitative			
ordinaire <i>ex : rangs, hiérarchie, niveaux</i>	cartes en symboles (variation de taille ou de couleurs)  <i>(Structures de soins hiérarchisés)</i>	Symboles linéaires (variation de taille ou de couleurs)  <i>(Flux classés selon leur importance (faible, moyen, fort))</i>	Plages variant selon la couleur  <i>(Niveaux de prévalence (faible, moyen, fort))</i>
Quantitative			
discrète <i>ex : Effectifs de population, nombre de cas</i>	cartes en symboles proportionnels  <i>(Taille de villes, nombre de recours dans des structures de soins)</i>	Symboles linéaires ou flèches proportionnels  <i>(Fréquentation axes, fréquences de recours)</i>	Symboles proportionnels ou semis de points  <i>Répartition population, effectifs de médecins</i>
continue <i>ex : Taux, pourcentages</i>	symboles ponctuels (variation de couleurs)  <i>Courbes de niveau</i>	Symboles linéaires (variation de couleurs) 	Plages variant selon la couleur 

Figure 2. Principales règles de sémiologie graphique (Rican, 2007)

Dans le domaine de l'environnement, la surveillance et la biosurveillance de la qualité des milieux peuvent donner lieu à des cartographies ponctuelles discrètes ou continues. Par exemple, la taille ou la couleur d'un symbole ponctuel attribué au point d'échantillonnage est proportionnelle à la concentration mesurée en ce point. Lors des campagnes environnementales, la récolte des échantillons peut être variable dans l'espace, ce qui peut provoquer des « vides d'information ». Afin de combler ces manques, des méthodes géostatistiques d'interpolation spatiale sont mises en œuvre pour estimer les indicateurs environnementaux en tous points de l'espace. Ces méthodes sont présentées dans la partie suivante.

Partie 2. L'interpolation spatiale

La plupart des bases de données issues de la surveillance physico-chimique et biologique de l'environnement peuvent constituer des informations géostatistiques. Cette information peut être définie par des mesures quantitatives réalisées au niveau d'un échantillon de points géolocalisés dans l'espace. Bien que les stratégies d'échantillonnage de mesures environnementales soient initialement construites à partir de plans de mailles régulières et homogènes, la logistique et le coût élevé des prélèvements et d'analyse des échantillons occasionnent fréquemment le recours à un tirage au sort d'une partie de ces points. Cette sélection au hasard fait que la répartition des échantillons sur le terrain devient irrégulière et hétérogène sur la zone étudiée (AFNOR NF ISO 10381-1, 2003 ; AFNOR NF X43-904, 2013). De plus, les observations issues de ces échantillons sont rarement indépendantes de l'espace et identiquement distribuées. Leurs traitements statistiques nécessitent donc une attention particulière. C'est pour ces raisons que des méthodes d'analyse spatiale ont été développées afin de générer des représentations cartographiques globales de ces indicateurs de qualité des milieux environnementaux et évaluer les principales tendances du phénomène observé (Ripley, 1981 ; Cressie, 1993).

L'interpolation spatiale est une méthode d'estimation statistique de données spatialisées. Son principe est qu'à partir d'observations ponctuelles mesurées et géoréférencées, elle fournit la valeur la plus probable du paramètre observé (appelé variable régionalisée) en tout point du domaine spatial étudié (Hengl, 2007). Le résultat est une production cartographique d'estimations au niveau de chaque point d'une grille régulière recouvrant la zone d'étude (figure 3). Par conséquent, plus le nombre de localisations estimées est élevé, plus la carte prendra l'apparence d'une surface lisse. Ce résultat peut être représenté selon des courbes de niveau, une palette de couleurs ou une perspective 3D. A l'inverse de l'extrapolation (estimation des valeurs à l'extérieur de la zone couverte par les observations), l'interpolation se consacre à la même zone que celle couverte par les observations (Burrough & McDonnell, 1998).

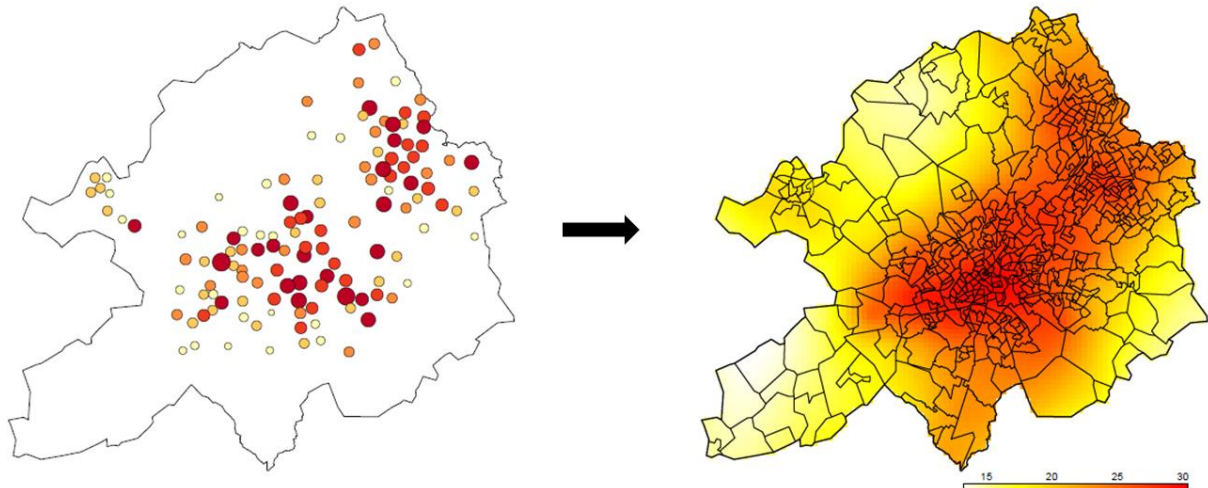


Figure 3. Campagne de mesure du NO₂ atmosphérique (en µg/m³) sur la communauté urbaine de Lille en été 2003 par le réseau ATMO Nord – Pas de Calais (modifié à partir de Declercq, 2007)

Pour prévoir la valeur d'un point non échantillonné à partir de plusieurs points d'observation, la formule mathématique simplifiée de l'interpolation spatiale peut être assimilée à une moyenne arithmétique pondérée des valeurs observées (Li & Heap, 2014) :

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i z(x_i)$$

où \hat{z} est la valeur estimée de la variable aléatoire au point d'intérêt x_0 , z est la valeur observée au point d'échantillon x_i , λ_i est le facteur de pondération attribué à ce point d'échantillon et n le nombre d'échantillons utilisés pour l'estimation (figure 4).

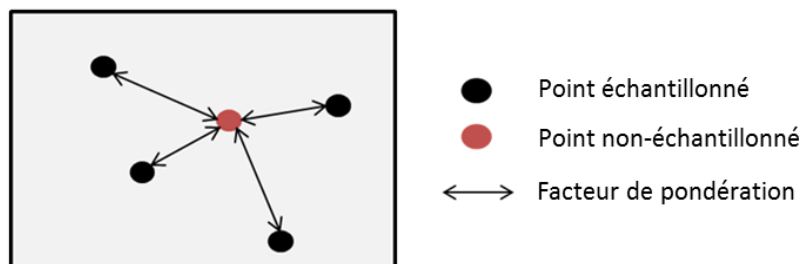


Figure 4. Schéma simplifié de l'interpolation spatiale

Les méthodes d'interpolation spatiale peuvent être classées selon deux grandes approches : l'approche déterministe et l'approche géostatistique. La principale différence de ces méthodes est la manière dont va être attribué le poids de chaque point d'observation dans le calcul d'estimation. D'autres critères peuvent également distinguer les méthodes d'interpolation. Par exemple, les méthodes d'estimation globales calculent la moyenne arithmétique de la variable régionalisée à partir de l'ensemble des observations mesurées sur le domaine spatial, alors que les méthodes d'estimation locales estiment cette moyenne à partir des observations présentes sur une portion du territoire proche. Par ailleurs, une méthode dite « exacte » conserve les valeurs des points d'observations mesurés dans le résultat cartographique, contrairement à une méthode dite « approchée » qui a tendance à lisser les valeurs aux points d'observation (Bossert, 2012). Dans les paragraphes suivants, nous expliquerons les deux grandes approches des méthodes d'interpolation spatiale (déterministe et géostatistique). L'approche géostatistique et notamment la méthode du krigeage souvent employée pour cartographier une variable environnementale seront plus détaillées.

a. Approche déterministe

Les méthodes déterministes d'interpolation reposent essentiellement sur des propriétés mathématiques et géométriques, sans tenir compte de la structure spatiale du phénomène. Parmi ces méthodes, nous pouvons citer les méthodes barycentriques (inverse des distances), les méthodes de partitionnement de l'espace par polygones de Thiessen ou cellules homogènes (le plus proche voisin, la triangulation) et les splines. La figure 5 propose un exemple de résultats cartographiques pour plusieurs de ces méthodes. Pour rappel, leur principe est de prévoir la valeur de la variable régionalisée en un point non échantillonné par une moyenne pondérée des valeurs observées. Pour l'approche déterministe, la pondération est uniquement fonction de la distance euclidienne entre le site d'observation et le site de prédiction. Les sites d'observation les plus proches ont ainsi une plus grande influence dans le calcul, alors qu'un poids faible (méthodes globales) ou nul (méthodes locales) est généralement attribué aux sites les plus éloignés. Il est par exemple possible de sélectionner uniquement les n points d'observation les plus proches pour le calcul d'estimation (Hengl, 2007).

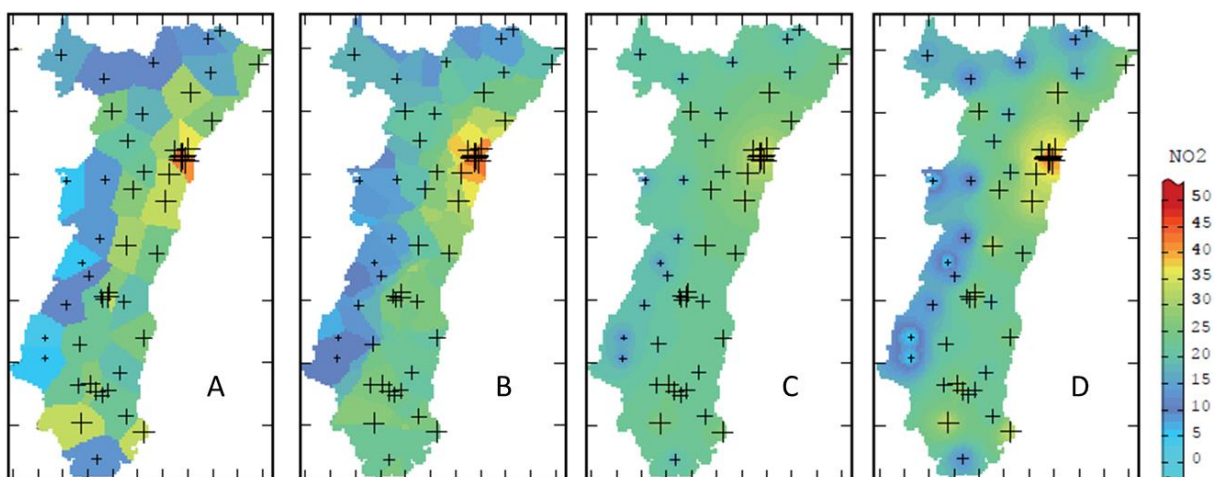


Figure 5. Estimation de la concentration en dioxyde d'azote en Alsace (A) par plus proche voisin, (B) par moyenne mobile, (C) par inverse des distances et (D) par inverse des distances au carré (modifié à partir de Lemarchand & Jeannée, 2009)

b. Approche géostatistique : le krigeage

Les méthodes probabilistes (ou stochastiques) d'interpolation sont appelées méthodes géostatistiques, puisqu'elles découlent directement de l'analyse géostatistique des données d'observation et incorporent le concept du phénomène naturel (Cressie, 1993 ; Goovaerts, 1997). En plus d'une structure déterministe, elles incluent des notions d'erreurs aléatoires pour représenter le comportement spatial d'un phénomène naturel (Hengl, 2007). Parmi les méthodes géostatistiques, il est possible d'utiliser les techniques de régression classique, de régression locale et de krigeage. Contrairement aux méthodes de régression, l'interpolation par krigeage prend en considération la structure de dépendance spatiale des données afin de minimiser l'erreur de prédiction. Cette dernière est donc considérée comme la méthode optimale d'interpolation pour les phénomènes environnementaux (Gratton, 2002). Les paragraphes suivants se focalisent sur les méthodes d'interpolation par krigeage.

L'initiateur du krigeage est l'ingénieur minier sud-africain D. Krige (1951), qui souhaitait déterminer la distribution spatiale de minerais à partir d'un ensemble de forages. Mais la théorie du krigeage de variables régionalisées et l'approche géostatistique ont réellement été développées par le mathématicien français Matheron (1962 ; 1963a ; 1963b). Dans cette méthode une pondération est attribuée à chaque point, qui est une fonction de plusieurs composantes : la distance entre la valeur observée et la valeur à estimer (comme pour l'approche déterministe), la structure spatiale de l'échantillonnage (présence de vides ou de clusters de points mesurés) et le comportement spatial du phénomène observé (variabilité spatiale brute ou lente, direction privilégiée, etc.). La première composante étant connue, elle va faciliter l'estimation de la structure et du comportement spatial du phénomène par le calcul d'un variogramme.

Le variogramme représente l'évolution de la variabilité du phénomène, c'est-à-dire la dissemblance entre les valeurs, en fonction de la distance entre deux points. C'est ici l'occasion d'introduire un phénomène clé en analyse spatiale, développé par Waldo Tobler dans sa première loi de géographie. Ce phénomène, appelé autocorrélation spatiale, indique que « *Everything is related to everything else, but near things are more related than far things* » (Tobler, 1970). Ainsi le variogramme expérimental évalue si en moyenne, la variance entre deux observations proches est petite et la variance entre deux observations éloignées est grande. Le calcul de variance γ pour chaque couple de points en fonction de la distance h qui les sépare est présenté ci-dessous et représente la moitié de la différence au carré.

$$\gamma(h) = \frac{1}{2}(Z(xi + h) - Z(xi))^2$$

Les différences de l'ensemble des couples de points sont ensuite moyennées par paliers de distance et forment le variogramme expérimental. Celui-ci est ensuite modélisé par une fonction aléatoire $Z(x)$, dont la réalisation est la variable régionalisée $z(x)$. La figure 6 présente un exemple de modèle de variogramme.

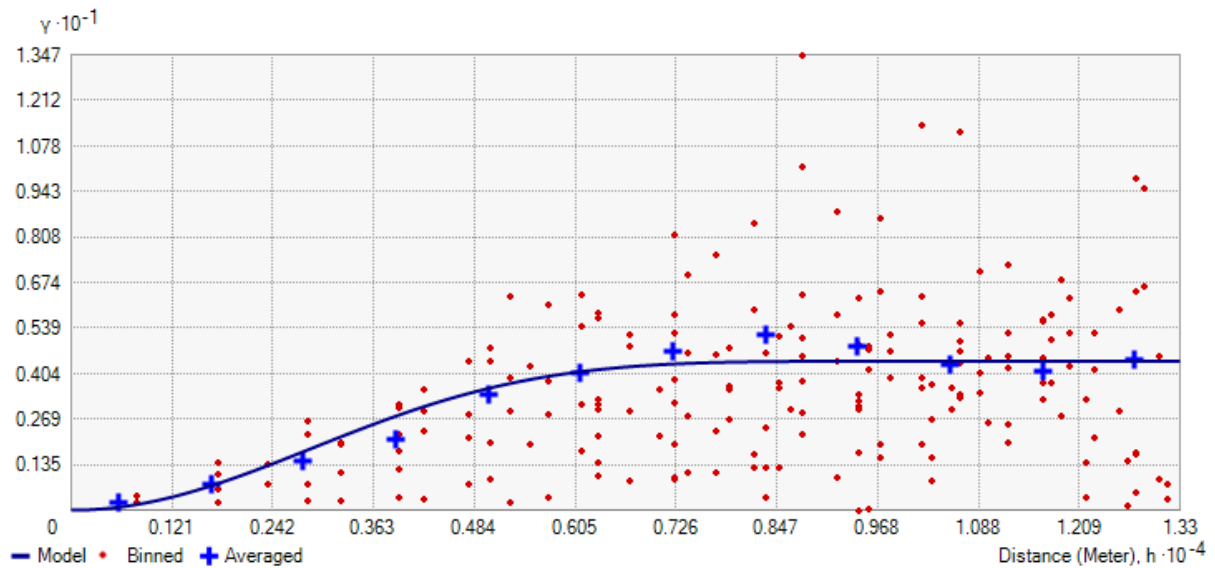


Figure 6. Exemple de variogramme expérimental (logiciel ArcGIS®) reprenant les valeurs des paires de points (rouge), leur moyenne (croix bleue) et le variogramme modélisé (ligne bleue)

En se basant sur les données d'observation et la modélisation du variogramme expérimental, le krigeage va donc attribuer un poids à chacun des sites mesurés à partir de la covariance entre ces points, en fonction de la distance qui les sépare. La fonction $Z(x)$ du modèle de krigeage sera alors la somme de deux composantes. La première composante $\mu(x)$ représente la structure déterministe du modèle, estimée à partir de la distance des n points voisins sélectionnés. La deuxième composante $\delta(x)$ représente la structure stochastique de dépendance spatiale (Baillargeon, 2005). Cette dernière est elle-même une combinaison de fonctions estimées à des niveaux emboîtés de dimensions, afin de modéliser les variations du phénomène à grande, petite et micro échelle.

$$Z(x) = \mu(x) + \delta(x)$$

Les méthodes de krigeage proposent également d'autres avantages. Il est par exemple possible de définir un poids plus ou moins important en fonction d'une direction (phénomène d'anisotropie). Cela permet par exemple de tenir compte de l'origine des vents dominants dans le cadre du suivi d'un panache de cheminée, ou le sens d'écoulement d'une rivière pour le suivi des sédiments (Merwade, 2009). Un autre avantage de certaines méthodes de krigeage (le co-krigeage) est l'intégration de variables auxiliaires, évoluant de la même manière dans l'espace que la variable régionalisée, pour améliorer le calcul du variogramme (Lemarchand & Jeannée, 2009). Mais la forte plus-value de cette méthode par rapport aux autres est que l'estimation de chaque point non échantillonné est accompagnée d'un calcul d'erreur de l'estimation associée. Comme l'illustre la figure 7, cette erreur de prédiction est globalement plus élevée au niveau des zones pauvres en échantillons et à proximité des points observés ayant des valeurs extrêmes. Un phénomène présentant une variabilité élevée dans l'espace entraîne des niveaux d'incertitudes plus importants.

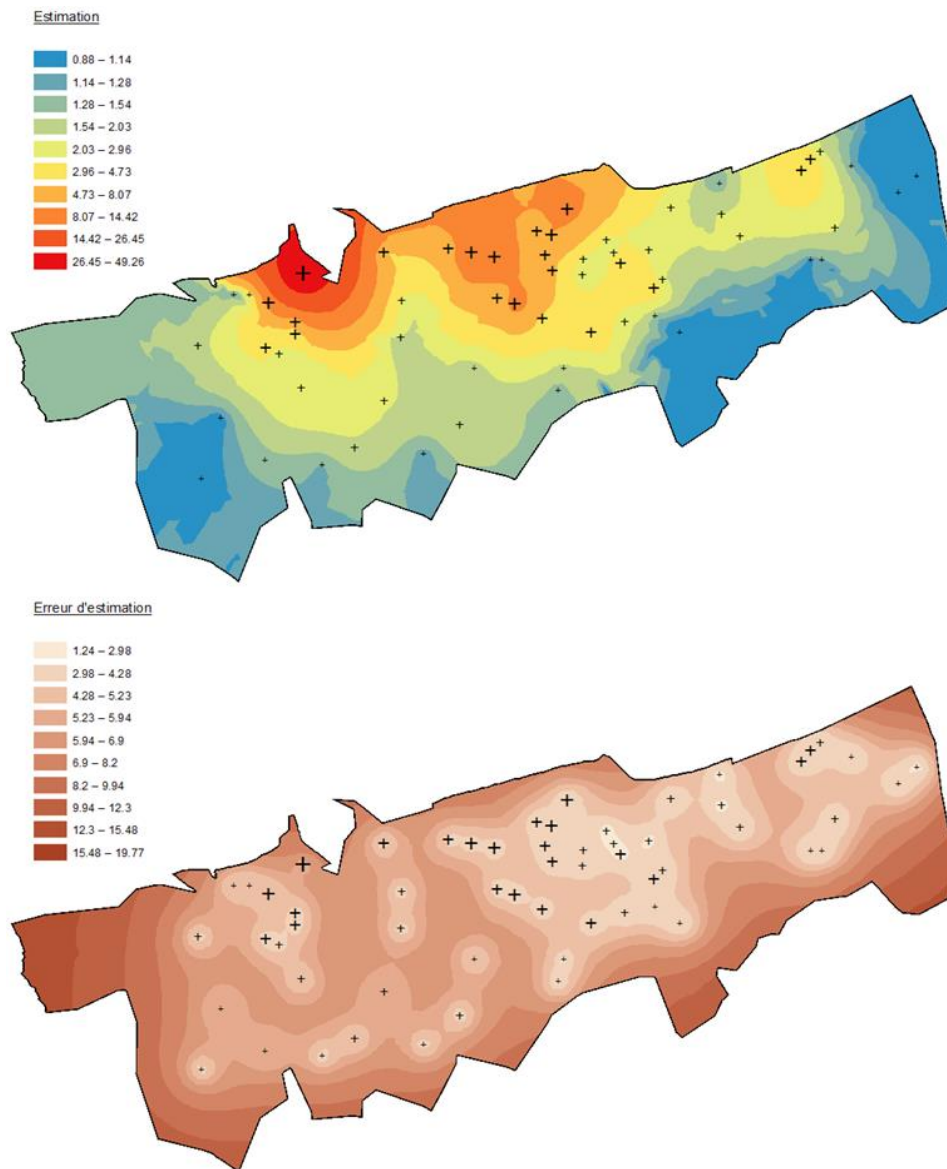


Figure 7. Résultats du krigeage (logiciel ArcGIS®). La cartographie de l'estimation de la variable aléatoire (en haut) et la cartographie de l'erreur d'estimation (en bas)

L'utilisation de techniques de validation croisée (Jackknife ou bootstrap) permet finalement de contrôler la conformité du modèle choisi par rapport à la réalité. Cela consiste à partager les observations de la variable régionalisée en deux groupes, l'un permet le calcul du modèle et l'estimation de la variable en tous points de l'espace, l'autre groupe est utilisé pour tester la correspondance de ce modèle aux valeurs réelles mesurées (Cressie, 1993).

LE KRIGEAGE EMPIRIQUE BAYESIEN (EBK)

Il existe différentes catégories de krigeage (simple, ordinaire, universel, etc.) qui se différencient notamment par la connaissance statistique de la variable régionalisée (Gratton, 2002). Très récemment, une méthode empirique de krigeage bayésien (EBK) a vu le jour et fut développée sous l'application Geostatistical Analyst d'ArcGIS (uniquement à partir de la version 10.1 du logiciel) par le géostatisticien Konstantin Krivoruchko. Son constat initial est que les méthodes classiques de krigeage assument que le variogramme modélisé est vrai pour les données observées. Or cette estimation est très sensible à la forme de distribution de ces données. En effet, le krigeage est peu performant pour des distributions non gaussiennes, ce qui est très fréquent pour les données

environnementales. De plus, même si des outils proposent de les transformer par des fonctions logarithmiques, cette hypothèse est souvent erronée et l'erreur d'estimation est sous-estimée. Il a alors proposé une méthode automatique pour rendre l'estimation du variogramme plus fiable et réaliste (Krivoruchko, 2012).

Le principe se différencie des autres méthodes de krigeage par la prise en considération de l'erreur d'estimation induite par le modèle du variogramme. Concrètement, ce n'est pas un seul variogramme qui est modélisé et utilisé, mais plusieurs. Le processus se déroule en quatre étapes :

- (i) Un échantillon de points d'observation est sélectionné au sein des données initiales. Un variogramme est modélisé à partir des valeurs de cet échantillon.
- (ii) A partir de ce variogramme, une nouvelle valeur est estimée pour chacun des points d'observation.
- (iii) A partir de ces nouvelles valeurs, un nouveau variogramme est modélisé et va estimer une nouvelle fois les valeurs pour chacun des points d'observation, ainsi que les erreurs d'estimations. Les opérations (ii) et (iii) sont renouvelées un certain nombre de fois pour plusieurs échantillons de (i) pouvant se chevaucher.
- (iv) Un poids est attribué à chaque variogramme estimé. Ce poids est calculé grâce à une méthode bayésienne qui évalue la performance du modèle à estimer les valeurs réelles des points d'observation. Le modèle final utilisé pour l'interpolation est une combinaison pondérée de l'ensemble de ces modèles (figure 8).

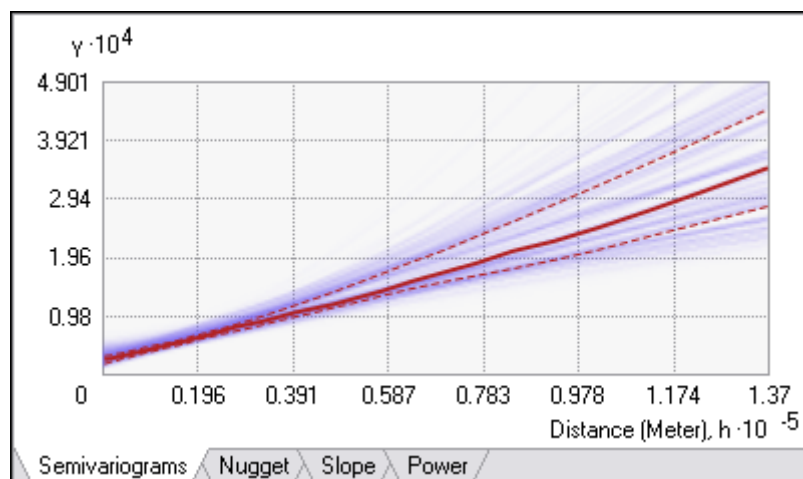


Figure 8. Spectre de variogrammes modélisés par krigeage empirique bayésien (en bleu) – les lignes rouges représentent la médiane (ligne continue), les 25^e et 75^e percentiles (pointillés) de la distribution des variogrammes (logiciel ArcGIS®)

La sélection des échantillons de points d'observation se fait de manière spatiale, en découpant le territoire en plusieurs sous-territoires contenant un nombre fixe de points, ces sous-territoires pouvant se chevaucher. Le spectre de variogrammes est produit pour chacun des échantillons. Par conséquent, en fonction de sa localisation, l'estimation finale de la valeur d'un point non observé est le résultat d'une somme pondérée de plusieurs spectres. Les spectres générés pour des territoires proches de ce point ont un poids plus important que les spectres générés pour des territoires éloignés. Krivoruchko assimile le résultat de l'EBK à l'œuvre « La corbeille de fruits » du peintre Giuseppe Arcimboldo, qui combine plusieurs modèles de fruits pour former une image représentant un visage (figure 9). A ce jour, seulement trois applications de cette méthode ont été répertoriées dans la littérature. Finzgar *et al.* (2014) ont employé l'EBK pour cartographier les concentrations en plomb et zinc dans les sols avant et après remédiation, Deng *et al.* (2013) pour le carbone organique des sols et Roberts *et al.* (2014) pour les concentrations atmosphériques en particules fines et ozone.

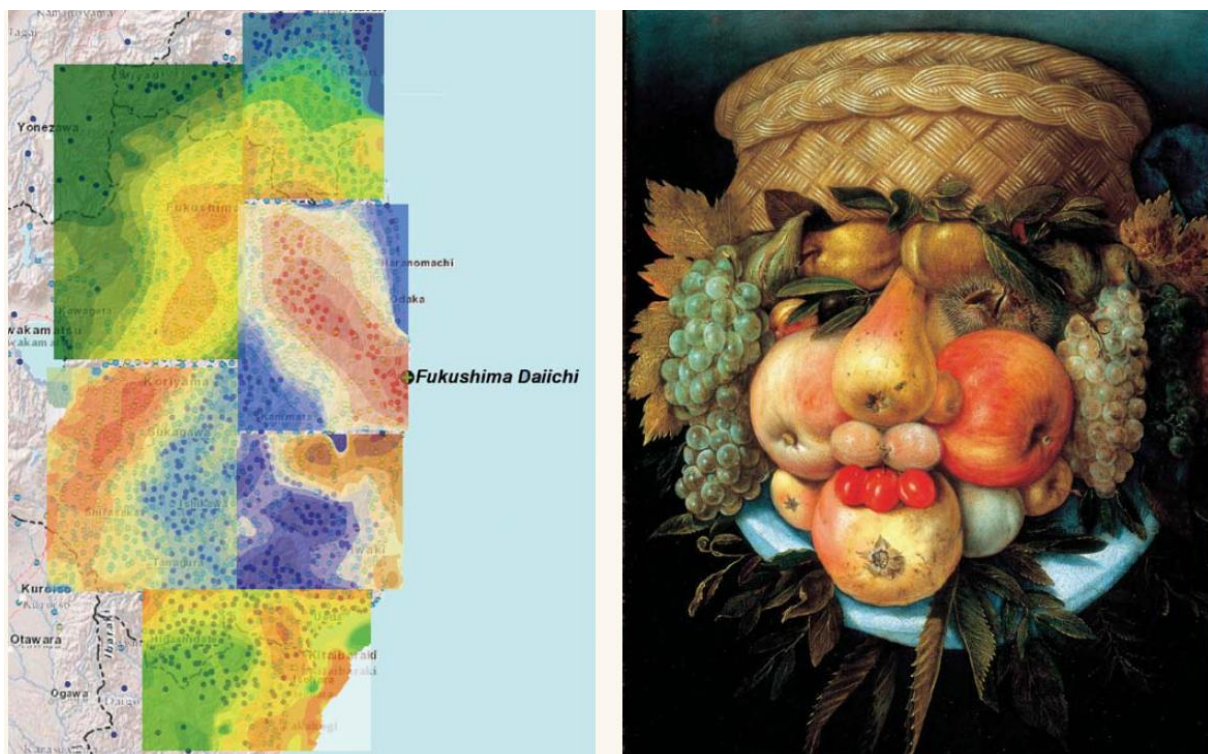


Figure 9. Résultat du krigeage empirique bayésien pour les concentrations en radiocésium dans les sols de la région de Fukushima Daiichi au Japon (à gauche : Krivoruchko, 2012) et « La corbeille de fruits » de Giuseppe Arcimboldo (à droite : ca. 1527–1593)

Dans leur récente revue bibliographique des méthodes d'interpolation spatiale en sciences environnementales, Li & Heap (2014) proposent plusieurs outils pratiques pour faciliter le choix de la méthode la plus appropriée à la situation d'étude et aux données d'observation. Les 25 méthodes répertoriées dans cette revue sont d'abord classées en fonction de 23 critères, selon une approche taxinomique et une analyse hiérarchique de clusters (détermine les ressemblances entre les méthodes). Au sein de ces critères, on retrouve par exemple les méthodes globales ou locales, exactes ou approchées, mais aussi le temps de calcul, le format des données en sortie, l'utilisation de covariables ou l'évaluation de l'erreur d'estimation. Un arbre de décision est ensuite défini en fonction de l'objectif d'estimation recherché, de la disponibilité et de la nature des données. Enfin, un tableau décrit la disponibilité de chaque méthode en fonction de différents logiciels existants (ArcGIS, R, Surfer, etc.). Cette revue propose également une synthèse des facteurs affectant la performance de ces méthodes, tels que la taille de l'échantillon, la densité de points, la forme de distribution statistique de la variable ou l'autocorrélation spatiale. Cette revue parue en 2014 nous a permis de confirmer le choix des méthodes utilisées dans le cadre de cette thèse.

Finalement, les applications de la géostatistique concernent non seulement la cartographie de variables environnementales, mais aussi l'analyse de données (variographie), l'intégration de différents types de données (dérive, co-krigeage) et la modélisation. Ces méthodes sont aussi employées pour réaliser des simulations, par exemple dans le cadre d'analyses de risques, afin d'évaluer la probabilité de dépassement d'une valeur seuil de concentration pour un polluant en tout point de l'espace, et ainsi identifier les populations à risque (Lemarchand & Jeannée, 2009).

Nous venons de détailler les méthodes spatialisées d'analyse de données environnementales. Ces données géostatistiques sont caractérisées par des informations quantitatives mesurées au niveau de points d'échantillonnage. En santé publique, les méthodes d'analyses spatiales sont basées sur

des données de format totalement différent. Les registres de santé répertorient de manière exhaustive les cas de maladies au sein de la population, sur une période et une étendue géographique donnée. Ainsi, l'information sous-jacente est de type binaire : malade / non-malade. Les outils d'épidémiologie spatiale sont alors employés pour évaluer les disparités d'incidence de ces maladies. Ils sont présentés dans la partie suivante.

Partie 3. L'épidémiologie spatiale

La cartographie des données de santé n'est pas une pratique récente (Bithell, 2000). L'un des exemples historiques les plus connus en épidémiologie spatiale est la carte de densité de décès liés au choléra dans le district de Soho de Londres, réalisée par John Snow (Snow, 1855). L'épidémiologie spatiale s'intéresse à la description et la compréhension des variations géographiques (et temporelles) de l'occurrence des maladies, basée sur des données observationnelles. Pour certaines pathologies, ces données sont recensées dans des registres, constitués suite au recueil exhaustif des cas au sein de la population générale. L'épidémiologie spatiale peut être mise en œuvre grâce aux outils que nous allons développer dans les paragraphes suivants. La cartographie des maladies examine l'hétérogénéité spatiale de l'incidence et la détection d'agrégats identifie des clusters atypiques d'évènements.

a. La cartographie des maladies

La cartographie des évènements de santé est un outil descriptif mis en œuvre pour évaluer les disparités spatiales d'incidence ou de mortalité des pathologies (Walter & Birnie, 1991 ; Elliott *et al.*, 2000 ; Waller & Gotway, 2004 ; Beale *et al.*, 2008 ; Beale *et al.*, 2010 ; Robertson *et al.*, 2010 ; Gorja *et al.*, 2011). Dans ce cadre, les évènements sont géoréférencés à partir de l'adresse du lieu de vie au moment de l'évènement, pour ensuite être rattachés à une unité géographique d'observation. Ainsi, pour des raisons souvent liées à un niveau de confidentialité élevé, les données relatant l'état de santé des populations sont la plupart du temps traitées de manière agrégée au sein d'unités spatiales (régions, départements, cantons, villes, quartiers, etc.).

En épidémiologie, le taux brut (représenté par le ratio du nombre total de nouveaux cas pour une période donnée sur l'effectif de population pour la même période) ou le taux spécifique (taux bruts calculés par classe d'âge et de sexe) sont classiquement utilisés pour décrire l'incidence d'une pathologie pour une population donnée (Bouyer *et al.*, 1993). Pour comparer cette occurrence d'évènements de santé au sein de plusieurs populations, appartenant par exemple à des unités géographiques différentes, ces indicateurs nécessitent une standardisation basée sur les facteurs de risque connus tels que l'âge et le sexe. En effet, une telle comparaison de l'incidence implique que les populations étudiées aient des structures d'âge et de sexe similaires, ce qui est très rarement le cas. Les taux standardisés d'incidence sont donc conçus pour intégrer ces différences. Cette standardisation peut également être effectuée selon d'autres facteurs de risques connus, comme la catégorie socio-professionnelle. Il existe deux catégories de standardisation : directe et indirecte.

La méthode de standardisation directe, ou méthode de population-type, aboutit au calcul d'un taux d'incidence standardisé (TIs), c'est-à-dire le taux d'incidence qu'aurait une population étudiée si elle avait la même structure qu'une population type de référence (Waller & Gotway, 2004). Le calcul se réalise en pondérant le taux spécifique d'incidence (taux ajusté sur les facteurs influents de la maladie) par la structure de la population de référence (dans les illustrations qui suivent, seul l'âge est considéré comme facteur influent) :

$$TIs_j = \sum_{i=1}^n P_i T_{ij}$$

Avec :

i : indice de classe d'âge

n : nombre de classes d'âge

j : unité géographique

P_i : part de la classe d'âge *i* dans la population de référence

T_{ij} : taux spécifique d'incidence observé pour la classe *i* de l'unité géographique *j*

Le TIs exprime une fréquence pour chaque unité géographique (ou population à comparer), le nombre de nouveaux cas par nombre de personnes-années (généralement par million d'habitant pour les maladies rares) recensés pendant la période d'étude. Deux TIs sont considérés comme significativement différents lorsque les intervalles de confiance (à 95% dans le cadre de ce travail) ne se recouvrent pas. La population de référence peut être composée de l'ensemble des unités géographiques à comparer, ou représenter la population nationale, continentale ou mondiale. Le choix de la population de référence dépend ainsi des objectifs et de la portée géographique de l'étude. Un indice comparatif d'incidence pour chaque unité spatiale peut être déduit du TIs en effectuant le rapport de celui-ci sur le TIs global de la zone d'étude. L'unité a une incidence significativement inférieure (ou supérieure) à l'incidence globale lorsque l'intervalle de confiance de l'indice comparatif ne contient pas la valeur 1.

La méthode de standardisation indirecte, ou méthode d'incidence-type, est plus classiquement utilisée pour la représentation cartographique des variations géographiques d'incidence (Waller & Gotway, 2004). Elle produit quant à elle un ratio d'incidence standardisé (SIR) en rapportant le nombre observé de cas au nombre attendu dans l'unité géographique. Le nombre de cas attendu est obtenu en appliquant le taux spécifique d'incidence d'une population type de référence à l'effectif spécifique de population de l'unité géographique :

$$SIR_j = \frac{O_j}{E_j} \quad \text{où} \quad E_j = \sum_{i=1}^n N_{ij} T_i$$

Avec :

i : indice de classe d'âge

n : nombre de classes d'âge

j : unité géographique

O_j : nombre total de cas observés (tous âges confondus) dans l'unité géographique *j*

E_j : nombre total de cas attendus dans l'unité géographique *j*

N_{ij} : effectif de la population correspondant à la classe d'âge *i* dans l'unité géographique *j*

T_i : taux spécifique d'incidence de la population de référence pour la classe d'âge *i*

Contrairement au TIs, le SIR ne représente pas une fréquence mais un indice comparatif de morbidité au sein de l'unité géographique, pendant la période d'étude. Un SIR supérieur à 1 indique une sur-incidence de la pathologie par rapport à la population de référence, alors qu'un SIR inférieur à 1 indique une sous-incidence de celle-ci. Un SIR de 1,3 implique une morbidité de 30% supérieure à celle observée pour la population de référence (et inversement pour un SIR de 0,7). Lorsque la valeur 1 n'est pas incluse dans l'intervalle de confiance du SIR, la population de la zone géographique en question présente une morbidité significativement plus élevée ou plus faible que la population de référence.

Ces deux indicateurs d'incidence standardisée sont complémentaires. Le TIs exprime une fréquence d'incidence pour laquelle nous allons comparer l'écart des taux observés entre les différentes unités géographiques (ou l'écart au taux d'incidence de la population de référence), alors que le SIR est une approximation du risque de développer la pathologie au sein de chaque unité géographique par rapport à la population de référence. Le TIs prenant directement en compte la structure de population, il est possible de comparer les taux entre les hommes et les femmes ou entre différentes classes d'âge (ce qui est moins évident pour le SIR), mais les résultats de cet indicateur sont très liés à la structure de la population de référence.

Les indicateurs d'incidence sont considérés comme très « instable » en cas de faibles variations du nombre de malades au sein d'unités géographiques. En effet, dans le cas de faibles effectifs de population ou de maladies rares, ils présentent une variabilité inversement proportionnelle au nombre de cas attendus dans la population étudiée (figure 10), ce qui peut engendrer l'observation de valeurs extrêmes dans les unités géographiques faiblement peuplées. Le résultat cartographique est dans ce cas bruité et difficilement interprétable en raison de la forte hétérogénéité des densités de population au sein du territoire.

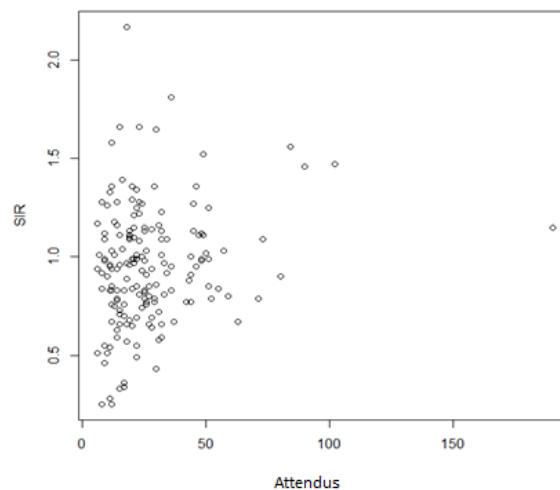


Figure 10. Instabilité du SIR en fonction du nombre de cas attendus (logiciel R)

Bien que le SIR fournit des résultats plus robustes que le TIs pour des petits effectifs de population, représenter cet indicateur de manière brute sur une carte implique de considérer les risques indépendamment d'une unité géographique à l'autre, sans tenir compte de l'information fournie par les unités voisines (Goria *et al.*, 2011). Pourtant, comme évoqué précédemment, le phénomène d'autocorrélation spatiale implique une similarité de la fréquence de la maladie dans des régions spatiales proches. Le risque de maladie d'une zone géographique n'est donc pas indépendant de celui des zones voisines.

Afin de présenter des cartographies pouvant être considérées comme des images fiables de la variabilité spatiale de la fréquence des maladies, ce phénomène de ressemblance est utilisé pour réduire la variance des estimations de risque, à travers des modèles de lissage de ces indicateurs (Clayton & Kaldor, 1987). Ces modèles ont été développés pour mieux estimer la structure spatiale sous-jacente des SIR et lisser le bruit observé dans les zones présentant un faible nombre de cas, en partageant l'information apportée par les unités voisines (Richardson, 1992 ; Elliott & Wartenberg, 2004). Plusieurs modèles de lissage sont recensés dans la littérature (Bithell, 2000 ; Best *et al.*, 2005 ; Goria *et al.*, 2011 ; Auchincloss *et al.*, 2012). Parmi les plus utilisés pour estimer les risques de maladies rares figurent les modèles hiérarchiques bayésiens, notamment celui de Besag, York et Mollié (1991), qui a l'avantage de proposer simultanément une structure spatiale globale et locale. Ce modèle fut utilisé dans le cadre de cette thèse pour évaluer les disparités d'incidence de l'insuffisance rénale chronique terminale en Nord – Pas de Calais.

b. La détection d'agrégats

Les techniques de cartographie des maladies sont employées pour évaluer l'hétérogénéité spatiale d'incidence des événements de santé. Bien qu'elles soient indispensables pour décrire la distribution géographique, elles ne permettent pas de détecter des zones d'incidence atypique, appelés clusters, et d'évaluer leur significativité. Un cluster est une concentration de cas « anormalement » faible ou élevée par rapport à celle attendue sur la zone (Goria *et al.*, 2011). De nombreuses méthodes, comme les statistiques de Moran (1950) et de Tango (1995 ; 2000), testent la tendance des événements à s'agréger dans l'espace en évaluant le niveau d'autocorrélation spatiale. Mais celles-ci ne fournissent pas d'information sur la localisation géographique des clusters. Pour identifier ces zones géographiques atypiques, les méthodes de balayage spatial peuvent être employées (Openshaw *et al.*, 1987 ; Besag & Newell, 1991 ; Kulldorff & Nagarwalla, 1995 ; Gomez-Rubio *et al.*, 2005). Parmi ces méthodes de détection zonale d'agrégats, la statistique de scan est la plus utilisée (Fritz *et al.*, 2013).

Développée par Martin Kulldorff dans la fin des années 1990, la statistique de scan détecte des clusters d'évènements spatiaux, temporels et spatio-temporels, sans biais de présélection. Cette détection peut être validée par un test de significativité pour chacun des clusters et ajustée par des facteurs confondants, comme l'âge et le sexe (Kulldorff 1997 ; Kulldorff *et al.* 1998 ; Glaz *et al.*, 2009). Le principe de cette statistique se déroule en deux étapes. Dans une première phase de détection, une fenêtre de forme et de taille variable balaye la zone d'étude (dans l'espace et/ou le temps) pour comptabiliser les événements observés à partir de chaque centroïde d'unité géographique (centre géographique représenté par des coordonnées XY). Pour chacun des emplacements et tailles de fenêtre, une fonction de probabilité (le risque relatif : RR) est alors calculée en considérant la distribution des cas de maladie selon une loi de Poisson. Le RR dépend à la fois du nombre de cas observés à l'intérieur de la fenêtre de scan et du nombre de cas observés à l'extérieur (figure 11). La deuxième phase est ensuite celle de l'inférence statistique. L'hypothèse nulle correspond à l'absence de cluster (le risque est homogène et constant sur l'ensemble de la zone et/ou de la période d'étude) et l'hypothèse alternative à la présence d'au moins un cluster atypique (le risque est différent à l'intérieur de la fenêtre par rapport à l'extérieur de la fenêtre). Un test du rapport de vraisemblance, consolidé par des simulations de Monte Carlo, calcule finalement une valeur de p pour chacun des clusters afin de déterminer son niveau de significativité. Cette analyse est réalisée sous le logiciel SaTScan™ (Kulldorff, 2010), dont les fichiers en sortie sont compatibles avec les SIG tel qu'ArcGIS®.

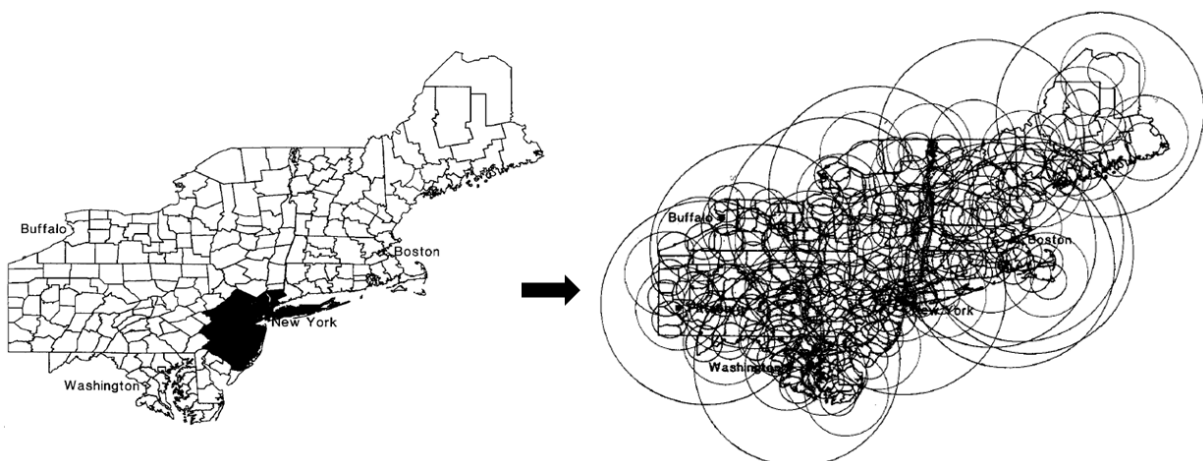


Figure 11. Un petit échantillon des nombreuses fenêtres de scan utilisées par SaTScan™ (modifié à partir de Kulldorff, 1999)

Les premières applications épidémiologiques de cette statistique furent consacrées à la leucémie infantile et à la mortalité par cancer du sein dans la région de New-York (Hjalmarsson *et al.* 1996; Kulldorff *et al.* 1997). Cette méthode a ensuite largement été appliquée à travers le monde. Le site internet proposant le téléchargement gratuit du logiciel recense par exemple plus de 200 publications concernant 40 thématiques différentes, comme la criminologie, la botanique ou encore l'archéologie (SaTScan, 2014).

Dans le Chapitre II, nous verrons également qu'une version plus élaborée de la statistique de scan permet d'estimer des risques non homogènes au sein de chaque cluster détecté (Kulldorff, 1999). En effet, la version isotonique de SaTScanTM peut déterminer plusieurs niveaux de risque à l'intérieur d'un même cluster grâce à une fonction de régression isotonique décroissante en fonction de la distance au centre du cluster (Barlow *et al.*, 1972). Cette fonction apporte une information supplémentaire, notamment pour l'identification précise des zones atypiques d'incidence dans le cas de clusters de grande taille.

Dans le domaine de l'épidémiologie spatiale, la mise en évidence de disparités d'incidence, ainsi que la détection de clusters atypiques supposent l'implication de facteurs de risques environnementaux. Les études spatialisées en santé – environnement sont alors mises en œuvre pour évaluer l'influence des contaminants suspectés. Celles-ci emploient des mesures physico-chimiques de surveillance de la qualité de milieux, mais peu d'entre elles intègrent des données de biosurveillance environnementale, qui renseignent sur les impacts biologiques des polluants environnementaux. De fait, les organismes sentinelles utilisés en biosurveillance sont des indicateurs de la qualité des milieux et représentent certainement un atout pour la recherche de facteurs étiologiques des maladies. Dans la partie suivante, nous verrons que les méthodes géostatistiques d'analyse spatiale peuvent notamment être appliquées aux données issues de la biosurveillance lichénique de la qualité de l'air.

Partie 4. La biosurveillance lichénique de la qualité de l'air

a. La biosurveillance végétale et fongique de la qualité de l'air

Les méthodes physico-chimiques de surveillance de la qualité des milieux mesurent les concentrations de composants présents dans l'environnement, permettant par la suite leur comparaison à des valeurs réglementaires. Cependant, ces techniques n'apportent pas d'information directe sur les effets des polluants sur les organismes. Quelles que soient les évolutions technologiques des systèmes de métrologie, la biosurveillance (basée sur des modèles fongiques ou végétaux) demeurera toujours la seule approche capable d'évaluer des effets biologiques d'une altération de l'environnement (Van Haluwyn *et al.*, 2011). La biosurveillance et ses différents concepts ont été formalisés par Garrec & Van Haluwyn (2002). La biosurveillance est alors définie comme « l'utilisation des réponses à tous les niveaux d'organisation biologique (moléculaire, biochimique, cellulaire, physiologique, tissulaire, morphologique, écologique) d'un organisme ou d'un ensemble d'organismes pour prévoir et/ou révéler une altération de l'environnement et pour en suivre son évolution ». Ces approches consistent à observer et mesurer de façon globale la réponse des organismes aux contaminants et intègre les effets du mélange des composés présents dans l'environnement, mais aussi les caractéristiques écologiques et climatiques du territoire (Van Haluwyn, 1998).

L'idée d'observer les effets de la pollution atmosphérique sur le compartiment végétal remonte à plusieurs siècles, avec la description de dégâts foliaires en 1661 par J. Evelyn, suite à la contamination de l'air par la combustion de charbon (Treshow & Bell, 2002). Mais le véritable concept de la biosurveillance végétale et fongique de la qualité de l'air date de la fin du XIX^e siècle où Nylander propose d'utiliser les lichens comme « hygiomètres » de la qualité de l'air, après avoir

constaté la progressive disparition des colonies à l'approche des villes, comme à Paris (Nylander, 1866). Depuis, la biosurveillance est employée pour répondre à différents objectifs, que ce soit l'analyse des distributions spatiales et temporelles des impacts des polluants ou l'identification et le suivi de sources ponctuelles de contamination. Elle est également de plus en plus intégrée aux méthodologies d'évaluations des risques sanitaires et écologiques et constitue un outil d'information du grand public et d'aide à la décision dans les politiques publiques (Cuny, 2012).

Cette approche n'est donc pas substitutive des techniques physico-chimiques de surveillance, mais intervient en complément de celles-ci (Cuny *et al.*, 2008). Elle présente divers avantages, puisqu'elle emploie de multiples modèles, traduisant des effets à plusieurs niveaux d'organisation biologique. Pour l'air, ce sont des végétaux supérieurs comme le ray-grass, le chou, le trèfle ou le tabac, des bryophytes ou encore des lichens (De Temmerman *et al.*, 2004 ; Van Haluwyn, 1993). La méthodologie peut se baser sur une approche *in situ*, également appelée approche intégrée ou passive, en observant les organismes présents naturellement dans l'environnement étudié. Elle peut aussi constituer une approche active, en introduisant dans le site d'étude des organismes cultivés dans des conditions contrôlées ou prélevés dans un site témoin (techniques de transplantation des lichens qui ne se cultivent pas par exemple), ou une approche de laboratoire en conditions entièrement contrôlées en termes de développement et d'exposition. La principale limite de cette méthode réside dans la nature biologique de l'indicateur, puisqu'elle dépend de la présence de l'organisme étudié en quantité suffisante, de sa sensibilité à d'autres facteurs (biotiques ou abiotiques) ou encore de sa compétitivité par rapport à d'autres organismes ou espèces (Cuny, 2012). Ainsi des facteurs écologiques peuvent avoir un impact suffisant pour conduire à une mauvaise estimation des effets des polluants étudiés.

La biosurveillance englobe quatre concepts : l'utilisation de biomarqueurs, la bioindication, la biointégration et la bioaccumulation.

L'utilisation de biomarqueurs : un biomarqueur est un « changement observable et/ou mesurable au niveau moléculaire, biochimique, cellulaire, physiologique qui révèle l'exposition présente ou passée d'un individu à au moins une substance chimique à caractère polluant » (Lagadic *et al.*, 1997). Il caractérise un effet de stress précoce, non visible et spécifique au niveau infra-individuel. Concept le plus récent dans le domaine de la biosurveillance végétale et encore en développement important, il repose par exemple sur le suivi de la fluorescence de la chlorophylle, de l'activité photosynthétique (Catalyud & Barreno, 2004 ; Crous *et al.*, 2006), des dosages d'enzymes (Rai & Agrawal, 2008), de l'intégrité membranaire ou de la génotoxicité (Rzepka & Cuny, 2011 ; Misik *et al.*, 2011).

La bioindication se situe au niveau individuel. Elle se fonde sur l'observation des effets cliniques et visibles des pollutions sur un organisme. Ces effets sont observables au niveau morphologique, tissulaire ou physiologique. Il s'agit par exemple de l'apparition de nécroses foliaires (Silva *et al.*, 2012) ou la modification de la croissance racinaire (Manier *et al.*, 2009).

La biointégration permet d'évaluer les effets à long terme sur les communautés par l'étude de la densité et de la diversité spécifique de populations au sein d'écosystèmes. Les organismes sentinelles utilisés sont généralement les végétaux supérieurs, les bryophytes et les lichens (Thimonier *et al.*, 1994 ; Takahashi & Miyajima, 2010).

A la différence des trois concepts précédents, la bioaccumulation ne mesure une réaction des organismes, mais se base sur le dosage de contaminants accumulés dans les résistances des tissus. Dans ce cas, l'organisme utilisé sert principalement de matrice pour le dosage de différents polluants suite à des mécanismes de transfert à partir de l'environnement et de sa distribution dans ses différents tissus. C'est une méthode extrêmement utilisée actuellement qui permet d'étudier l'imprégnation de l'environnement par les polluants et d'en suivre l'évolution spatiale et temporelle (Larsen *et al.*, 2007 ; Cloquet *et al.*, 2009 ; Gerdol *et al.*, 2014). Les concentrations tissulaires sont dans ce cas le reflet de ce qui se trouve dans l'environnement et correspondent à la part biodisponible qui, *in fine*, reste dans l'organisme. Ces quantités ne sont donc pas directement celles du milieu environnant.

b. La biosurveillance de la qualité de l'air à l'aide des lichens épiphytes

LES LICHENS

Parmi les organismes utilisés pour la biosurveillance de la qualité de l'air, les lichens épiphytes sont les représentants du règne fongique les plus couramment employés. Ils sont formés par l'association symbiotique entre un champignon et une algue verte (dans 90% des cas) ou une cyanobactérie (dans 10% des cas). Leur appareil végétatif est un thalle (dépourvu de système racinaire, de tiges ou de feuilles), celui-ci supportant également les organes de reproduction sexuées (apothécie, périthèce) et végétatives (isidies, soralies..). Les lichens se développent lentement (leur croissance varie entre moins d'un millimètre et un centimètre par an selon les espèces et le milieu de vie) et sont exclusivement dépendants de l'atmosphère pour leur nutrition (Loppi & Nascimbene, 2010). De fait, le prélèvement de nutriments se réalise majoritairement à partir de l'atmosphère au travers de la surface du thalle (cortex supérieur : Nieboer *et al.*, 1978). Contrairement aux végétaux supérieurs, ils ne possèdent ni cuticule cireuse, ni stomate qui pourraient moduler leur exposition aux polluants atmosphériques. Ils sont par conséquent très sensibles aux substances présentes dans l'air ambiant. Les polluants sont ainsi facilement absorbés par les lichens et perturbent profondément leur métabolisme (Conti & Cecchetti, 2001), entraînant parfois des effets morphologiques visibles. Les effets de la pollution atmosphérique peuvent aller jusqu'à modifier la composition des communautés lichéniques en fonction de la sensibilité différentielle des espèces. Du fait de leur longévité, de leur activité biologique annuelle et dotés d'une paroi fongique offrant une multitude de sites de fixation possible des polluants, les lichens sont aussi d'excellents bioaccumulateurs (Loppi *et al.*, 1997). Cette particularité permet la mesure de concentrations biologiques de polluants bioaccumulables, tels que les ETM, les radioéléments et les polluants organiques persistants. Les lichens constituent ainsi d'excellents indicateurs de la pollution atmosphérique de fond, au titre de bio-intégrateurs et bioaccumulateurs (Gadsdon *et al.*, 2010).

INTERETS DES LICHENS DANS L'ANALYSE SPATIO-TEMPORELLE DE LA QUALITE DE L'AIR

Compte tenu de leur forte propension à coloniser des milieux biogéographiques très variés, même dans des conditions extrêmes (Agnan, 2013), ils constituent un matériel d'observation de choix dans l'évaluation spatio-temporelle des effets de la qualité de l'air sur le vivant (Loppi *et al.*, 2004). La disparition ou apparition d'une espèce, la modification de son abondance relative ou la modification de la structure des peuplements lichéniques d'un écosystème signifient une modification de l'environnement, souvent induite par l'introduction d'un polluant dans le milieu et traduisent un état de la dégradation de la qualité de l'air (Echaubard, 1995). Ainsi, le suivi spatial et temporel employant la cartographie est l'application la plus historique et la plus importante de la biointégration lichénique. Comme cité précédemment, Nylander évoquait la raréfaction d'espèces lichéniques à proximité des grandes villes, comme Paris, et rapprochait ce phénomène à une mauvaise qualité de l'air (Nylander, 1866). C'est environ un siècle plus tard qu'est apparue la première méthode formalisée de biosurveillance lichénique, avec l'élaboration d'une échelle de correspondance entre environ 80 espèces de lichens et les teneurs moyennes hivernales en dioxyde de soufre (figure 12 ; Hawksworth & Rose, 1970). A cette époque et notamment dans le Nord-Ouest de l'Europe, la pollution atmosphérique était essentiellement d'origine industrielle et le dioxyde de soufre, très abondant, en était le traceur.

ZONES	ESPECES DE LICHENS	SO ₂ (µg·m ⁻³)
0	Epiphytes absents.	
1	<i>Pleurococcus viridis</i> limité à la base du tronc.	environ 170
2	<i>Pleurococcus</i> s.l. s'étend sur le tronc ; <i>Lecanora conizaeoides</i> est limité à la base.	environ 150
3	<i>Lecanora conizaeoides</i> s'étend sur le tronc ; <i>Lepraria incana</i> devient fréquent à la base.	environ 125
4	<i>Hypogymnia physodes</i> et/ou <i>Parmelia saxatilis</i> ou <i>Parmelia sulcata</i> apparaissent à la base. <i>Lecidea scolaris</i> , <i>Lecanora expallens</i> et <i>Chaenotheca ferruginea</i> souvent présents.	environ 70
5	<i>Hypogymnia physodes</i> ou <i>P. saxatilis</i> s'étendent jusqu'à 2,5 m ou plus ; <i>P. glabratula</i> , <i>P. subrudecta</i> , <i>Parmeliopsis ambigua</i> et <i>Lecanora chlorotera</i> font leur apparition ; <i>Calicium viride</i> , <i>Lepraria candelaris</i> , <i>Pertusaria amara</i> peuvent apparaître . Si <i>Ramalina farinacea</i> et <i>Evernia prunastri</i> sont présents, ils sont limités à la base ; <i>Platismata glauca</i> peut être présent sur les branches horizontales.	environ 60
6	<i>Parmelia caperata</i> présent au moins à la base, riches communautés de <i>Pertusaria</i> (<i>P. albescens</i> , <i>P. hymenea</i>) et de <i>Parmelia</i> (<i>P. biliacea</i> , <i>P. exasperatula</i>) ; <i>Graphis elegans</i> , <i>Pseudevernia furfuracea</i> et <i>Alectoria fuscescens</i> présents dans les régions montagneuses.	environ 50
7	<i>Parmelia caperata</i> , <i>P. revoluta</i> , <i>P. biliacea</i> , <i>P. exasperatula</i> s'étendent sur le tronc ; apparition de <i>P. hemisphaerica</i> , <i>Usnea subfloridana</i> , <i>Rinodina roboris</i> et <i>Arthonia imposita</i> .	environ 40
8	<i>Usnea ceratina</i> , <i>Parmelia perlata</i> ou <i>P. reticulata</i> apparaissent ; extension de <i>Rinodina roboris</i> ; <i>Normandina pulchella</i> et <i>U. rubigena</i> généralement présents	environ 35
9	<i>Lobaria pulmonaria</i> , <i>L. amplissima</i> , <i>Pachyphiale cornea</i> , <i>Dimerella lutea</i> ou <i>Usnea florida</i> présents ou lichens crustacés très bien développés avec généralement plus de vingt cinq espèces sur des arbres bien illuminés.	environ 30
10	<i>Lobaria amplissima</i> , <i>L. scrobiculata</i> , <i>Siccia limbatata</i> , <i>Pannaria</i> sp., <i>Usnea articulata</i> , <i>Usnea filipendula</i> ou <i>Teloschistes flavicans</i> .	pur

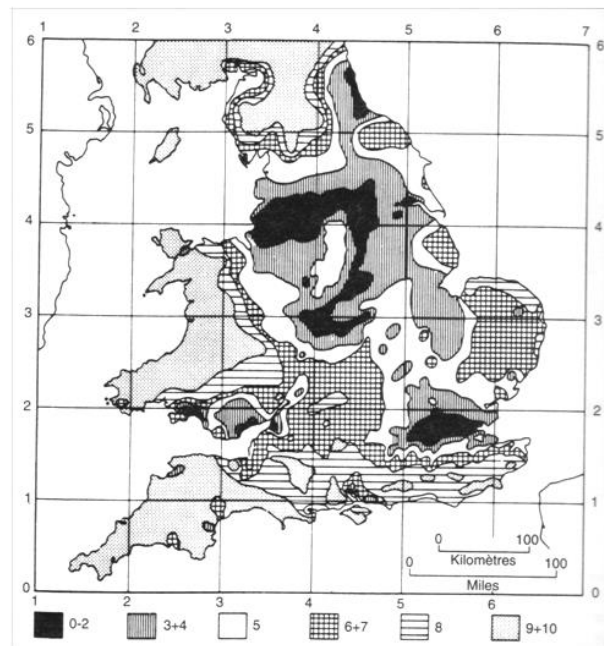


Figure 12. Echelle d'estimation de la qualité de l'air en Angleterre et au Pays de Galles à partir des lichens observés sur des écorces non eutrophisées et cartographie réalisée à partir de cette échelle (Hawksworth et Rose, 1970) (Gavériaux, 2005)

Cette méthode donna lieu à la première cartographie de la pollution atmosphérique à l'échelle d'un pays (courbes d'isopollution en SO₂ sur l'ensemble de la Grande-Bretagne) et sera appliquée en France par Chantal Van Haluwyn (figure 13) et Michel Lerond pour le quart Nord-Ouest de la France (Delzenne-Van Haluwyn, 1973 ; Lerond, 1981). Par la suite, Van Haluwyn et Lerond (1986) élaboreront une méthodologie permettant de produire, pour une région donnée, un « éco-diagnostic de la qualité de l'air ». Elle reposait non plus sur une échelle de concentration en SO₂, mais sur l'observation phytosociologique des groupements de lichens se fragmentant proportionnellement au niveau global de pollution atmosphérique. Les espèces épiphytes (se développant sur un végétal) et plus précisément corticoles (se développant sur les écorces des troncs d'arbres) ont été utilisées dans l'ensemble des méthodes car ce sont les moins sujettes à l'influence de leur substrat. De nombreux travaux ont depuis été engagés au plan international, employant des méthodologies diversifiées afin de mettre en évidence la dégradation des communautés lichéniques par les polluants atmosphériques (Conti & Cecchetti, 2001 ; Gombert *et al.*, 2004 ; Remy, 2011 ; Llop *et al.*, 2012). Récemment, la méthode française de l'IBLE (Indice Biologique de Lichens Epiphytes) fut normalisée par l'AFNOR en 2008 (Norme NF-X 43-903) et est en cours de normalisation au niveau européen.

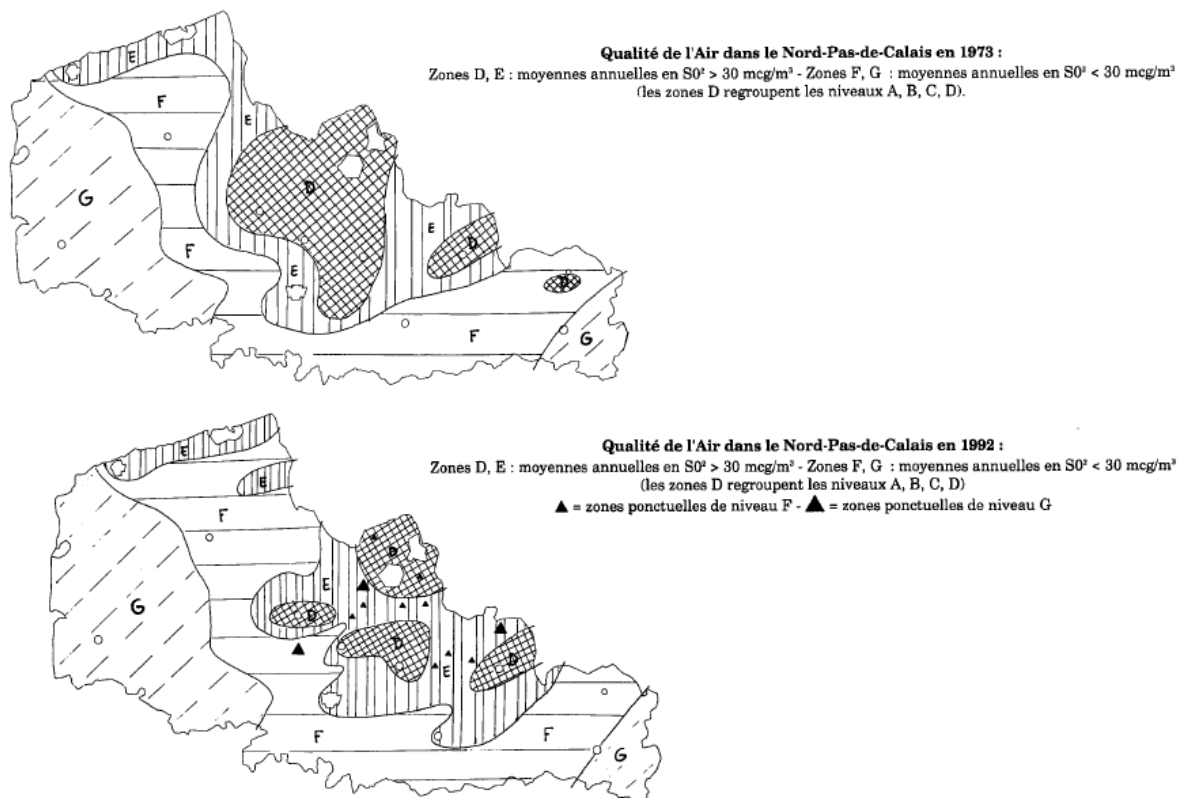


Figure 13. Evolution de la qualité de l'air entre 1973 et 1992 dans le Nord – Pas de Calais (Van Haluwyn, 1993)

Le concept des lichens bioaccumulateurs a connu un essor plus tardif (Ferry *et al.*, 1973 ; Bargagli, 1998). Les contaminants de l'environnement, et notamment les éléments traces métalliques, se déposent à la surface des lichens sous forme de particules sèches ou lors de précipitations sous forme humide, puis sont absorbés et accumulés dans les thalles (Garty, 1993, 2001). Les dosages de ces éléments dans les tissus lichéniques reflètent ainsi des expositions récentes à la pollution atmosphérique (Rzepka & Cuny, 2008). Plusieurs auteurs ont d'ailleurs démontré que les concentrations en ETM mesurées chez les lichens sont directement corrélés à celles retrouvées dans l'environnement (Andersen *et al.*, 1978 ; Sloof & Wolterbeek, 1991 ; Bari *et al.*, 1998), mais également proportionnelles à la période d'exposition (Bergamaschi *et al.*, 2007).

Ces techniques sont la plupart du temps employées pour le suivi des éléments traces métalliques à proximité de sources d'émissions routières et industrielles (Cuny *et al.*, 2001 ; Garty, 2001 ; Loppi *et al.*, 2004 ; Aznar *et al.*, 2008 ; Boamponsem *et al.*, 2010), ainsi que le suivi spatio-temporel de leur imprégnation (Sloof & Wolterbeek, 1991 ; Freitas *et al.*, 1999 ; Loppi & Pirintzos, 2003 ; Loppi *et al.*, 2004 ; Scerbo *et al.*, 1999 ; 2003 ; Fuga *et al.*, 2008 ; Balabanova *et al.*, 2012). Au sein de la région Nord - Pas de Calais, une étude menée sur le littoral dunkerquois montrait qu'en 1995, l'ensemble du territoire présentait des niveaux de concentrations en arsenic, cadmium, plomb, zinc, nickel et vanadium supérieures à la fois au bruit de fond régional et aux limites classiquement retenues par la littérature scientifique. Les contaminations plus importantes étaient situées à proximité des zones industrielles. Ces cartographies ont ensuite été comparées à celles des dosages réalisés sur le même territoire en 2002 afin de suivre l'évolution des contaminations métalliques (Figure 14 : Cuny *et al.*, 2004). Ces méthodes font désormais l'objet d'une normalisation française, prochainement étendue à l'Europe (AFNOR, 2013 : NF X43-904).

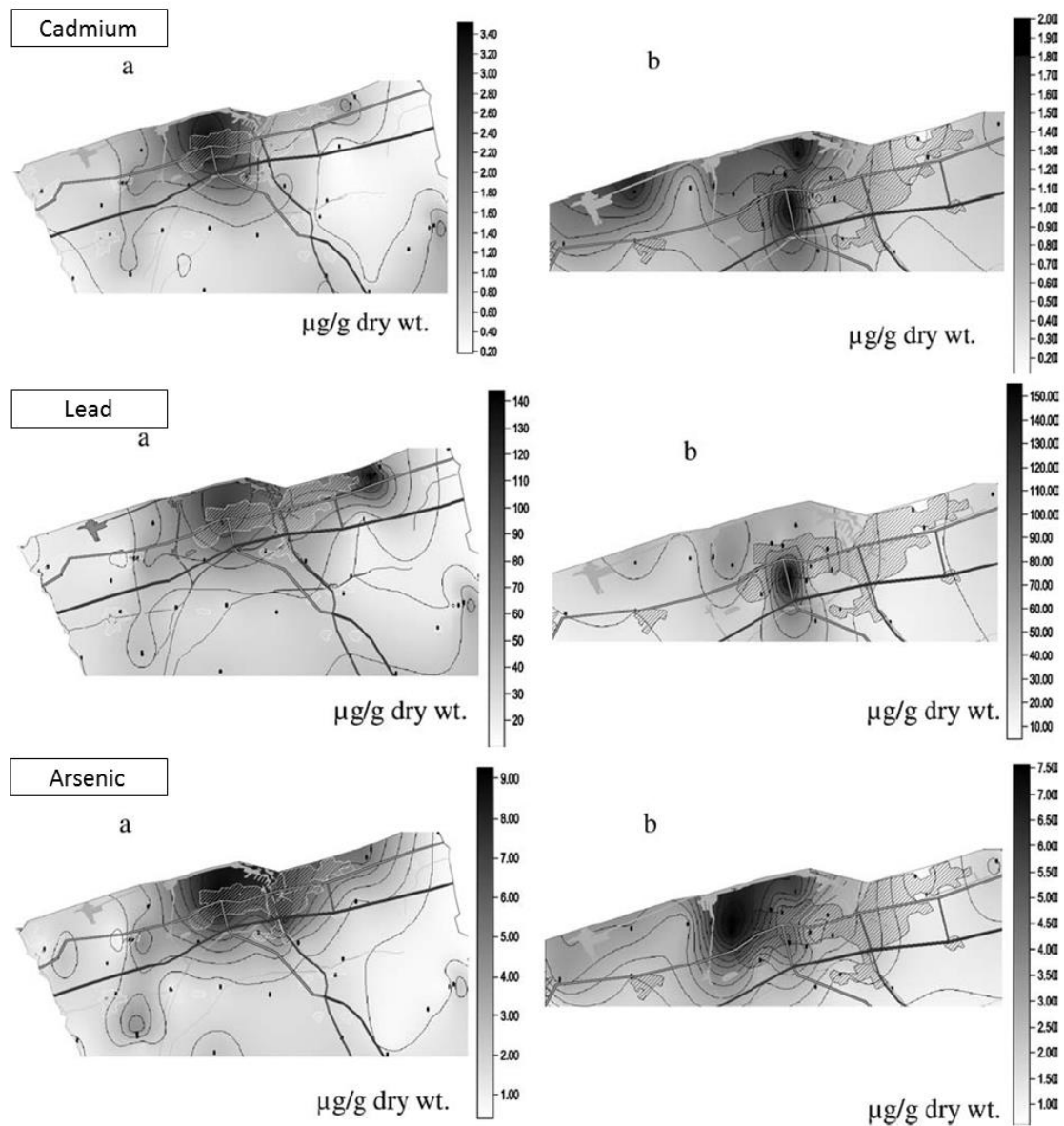


Figure 14. Distribution des concentrations en cadmium, plomb et arsenic chez le lichen *Xanthoria parietina* sur la région de Dunkerque (a) en 1995 et (b) en 2002 (modifié à partir de Cuny *et al.*, 2004)

Comme le montre la figure précédente, des méthodes géostatistiques d'interpolation spatiale sont utilisés dans ce cadre pour établir les cartographies d'imprégnation. La méthode la plus utilisée est le krigeage (Cuny *et al.*, 2002 ; Pignata *et al.*, 2002 ; Otvos *et al.* 2003 ; Loppi *et al.*, 2004). A la différence du concept de biointégration, la bioaccumulation est mesurée chez une sélection d'espèces lichéniques largement disponibles sur les territoires d'études.

Selon Conti et Cecchetti (2001), les principales caractéristiques que doit avoir une espèce pour être utilisée comme bioaccumulatrice sont :

- accumuler le polluant sans être tué par les concentrations avec lesquelles il entre en contact ;
- avoir une vaste répartition géographique ;
- être abondant ;
- être disponible tout au long de l'année en quantité suffisante pour être facilement récolté ;
- être facile à récolter, résistant aux conditions de laboratoire.

L'UTILISATION DES LICHENS EPIPHYTES DANS UN CONTEXTE EPIDEMIOLOGIQUE

Malgré l'intérêt de ces méthodes de biosurveillance, leur utilisation dans le domaine de l'épidémiologie et de la santé environnementale est assez récente. Beaucoup de travaux se basent sur les mesures physico-chimiques pour étudier les effets des polluants sur la santé, mais très peu utilisent les mesures d'effets biologiques réalisées chez les végétaux ou les lichens (Cuny, 2012).

Cette réticence peut résider dans la difficulté de positionnement de la biosurveillance dans les études santé – environnement. En effet il est difficile d'extrapoler à l'Homme des effets observés chez des organismes végétaux et fongiques. L'Homme est exposé à son environnement de manière plus complexe (voies d'entrées multiples, mobilité, activité professionnelle, etc.) et possède un système de métabolisation différent. La biosurveillance ne peut donc pas être utilisée pour mesurer des impacts sanitaires, mais se rapproche plus du domaine de l'évaluation des risques, en tant qu'indicateur biologique de potentiel (éco)toxique du milieu (Van Haluwyn *et al.*, 2011).

Certaines études illustrent en effet qu'il est réaliste d'envisager un lien entre évaluation de l'imprégnation de l'environnement, évaluation de l'imprégnation des populations et incidence de pathologies. L'un des premiers travaux de grande ampleur concernent la pollution atmosphérique globale en Vénétie (Nord de l'Italie). Cislighi & Nimis (1997) ont comparé la cartographie de la mortalité par cancer pulmonaire avec celle des indices de biodiversité lichénique. Les résultats montrent qu'il existe une correspondance importante entre les deux cartographies obtenues (figure 15). Les zones présentant un impact important de la pollution atmosphérique sur la diversité lichénique sont également celles pour lesquelles la mortalité par cancer du poumon est élevée.

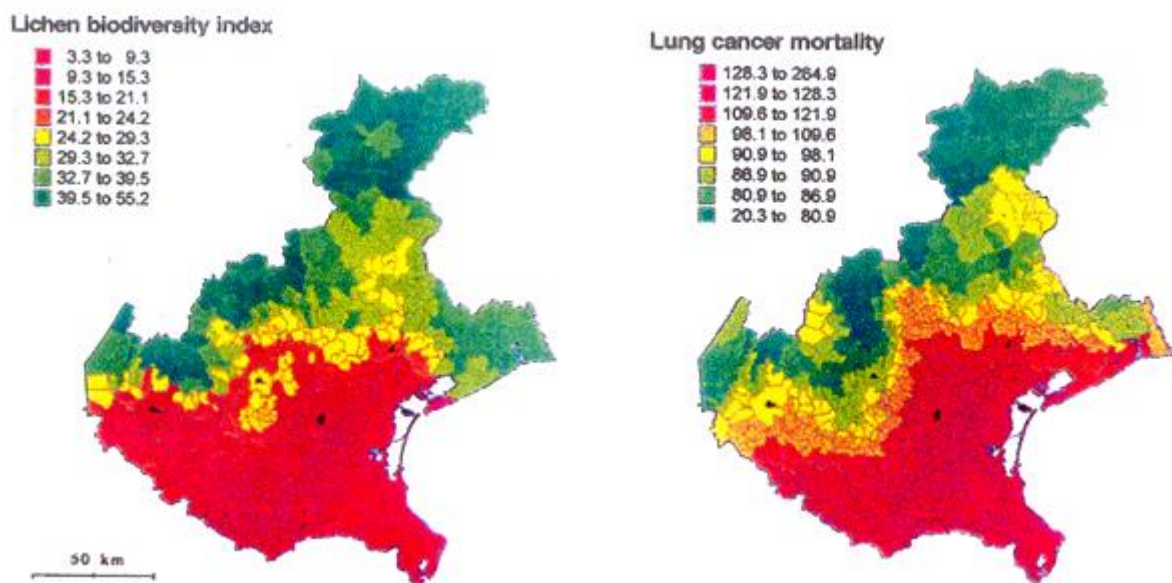


Figure 15. Cartographies des indices de diversité lichénique (à gauche) et de la mortalité par cancers pulmonaires chez les hommes de moins de 55 ans (à droite) en Vénétie, Nord de l'Italie (Cislighi & Nimis, 1997)

Concernant les métaux, les premiers travaux d'envergure ont été réalisés en Europe centrale. Wappelhorst *et al.* (2000) ont mis en évidence des corrélations positives significatives, par exemple entre les concentrations en cérium, fer, gallium et germanium chez les mousses et l'incidence des tumeurs de la trachée, des bronches, des poumons et des pathologies de la peau et des tissus sous cutanés. De même, les concentrations de thallium étaient positivement corrélées avec les incidences des pathologies cardio-vasculaires (l'hypertension artérielle, infarctus du myocarde, pathologies ischémiques) et des broncho-pneumopathies chroniques obstructives. Pour ces auteurs, les corrélations positives témoignent que l'augmentation des dépôts des éléments concernés dans

l'environnement provoque ou favorise l'apparition de ces pathologies. Wolterbeek & Verburg (2004) ont également observé des corrélations entre les dépôts atmosphériques métalliques mesurés chez les mousses et plusieurs indicateurs de santé (figure 15).

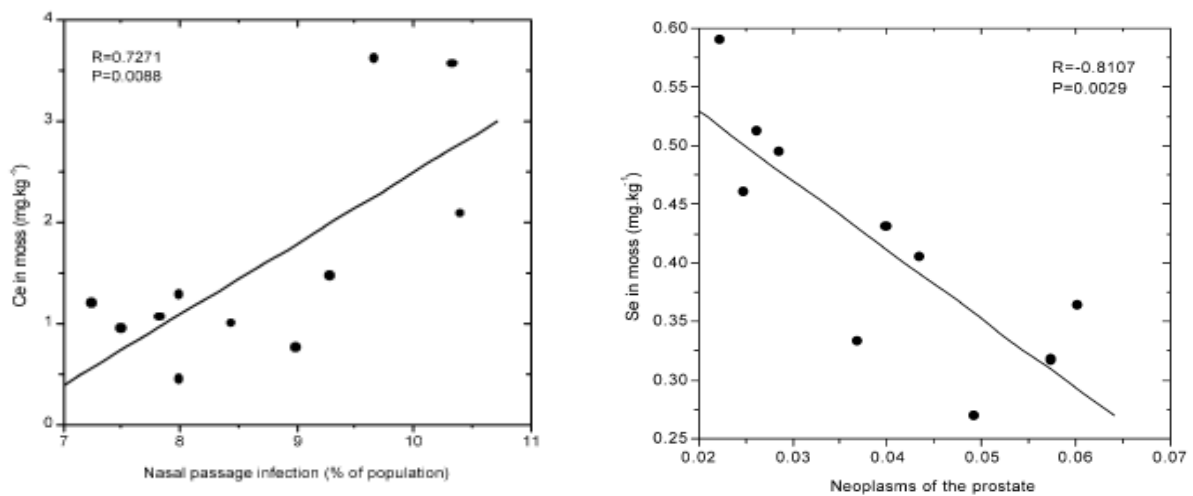


Figure 16. Corrélation entre les concentrations de cérium et les infections nasales (à gauche) et les concentrations de sélénium et le cancer de la prostate (à droite). Concernant le sélénium, des relations similaires ont été décrites avec les pathologies du système circulatoire et digestif (Wolterbeck & Verburg, 2004)

Dans une étude consacrée à la bioaccumulation des ETM dans les lichens et les mousses en Ecosse, Gailey & Lloyd (1993) démontraient le potentiel informatif de la biosurveillance sur l'évolution spatiale et temporelle de la pollution atmosphérique et son aide à l'interprétation des disparités de pathologies respiratoires. Plus récemment, des études se sont intéressées aux relations entre la diversité lichénique ou leur imprégnation par les métaux et l'apparition de diverses pathologies, mais mettent en exergue à la difficulté d'associer une pathologie à un élément en particulier (Fuga *et al.*, 2008 ; Sarmiento *et al.*, 2008 ; Carreras *et al.*, 2009). Ces études justifient néanmoins les apports de la biosurveillance et l'utilisation d'organismes sentinelles dans le domaine des risques sanitaires, ainsi que les potentialités de leur représentation cartographique par les SIG. En s'appuyant sur des cartographies de diversité ou d'imprégnation lichénique, il est par exemple envisageable de définir des zones à risque pour les populations, argument important pour mener par la suite des investigations épidémiologiques au sein de ces zones.

Dans le cadre de nos recherches, nous nous sommes intéressés aux mesures de concentrations en différents ETM, effectuées chez des lichens épiphytes. L'analyse spatiale de ces mesures aboutit à l'élaboration de 18 cartographies lichéniques. Ces cartographies sont destinées à être confrontées aux cartographies de santé. Pour certaines pathologies, les facteurs de risques environnementaux sont peu ciblés. Afin d'identifier un possible rôle de l'environnement, nous nous sommes alors intéressés à la genèse d'un indicateur capable de caractériser l'état global de contamination par les métaux. Pour ce faire, nous nous sommes inspirés des indices intégrés de qualité des milieux, détaillés dans la partie suivante.

Partie 5. Les indices intégrés de qualité des milieux

Dans la majeure partie des études en santé – environnement, la description qualitative des milieux environnementaux et, par conséquent, le lien vers les effets sanitaires font l’objet d’une interprétation substance par substance et ne procurent pas d’indice intégrateur de l’ensemble des contaminants. Pourtant, l’Homme se situe dans un contexte récurrent de pollutions complexes, du fait notamment de la multiplicité des activités anthropologiques émettrices. Il est quotidiennement exposé à son environnement d’une manière globale, et donc à l’ensemble des contaminants environnementaux qu’il contient. Ces constats nous amènent à réfléchir en termes de pollution générale et intégrée. En ce sens, bien qu’il soit encore difficile de générer un indicateur permettant de caractériser la qualité globale d’un milieu, tous polluants confondus, certains indices intégrant plusieurs contaminants de manière concomitante ont été développés. Dans le cadre de nos travaux de recherche, nous nous sommes intéressés aux indices de pollution intégrant les concentrations de plusieurs polluants.

Plusieurs synthèses bibliographiques ont été réalisées sur la base des indices intégrés, utilisés pour caractériser les différents milieux environnementaux. L’une des premières synthèses répertoriées est un ouvrage sur les indices intégrés (physiques, biologiques et chimiques) de qualité de l’air et de l’eau (Ott, 1978). Plus récemment, les articles de Cairncross *et al.* (2007) et de Bharti & Katyal (2011) offrent une vue d’ensemble de ces derniers. Concernant la qualité des sols et des sédiments, plusieurs articles élaborent une synthèse des indices intégrés proposés depuis les 40 dernières années (Caeiro *et al.*, 2005 ; Qingjie *et al.*, 2008 ; Zhao *et al.*, 2012).

Grâce à cette littérature, les indices intégrés de qualité des milieux environnementaux peuvent être catégorisés selon plusieurs critères. Le premier critère concerne la méthode de construction. Deux étapes distinctes de calcul sont généralement nécessaires. La première consiste à générer un sous-indice (P) pour chacun des paramètres i mesurés, afin de leur attribuer une échelle commune. Trois méthodes ont été répertoriées :

- le calcul d’un ratio représentant le rapport de la concentration mesurée (C_i) sur une valeur seuil de référence (B_i) ;

$$\rightarrow P_i = C_i / B_i$$

- la définition de classes de concentrations ;

$$\rightarrow$$

C_i (en $\mu\text{g}/\text{m}^3$)	$C_i < 10$	$10 \leq C_i < 20$...	$C_i > 100$
P_i (classe)	1	2	...	10

- la pondération de la concentration mesurée par un indicateur T_i (cet indicateur est généralement attribué en fonction du potentiel toxique ou écotoxique de la substance en question ou de l’importance perçue de cette substance dans le milieu).

$$\rightarrow P_i = C_i \times T_i$$

La deuxième étape de calcul représente ensuite le caractère multiparamétrique de l’indice intégré. Des méthodes simples d’agrégation ont été identifiées, telles que la somme, le produit, la moyenne des sous-indices ou le choix du paramètre ayant le sous-indice le plus critique en termes de contamination. Mais il existe aussi d’autres méthodes plus complexes intégrant par exemple des facteurs de pondération, des fonctions mathématiques de puissances et de racines. Le tableau 2 illustre la diversité des méthodes de calcul.

Ces indices peuvent aussi se différencier selon leur objectif intrinsèque. Trois types ont ainsi été identifiés. Les indices de contamination apportent une information sur le niveau global de

contamination du milieu en termes de concentrations. Les indices d'enrichissement renseignent sur l'état d'enrichissement du milieu en rapport à une valeur seuil, tels que le bruit de fond ou la norme environnementale. Enfin, les indices de risque expriment un potentiel toxique ou écotoxique du milieu, notamment par l'utilisation de valeurs de références sanitaires ou écologiques (tableau X).

Tableau 2. Diversité des indices intégrés de qualité des milieux environnementaux – avec C_i (concentration mesurée du paramètre i), B_i (valeur de fond, norme, etc.), T_i (facteur de toxicité ou écotoxicité), P_i (indice de pollution du paramètre i), W_i (poids attribué au paramètre i), $MaxP_i$ (P_i maximal observé sur l'ensemble des paramètres intégrés), \bar{P}_i (moyenne des P_i) et n (nombre de paramètres intégrés)

Type d'indice	Méthode de calcul		Exemple d'indice dans la littérature
	Sous-indice	Indice intégré	
Contamination	Classes de concentrations (6 ETM) P_i de 1 (faible) à 6 (élevée)	Somme des P_i ICMM de 6 (faible) à 36 (élevée)	Indice de Charge Multi-Métallique (ICMM) Duigou & Baize, 2010
	Classes de concentrations (4 polluants) P_i de 1 (très bon) à 10 (très mauvais)	P_i maximal observé Indice ATMO de 1 à 10	Indice ATMO
Enrichissement	$P_i = C_i/B_i$	$DC = \sum_{i=1}^n P_i$	Degree of Contamination (DC) Hakanson, 1980
	$P_i = C_i/B_i$	$IPI = \frac{\sum_{i=1}^n P_i}{n}$	Integrated Pollution Index (IPI) Chen <i>et al.</i> , 2005
	$P_i = C_i/B_i$	$PLI = \sqrt[n]{P_1 \times P_2 \times \dots \times P_n}$	Pollution Load Index (PLI) Tomlinson <i>et al.</i> , 1980
	$P_i = C_i/B_i$	$LPI = \sum_{i=1}^n W_i \times P_i$	Leachate Pollution Index (LPI) Kumar & Alappat, 2004
	$P_i = C_i/B_i$	$NPI = \sqrt{(MaxP_i^2 + \bar{P}_i^2)/2}$	Nemerow Pollution Index (NPI) Cheng <i>et al.</i> , 2007
Risque	$P_i = C_i \times T_i$	$RI = \sum_{i=1}^n P_i$	Risk Index (RI) Hakanson, 1980

Finalement, ces indices se diversifient en fonction du milieu environnemental pour lequel ils ont été développés. Les paragraphes suivants se consacreront à leur synthèse par type de matrice.

Concernant le milieu atmosphérique, les indices intégrés ont dans un premier temps été développés pour exprimer un potentiel impact sanitaire de la qualité de l'air (Murena, 2004). Pour cela, la majeure partie (Air Quality Indices : AQI ou Air Pollution Index : API) se base sur les méthodes développées aux Etats-Unis et en Grande-Bretagne par les instances agréées pour la surveillance journalière de la qualité de l'air des villes (US EPA, 1998 & 1999 ; UK National Air Quality Information Archive ; Joseph, 2002). Ces instances attribuent un sous-indice à chacun des polluants atmosphériques réglementés (SO_2 , PM_{10} , NO_2 , CO , O_3 et dans certains cas les $PM_{2.5}$) à partir d'une échelle de concentrations composée de classes. Ces classes sont définies en fonction de valeurs à partir desquelles un effet sanitaire peut être observé (de très bon : risque faible à très mauvais : risque élevé). Le sous-indice le plus élevé parmi les polluants mesurés constitue alors l'AQI (Maynard and Coster, 1999). De fait, même s'il renvoie à un effet sanitaire, cet indice représente un niveau de contamination du milieu. En France, c'est l'Indice ATMO qui fait référence (Fédération ATMO France, 2014). Déterminé de manière journalière à partir de quatre contaminants indicateurs de la pollution atmosphérique (SO_2 , NO_2 , O_3 , PM_{10}) et sur la base de seuils réglementaires définis par l'arrêté du 22 juillet 2004 relatif aux indices de qualité de l'air, l'Indice ATMO croît de 1 (très bon) à 10 (très mauvais). Le tableau 3 suivant synthétise sa méthode de calcul.

Tableau 3. Indice ATMO synonyme d'un état de la qualité de l'air en France (Fédération ATMO France, 2014)

Valeurs réglementaires (concentrations en µg/m ³)				
Indice	Poussières en suspension PM10	Dioxyde d'azote NO ₂	Ozone O ₃	Dioxyde de Soufre SO ₂
10 - très mauvais	80+	400+	240+	500+
9 - mauvais	65-79	275-399	210-239	400-499
8 - mauvais	50-64	200-274	180-209	300-399
7 - médiocre	42-49	165-199	150-179	250-299
6 - médiocre	35-41	135-164	130-149	200-249
5 - moyen	28-34	110-134	105-129	160-199
4 - bon	21-27	85-109	80-104	120-159
3 - bon	14-20	55-84	55-79	80-119
2 - très bon	07-13	30-54	30-54	40-79
1 - très bon	0-6	0-29	0-29	0-39

Ces indices calculés en routine se focalisent sur les substances réglementées de l'atmosphère et sont essentiellement destinés à sensibiliser et informer le public sur le niveau de qualité de l'air. En dépit de leur généralisation, aucune méthode de construction n'est actuellement reconnue au niveau international. En effet, celles-ci diffèrent selon plusieurs points en fonction des instances. L'article de Cairncross *et al.* (2007) porte sur la comparaison des méthodes de calcul pour 19 pays. Les auteurs ont notamment identifié plusieurs facteurs susceptibles d'interférer dans leur interprétation :

- les paramètres intégrés au calcul ;
- la valeur de concentration utilisée (moyenne court terme ou long terme, pic de concentration, etc.) ;
- la sélection des valeurs seuils (normes environnementales, valeurs de référence sanitaire) ;
- la définition d'une échelle relative de risque potentiel ;
- la fonction d'agrégation (sous-indice le plus élevé, somme des sous-indices, etc.).

En dehors du cadre du suivi journalier de la qualité de l'air par les instances agréées, très peu d'études se sont intéressées à l'intégration d'autres contaminants, tels que les métaux ou les composés organiques volatils (Bezuglaya *et al.*, 1993 ; Qiao *et al.*, 2013). Les indices alors calculés ne sont pas construits à partir de classes de concentrations, mais plutôt par la somme ou le produit de ratios de pollution, afin d'exprimer un niveau d'enrichissement du milieu atmosphérique par rapport à une norme environnementale.

Une recherche bibliographique a enfin été menée concernant l'application d'indices intégrés de pollution dans le cadre de la biosurveillance de la qualité de l'air, dans le but de mesurer le niveau d'imprégnation global d'organismes sentinelles par les contaminants. Boamponsem *et al.* (2010) et Salo *et al.* (2012) ont par exemple utilisé un indice de charge multi-métallique des sédiments, le Pollution Load Index (PLI : Tomlinson *et al.*, 1980), pour évaluer le niveau de charge accumulé dans les lichens. Madejon *et al.* (2006) et Guéguen *et al.* (2012) se sont pour leur part intéressés au même indicateur, respectivement dans les feuilles ou fruits et dans les écorces d'arbres.

Le milieu aquatique fait l'objet du plus grand nombre d'articles concernant l'élaboration d'indices de qualité (Water Quality Index : WQI) ou de pollution (Water Pollution Index : WPI) depuis plusieurs dizaines d'années (Horton, 1965 ; Brown *et al.*, 1970 ; Landwehr & Deininger, 1976 ; Bhargava, 1983 ; Gijanovic, 1999 ; Swamee & Tyagi, 2000 ; Rajagopal *et al.*, 2010). Béron *et al.* (1982) estimaient déjà à plus d'une centaine le nombre de publications scientifiques proposant chacune un nouvel indice.

L'ensemble de ces indices se distingue principalement par la fonction mathématique d'agrégation utilisée. Singh *et al.* (2008) recensent et critiquent 14 de ces fonctions. Ils définissent pour cela plusieurs critères permettant de sélectionner la fonction la plus adéquate au contexte d'étude. Selon ces auteurs, deux phénomènes sont tout d'abord à éviter : l'ambiguïté (surestimation de l'indice agrégé), lorsque l'indice excède la valeur critique définie alors qu'aucun des paramètres intégrés

n'excède de valeur seuil, ou à l'inverse l'éclipse (sous-estimation de l'indice), lorsque l'indice présente une valeur trop faible par rapport aux valeurs des paramètres. L'indice se doit ensuite d'être sensible aux variations d'un paramètre. Finalement, les méthodes les plus simples, transparentes et faciles à interpréter sont à préconiser.

Bharti & Katyal (2011) ont réalisé une revue historique conséquente et une analyse critique des indices de qualité de l'eau les plus utilisés. Ils isolent finalement 8 WQI perçus comme de simples, basiques et performants indicateurs de qualité des eaux. Parmi eux figure notamment l'Overall Index of Pollution (OIP : Sargaonkar & Deshpande, 2003), qui a servi de base pour la construction d'un indice multi-métallique lors de cette thèse.

Les indices intégrés de la qualité de l'eau se différencient également par le nombre et la nature des paramètres intégrés, qu'ils soient physiques, chimiques ou biologiques (Smith, 1990). Ils peuvent ainsi combiner les trois types de paramètres, incluant par exemple plusieurs métaux, le sodium, les nitrates, les chlorures, les fluorures, l'oxygène dissout, le pH, la dureté, la turbidité, les matières en suspensions et les bactéries. Une telle variété de paramètres nécessite d'attribuer un facteur de pondération à chaque paramètre, en fonction de sa contribution à la pollution (Singh *et al.*, 2008). Cette pondération peut varier selon l'objectif de qualité évalué en fonction de l'usage (eau de baignade, eau potable, cours d'eau, etc.) et fait encore l'objet de controverses (Kumar & Alappat, 2004 ; Singh *et al.*, 2008).

En France, la qualité des cours d'eau est évaluée à l'aide de l'indice SEQ eau (Système d'Évaluation de la Qualité des cours d'eau). Pour une série de paramètres physico-chimiques et biologiques, une grille associe les valeurs à 5 classes de qualité. Cette grille permet de définir si l'eau est apte ou non en fonction de l'usage. Dans la même logique que l'indice ATMO, la qualité de l'eau est déterminée par le paramètre le plus déclassant, sous condition qu'il le soit dans au moins 10% des prélèvements (Eau France, 2014).

Enfin la qualité matrice solide englobe la qualité des sols et des sédiments. Les indices intégrés des matrices solides s'adressent préférentiellement aux contaminations multi-métalliques et aux PCB (Polychlorobiphényles). Les premiers indices ont été développés dans des travaux de suivi de la contamination des sédiments. Le DC (Degree of Contamination : Hakanson, 1980) et le PLI (Pollution Load Index : Tomlinson *et al.*, 1980) caractérisaient le niveau de charge globale en contaminants, tandis que le RI (Potential ecological Risk Index : Hakanson, 1980) évaluait le potentiel risque écologique de la contamination (ces trois indices sont présentés dans le tableau 3 précédent). Ils ont depuis été repris et adaptés par plusieurs auteurs et peuvent se distinguer par les contaminants intégrés, les valeurs seuils utilisées ou la méthode d'agrégation employée (Chapman, 1990 ; Usero *et al.*, 1996 ; Long & MacDonald, 1998 ; Fairey *et al.*, 2001 ; Singh *et al.*, 2002 ; Ikem *et al.*, 2003 ; Chen *et al.*, 2005 ; Cheng *et al.*, 2007).

Qingjie *et al.* (2008) recensent 8 méthodes mathématiques d'agrégation différentes, catégorisées en deux groupes. Le premier groupe de fonctions intégrant les sommes et moyennes est adapté aux sous-indices présentant une distribution normale. Le deuxième groupe concerne les sommes géométriques et moyennes géométriques, fonctions utilisés pour des sous-indices présentant une distribution log-normale. Caeiro *et al.* (2005) ont recensé et classé 8 indices intégrés de qualité de sols en leur attribuant un score de performance selon six critères (simplicité, représentativité, crédibilité, compatibilité, sensibilité / robustesse, niveau d'incertitudes). Le Marine Sediment Pollution Index (MSPI : Shin & Lam, 2001) du fait de sa sensibilité / robustesse et du faible niveau d'incertitude et le New Pollution Index (PIN, défini par les auteurs) du fait de sa simplicité et de sa comparabilité semblent être les plus performants, contrairement au Metal Pollution Index (MPI : Usero *et al.*, 1996). Enfin Zhao *et al.* (2012) s'intéressent plutôt à l'apport de 4 indices multi-métalliques dans l'évaluation des risques environnementaux et conseillent la prise en compte de la spéciation des éléments.

Les synthèses bibliographiques de Caeiro *et al.* (2005) et Zhao *et al.* (2012) répertorient un total de 13 indices intégrés de la qualité de sols et sédiments existants. Quatre de plus ont été recensés lors de notre recherche (Singh *et al.*, 2002 ; Chen *et al.*, 2005 ; Cheng *et al.*, 2007 ; Duigou & Baize, 2010).

L'un d'eux est d'ailleurs construit de manière totalement différente de ce qui a pu être proposé auparavant (Duigou & Baize, 2010). L'Indicateur de Charge Multi-Métallique (ICMM), appliqué aux sols agricoles, répartit dans un premier temps les concentrations de chacun des 6 éléments (Cd, Cr, Cu, Ni, Pb et Zn) selon 6 classes de concentrations, prédéfinies à l'échelle nationale française (de 1 : valeurs faibles à 6 : valeurs fortes). L'ICMM est ensuite défini comme la somme des numéros de classes. C'est donc un nombre entier qui ne peut être inférieur à 6 ni excéder 36.

Comme pour le milieu atmosphérique, certains de ces indices intégrés ont été appliqués dans le cadre de la biosurveillance de la qualité des sédiments pour mesurer le niveau d'imprégnation global d'organismes sentinelles (algues et crustacés) par les métaux (Tomlinson *et al.*, 1980 ; Moschino *et al.*, 2012). En revanche, aucune application aux organismes des sols n'a été recensée à ce jour.

Cette synthèse montre que l'application des indices intégrés n'est pas homogène pour l'ensemble des milieux environnementaux. Les matrices aquatiques et solides font l'objet d'indices très diversifiés de par les paramètres intégrés et les méthodes de calcul. Au contraire, les indices intégrés développés pour la surveillance en routine de la qualité de l'air sont relativement homogènes, dépourvus de formule mathématique et basés sur les concentrations des polluants réglementés. C'est d'ailleurs à partir de l'indice sédiment de Tomlinson *et al.* (1980) que se basent les calculs d'imprégnation globale des organismes sentinelles de la qualité de l'air. Il ressort également de cette synthèse qu'aucune fonction d'agrégation ne semble parfaite ou même privilégiée, puisqu'elles présentent toutes des avantages et inconvénients. De fait, le choix de l'opération mathématique d'agrégation dépend souvent du contexte de l'étude.

L'indice multimétallique utilisé pour la biosurveillance lichénique dans le cadre de cette thèse (le Ratio d'Imprégnation Moyen : RIM), est un indice d'enrichissement du milieu évalué en rapport aux concentrations naturelles retrouvées dans les lichens. Il est inspiré de l'indice OIP (Overall Index of Pollution) développé par Sargaonkar & Deshpande (2003) pour la surveillance des rivières en Inde, dont la méthode de calcul a été reprise par Chen *et al.* (2005) avec l'IPI (Integrated Pollution Index) pour évaluer la qualité des sols de surface en Chine.

La plupart des indices développés dans le cadre du suivi d'un milieu en particulier sont en théorie applicables à l'ensemble des milieux environnementaux. D'autres ont même été construits pour intégrer une contamination multi-milieux. Ces indices composites multi-milieux, tenant compte la contamination de plusieurs compartiments environnementaux (air, eau, sols), sont peu communs (Golge *et al.*, 2013). Rojanschi (1991) proposait un indice de pollution global (Ipg) pour évaluer les impacts environnementaux d'implantations industrielles. Cet indice fut repris et agrémenté par plusieurs auteurs (Macoveanu, 2005; Zaharia & Surpateanu, 2006 ; Zaharia & Murarasu, 2009). Pour chacun des milieux, un score de 1 (mauvaise qualité) à 10 (qualité naturelle) est attribué à plusieurs paramètres mesurés à partir de classes de concentrations. Le score du milieu est ensuite calculé comme étant la moyenne des scores de chacun des paramètres. L'indice de pollution globale est ensuite calculé à partir d'une représentation graphique des scores de chaque milieu. Ainsi, l'Ipg représente le niveau de dégradation du milieu par rapport à un état naturel. Plus récemment, Golge *et al.* (2013) proposent une simple moyenne arithmétique des indices calculés pour chaque milieu.

Bien que les indices intégrés représentent une vulgarisation de la qualité environnementale, ils peuvent être considérés comme de puissants outils de traitement et d'analyse globale, à partir d'informations environnementales brutes (Caeiro *et al.*, 2005 ; Golge *et al.*, 2013). En effet, en plus des nombreuses mesures de concentrations, un indice intégré caractérise à lui seul le statut général de qualité environnementale. Ainsi, plusieurs auteurs citent l'intérêt particulier de ces indices pour les gestionnaires de territoires, puisqu'ils peuvent aider au suivi spatial et temporel de la qualité des milieux environnementaux (Pykh *et al.*, 2000).

Finalement, l'utilisation concomitante des indices intégrés et des outils SIG constitue une solution optimale et incontournable pour caractériser la qualité des milieux environnementaux. En effet, de par leurs fonctions synthétique et globale, de tels indices peuvent agréments l'analyse SIG en santé

environnementale. Mais il est aussi important de constater que les outils SIG sont conçus pour aider à la construction et au renforcement de tels indices. Dans le cadre de contaminations multi-paramètres, les géostatistiques peuvent être employées pour estimer les valeurs de concentrations en polluants au niveau de points d'échantillonnage pour lesquels tous les paramètres n'ont pas été mesurés et ainsi pallier la forte hétérogénéité des bases de données environnementales. Cette procédure est également valable dans le cas d'analyses multi-milieux, par exemple pour estimer les valeurs d'indices de qualité de l'air aux coordonnées géographiques de stations de mesure de la qualité des eaux.

Les outils et méthodes présentées dans cette introduction générale permettent donc de générer des indicateurs de caractérisation des milieux environnementaux et de santé des populations. La finalité d'une étude en santé – environnement étant la mise en relation de ces deux informations, la dernière partie de ce chapitre concerne les études écologiques géographiques, mises en œuvre pour évaluer les liens entre les indicateurs sanitaires, environnementaux et socioéconomiques.

Partie 6. Les études de corrélation écologique

A la suite de la mise en évidence de variations géographiques d'incidence de maladies, l'épidémiologie spatiale s'intéresse aux facteurs explicatifs de ces disparités, notamment à travers des études de corrélations écologiques. Ces études visent à évaluer les covariations spatiales et temporelles de facteurs de risques environnementaux et socio-économiques en lien avec les événements de santé. Le terme écologique renvoie ici à la notion de groupe, c'est-à-dire que la corrélation n'est pas évaluée à l'échelle de l'individu, mais à l'échelle de groupes de populations, agrégés au sein d'unités géographiques, ou au sein de périodes temporelles (Best *et al.*, 2001 ; Wakefield & Salway, 2001).

Ces études sont plus faciles et rapides (également moins onéreuses) à mettre en œuvre que d'autres catégories d'études épidémiologiques (cohortes, cas-témoins), puisqu'elles n'utilisent pas la mesure de facteurs individuels. Elles s'intéressent au contraire aux indicateurs environnementaux, sanitaires et sociaux attribués à un groupe d'individus. Ces indicateurs nécessitent d'être disponibles sur la même zone pour un groupe de population et recueillis de la même manière pour les différents groupes comparés (Goria *et al.*, 2011). Lors de la réalisation d'une étude écologique, la corrélation est par exemple évaluée par un risque relatif de développer l'évènement de santé, attribué aux groupes de population en fonction du niveau d'exposition au facteur de risque étudié. L'intervalle de confiance du RR permet dans ce cas de mesurer le niveau de significativité de la relation (Elliott *et al.*, 2000).

Parmi les études réalisées sur l'Homme, celles à l'échelle de groupes d'individus concluent à une importance causale la plus faible, du fait de l'existence de biais écologiques (Sackett *et al.*, 1985 ; Wakefield, 2008). Ces biais sont principalement liés au format agrégé des données populationnelles utilisées, ne permettant pas la distinction des variations d'exposition aux facteurs de risques au sein d'un même groupe. Ils seront détaillés lors de la discussion générale. De manière moins spécifique à ce type d'étude, la temporalité de l'effet sanitaire observé suite à l'exposition à un facteur de risque est souvent difficile à intégrer, notamment pour les pathologies chroniques comme les cancers (Richardson, 2000). Cela peut notamment s'expliquer par le nombre restreint de données environnementales disponibles sur le long terme.

Dans ce chapitre d'introduction générale, nous avons pu apercevoir la diversité et la puissance des outils SIG en termes d'analyse de données spatialisées. La cartographie des maladies évalue l'hétérogénéité spatiale des événements de santé à partir d'indicateurs standardisés d'incidence. La détection de clusters identifie les zones pour lesquelles on observe une concentration atypique d'événements en termes de surincidence ou de sous-incidence. Ces méthodes sont développées sur la base de registres de maladies. Dans le domaine de l'environnement, les méthodes géostatistiques d'interpolation, notamment le krigeage, sont employées pour estimer la valeur d'indicateurs environnementaux en tous points de l'espace à partir de campagnes d'échantillonnage. Ces méthodes peuvent notamment être appliquées aux données issues de la biosurveillance de la qualité des milieux environnementaux. Finalement les études écologiques géographiques étudient, à l'échelle de groupes d'individus, l'évolution conjointe ou non de données sanitaires, environnementales et socio-économiques.

Au moment de poser le contexte de ce travail de recherche, nous avons constaté que l'état de santé des populations est certainement impacté par la qualité du milieu environnemental dans lequel elles demeurent, mais que ce lien environnement – santé est difficile à caractériser. Si les outils d'analyse existent, l'un des obstacles à l'identification de ce lien à l'échelle des populations est sans doute la forte hétérogénéité des bases de données populationnelles et environnementales. En santé publique, les registres de maladies sont des ressources d'information robustes et exhaustives, généralement bâties sur de longues périodes et de grandes aires géographiques. Ces registres sont de plus élaborés sous la tutelle d'agences nationales de santé, gage de fiabilité (Agence de biomédecine, Institut National de Veille Sanitaire (InVS)). Ce constat est aussi valable pour les indicateurs sociodémographiques, produits par l'INSEE de manière annuelle à l'échelle nationale et à plusieurs niveaux d'imbrication administratifs (du quartier à la région). En environnement, il est au contraire plus complexe d'obtenir de l'information brute concernant la contamination des milieux, à la fois précise et disponible à grande échelle (spatiale et temporelle). Bien que de telles données existent, très peu sont aujourd'hui recensées au sein de bases de données publiques.

Par ailleurs, ces bases de données sont de nature très différente. Les informations environnementales sont de nature géostatistique, alors que les informations populationnelles sont de nature agrégée. Les bases de données environnementales ne sont pas conçues au départ pour être mise en relation écologique avec des indicateurs sanitaires ou socio-économiques. Une étape de transformation est donc nécessaire.

Ainsi, les bases de données des thématiques environnementales, sanitaires et sociodémographiques sont très éloignées en termes de développement, de précision, d'accessibilité et de format. L'un des objectifs de la thèse fut donc de produire de la donnée environnementale spatialisée, en vue d'évaluer les inégalités environnementales et sociales de santé. Cela passe notamment par l'analyse spatiale de bases de données issues de la biosurveillance lichéniques de la qualité de l'air (Chapitre I) et de la contamination des eaux souterraines (Chapitre II).

REFERENCES BIBLIOGRAPHIQUES

- AFNOR. 2003. NF ISO 10381-1. Qualité du sol - Échantillonnage - Partie 1 : lignes directrices pour l'établissement des programmes d'échantillonnage.
- AFNOR. 2008. NF X43-903 : Biosurveillance de l'air - Détermination d'un indice biologique de lichens épiphytes (IBLE). 19p.
- AFNOR. 2013. NF X 43-904 : Biosurveillance passive de la qualité de l'air à l'aide des lichens autochtones : de la récolte à la préparation des échantillons.
- Agnan Y. 2013. Bioaccumulation et bioindication par les lichens de la pollution atmosphérique actuelle et passée en métaux et en azote en France : sources, mécanismes et facteurs d'influence. Thèse de Doctorat Université de Toulouse. 306p.
- Andersen A, Hovmand M, Johnsen IB. 1978. Atmospheric heavy metal deposition in the Copenhagen area. *Environ. Pollut*, 17 : 133-151.
- Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. 2012. A review of spatial methods in epidemiology, 2000-2010. *Annu Rev Public Health*, 33: 107-122.
- Aznar JC, Richer-Lafleche M, Cluis D. 2008. Metal contamination in the lichen *Alectoria sarmentosa* near the copper smelter of Murdochville, Quebec. *Environ. Pollut*, 156-76.
- Baillargeon S. 2005. Le krigeage : revue de la théorie et application à l'interpolation spatiale de données de précipitations. Mémoire Faculté des études supérieures de l'Université Laval. 128p.
- Balabanova B, Stafilov T, Sajn R, Baeeva K. 2012. Characterisation of Heavy Metals in Lichen Species *Hypogymnia physodes* and *Evernia prunastri* due to Biomonitoring of Air Pollution in the Vicinity of Copper Mine. *Int. J. Environ. Res.*, 6(3): 779-794.
- Bargagli, R. 1998. Trace Elements in Terrestrial Plants, an Ecophysiological Approach to Biomonitoring and Biorecovery. Springer, Berlin, 324p.
- Bari A, Minciardi M, Troiani F, Bonotto F, Paonessa F. 1998. Lichens and mosses in air quality monitoring: a biological model proposal. *Govt. Reports Announcements & Index*, Issue 16.
- Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD. 1972. Statistical inference under restrictions: The theory and application of isotonic regression. Wiley: England.
- Beale L, Abellan JJ, Hodgson S, Jarup L. 2008. Methodologic Issues and Approaches to Spatial Epidemiology. *Environmental Health Perspectives*, Vol 116 (8): 1105-1110.
- Beale L, Hodgson S, Abellan JJ, LeFevre S, Jarup L. 2010. Evaluation of spatial relationships between health and the environment: the Rapid Inquiry Facility. *Environ Health Perspect*, 118(9): 1306–1312.
- Bergamaschi L, Rizzio E, Giaveri G, Loppi S, Gallorini M. 2007. Comparison between the accumulation capacity of four lichen species transplanted to a urban site. *Environmental Pollution*, 148: 468-476.
- Bergeron M. 1992. Vocabulaire de la géomatique. Office de la langue française, Québec. 41 p.
- Béron P, Valiquette L, Patty G, Briere F. 1982. Water quality indices. *Trib. Cebedeau*, 35: 385-391.
- Besag J & Newell J. 1991. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society, Series A*, 154: 143–155.
- Besag J, York J, Mollié A. 1991. Bayesian image restoration, with two applications in spatial statistics (with Discussion). *Annals of the Institute of Statistical Mathematics*, 43(1): 1-59.

- Best NG, Cockings S, Bennett JE, Wakefield JC, Elliott P. 2001. Ecological regression analysis of environmental benzene exposure and childhood leukaemia: sensitivity to data inaccuracies, geographical scale and ecological bias. *Journal of the Royal Statistical Society, Series A* (164): 155-74.
- Best N, Richardson S, Thomson A. 2005. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res*, 14(1): 35-59.
- Bezuglaya EY, Shchutskaya AB, Smirnova IV. 1993. Air pollution index and interpretation of measurements of toxic pollutant concentrations. *Atmospheric Environment*, Vol. 27A, No. 5: 773-779.
- Bhargava DS. 1983. Use of a water quality index for river classification and zoning of Ganga river. *J. Environ. pollut., Series B*, 6: 51-67.
- Bharti N & Katyal D. 2011. Water quality indices used for surface water vulnerability assessment. *International Journal of Environmental Sciences*, Vol. 2 (1): 154-173.
- Bithell JF. 2000. A classification of disease mapping methods. *Statist. Med*, 19: 2203-2215.
- Boamponsem LK, Adam JI, Dampare SB, Nyarko BJB, Essumang DK. 2010. Assessment of atmospheric heavy metal deposition in the Tarkwa gold mining area of Ghana using epiphytic lichens. *Nuclear Instruments and Methods in Physics Research, B* 268: 1492-1501.
- Bosser P. 2012. Interpolation spatiale. *Ouvrage Ecole Nationale des Sciences Géographiques*. 53p.
- Bouyer I, Hémon D, Cordier S, Derriennic F, Stüker I, Stengel B, Clavel J. 1993. *Epidémiologie : principes et méthodes quantitatives*. Les Éditions INSERM. 498p.
- Brown RM, McClelland NI, Deininger RA, Tozer RG. 1970. A water quality index—Do we dare? *Water Sewage Works*, 117(10): 339-343.
- Burrough PA & McDonnell RA. 1998. *Principles of geographical information systems*. Oxford University Press, Oxford.
- Caeiro S, Costa MH, Ramos TB, Fernandes F, Silveira N, Coimbra A, Medeiros G, Painho M. 2005. Assessing heavy metal contamination in Sado Estuary sediment: An index analysis approach. *Ecological Indicators*, 5: 151-169.
- Cairncross EK, John J, Zunckel M. 2007. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. *Atmos. Environ*, 41: 8442-8454.
- Carreras HA, Wannaz ED, Pignata ML. 2009. Assessment of human health risk related to metals by the use of biomonitors in the province of Cordoba, Argentina. *Environmental Pollution*, 157 (1): 17-122.
- Catalyud A & Barreno E. 2004. Response to ozone in two lettuce varieties on chlorophyll a fluorescence, photosynthetic pigments and lipids peroxidation. *Plant Physiology and Biochemistry*, 42: 549-555.
- Chapman P. 1990. The sediment quality Triad approach to determining pollution-induced degradation. *Sci. Total Environ*, 97/98: 815-825.
- Chen TB, Zheng YM, Lei M, Huang ZC, Wu HT, Chen H, Fan KK, Yu K, Wu X, Tian QZ. 2005. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere*, 60: 542-551.
- Cheng JL, Shi Z, Zhu YW. 2007. Assessment and mapping of environmental quality in agricultural soils of Zhejiang Province, China. *Journal of Environmental Sciences*, 19: 50-54.
- Cislaghi C & Nimis PL. 1997. Air pollution and lung cancer. *Nature*, 387: 463-464.

- Clarke KC. 2001. Getting Started with Geographic Information Systems (3rd ed.). Upper Saddle River, NJ: Prentice Hall. 352p.
- Clayton D & Kaldor J. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43: 671-681.
- Cloquet C, De Muyneck D, Signoret J, Vanhaecke F. 2009. Urban/peri-urban aerosol survey by determination of the concentration and isotopic composition of Pb collected by transplanted lichen *Hypogymnia physodes*. *Environmental Science and Technology*, 43 (3): 623-9.
- Conti ME, Cecchetti G. 2001. Biological monitoring: lichens as bioindicators of air pollution assessment: a review. *Environ Pollut*, 114: 471-92.
- Cressie NA. 1993. *Statistics for Spatial Data* (revised edition). John Wiley & Sons, Inc., New York.
- Crous KI, Vandermeiren K, Ceulemans R. 2006. Physiological responses to cumulative ozone uptake in two white clover (*Trifolium repens* L. cv. Regal) clones with different ozone sensitivity. *Environmental and Experimental Botany*, 58(1-3): 169-179.
- Cuny D, Van Haluwyn C, Pesch R. 2001. Biomonitoring of trace elements in air and soil compartments along the major motorway in France. *Water, Air, and Soil Pollution*, 125: 273-289.
- Cuny D, Pignata ML, Kranner I, Beckett R. 2002. Biomarkers of pollution-induced oxidative stress and membrane damage in lichens. In : Nimis PL, Scheidegger C, Wolseley PA. *Monitoring with lichens – Monitoring lichens*. Kluwer, The Netherlands.
- Cuny D, Davranche L, Thomas P, Kempa M, Van Haluwyn C. 2004. Spatial and temporal variations of trace element contents in *Xanthoria parietina* thalli collected in a highly industrialized area in Northern France as an element for a future epidemiological study. *Journal of Atmospheric Chemistry*, 49: 391-401.
- Cuny D, Rzepka MA, Van Haluwyn C. 2008. Introduction à la biosurveillance végétale et fongique de la qualité de l'air, concept et intérêt. *Spectra Analyse*, 264: 31-35.
- Cuny D. 2012. La biosurveillance végétale et fongique de la pollution atmosphérique : concepts et applications. *Annales Pharmaceutiques Françaises*, 70: 182—187.
- De Blomac F, Gal R, Hubert M, Richard D, Turret C. 1994. *Arc/Info, concepts et applications en géomatique*. Paris, Hermès, 256 p.
- De Temmerman L, Bell JNB, Garrec JP, Klumpp A, Krause GHM, Tonneijck AEG. 2004. Biomonitoring of air pollutants with plants. In Klumpp A, Ansel W, Klumpp G, editors. *Urban Air Pollution, bioindication and environmental awareness*. Göttingen: Cuvillier Verlag: 337—74.
- Declercq C. 2007. Géographie de la santé : le point de vue de l'épidémiologiste. Séminaire ORS Île de France « De la mesure des expositions à l'évaluation des conséquences pour la santé : le traitement spatialisé des données ». Paris, 7 septembre 2007. Communication orale.
- Delzenne-Van Haluwyn C. 1973. Contribution à l'étude de la distribution des lichens épiphytes dans le Nord de la France : application au problème de la pollution atmosphérique. Thèse de doctorat, Faculté de Pharmacie, Lille: 162 p.
- Deng F, Minasny B, Knadel M, McBratney A, Heckrath G, Greve MH. 2013. Using vis-NIR spectroscopy for monitoring temporal changes in soil organic carbon. *Soil Sci*, 178 (8): 389-99.

- Duigou N & Baize D. 2010. Nouvelle collecte nationale d'analyses d'éléments en traces dans les sols (horizons de surface) - (Cd, Cr, Cu, Hg, Ni, Pb, Se, Zn). Contrat ADEME. 284 p.
- Echaubard M. 1995. Les animaux comme bioindicateurs biologiques de pollutions. *Marqueurs Biologiques de Pollution*. Chinon, 335-338.
- Elliott P & Wartenberg D. 2004. *Spatial Epidemiology: Current Approaches and Future Challenges*. *Environmental Health Perspectives*, Vol 112: 998-1006.
- Elliott P, Wakefield JC, Best NG, Briggs DJ. 2000. *Spatial epidemiology: methods and applications*. In: Elliott P, Wakefield JC, Best NG, Briggs DJ, (dir.). *Spatial epidemiology: methods and applications*. Oxford: Oxford University Press, 3-14.
- Fairey R, Long ER, Roberts CA, Anderson BS, Phillips BM, Hunt JW, Puckett HR, Wilson CJ. 2001. An evaluation of methods for calculating mean sediment quality guideline quotients as indicators of contamination and acute toxicity to amphipods by chemical mixture. *Environ. Toxicol. Chem.*, 20: 2276–2286.
- Fernandez-Falcon E, Strittholt JR, Alobaida AI, Schmidley RW, Bossler JD, Ramirez JR. 1993. A Review of Digital Geographic Information Standards for the State/Local User. *URISA Journal*, Vol 5 (2): 21-27.
- Ferry BW, Baddeley MS, Hawksworth DL. 1973. *Air Pollution and Lichens*. Athlone Press, London: 389 pp.
- Finzgar N, Jez E, Volgar D, Lestan D. 2014. Spatial distribution of metal contamination before and after remediation in the Meza Valley, Slovenia. *Geoderma*, 217–218: 135–143.
- Freitas MC, Reis MA, Alves LC, Wolterbeek HT. 1999. Distribution in Portugal of some pollutants in the lichen *Parmelia sulcata*. *Environmental Pollution*, 106 (2): 229-235.
- Fritz CE, Schuurman N, Robertson C, Lear S. 2013. A scoping review of spatial cluster analysis techniques for point-event data. *Geospatial Health*, 7(2): 183-198.
- Fuga A, Saiki M, Marcelli MP, Saldiva PHN. 2008. Atmospheric pollutants monitoring by analysis of epiphytic lichens. *Environmental Pollution*, 151: 334-340.
- Gadal S. 2008. *Les systèmes d'informations géographiques libres*. Agence universitaire de la francophonie. Ouvrage numérique.
- Gadsdon S, Dagley JR, Wolseley PA, Power SA. 2010. Relationships between lichen community composition and concentrations of NO₂ and NH₃. *Environmental Pollution*, 158: 2553-2560.
- Gailey FAY & Lloyd OLI. 1993. Spatial and temporal patterns of airborne metal pollution: The value of low technology sampling to an environmental epidemiology study. *Sci. Total Environ.* 113: 201–219.
- Garrec J-P, Van Haluwyn C. 2002. *Biosurveillance végétale de la qualité de l'air*. Tec & Doc, Paris.
- Garty J. 1993. Lichens as biomonitors for heavy metal pollution. in B. Market (ed.), *Plants as Biomonitors, Indicators for Heavy Metals in the Terrestrial Environment*, VCH, Weinheim: 193–263.
- Garty J. 2001. Biomonitoring atmospheric heavy metals with lichens: Theory and application. *Crit. Rev. Plant Sci*, 20(4): 309–371.
- Gaveriaux J-P. 1995. *Lichens et la bioindication de la qualité de l'air (les)*. Avion : Lycée Pablo Picasso. 52 p.
- Gerdol R, Marchesini R, Iacuminb P, Brancaleoni L. 2014. Monitoring temporal trends of air pollution in an urban area using mosses and lichens as biomonitors. *Chemosphere*, 108: 388-395.

- Giljanovic NS. 1999. Water quality evaluation by index in Dalmatia. *Water Res.*, 33(16): 3423–3440.
- Glaz J, Pozdnyakov V, Wallenstein S. 2009. *Scan Statistics Methods and Applications*. Statistics for Industry and Technology, Birkhauser, 394p.
- Golge M, Yenilmez F, Aksoy A. 2013. Development of pollution indices for the middle section of the Lower Seyhan Basin (Turkey). *Ecological Indicators*, 29: 6–17.
- Gombert S, Asta J, Seaward MRD. 2004. Assessment of lichen diversity by index of atmospheric purity (IAP), index of human impact (IHI) and other environmental factors in an urban area (Grenoble, southeast France). *Sci Total Environ*, 324: 183-99.
- Gomez-Rubio G, Ferrandiz-Ferragud J, Lopez-Quilez A. 2005. Detecting clusters of disease with R. *J Geograph Syst*, 7: 189–206.
- Goovaerts P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Goria S, Stempfelet M, de Crouy-Chanel P. 2011. *Introduction aux méthodes statistiques et aux systèmes d'information géographique en santé environnement– Application aux études écologiques*. Saint-Maurice: Institut de veille sanitaire : 65 p.
- Gratton Y. 2002. Le krigeage : la méthode optimale d'interpolation spatiale. *Les articles de l'Institut d'Analyse Géographique*, 4p.
- Guéguen F, Stille P, Geagea ML, Boutin R. 2012. Atmospheric pollution in an urban environment by tree bark biomonitoring – Part I: Trace element analysis. *Chemosphere*, 86: 1013–1019.
- Hakanson L. 1980. An ecological risk index for aquatic pollution control. A sedimentological approach. *Water Res*, 14: 975–1001.
- Hawksworth DL & Rose F. 1970. Qualitative scale for estimating sulphur dioxide air pollution in England and Wales using epiphytic lichens. *Nature*, 227: 145-8.
- Hengl T. 2007. *A Practical Guide to Geostatistical Mapping of Environmental Variables*. Office for Official Publication of the European Communities, Luxembourg, 143p.
- Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. 1996. Childhood leukemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. *Statistics in Medicine*, 15: 707-715.
- Horton RK. 1965. An index number system for rating water quality. *J. Water Pollut. Control Fed.*, 37(3): 300–306.
- Ikem A, Egiebor NO, Nyavor K. 2003. Trace elements in water, fish and sediment from Tuskegee Lake, Southeastern USA. *Water Air Soil Pollut*, 149: 51–75.
- Joseph WD. 2002. Comparative assessment and harmonization of the US EPA air quality index (AQI) with related air quality and pollutant standard indices in other nations—phase 2. Work assignment 5-15 of Contract no. 68-D-98-030.
- Krige DG. 1951. A statistical approach to some mine valuations problems at the Witwatersrand. *J. Chem. Metall. Min. Soc. S. Afr*, 52: 119-139.
- Krivoruchko K. 2012. Empirical Bayesian Kriging implemented in ArcGIS Geostatistical Analyst. *ArcUser*, 15 (4): 6-10.
- Kulldorff M & Nagarwalla N. 1995. Spatial disease clusters: Detection and inference. *Statistics in Medicine*, 14: 799–810.
- Kulldorff M. 1997. A spatial scan statistic. *Communication in Statistics - Theory and Methods*, 26: 1481–1496.

- Kulldorff M, Feuer EJ, Miller BA, Freedman LS. 1997. Breast cancer in northeastern United States: A geographical analysis. *American Journal of Epidemiology*, 146: 161-170.
- Kulldorff M, Athas W, Feuer E, Miller B, Key C. 1998. Evaluating cluster alarms: A space-time scan statistic and brain cancer in Los Alamos. *American Journal of Public Health*, 88: 1377-1380.
- Kulldorff M. 1999. An isotonic spatial scan statistic for geographical disease surveillance. *Journal of the National Institute of Public Health*, 48: 94–101.
- Kulldorff M. 2010. SaTScan User Guide for version 9.0. 109p.
- Kumar D & Alappat BJ. 2004. Selection of the appropriate aggregation function for calculating leachate pollution index. *Pract. Period. Hazard. Toxic Radioact. Waste Manage.*, 8(4): 253–264.
- Lagadic L, Caquet T, Amiard JC. 1997. Biomarqueurs en écotoxicologie : principes et définitions. Dans Lagadic L, Caquet, Amiard JC, Ramade F. *Biomarqueurs en écotoxicologie – Aspects fondamentaux*. Masson, Paris: 1-9.
- Landwehr JM & Deininger RA. 1976. A comparison of several water quality indexes. *J. Water Pollut. Control Fed.*, 48(5): 954–958.
- Larsen RS, Bell JNB, James PW, Chimonides PJ, Rumsey FJ, Tremper A, Purvis OW. 2007. Lichen and bryophyte distribution on oak in London in relation to air pollution and bark acidity. *Environmental Pollution*, 146: 332-340.
- Lemarchand O, Jeannée N. 2009. Méthodes de cartographie et approche géostatistique – La cartographie de la pollution au dioxyde d’azote en Alsace. *Cahier des thèmes transversaux ArScAn*, 9: 203-214.
- Lerond M. 1981. Les lichens épiphytes en Normandie orientale, distribution, sociologie et application à la cartographie de la pollution atmosphérique, *Actes Muséum de Rouen*: 1-295.
- Li J & Heap AD. 2014. Spatial interpolation methods applied in the environmental sciences: a review. *Environmental Modelling & Software*, 53: 173-189.
- Llop E, Pinho P, Matos P, Pereira MJ, Branquinho C. 2012. The use of lichen functional groups as indicators of air quality in a Mediterranean urban environment. *Ecol Indicator*, 13: 215-21.
- Long ER, MacDonald DD. 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Hum. Ecol. Risk Assess.*, 4: 1019–1039.
- Loppi S & Nascimbene J. 2010. Monitoring H₂S air pollution caused by the industrial exploitation of geothermal energy: The pitfall of using lichens as bioindicators. *Environmental Pollution*, 158: 2635-2639.
- Loppi S, Frati L, Paoli L, Bibabli V, Rossetti C, Bruscoli C, Corsini A. 2004. Biodiversity of epiphytic lichens and heavy metal contents of *Flavoparmelia caperata* thalli as indicators of temporal variations of air pollution in the town of Montecatini Terme (central Italy). *Science of the Total Environment*, 326: 113–122.
- Macoveanu M. 2005. *Methods and techniques of ecological impact assessment (in Romanian)*. Iasi, Romania: Ecozone Publishing House, 2nd Edition.
- Madejon P, Maranon T, Murillo JM. 2006. Biomonitoring of trace elements in the leaves and fruits of wild olive and holm oak trees. *Science of the Total Environment*, 355: 187– 203.

- Manier N, Deram A, Lecurieux F, Marzin D. 2009. Comparison between new wild plant *Trifolium repens* and *Vicia faba* on their sensitivity in detecting the genotoxic potential of heavy metal solutions and heavy metal-contaminated soils. *Water, Air, and Soil Pollution*, 202 (1-4): 343-352.
- Matheron G. 1962. *Traité de géostatistique appliquée, Tome I. Mémoires du Bureau de Recherches Géologiques et Minières, No.14.* Editions Technip, Paris.
- Matheron G. 1963a. Principles of geostatistics. *Economic Geology*, 58: 1246–1266.
- Matheron G. 1963b. *Traité de géostatistique appliquée, II : Le Krigeage. Mémoires du Bureau de Recherches Géologiques et Minières, No.24.* Editions B. R. G. M., Paris.
- Maynard RL, Coster SM. 1999. Informing the public about air pollution. In: Holgate ST *et al.* (Eds.) *Air Pollution and Community Health.* Academic Press, Sydney: 1019–1033. (ISBN: 0-12-352335-4).
- Merwade V. 2009. Effect of spatial trends on interpolation of river bathymetry. *Journal of Hydrology*, 371: 169–181.
- Misík M, Ma TH, Nersesyan A, Monarca S, Kim JK, Knasmueller S. 2011. Micronucleus assays with *Tradescantia pollen* tetrads: an update. *Mutagenesis*, 26(1): 215-21.
- Moran P. 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2): 17–23.
- Moschino V, Delaney E, Da Ros L. 2012. Assessing the significance of *Ruditapes philippinarum* as a sentinel for sediment pollution: Bioaccumulation and biomarker responses. *Environmental Pollution*, 171: 52-60.
- Murena F. 2004. Measuring air quality over large urban areas: development and application of an air pollution index at the urban area of Naples. *Atmospheric Environment*, 38: 6195–6202.
- Niober EA, Richardson DHS, Tomassini FD. 1978. Mineral Uptake and Release by Lichens: An Overview. *Bryologist*, 81: 226-246.
- Nuckols JR, Ward MH, Jarup L. 2004. Using geographic information systems for exposure assessment in environmental epidemiology studies. *Environ Health Perspect*, 112 (9): 1007-15.
- Nylander W. 1866. Les lichens du Jardin du Luxembourg. *Bulletin de la Société botanique de France*, 13: 364-372.
- Openshaw S, Charlton M, Wymer C, Craft AW. 1987. A mark I geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information Systems*, 1: 335–358.
- Ott W. 1978. *Environmental Indices—Theory and Practice.* Ann Arbor Science, Michigan, USA, 371p.
- Otvös E, Pazmandi T, Tuba Z. 2003. First national survey of atmospheric heavy metal deposition in Hungary by the analysis of mosses. *The Science of the Total Environment*, 309: 151-160.
- Pignata ML, Gudino GL, Wannaz ED, Pla RR, Gonzalez CM, Carreras HA, Orellana L. 2002. Atmospheric quality and distribution of heavy metals in Argentina employing *Tillandsia capillaris* as a biomonitor. *Environmental Pollution*, 120: 59-68.
- Pornon H. 2007. Bilan et perspectives de 20 années de Géomatique. *Géomatique Expert*, 57: 36-46.
- Pykh YA, Kennedy ET, Grant WE. 2000. An overview of systems analysis methods in delineating environmental quality indices. *Ecol. Model.*, 130: 25–38.

- Qiao Q, Huang B, Zhang, C, Piper JDA, Pan Y, Sun Y. 2013. Assessment of heavy metal contamination of dustfall in northern China from integrated chemical and magnetic investigation. *Atmospheric Environment*, 74: 182-193.
- Qingjie G, Jun D, Yunchuan X, Qingfei W, Liqiang Y. 2008. Calculating Pollution Indices by Heavy Metals in Ecological Geochemistry Assessment and a Case Study in Parks of Beijing. *Journal of China University of Geosciences*, 19 (3): 230–241.
- Rai R & Agrawal M. 2008. Evaluation of physiological and biochemical responses of two rice (*Oryza sativa* L.) cultivars to ambient air pollution using open top chambers at a rural site in India. *Science of the Total Environment*, 407 (1): 679-691.
- Rajagopal B, Shrihari S, Dwarakish GS. 2010. Global Water Quality Indices for River Gurpur, Karnataka State, India. *International Journal of Earth Sciences and Engineering*, 3 (6): 833-840.
- Remy C. 2011. Utilisation de la modification de la biodiversité lichénique sur résineux pour détecter la pollution azotée d'origine routière. *Pollut Atmos, Numéro spécial « Biosurveillance de la pollution atmosphérique »*: 31-4.
- Rican S. 2007. Représentation cartographique des données sanitaires. Séminaire ORS Île de France « De la mesure des expositions à l'évaluation des conséquences pour la santé : le traitement spatialisé des données ». Paris, 7 septembre 2007. Communication orale.
- Richardson S. 1992. Modélisation statistique des variations géographiques en épidémiologie. *Rev Epidemiol Sante Publique*, 40: 33-45.
- Richardson S. 2000. Problèmes méthodologiques dans les études écologiques santé–environnement. *Life Sciences*, 323: 611–616.
- Ripley BD. 1981. *Spatial Statistics*. New York: Wiley.
- Roberts JD, Voss JD, Knight B. 2014. The association of ambient air pollution and physical inactivity in the United States. *Plos One*, 9 (3): e90143.
- Robertson C, Nelson TA, MacNab YC, Lawson AB. 2010. Review of methods for space–time disease surveillance. *Spatial and Spatio-temporal Epidemiology*, 1: 105–116.
- Rojanschi V. 1991. Global estimation methodology for ecosystem state. *The Environment (in Romanian: Mediul înconjurator)*, 2: 45-52.
- Rzepka MA & Cuny D. 2008. Biosurveillance végétale et fongique des éléments traces métalliques atmosphériques. *Air Pur*, 75: 66–77.
- Rzepka MA & Cuny D. 2011. Recherches préliminaires sur la mise en place d'une biosurveillance végétale de la pollution atmosphérique en zone industrielle complexe : applications du ray-grass dans la zone industrialoportuaire de Dunkerque. *Acta Botanica Gallica*, 158(3): 387-400.
- Sackett DL, Haynes RB, Guyatt GH, Tugwell P. 1985. *Clinical Epidemiology, a Basic Science for Clinical Medicine*. Little, Brown and Co, Boston, 1985.
- Salo H, Bucko MS, Vaahtovuori E, Limo J, Mäkinen J, Pesonen LJ. 2012. Biomonitoring of air pollution in SW Finland by magnetic and chemical measurements of moss bags and lichens. *Journal of Geochemical Exploration*, 115: 69–81.
- Sargaonkar, A & Deshpande V. 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environmental Monitoring and Assessment*, 89: 43-67.

- Sarmiento S, Wolterbeek HT, Verburg TG, Freitas MC. 2008. Correlating element atmospheric deposition and cancer mortality in Portugal: data handling and preliminary results. *Environmental Pollution*, 151: 314-351.
- Scerbo R, Possenti L, Lampugnani L, Ristori T, Barale R, Barghigiani C. 1999. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Livorno Province (Tuscany, Italy). *The Science of the Total Environment*, 241: 91-106.
- Scerbo R, Ristori T, Possenti L, Lampugnani L, Barale R, Barghigiani C. 2002. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Pisa Province (Tuscany, Italy). *The Science of the Total Environment*, 286: 27-40.
- Shin PKS & Lam WKC. 2001. Development of a marine sediment pollution index. *Environ. Pollut.*, 113: 281–291.
- Silva DT, Meirelles ST, Moraes RM. 2012. Relationship between ozone, meteorological conditions, gas exchange and leaf injury in *Nicotiana tabacum* Bel-W3 in a sub-tropical region. *Atmospheric Environment*, 60: 211-216.
- Singh M, Müller G, Singh IB. 2002. Heavy metals in freshly deposited stream sediments of rivers associated with urbanization of the Ganga plain, India. *Water, Air, and Soil Pollution*, 141: 35–54.
- Singh RP, Nath S, Prasad SC, Nema AK. 2008. Selection of Suitable Aggregation Function for Estimation of Aggregate Pollution Index for River Ganges in India. *Journal of Environmental Engineering*, 134 (8): 689-701.
- Sloof JE & Wolterbeek HT. 1991. National trace-element air pollution monitoring survey using epiphytic lichens. *Lichenologist*, 23(2): 139-165.
- Smith DG. 1990. A better water quality indexing system for rivers and streams. *Water Res.*, 24(10): 1237–1244.
- Snow J. 1855. *On the Mode of Communication of Cholera*. 2nd ed. London: Churchill.
- Swamee PK & Tyagi A. 2000. Describing water quality with aggregate index. *J. Environ. Eng.*, 126(5): 451–455.
- Takahashi T & Miyajima Y. 2010. Effects of roads on alpine and subalpine plant species distribution along an altitudinal gradient on Mount Norikura, central, Japan. *Journal of Plant Research*, 123: 741-749.
- Tango T. 1995. A class of tests for detecting 'general' and 'focused' clustering of rare diseases. *Stat Med*, 14(21-22): 2323-34.
- Tango T. 2000. A test for spatial disease clustering adjusted for multiple testing. *Stat Med*, 19(2): 191-204.
- Thimonier A, Dupouey JL, Bost F, Becker M. 1994. Simultaneous eutrophication and acidification of a forest ecosystem in North-East France. *New Phytologist*, 126: 533-9.
- Tobler W. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46: 234–40.
- Tomlinson DL, Wilson JG, Harris CR, Jeffrey DW. 1980. Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgoländer Meeresunters*, 33: 566-575.
- US EPA. 1998. Air quality index reporting; proposed rules. *Federal Register*, 63 (236).

- US EPA. 1999. Air quality index reporting; final rule. Federal Register, 64 (149).
- Usero J, Gonzalez-Regalado E, Gracia I. 1996. Trace metals in the bivalve mollusc *Chamelea gallina* from the Atlantic Coast of Southern Spain. Mar. Pollut. Bull., 32 (3): 305–310.
- Van Haluwyn C. 1993. Guide des lichens. Paris: Masson.
- Van Haluwyn C. 1998. La bioindication : applications dans notre région. Air Pur, 55: 23-25.
- Van Haluwyn C, Cuny D, Garrec J-P. 2011. Introduction générale - Définitions, concepts et enjeux généraux de la biosurveillance de la qualité de l'air. Pollution Atmosphérique, Numéro spécial : 7-13.
- Vine MF, Degnan D, Hanchette C. 1997. Geographic Information Systems : Their use in Environmental Epidemiologic Research. Environmental Health Perspectives, 105 (6): 598-605.
- Wakefield JC & Salway R. 2001. A statistical framework for ecological and aggregate studies. Journal of the Royal Statistical Society, series A (164): 119-37.
- Wakefield JC. 2008. Ecologic Studies Revisited. Annu. Rev. Public. Health., 29: 75-90.
- Waller LA & Gotway CA. 2004. Applied spatial statistics for public health data. Wiley, Hoboken. 494p.
- Walter SD & Birnie SE. 1991. Mapping mortality and morbidity patterns: an international comparison. Int J Epidemiol, 20: 678–689.
- Wappelhorst O, Kühn I, Oehlmann J, Markert B. 2000. Deposition and disease: a moss monitoring project as an approach to ascertaining potential connections. The Science of the Total Environment, 249: 243-256.
- Wolterbeek HT & Verburg TG. 2004. Atmospheric metal deposition in a moss data correlation study with mortality and disease in the Netherlands. The Science of the Total Environment, 319: 53-64.
- Zaharia C & Murarasu I. 2009. Environmental impact assessment induced by an industrial unit of basic chemical organic compounds synthesis using the alternative method of global pollution index. Environ. Eng. Manag. J., 8: 107–112.
- Zaharia C & Surpateanu M. 2006. Environmental impact assessment using the method of global pollution index applied for heat and power co-generation plant. Environ. Eng. Manag. J., 5: 1141–1152.
- Zhao S, Feng C, Yang Y, Niu J, Shen Z. 2012. Risk assessment of sedimentary metals in the Yangtze Estuary: New evidence of the relationships between two typical index methods. Journal of Hazardous Materials, 241– 242: 164– 172.

Sites internet :

- ESRI : www.esrifrance.fr (consulté le 08/04/2014).
- Eau France : www.eaufrance.fr (consulté le 11/03/2014).
- Fédération ATMO France : www.atmo-france.org (consulté le 11/03/2014).
- SaTScan : www.satscan.org (consulté le 03/04/2014).
- UK Air Quality Archive : www.airquality.co.uk (consulté le 11/03/2014).

MATERIEL ET METHODES

Dans ce chapitre consacré aux matériels et méthodes de la thèse, nous présenterons dans un premier temps la région Nord-Pas de Calais sur laquelle s'est déclinée notre thèse (1^{ère} partie) et les ressources utilisées (2^e partie). Les trois parties suivantes consisteront à définir les matériels et méthodes employés pour chacun des axes de recherche de la thèse.

Partie 1. La région Nord – Pas de Calais

Pour faciliter la mise en œuvre d'études écologiques en santé environnementale, plusieurs critères peuvent être identifiés :

- une densité de population conséquente, afin de maximiser la puissance statistique des indicateurs de santé ;
- un contraste en termes d'occupation du sol, avec des milieux urbains, industriels et ruraux, ce qui favorise la diversité et l'identification des sources de contamination des milieux ;
- des indicateurs environnementaux, sociodémographiques et sanitaires variables, pouvant générer la présence de points noirs

La région Nord – Pas de Calais présente l'ensemble de ces critères. Elle semble donc un territoire approprié pour des études intégrant l'état de dégradation des milieux, ainsi que l'état de santé et le statut socio-économique des populations. Les paragraphes suivants développent les différents critères de choix recensés pour cette région.

Quatrième région de France pour sa population avec 4,052 millions d'habitants au 1^{er} janvier 2013 répartis au sein de 1546 communes (soit 6,4 % de la population métropolitaine), le Nord - Pas-de-Calais est surtout positionné en deuxième position pour sa densité de population avec en moyenne 325 habitants par km² (soit trois fois la moyenne nationale) pour seulement 2,3 % de l'espace métropolitain (INSEE, 2014). Les départements du Nord et du Pas de Calais représentent respectivement 64 % et 36 % de la population régionale, avec une densité de population très élevée pour le Nord (449 hab/km²) et deux fois plus faible pour le Pas de Calais (219 hab/km²) (figure 17).

Cette forte densité implique un territoire très urbanisé, avec des aires urbaines recouvrant 58% du territoire (contre 35% en France provinciale). L'aire urbaine de Lille, qui est l'une des zones d'intérêt du premier axe de recherche (chapitre I), représente plus d'un quart de la population régionale avec plus d'un million d'habitants. En dehors de l'Île de France, le Nord – Pas de Calais est la région la plus artificialisée, avec moins de 10% de territoire à l'état naturel (données issues du recensement Teruti-Lucas de 2010 concernant l'utilisation du territoire : Agreste, 2014). Cette artificialisation est non seulement marquée par les grandes aires urbaines, mais aussi par des surfaces importantes de cultures, la présence d'un réseau dense d'infrastructures de transport (elle possède une position géographique stratégique de carrefour au sein de l'Europe) et surtout une concentration industrielle passée qui a façonné son territoire depuis le début du XIX^e siècle et dont l'activité reste encore importante (PRSE2, 2011). La figure 18 présente l'occupation du sol en 2009 sur la région Nord – Pas de Calais selon la classification Corine Land Cover.

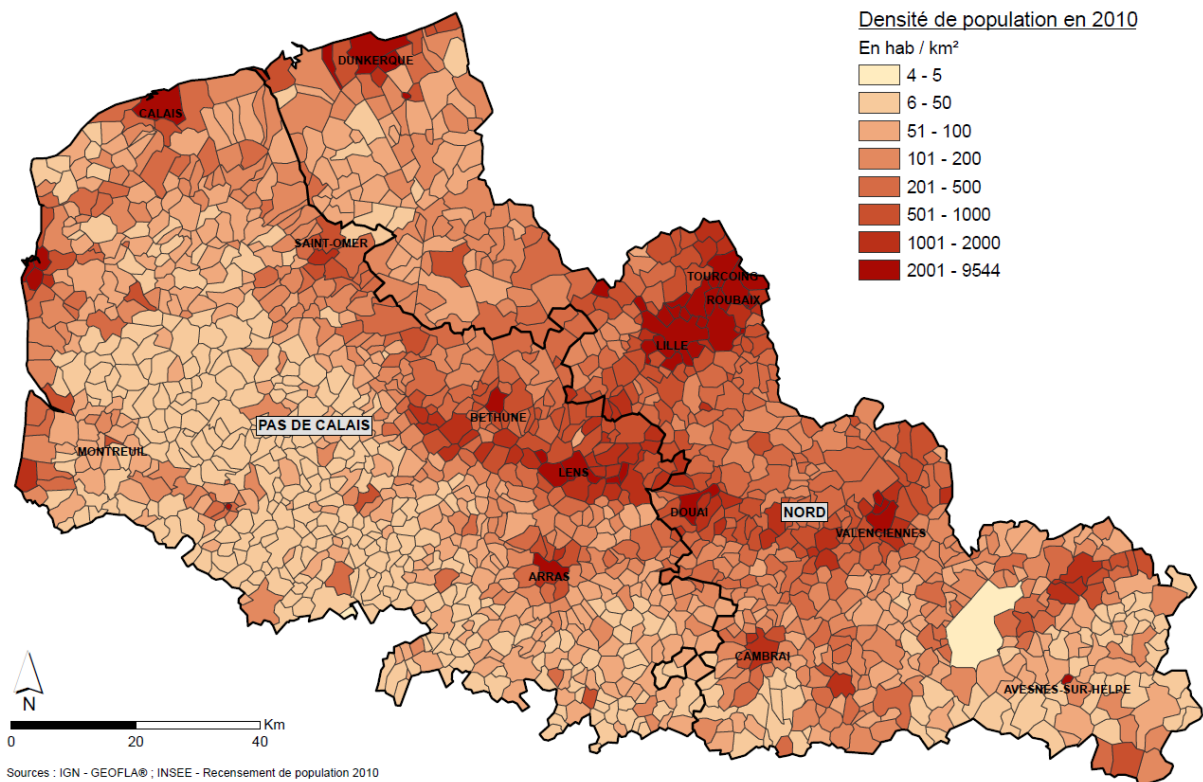


Figure 17. Densité de population par commune en Nord – Pas de Calais (logiciel ArcGIS®)

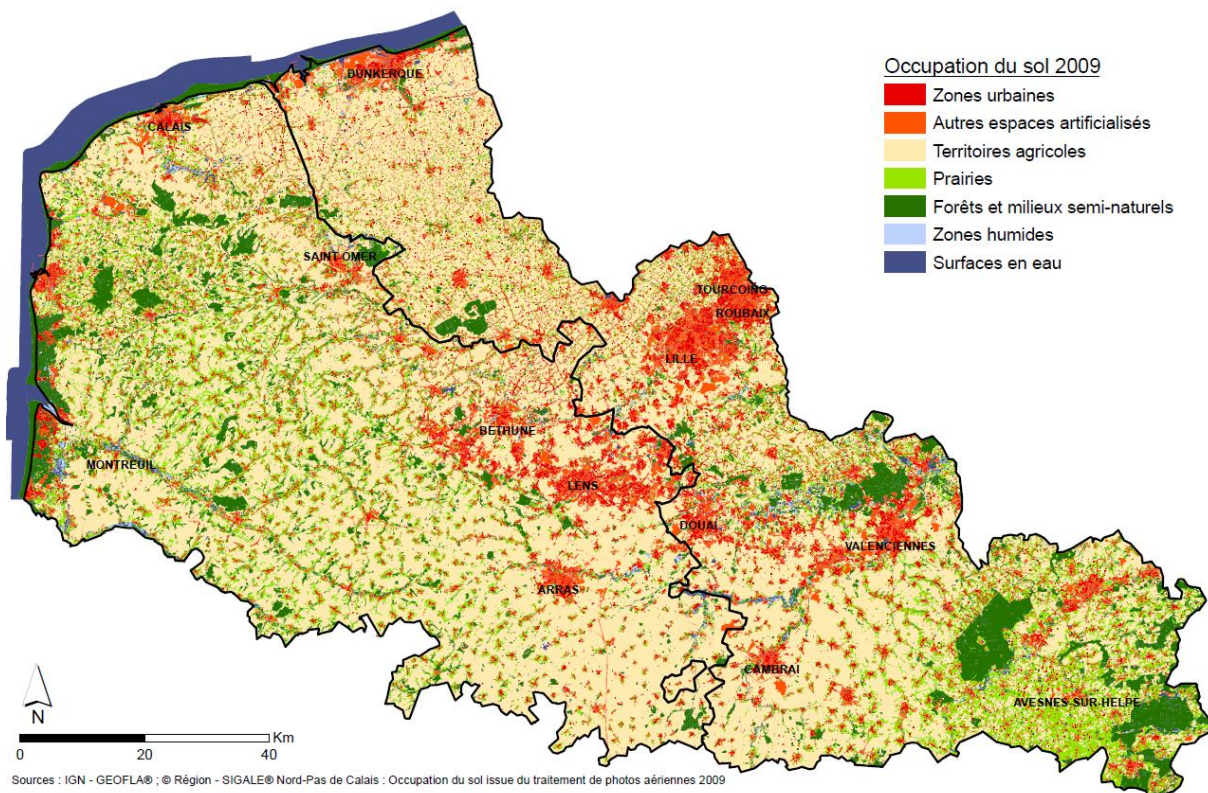


Figure 18. Occupation du sol en Nord – Pas de Calais (logiciel ArcGIS®)

D'un point de vue environnemental, le passé industriel se traduit par la présence de la moitié des friches industrielles de France (majoritairement regroupées sur les agglomérations de Lille – Roubaix – Tourcoing, Valenciennes et le bassin minier : Decocq, 2010), 15% des sites et sols pollués nationaux (648 sites pollués ou potentiellement pollués d'origine industrielle, dont 77% sont situés sur le département du Nord : inventaire BASOL, 2013) et 6,7% des sites potentiellement pollués d'origine industrielle, commerciale et de service (16 800 sites, dont 75% pour le Nord : inventaire BASIAS, 2013) (figure 19). Les activités de toutes natures de ces sites (chimie, textile, briqueterie, industrie pétrolière, sidérurgie, usine à gaz, etc.) ont notamment provoqué une pollution des sols très variée : éléments traces métalliques, hydrocarbures aromatiques polycycliques (HAP), polychlorobiphényles (PCB), pesticides organochlorés, dioxines, furanes, etc. (DREAL, 2014a).

Le nombre de sites industriels actuellement en activité est encore élevé en région (13 500 en 2011 : INSEE, 2014), malgré une forte diminution au cours des dernières décennies. Le Nord – Pas de Calais compte près de 1700 Installations Classées pour la Protection de l'Environnement (ICPE) soumises au régime d'autorisation administrative et 74 sites classés SEVESO, dont 50 au régime seuil haut (MEDDE : Inspection des installations classées, 2014). En termes de densité d'installations, la région se situe au premier rang en France. Depuis la fermeture des quelques 600 puits du bassin minier à la fin du XX^e siècle, c'est le littoral dunkerquois qui présente la plus grande concentration de sites industriels de la région, avec plus d'une centaine d'établissements classés ICPE (régime d'autorisation) sur les communes de Dunkerque, Grande-Synthe, Loon-Plage et Gravelines. A ce titre, le bassin dunkerquois fut également choisi comme zone d'intérêt pour le premier axe de recherche (chapitre I). Malgré une nette diminution, l'activité industrielle de l'ancien bassin minier reste tout de même importante.

Le grand nombre de sites pollués, la forte densité de population et l'activité industrielle de la région sont susceptibles d'impacter la qualité des milieux environnementaux. En ce qui concerne le milieu aquatique, la qualité physico-chimique et biologique des eaux souterraines et supérieures reste préoccupante et les masses d'eau sont sensibles à toute pollution (DREAL, 2014a). Seules 17 % des masses d'eau de surface du Nord – Pas de Calais présentent un bon état écologique en 2011. Cette qualité insatisfaisante est notamment liée à la présence constante et importante de HAP (94% des masses d'eau) et de pesticides (25% des masses d'eau) et à la forte artificialisation des cours d'eau (figure 20). La tendance de cet état est tout de même à l'amélioration suite aux objectifs fixés par la Directive Cadre Eau et le risque de non-atteinte du bon état écologique d'ici 2015 est moins préoccupant qu'à l'échelle nationale (INSEE, 2013).

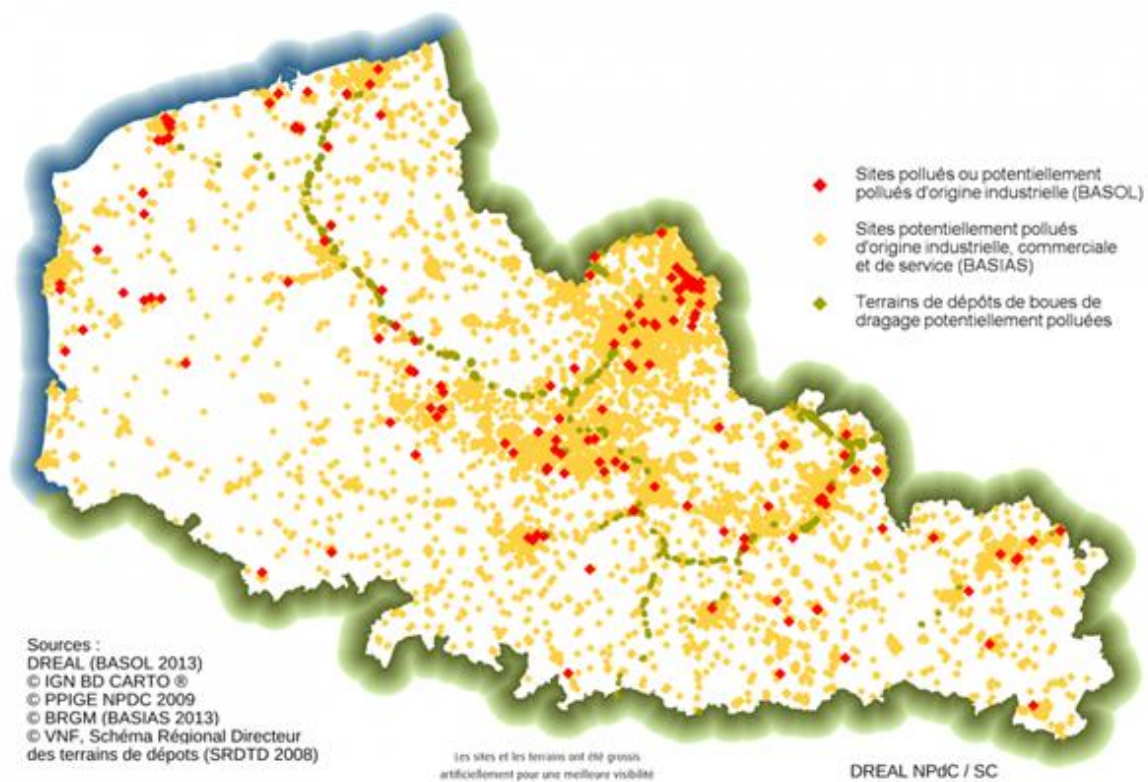


Figure 19. Localisation des sols potentiellement pollués inventoriés (DREAL, 2014a)

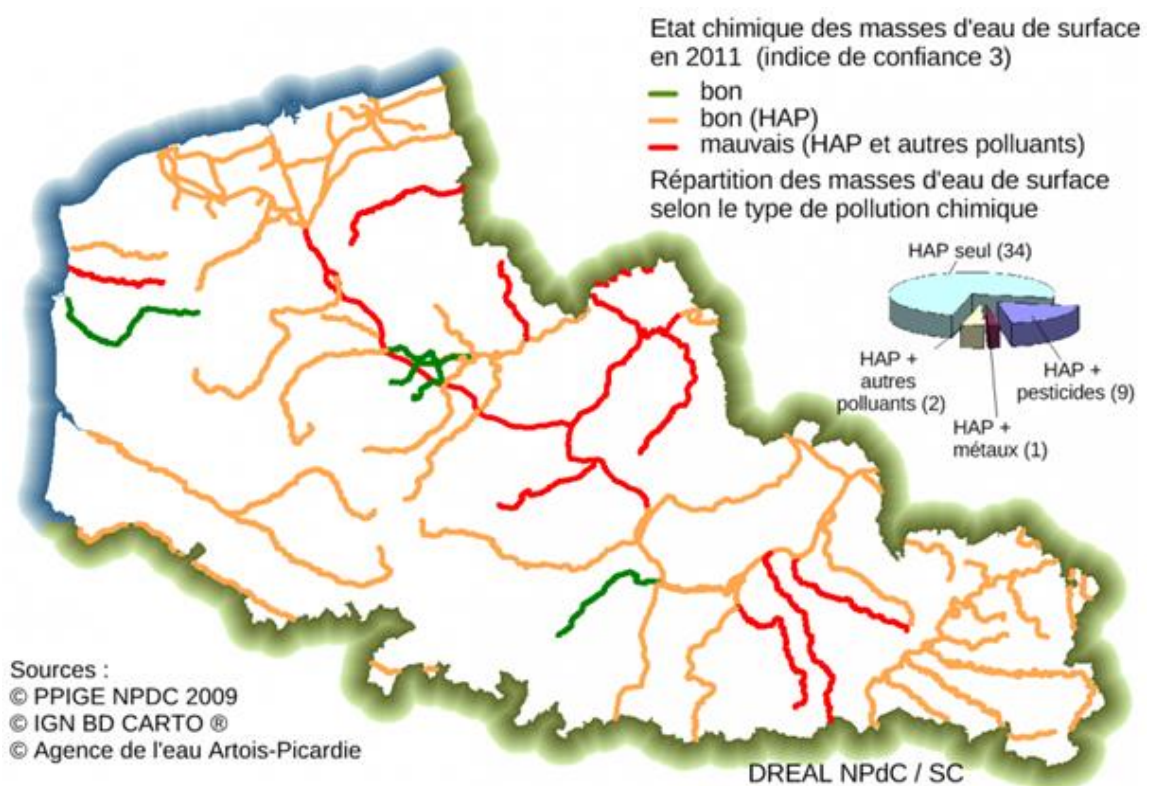


Figure 20. État chimique des eaux de surface en 2011 (DREAL, 2014a)

De manière atypique en France, près des trois quarts des prélèvements d'eau destinés à alimenter la population sont issus du sous-sol (ADES, 2014). La région bénéficie en effet d'une grande richesse en eaux souterraines. La nappe de la craie recouvre 80% du territoire régional et contient des volumes importants d'eau situés à seulement quelques dizaines de mètres de profondeur (DREAL, 2014a). Elle est majoritairement libre (non recouverte par une couche imperméable) ce qui la rend très vulnérable aux pollutions. Dans la plaine des Flandres, cette nappe est même non aquifère, c'est-à-dire en affleurement avec la surface. Actuellement, 9 des 15 masses d'eau souterraine sont déclarées en mauvais état qualitatif (toutes issues de la nappe de la Craie). Les paramètres anthropiques responsables et en hausse ces dernières années sont les nitrates et les pesticides. Des paramètres naturellement présents dans les roches, tels que le nickel et le sélénium sont également surveillés.

La pollution de l'air en région Nord – Pas de Calais est importante, du fait du grand nombre de sources d'émissions locales (industries, axes routiers majeurs, forte densité d'habitats et grandes aires de cultures) et limitrophes (trafic maritime, Grande Bretagne, Europe de l'Est) (DREAL, 2014b). Ces dernières années, les normes réglementaires de concentrations en polluants dans l'air extérieur (moyennes journalières) sont régulièrement dépassées (plus de 35 jours par an : figure 21) pour les particules fines (PM_{10}), et le dioxyde d'azote (NO_2). Ces dépassements sont l'une des causes d'un contentieux de l'a France avec l'Europe. Un Plan de Protection de l'Atmosphère (PPA), ayant pour objectif de réduire les concentrations en polluants dans l'air ambiant sous des valeurs assurant le respect de la santé des populations (valeurs réglementaires définies dans le Code de l'Environnement) vient à ce titre d'être approuvé par les préfets de région (DREAL, 2014b).

Concernant les indicateurs socio-économiques, le Nord-Pas-de-Calais présente en 2011 le revenu fiscal médian par unité de consommation le plus faible de France (16 800€), avec un taux de pauvreté de 19,7%, figurant parmi les trois premières régions métropolitaines. Les taux d'allocataires du revenu de solidarité active (RSA) et de la couverture maladie universelle de base (CMU) sont les plus élevés de France (INSEE, 2014). Au sein de la région, de fortes disparités de niveau de précarité sociale sont observées. Ainsi les régions de Lille – Roubaix – Tourcoing, Dunkerque, Maubeuge, Calais et de l'ancien bassin minier présentent les plus hauts niveaux de défaveur sociale (figure 22).

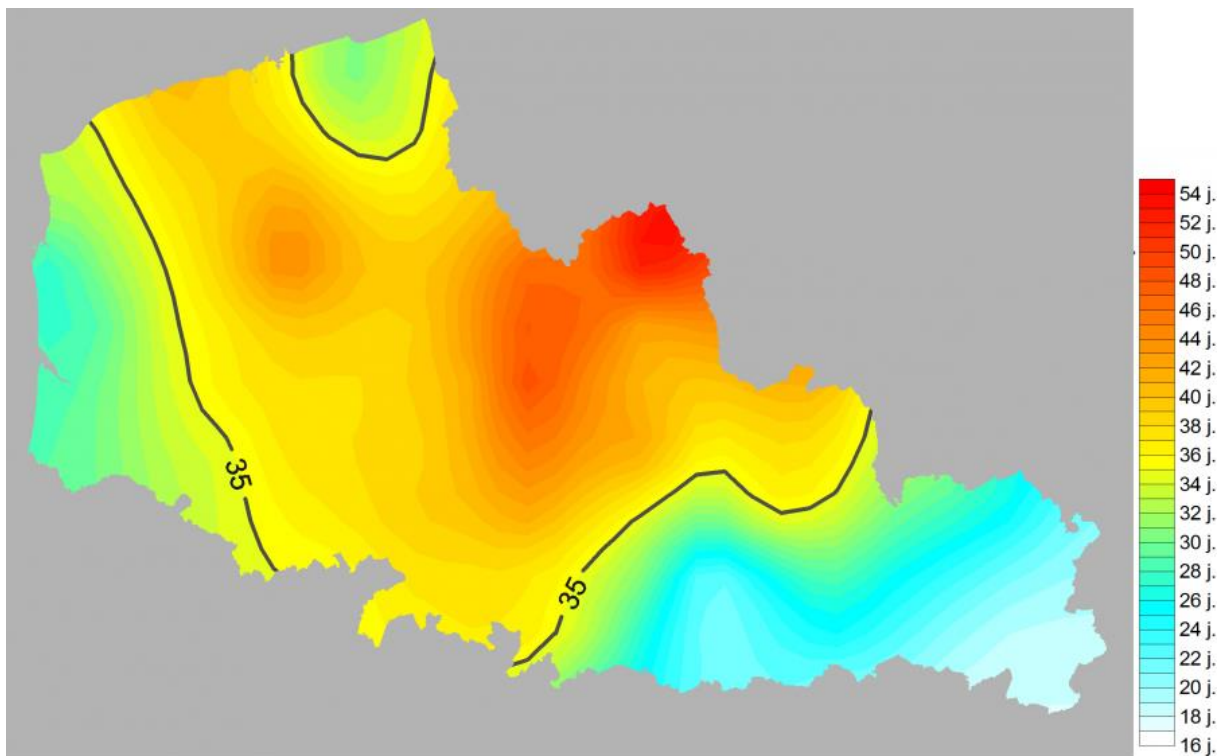


Figure 21. Nombre de jours de dépassements de la valeur limite journalière de $50 \mu\text{g}/\text{m}^3$ pour les PM10 en 2011 (ATMO Nord – Pas de Calais)

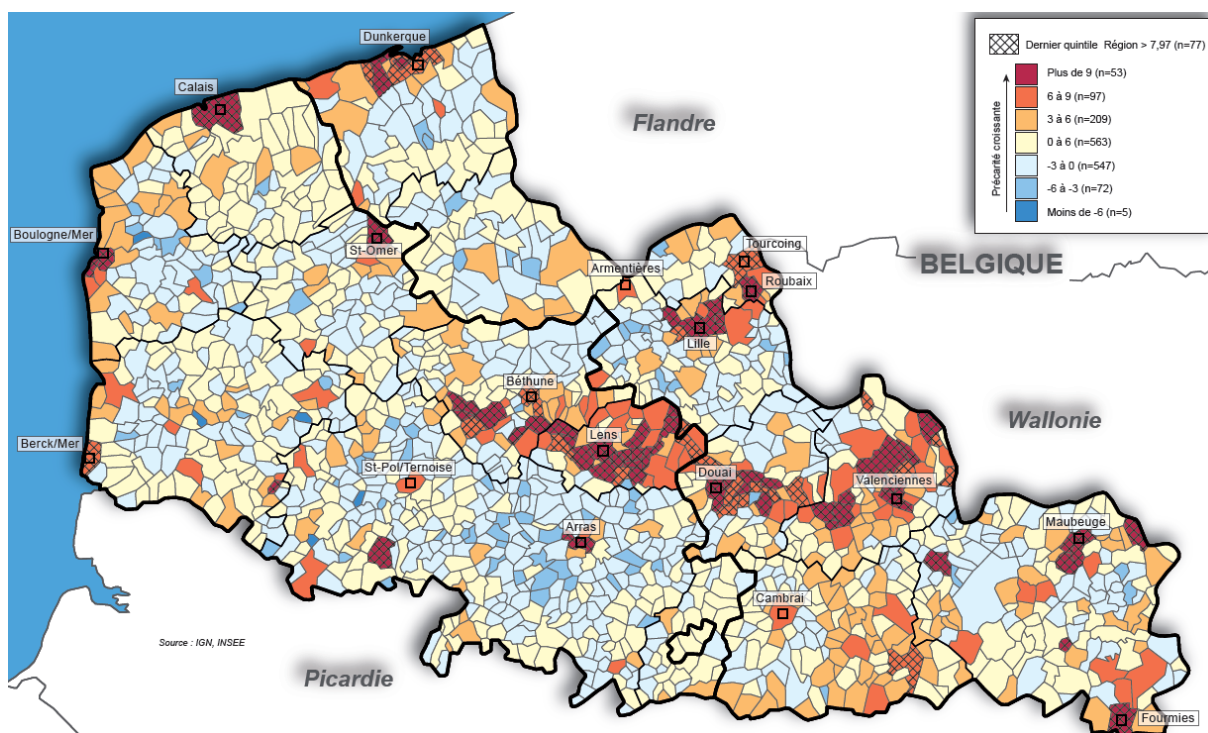


Figure 22. Indice de défaveur sociale de Townsend par commune en 2006 (ORS Nord – Pas de Calais, 2011)

L'espérance de vie à la naissance en Nord – Pas de Calais est la plus faible des régions métropolitaines (74 ans pour les hommes et 82 ans pour les femmes) et la mortalité prématurée (avant 65 ans) y est la plus importante, notamment chez les hommes. Cette surmortalité est principalement liée aux comportements à risques (alcoolisme, tabagisme) et à l'inégale répartition de l'offre de soins sur le territoire (ORS Nord – Pas de Calais, 2014). D'importantes inégalités de mortalité et morbidité apparaissent au sein de la région. Les bassins de Lens et Valenciennes présentent les taux les plus élevés, alors que les bassins de Lille, d'Artois et des Flandres se rapprochent plus de la moyenne nationale (figure 23).

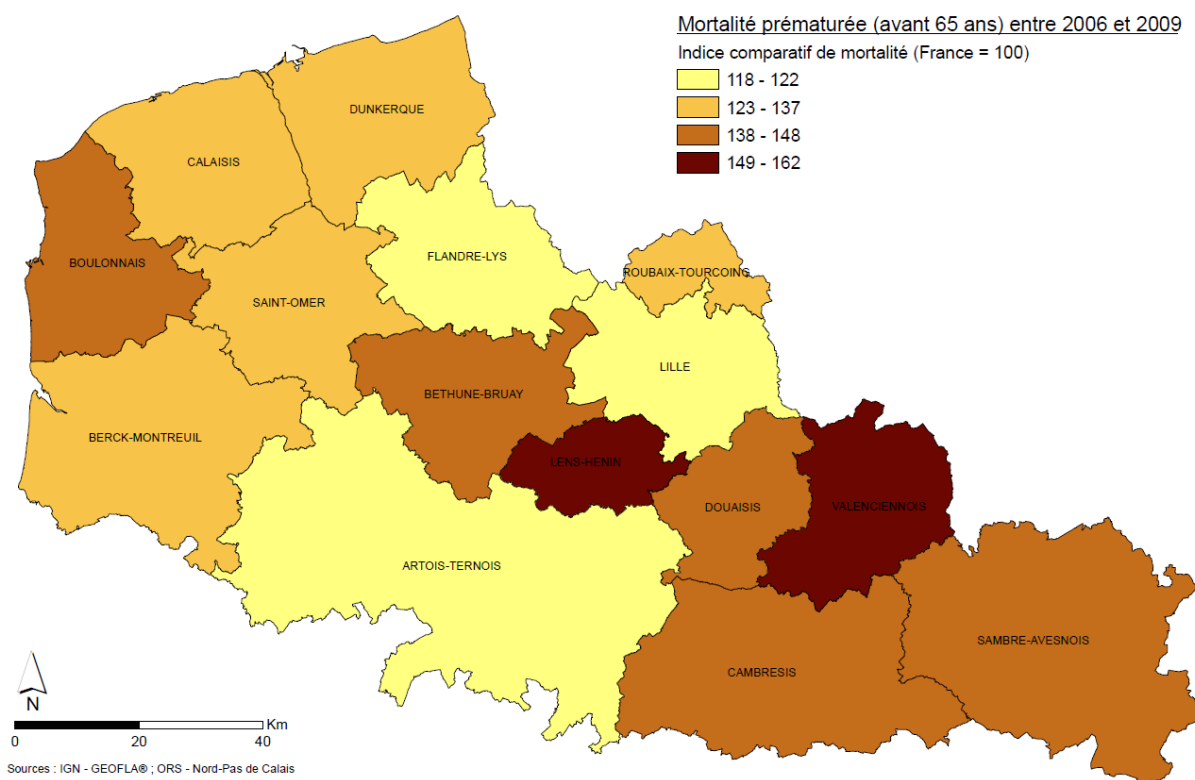


Figure 23. Mortalité prématurée (avant 65 ans) par zone d'emploi, deux sexes confondus – un ICM supérieur à 100 indique une surmortalité par rapport à la population de référence – un ICM de 130 implique une surmortalité de 30% (logiciel ArcGIS®)

Du fait de la forte dégradation de ses milieux environnementaux, des disparités importantes en termes de morbidité et mortalité et des hauts niveaux de précarité sociale, la région Nord – Pas de Calais constitue donc un territoire propice pour les études d'inégalités environnementales et sociales de santé.

Pour la réalisation d'études en santé environnementale, après avoir défini le territoire d'étude, il est nécessaire d'établir un bilan des ressources disponibles. A ce titre, les bases de données environnementales, sanitaires et sociodémographiques disponibles sur le territoire du Nord – Pas de Calais ont été recensées et sont présentées dans la partie suivante.

Partie 2. Ressources

Dans cette partie nous dressons la liste des sources de données environnementales, sanitaires et sociodémographiques disponibles sur le territoire de la région Nord – Pas de Calais. Une attention particulière est portée sur l’accessibilité de ces bases, ainsi que leur contenu.

a. Données sanitaires

Nous nous sommes dans un premier temps intéressé aux sources de données renseignant sur l’état de santé des populations, disponibles sur le territoire de la région Nord – Pas de Calais. Le tableau 4 dresse la liste des bases récoltées, ainsi que leurs étendues spatiales et temporelles. Les équipes responsables de ces bases ont toutes été rencontrées afin d’établir des partenariats scientifiques dans le cadre de recherches en santé – environnement. Ces bases concernent 4 registres de maladies : les cancers toutes localisations confondues sur la zone de proximité de Lille (Registre des cancers de Lille), l’IRCT sur la région Nord – Pas de Calais (Réseau REIN Néphronor), les MICI sur 4 départements français dont le Nord – Pas de Calais (Registre EPIMAD) et les maladies cardio-vasculaires sur la Communauté Urbaine de Lille (Registre MONICA de Lille). En plus de ces registres de santé, des cohortes de patients ont été suivies. Une première cohorte concerne les patients ayant consulté l’Institut de Biologie de la Reproduction (IBDR) du CHRU de Lille (Dr. Valérie Mitchell, EA 4308) pour des troubles de de la fertilité masculine. A ce titre, plusieurs bases sont en cours de création. Elles concernent des questionnaires administrés à ces patients ainsi qu’à de jeunes pères et les spermogrammes établis lors de la consultation. Les partenariats concernent deux autres cohortes suivies en population générale. La cohorte IMEPOGE s’intéresse aux dosages d’ETM dans les fluides biologiques (sang et urines) et la cohorte ELISABET aux troubles ventilatoires obstructifs, notamment par des explorations respiratoires (mesures effectuées dans l’air expiré) et des dosages biologiques de métaux.

Tableau 4. Ressources de données sanitaires récoltées lors de la thèse

Ressource	Etendue spatiale	Etendue temporelle	Responsable
Registre des cancers 7000 cas de cancers	Zone de Proximité de Lille	2005 et 2008	Dr K. Ligier
Registre REIN (Néphronor) 5000 cas d'IRCT	Nord - Pas de Calais	2005 - 2012	Dr F-X. Glowacki
Registre EPIMAD 8000 cas de MICI	Nord - Pas de Calais Somme Seine-Maritime	1990 - 2012	Dr C. Gower-Rousseau
Registre MONICA Cas de pathologies cardio-vasculaires	Lille Métropole Communauté Urbaine	1995 - 2010	Pr P. Amouyel
Etude fertilité masculine 1333 Questionnaires administrés (dont 250 jeunes pères) 1083 spermogrammes	Nord - Pas de Calais	2012 - 2013	Dr Valérie Mitchell
Etude IMEPOGE Imprégnation par les ETM 2000 sujets	Nord - Pas de Calais	2008 - 2010	Dr C. Nisse Dr A. Leroyer
Etude ELISABET Mesures respiratoires Imprégnation par les ETM 4200 sujets	Lille Métropole Communauté Urbaine Communauté Urbaine de Dunkerque	2011 - 2013	Dr L. Dauchet

Ainsi, deux formats de données sont disponibles. Les registres de santé recensent de manière exhaustive les cas de maladies sur une étendue géographique et temporelle donnée. A partir de ces informations et des recensements de population, des indicateurs d’incidence sont calculés, de manière agrégée pour des unités administratives (commune, canton, etc.). Les cohortes de sujets

apportent quant à elles des données de biosurveillance individuelles des populations, telles que des dosages biologiques de marqueurs d'exposition ou d'effets sanitaires. Dans le cadre de cette thèse, trois de ces bases ont été étudiées : la base d'imprégnation par les métaux en population générale IMEPOGE (chapitre I), le registre REIN du Nord – Pas de Calais (chapitre II) et le registre EPIMAD (chapitre III).

b. Données environnementales

Concernant la caractérisation des milieux environnementaux, un rapport du réseau coopératif de recherche sur les déchets (rapport n°09-0812/1A soutenu par l'ADEME : Algros & Charissou, 2010) dresse un état des lieux des bases de données existantes en France. Les principales conclusions de ce rapport sont la prédominance de bases de données de caractérisation des sols, qui concernent 10 des 14 bases recensées et la focalisation sur les mesures de concentrations en micropolluants métalliques. Ce rapport souligne également de fortes disparités entre les différents milieux en termes d'accessibilité aux données et de précision du contenu (donnée brute ou prétraitée). Le milieu aquatique semble être le moins contraignant, suivi du sol et de l'air.

Sur la base de ce rapport, un état des lieux des données environnementales existantes fut réalisé lors de cette thèse, avec pour critères le territoire de la région Nord – Pas de Calais et la contamination des milieux. Le tableau 5 dresse la liste de ces ressources.

Les conclusions de notre recherche sont similaires à celles établies par le rapport RECORD en termes de quantités, d'accessibilité et de pertinence des sources d'informations. Le sol est le milieu le plus représenté, suivi de l'eau et enfin de l'air. L'accessibilité aux données environnementales dépend de leur valeur scientifique intrinsèque, ainsi que de leur sensibilité. En effet, le secret statistique limite souvent la diffusion publique de données issues des campagnes d'analyses. C'est notamment le cas pour les sols.

Tableau 5. Ressources de données environnementales existantes sur le territoire de la région Nord – Pas de Calais

Ressource	Détenteur	Milieu	Contenu	Accessibilité	Obtention	Format	Donnée polluants
Inventaire National Spatialisé des émissions	CITEPA	Air	Emissions atmosphériques	Non publique	Partenariat en cours	Base	Traitée
Inventaire Régional Spatialisé des émissions	ATMO NPC	Air	Emissions atmosphériques	Non publique	Partenariat en cours	Base	Traitée
Base de Données de la Qualité de l'Air	ATMO NPC	Air	Stations de mesures	En ligne	Sur commande	Base	Brute
Biosurveillance de la qualité de l'air	LSVF	Air	Bioindication / Bioaccumulation	Interne	Interne	Bases	Brute
Inventaire BASOL	BRGM	Sol	Sites et sols pollués	En ligne	Téléchargement	Base	Aucune
Réseau de Mesure de la Qualité des Sols	INRA	Sol	Campagnes d'analyses	En ligne	Téléchargement	Cartographie	Traitée
Banque de Données d'Analyses des Terres	INRA	Sol	Campagnes d'analyses	En ligne	Téléchargement	Cartographie	Traitée
Banque de Données ETM	ADEME / INRA	Sol	Campagnes d'analyses	En ligne	Téléchargement	Cartographie	Traitée
Diagnostics de pollution des sites	LMCU / Ville de Lille / Tauw	Sol	Campagnes d'analyses	Non publique	Partenariat en cours	Rapports	Brute
Campagnes d'analyses des sols	ADEME / ISA Lille EA 4415	Sol	Campagnes d'analyses	Non publique	Non	Bases	Brute
Biosurveillance de la qualité des sols	LSVF	Sol	Bioindication / Bioaccumulation	Interne	En cours	Base	Brute
Accès aux Données des Eaux Souterraines	Eau France / Agences de l'eau	Eau	Piézomètres / qualitomètres	En ligne	Téléchargement	Base	Brute
Surveillance des eaux de distribution	Eaux du Nord	Eau	Non connu	Non publique	Non	Non connu	Non connu
Surveillance sanitaire des eaux de distribution	ARS NPC	Eau	Campagnes d'analyses	En ligne	Téléchargement	Base	Traitée
Surveillance des eaux de surface	Eau France / Agences de l'eau	Eau	Stations de mesures	Non publique	Non	Base	Brute
Surveillance de l'eau potable	Ministère de la Santé	Eau	Campagnes d'analyses	Publique	Non	Rapport	Traitée
Industrie au Regard de l'Environnement	DREAL NPC	Tous	Emissions par industrie	En ligne	Téléchargement	Rapports	Brute
Registre Français des Emissions Polluantes	MEDDE / INERIS	Tous	Emissions par industrie	En ligne	Téléchargement	Base	Brute
Inventaire BASIAS	BRGM	Tous	Activités polluantes	En ligne	Téléchargement	Base	Aucune

Ainsi, même si une dizaine de bases de données de qualité des sols existent, aucune n'est accessible sous la forme d'un tableau de valeurs brutes de concentrations, mesurées au niveau d'un échantillon de points. La plupart d'entre elles sont diffusées sous forme de cartographies. Par exemple, le Réseau de Mesures de la Qualité de Sols (RMQS), propose une carte des concentrations en cadmium des sols agricoles de surface à l'échelle nationale, selon une grille comportant 2200 mailles de 16 km² (figure 24). Les seules bases de données disponibles sous forme de téléchargement sont les inventaires BASIAS et BASOL. Mais aucun ne répertorie les résultats des campagnes d'analyses de sols. BASIAS dresse la liste des activités potentiellement polluantes et BASOL les sites et sols pollués ou potentiellement pollués. Enfin les campagnes d'analyses des sols menées dans le cadre de recherches scientifiques font encore l'objet de publications et ne sont donc pas diffusées. Afin de pallier ce manque de données brutes de qualité des sols pouvant être intégrées à nos études en santé – environnement, des partenariats avec la Ville de Lille, la Communauté Urbaine de Lille et le bureau d'étude Tauw France ont été entamés. Ils concernent la récolte des données issues des études de sols réalisées dans le cadre de projets de réaménagement de sites. En parallèle, les données issues des études de biosurveillance de la qualité des sols réalisées par le LSVF sont en cours de saisie.

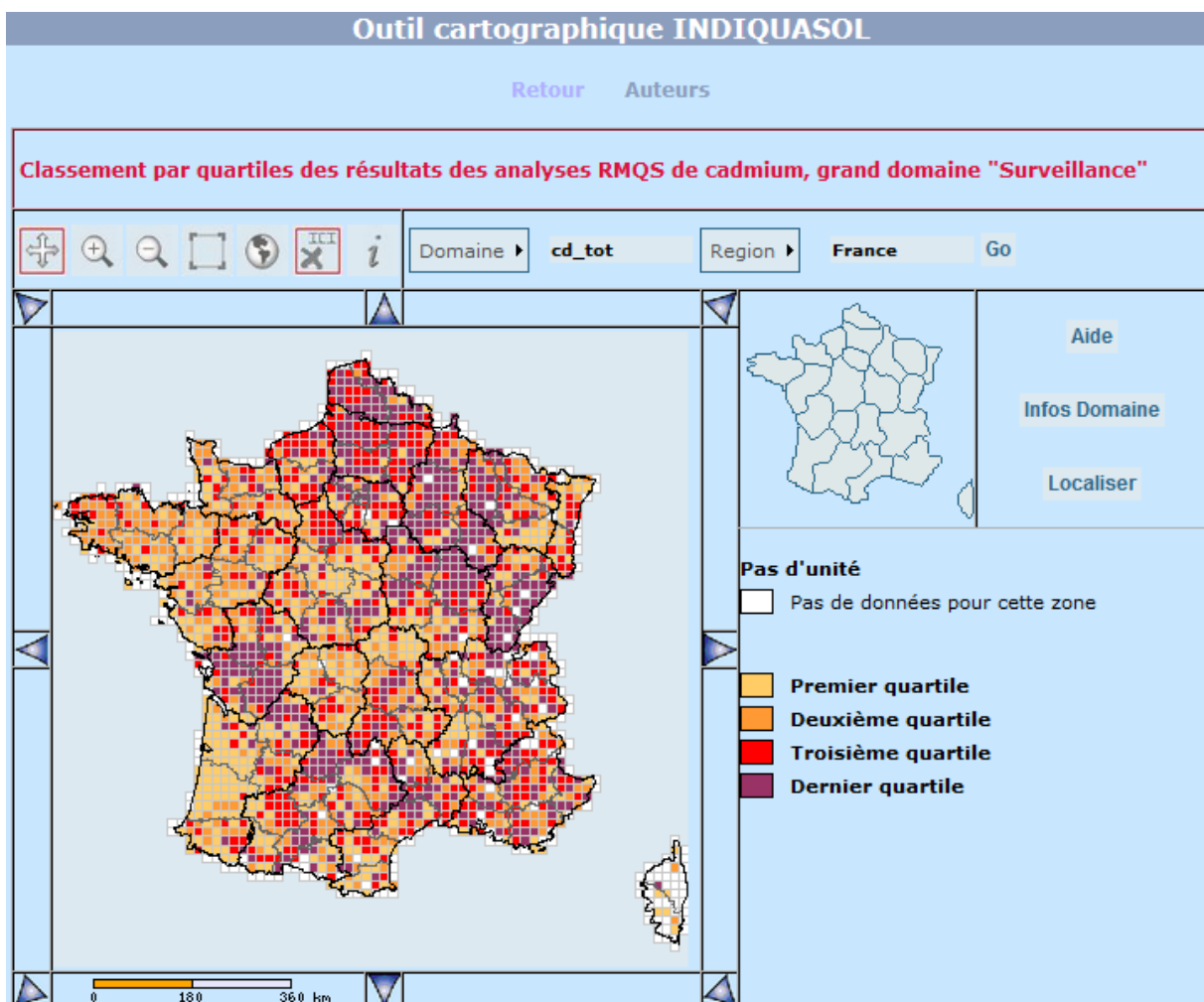


Figure 24. Diffusion cartographique du RMQS (INDIQUASOL, 2014)

Pour le milieu aquatique, le portail Eau France représente une ressource conséquente de données sur l'eau superficielle et souterraine. Cependant, seul le portail d'Accès aux Données des Eaux Souterraines (ADES) propose l'export de données brutes à différentes échelles spatiales et temporelles. Les données issues de la banque ADES ont été utilisées dans le 3^e axe de recherche (chapitre III). Pour les eaux de distribution, les résultats du contrôle sanitaire de la qualité de l'eau potable sont diffusés sur internet par le Ministère chargé de la Santé à l'échelle de la commune. Mais l'export des informations n'est pas possible pour un ensemble de communes, ce qui complexifie la récolte des données. De plus, les paramètres de potabilités renseignés ne concernent pas les micropolluants environnementaux.

Enfin, trois principaux types de données sont disponibles pour le compartiment atmosphérique. Les émissions en polluants atmosphériques sont estimées par le CITEPA à l'échelle nationale et par les AASQA à l'échelle régionale. En Nord – Pas de Calais, les données de plus de 40 polluants atmosphériques sont répertoriées par secteur d'activités, selon une maille communale ou maille de 3km. Celles-ci n'étant pas rendues publiques, un partenariat avec le réseau ATMO NPC vient d'être entamé pour leur transmission. Le deuxième type de données représente les immissions atmosphériques, c'est-à-dire les concentrations ambiantes. La banque de données de la qualité de l'air (BDQA) recense les données issues des stations de mesure et est accessible sur commande. Cependant, les stations de mesures étant réparties de manière très disparate sur le territoire, ces données ne peuvent être utilisées dans une étude spatialisée. Enfin des données de biosurveillance de la qualité de l'air sont générées par le LSVF depuis les années 1970 et font l'objet de diverses études spatialisées (voir Introduction générale, Partie 4). Une partie de ces données est l'objet du premier axe de recherche de la thèse (chapitre I).

c. Données sociodémographiques

L'étude des disparités spatiales d'incidence de pathologies nécessite l'intégration de données démographiques, notamment pour le calcul des indicateurs standardisés d'incidence. Les données des recensements de populations effectués par l'INSEE à différents niveaux administratifs (quartiers, communes, cantons) ont donc été récoltées (INSEE, 2014). Les recensements de population concernent les années 1990, 1999, 2006, 2008, 2009 et 2010.

Afin d'évaluer les inégalités environnementales et les inégalités sociales de santé, des indicateurs composites de précarité sociale ont été utilisés dans cette thèse. La précarité sociale est considérée dans ces travaux comme « un état observable et démontrable de désavantage relatif face à la communauté locale ou l'ensemble de la société à laquelle appartient l'individu, la famille ou le groupe » (Townsend, 1987). Le terme « défaveur sociale » est ici employé comme la traduction du mot anglais « deprivation ». Plusieurs indicateurs de défaveur ont été recensés lors de la thèse. Ces indices sont construits sur la base des données sociodémographiques de l'INSEE. Les indices de Townsend (1987) et Carstairs (Carstairs & Morris, 1991) sont les plus couramment utilisés. Parmi les plus récents, nous pouvons citer l'indice de Havard (Havard *et al.*, 2008), calculé à l'échelle de l'Iris (Ilôts regroupés pour l'information statistique) sur la Communauté Urbaine de Strasbourg. Sur le territoire du Nord – Pas de Calais, l'Indice de Défaveur Localisé (IDL) fut récemment développé pour les Iris par le laboratoire TVES ULCO (Dr. A.P Hellequin et Rachel Bavdek). Celui-ci intègre huit grandes thématiques : les formes d'emploi, l'exclusion de l'emploi, l'éducation, le lien social, le revenu et enfin le logement.

Parmi ces indicateurs, l'indice de Townsend a été employé dans le 2^e axe de recherche (chapitre II). Un partenariat avec le laboratoire TVES ULCO a permis l'utilisation de l'IDL dans le cadre du premier axe de recherche (chapitre I).

En conclusion de cette partie consacrée aux ressources, nous avons pu constater un retard des bases de données environnementales par rapport aux registres de santé ou aux recensements des populations. En effet les données populationnelles sont recensées au sein de bases de données standardisées et établies depuis plusieurs années, ce qui offre une puissance statistique satisfaisante lors de leur analyse. Bien que protégées par la loi Informatique et Libertés, nous n'avons rencontré aucune difficulté d'accès à ces données, sous la condition d'autorisation par la Commission Nationale de l'Informatique et des Libertés (CNIL).

Concernant les ressources environnementales, il existe une grande diversité de ressources, celles-ci étant souvent volumineuses. Pourtant (contrairement à ce qui était attendu au début de la thèse), peu de données brutes sont actuellement disponibles en accès libre. Les données environnementales de contamination des milieux ne sont en effet pas concernées par la Directive INSPIRE⁵ et font la plupart du temps l'objet de confidentialité.

Parmi l'ensemble des ressources recensées, le chapitre I traitera des données de biosurveillance lichénique du LSVF, de celles de l'étude IMEPOGE, ainsi que de l'IDL. Le chapitre II s'intéressera au registre Néphronor d'IRCT et à l'indice de Townsend. Enfin dans le chapitre III, les données de la banque ADES seront analysées, en vue d'une confrontation aux cartographies d'incidence des MICI. Les parties suivantes correspondent aux matériels et méthodes de chacun des chapitres de la thèse.

Partie 3. Matériels et méthodes du chapitre I

Ce chapitre a pour premier objectif d'évaluer les disparités spatiales d'imprégnation de l'environnement par les ETM au niveau de trois bassins de vie de la région Nord – Pas de Calais. Les données de biosurveillance lichénique seront pour cela cartographiées à l'aide des outils géostatistiques. Le deuxième objectif est d'évaluer les inégalités environnementales, par la mise en relation des cartographies lichéniques avec un indicateur de défaveur sociale, développé à l'échelle de l'Iris. Enfin le troisième objectif est d'évaluer les inégalités environnementales et sociales de santé en intégrant à l'analyse les données de biosurveillance humaine, de dosages en ETM dans le sang et les urines.

a. Matériels

LA BIOSURVEILLANCE LICHENIQUE

La biosurveillance lichénique est utilisée depuis de nombreuses années par le LSVF, en collaboration avec l'APPA-NPC, pour caractériser et suivre l'évolution de la qualité de l'air dans l'espace et le temps (Delzenne-Van Haluwyn, 1973 ; Hardouin, 1983 ; Van Haluwyn, 1994 ; Van Haluwyn & Cottard, 1995 ; Cuny *et al.*, 2003 ; Cuny *et al.*, 2004). Dans le cadre de l'institut de recherche en environnement industriel (IRENI), les lichens ont été employés de manière passive pour mettre en évidence les niveaux d'imprégnation de l'environnement atmosphérique par les ETM. Des dosages de 18 métaux ont été réalisés dans les thalles du lichen *Xanthoria parietina*, prélevés au niveau de trois bassins de vie de la région Nord – Pas de Calais en 2009.

Les trois zones d'étude sont le bassin de Dunkerque (28 communes, 220 000 habitants), possédant une importante activité industrielle, le bassin de Lille (85 communes, 1 105 000 habitants), qui se caractérise d'avantage par un habitat dense et de très nombreuses infrastructures routières importantes et le bassin de Maubeuge (40 communes, 120 000 habitants), où le milieu rural est dominant. A partir d'un maillage kilométrique établi sur chacun des trois bassins, un tirage au sort

⁵ La directive INSPIRE, élaborée par la Direction générale de l'environnement de la Commission européenne, vise à établir en Europe une infrastructure de données géographiques pour assurer l'interopérabilité entre bases de données et faciliter la diffusion, la disponibilité, l'utilisation et la réutilisation de l'information géographique en Europe.

des mailles de prélèvement a été effectué (NF X43-904 AFNOR, 2013). La force d'échantillonnage est de 60 mailles pour Dunkerque, 120 mailles pour Lille et 40 mailles pour le Maubeuge.

Le lichen *Xanthoria parietina* est une espèce nitrophile très courante dans la région et a un rythme de croissance d'environ 1 cm par an. Des thalles de 6 cm de diamètre ont été récoltés, ce qui correspond donc à une exposition sur les 6 dernières années environ. Pour chacune des mailles, trois échantillons de thalles ont été prélevés sur trois arbres différents et de même essence.

Les 18 éléments traces recherchés dans les thalles de lichens étaient les suivants : l'aluminium (Al), l'antimoine (Sb), l'arsenic (As), Le béryllium (Be), le cadmium (Cd), le cobalt (Co), le chrome (Cr), le cuivre (Cu), le manganèse (Mn), le mercure (Hg), le nickel (Ni), le plomb (Pb), le palladium (Pd), le Platine (Pt), le Rhodium (Rh), le titane (Ti), le vanadium (V), le zinc (Zn). L'aluminium, le manganèse, le titane et le zinc ont été dosés par ICP-AES (spectroscopie d'émission optique avec plasma induit par haute fréquence). Les autres éléments ont été dosés par ICP-MS (spectrométrie de masse avec plasma à couplage inductif). L'échantillon européen de référence (numéro d'identification 482 – lichen *Pseudevernia furfuracea* contaminé par des métaux) a été utilisé dans le contrôle qualité des dosages. Les concentrations résultantes mesurées représentent « les traces » accumulées dans l'environnement et non directement les concentrations atmosphériques.

Le centre de chaque maille de récolte a été géoréférencé à partir de coordonnées GPS (protocole WGS84). La base de données fut construite sous Excel et comporte 220 lignes correspondant aux mailles et 20 variables (les 18 ETM et les coordonnées géographiques).

L'INDICE DE DEFAVEUR LOCALISE

Dans le cadre du projet VULDEPO2 (vulnérabilité des populations à la pollution atmosphérique), également lancé par l'IRENI, un indicateur de défaveur sociale fut développé afin d'évaluer les inégalités sociales à très fine échelle sur la région. L'Indice de Défaux Localisé fut construit par le laboratoire TVES de l'ULCO. L'unité géographique définie pour cet indicateur est l'Iris. L'IDL a pour objectif de permettre une identification territorialisée des populations les plus vulnérables sur un plan socio-économique.

Contrairement aux indicateurs statistiques (construits sur la base d'une analyse en composante principale par exemple), l'IDL est constitué à partir d'une approche éco-sociologique, basant le choix des variables selon un critère bibliographique. Il est inspiré d'un indicateur reconnu, institutionnalisé et facilement appropriable : l'Indicateur de Santé Sociale des régions (ISS : Jany-Catrice & Zotti, 2009). Pour cela, il synthétise diverses variables, elles-mêmes regroupées en 6 grands thèmes illustratifs des conditions de vie des habitants d'un territoire, à savoir les formes d'emploi, l'exclusion de l'emploi, l'éducation, le lien social, le revenu et enfin le logement. L'IDL vise à situer un Iris par rapport à l'ensemble des Iris d'un espace de référence. Il s'agit donc d'un indicateur relatif et transposable à d'autres territoires, puisqu'une unité considérée comme défavorisée dans un territoire A (par exemple une communauté urbaine) ne le sera pas nécessairement dans un territoire B (une région). Afin de valider l'IDL, un test de corrélation avec les indices les plus utilisés à cette échelle (Townsend, Carstairs, Havard) fut effectué. Le niveau de corrélation est proche de 90%. L'IDL a donc été estimé à partir du recensement de population de 2009 de l'INSEE, pour les Iris des bassins de Dunkerque, de Lille et de Maubeuge.

IMPREGNATION PAR LES METAUX EN POPULATION GENERALE (IMEPOGE)

La base de données IMEPOGE fut générée par les Dr C. Nisse et A. Leroyer (EA4483 Université Lille 2 - Centre de Recherches en Santé - Travail - Environnement – CHRU de Lille), en parallèle des dosages lichéniques dans le cadre du programme scientifique IRENI. Elle est issue du suivi de l'imprégnation par des polluants environnementaux de populations fréquentant les centres d'examen de santé du Nord – Pas de Calais. Cette étude a pour objectif initial d'établir des valeurs de base d'imprégnation en population générale de la région Nord – Pas de Calais, afin :

- d'identifier les sources d'exposition contribuant à la variation des concentrations ;
- de positionner les niveaux d'exposition par rapport à ceux d'autres pays ;
- de guider l'interprétation des indicateurs d'exposition en milieu professionnel.

En effet, la mesure de substances chimiques ou de leurs métabolites dans les liquides biologiques, comme indicateur d'exposition, permet de donner des informations sur l'exposition récente ou la dose accumulée dans l'organisme, en fonction des indicateurs (Nisse, 2010).

Des prélèvements sanguins et urinaires ont été réalisés sur un échantillon de 2000 adultes (20-59 ans) répartis sur la région Nord – Pas de Calais pour doser la concentration de différents métaux. Dans ce cadre, des données biométriques (taille, poids, tension artérielle, etc.), environnementales (coordonnées du lieu de vie, habitudes alimentaires et tabagiques, habitat, activités professionnelles et extra-professionnelles), socio-économiques (diplômes, consommation médicale, accès à la propriété, etc.), biologiques (bilans sanguins, hépatiques, etc.) et d'antécédents médicaux ont été recueillies pour chaque participant. Un total de 12 métaux ont été recherchés dans les échantillons sanguins et urinaires : Al, As, Cd, Co, Cr, Hg, Mn, Ni, Pb, Tl (Thallium), V, Zn.

Une première cartographie de la répartition spatiale des participants (géoréférencés selon l'adresse du lieu d'habitation) nous indiquait la présence de sujets sur les bassins de Dunkerque et Lille. En revanche, aucun point n'était situé sur le bassin de Maubeuge. Les participants domiciliés sur les deux premiers bassins de vie ont donc été sélectionnés, afin d'évaluer les inégalités environnementales et sociales de santé. Parmi les 12 ETM dosés, seul le thallium n'a pas été dosé chez les lichens. Cette variable fut retirée notre étude.

LES DONNEES ADMINISTRATIVES

Pour cartographier les indicateurs environnementaux, sanitaires et sociodémographiques, les contours des bassins furent obtenus à partir de l'agrégation des contours de communes qui les composent. Le fichier shape GEOFLA® de l'IGN contient la description de l'ensemble des unités administratives du territoire national : communes, cantons, arrondissements, départements, régions. Les contours des Iris du Nord et du Pas de Calais sont issus du fichier shape Contours...Iris®, coédité par l'INSEE et l'IGN (Institut National de l'Information Géographique et Forestière). Les Iris ont été définis par l'INSEE pour les besoins des recensements à l'échelle infra-communale des communes de plus de 10 000 habitants et la plupart des communes de 5 000 à 10 000 habitants. Ils regroupent en moyenne 2 000 habitants. Le fichier Contours...Iris® permet de faire le lien entre les données cartographiques et les données statistiques de l'INSEE.

b. Stratégie expérimentale

Dans une première partie de ce chapitre, les outils cartographiques ont été mis en œuvre afin d'évaluer les disparités spatiales d'imprégnation de l'environnement. Les données lichéniques ont été cartographiées de manière ponctuelle, puis estimées à l'aide d'une méthode géostatistique d'interpolation spatiale. Les cartes ont été générées pour chacun des 18 éléments traces de manière individuelle, mais également de manière globale grâce à un indice intégré d'imprégnation multimétallique. Cet indice est présenté par la suite.

Dans la deuxième partie consacrée aux inégalités environnementales, les résultats des cartographies lichéniques ont été mis en relation avec l'IDL, grâce à une étude écologique. Nous avons vu précédemment que les méthodes d'interpolation spatiale fournissent, à partir de données géostatistiques, une surface continue représentant la valeur prédite de l'indicateur en tout point de la zone étudiée. Cette surface permet d'évaluer l'hétérogénéité spatiale d'un indicateur de qualité des milieux. Mais l'information qu'elle contient ne peut pas être rapprochée sous cette forme à l'indicateur de défaveur sociale, puisque celui-ci est représenté par une information agrégée par unité géographique. Il est donc nécessaire que ces données de différentes natures soient exprimées dans les mêmes configurations. Les données environnementales ont donc été adaptées et transformées en format géostatistique : une concentration moyenne en ETM par Iris.

Enfin la troisième partie intègre les données de biosurveillance humaine. Nous disposons dans ce cas de données individuelles géolocalisées à l'adresse. Afin d'évaluer les inégalités environnementales et sociales de santé, les outils SIG (boîte à outils d'ArcGIS® Geostatistical Analyst) ont été employés pour estimer les concentrations environnementales au niveau des lieux de vie de chacun des sujets

intégrés à l'étude. L'IDL de l'Iris d'habitation a également été attribué à chacun des sujets à l'aide d'une jointure spatiale. La base finale générée contient pour chacun des individus les caractéristiques individuelles, les concentrations en métaux dans les fluides biologiques, les concentrations en métaux dans l'environnement proche et l'IDL de l'Iris d'habitation. Une analyse multivariée a ainsi pu être mise en œuvre. Parmi les métaux mesurés chez l'Homme et chez les lichens, 11 sont communs aux deux bases. L'indice d'imprégnation multimétallique des lichens fut redéfini à partir de ces 11 éléments. Afin d'évaluer le niveau d'imprégnation global des populations, le même indice fut calculé pour le sang et les urines.

c. Méthodes

CARTOGRAPHIES ENVIRONNEMENTALES

Pour évaluer les variations spatiales du niveau d'imprégnation de l'environnement par les ETM, des cartographies ponctuelles et géostatistiques ont été réalisées grâce au logiciel ArcGIS® (version 10.1 : ESRI France). Il n'existe pas de normes établies de concentration en éléments traces dans les lichens. De fait, afin d'interpréter conjointement les résultats obtenus pour chaque élément, nous avons utilisé comme valeurs de référence, des gammes de concentrations retrouvées dans la littérature, notamment les travaux de Van Haluwyn & Cuny (1997), Bargagli (1998), Scerbo *et al.* (1999 et 2002) et Cuny *et al.* (2004 ; 2010). Ces valeurs sont présentées dans le tableau ci-dessous (tableau 6). Des classes de niveau d'imprégnation ont été définies à partir de ces valeurs de référence, afin d'obtenir une échelle de valeurs commune lors de l'analyse conjointe des cartes.

Tableau 6. Valeurs référence (en mg/kg poids sec) des concentrations en ETM dans les lichens

Element	Al	As	Be	Cd	Co	Cr	Cu	Hg	Mn	Ni	Pb	Pd	Pt	Rh	Sb	Ti	V	Zn
Valeur référence	500	0,5	3	0,2	7	4	15	0,2	30	10	15	15	15	15	0,3	5	12	50

L'outil de krigeage empirique bayésien (EBK) de l'extension Geostatistical Analyst (disponible uniquement à partir de la version 10.1 d'ArcGIS®) fut utilisé pour estimer les valeurs de concentrations en ETM en tout point de l'espace des bassins (Krivoruchko, 2012). L'estimation des multiples variogrammes et simulations fut réalisée séparément pour chacun des bassins, afin de ne pas tenir compte des variations de grande distance (54 variogrammes estimés). Au final, 36 cartographies monométalliques (18 ponctuelles et 18 géostatistiques) ont été produites.

INDICE INTEGRE D'IMPREGNATION MULTIMETALLIQUE

Pour compléter les 18 cartographies monométalliques et identifier les points noirs de contamination, un indice intégré d'imprégnation fut généré. Il s'agit d'un indice d'enrichissement du milieu, représentant le niveau d'imprégnation par rapport aux concentrations naturelles retrouvées dans les lichens. Il est inspiré de l'indice OIP (Overall Index of Pollution) développé par Sargaonkar & Deshpande (2003) pour la surveillance des rivières en Inde, dont la méthode de calcul a été reprise par Chen *et al.* (2005) avec l'IPI (Integrated Pollution Index) pour évaluer la qualité des sols de surface en Chine.

Pour chaque élément dosé en un point d'échantillonnage, le ratio de la concentration mesurée (C_i) sur la valeur de référence (B_i) est dans un premier temps calculé. La moyenne de ces 18 ratios représente l'indice. Cet indice sera nommé par la suite le ratio d'imprégnation moyen (RIM).

$$RIM = \frac{\sum C_i/B_i}{18}$$

Pour le calcul des RIM sanguin et urinaire, les valeurs de référence sont celles déterminées lors de l'étude IMEPOGE. Elles représentent le percentile 95 des résultats globaux de l'enquête régionale Nord – Pas de Calais (Nisse, 2010).

AGREGATION DES DONNEES ENVIRONNEMENTALES

Pour mettre en relation les mesures d'imprégnation lichénique et l'IDL, une transformation des données environnementales était nécessaire. Nous avons pour cela généré un maillage régulier constitué de carrés de 100m de côté sur chacun des bassins. Ce maillage a ensuite été superposé à la couche géostatistique lichénique pour attribuer une valeur de concentration métallique à chacune des mailles. Une valeur moyenne de concentration est enfin attribuée à chaque Iris, à partir des valeurs des mailles contenues dans l'unité. Nous obtenons ainsi une information agrégée au sein de chaque Iris, qui peut être mise en relation statistique avec l'IDL. Cette opération fut réalisée pour chacun des 18 éléments et pour le RIM.

EVALUATION DES INEGALITES ENVIRONNEMENTALES

Les inégalités environnementales ont été estimées sous XLSTAT, à l'aide de corrélations écologiques. Afin de définir si les populations les plus défavorisées vivent dans un environnement imprégné par les ETM, plusieurs tests statistiques ont été mis en œuvre. Le coefficient de corrélation de Spearman a dans un premier temps été calculé. Celui-ci teste la corrélation entre les rangs attribués aux Iris en fonction des valeurs environnementales de l'IDL. Les Iris ont ensuite été catégorisés en quintiles en fonction de la valeur de l'IDL. Le premier quintile représente les Iris les plus défavorisés, et le dernier quintile les Iris les plus favorisés. Au sein de chaque quintile, les valeurs environnementales des Iris ont été répertoriées. Le test de Dunn a finalement été utilisé pour comparer les quintiles deux à deux, en fonction de ces valeurs environnementales.

EVALUATION DES INEGALITES ENVIRONNEMENTALES ET SOCIALES DE SANTE

Pour évaluer la place de l'environnement dans l'imprégnation des populations par les ETM, une analyse de corrélations entre les trois matrices (lichens, sang, urines) a été réalisée, en intégrant les facteurs confondants mis en évidence (tabac, âge). La matrice de corrélation de Spearman a pour cela été générée pour mesurer le niveau de corrélation entre les variables. Le coefficient de Spearman ne mesure pas la corrélation entre les valeurs brutes de concentrations des individus, mais entre les rangs qu'elles occupent au sein de la distribution. Les analyses statistiques ont également été réalisées à l'aide du logiciel XLSTAT.

Partie 4. Matériels et méthodes du chapitre II

Dans ce deuxième chapitre, le premier objectif est d'évaluer l'hétérogénéité spatiale de l'IRCT au sein de la région Nord – Pas de Calais, à partir d'un registre de santé. Des indicateurs standardisés d'incidence seront calculés à l'échelle des cantons et représentés à l'aide d'une méthode de cartographie des maladies. Le deuxième objectif est d'évaluer l'influence du statut socioéconomique des populations, ainsi que d'autres facteurs confondants dans l'apparition de ces disparités, à l'aide d'une étude écologique. Enfin le troisième objectif est de savoir si ces disparités s'expriment par la présence de clusters atypiques d'IRCT, testés grâce à la statistique de scan.

a. Matériels

REGISTRE REIN (NEPHRONOR)

Le Réseau Epidémiologie et Information en Néphrologie (REIN), est un registre de santé mené à l'échelle nationale par l'Agence de Biomédecine. Il concerne l'ensemble des patients en traitement de suppléance pour une insuffisance rénale chronique, par dialyse ou transplantation rénale, en France métropolitaine et dans les DOM-TOM (Couchoud *et al.*, 2006). Tous les nouveaux patients ayant atteint le stade terminal de l'insuffisance rénale chronique sont enregistrés dans ce système d'information. Le Registre Nephronor est la déclinaison régionale de REIN et recense l'ensemble des patients de la région Nord – Pas de Calais. Il existe depuis janvier 2005, sous la direction du Dr. F. Glowacki. Ce registre repose sur l'enregistrement continu et exhaustif d'informations pour l'ensemble des patients. Notre étude est basée sur les patients incidents entre 2005 et 2012. Nous nous sommes intéressés à la commune de résidence au moment du premier traitement de suppléance, la néphropathie initiale, la capacité d'épuration du système rénal à l'initiation (le débit de filtration glomérulaire), le nombre de consultations auprès d'un néphrologue, ainsi que l'âge et le sexe.

DONNEES DEMOGRAPHIQUES

Les données de population utilisées pour calculer les indicateurs d'incidence sont issues des recensements de population de l'INSEE. Seuls les recensements à la commune des années 2006, 2008, 2009 et 2010 sont disponibles. Les données de population pour l'année 2005 ont été considérées comme identiques au recensement de 2006. Pour les années 2007 et 2009, ce sont respectivement les recensements de 2006 et 2008 qui ont été utilisés. Enfin, le recensement de 2010 a servi de modèle pour les années 2011 et 2012. Les données communales ont finalement été agrégées au canton correspondant.

INDICE DE TOWNSEND

Pour évaluer le lien entre le niveau d'incidence de l'IRCT et la défaveur sociale à l'échelle des cantons, l'indice de Townsend (Townsend, 1987) a été utilisé. Cet indice est composé à partir de quatre variables : la proportion de chômeurs dans la population active ($V1$), la proportion de résidences principales occupées par plus d'une personne par pièce ($V2$), la proportion de résidences principales dont le ménage occupant n'est pas propriétaire ($V3$) et la proportion de ménages sans voiture ($V4$). La formule de construction de l'indice (I) est fournie ci-dessous. Les données issues du recensement de population communal de 2009, agrégé au canton, ont été utilisées pour le calcul de cet indice. Un indice faible indique une population favorisée et un indice élevé une population défavorisée.

$$I = \frac{\{\log(V1 + 1) - moyenne[\log(V1 + 1)]\}}{\text{écart} - type[\log(V1 + 1)]} + \frac{\{\log(V2 + 1) - moyenne[\log(V2 + 1)]\}}{\text{écart} - type[\log(V2 + 1)]} \\ + \frac{\{V3 - moyenne(V3)\}}{\text{écart} - type(V3)} + \frac{\{V4 - moyenne(V4)\}}{\text{écart} - type(V4)}$$

b. Méthodes

CARTOGRAPHIE DE L'INCIDENCE D'IRCT A L'ECHELLE DES CANTONS

Pour cartographier la répartition spatiale des cas d'IRCT (toute néphropathie initiale confondue), l'ensemble des cas ont été classés en fonction de classes d'âge quinquennales (0-4 ans ; 5-9ans ; ... ; 80-84 ans ; plus de 85 ans) et du sexe. Les données de population des cantons ont également été classées de la même manière. L'unité spatiale choisie pour représenter l'incidence d'IRCT est le canton-ville. Le Nord – Pas de Calais compte 170 cantons, ce qui offre une bonne résolution spatiale. De plus, le nombre de cas (5362) semble, de prime abord, trop faible pour calculer ces indicateurs à

plus fine échelle (communes ou Iris). Les patients ont été géocodés à partir du code INSEE de la commune de résidence (attribution des coordonnées géographiques « XY » du centroïde⁶ de la commune) et agrégés au sein de chaque canton, grâce au logiciel ArcGIS® v10.1.

Les taux d'incidence standardisée (TIs) et les ratios d'incidence standardisée (SIR) ont été calculés pour chacun des cantons, selon les formules présentées ci-dessous. La population de référence choisie pour les calculs est celle de la zone d'étude (le Nord – Pas de Calais). Le modèle hiérarchique bayésien (BYM) proposé par Besag *et al.* (1991) a enfin été utilisé en collaboration avec l'équipe du Pr. Alain Duhamel (EA 2694, Université Lille 2) pour lisser les SIR calculés et les cartographies ont été réalisées sous ArcGIS® v10.1.

$$TIs_j = \sum_{i=1}^n P_i T_{ij}$$

$$SIR_j = \frac{O_j}{E_j} \quad \text{où} \quad E_j = \sum_{i=1}^n N_{ij} T_i$$

Avec :

i : indice de classe d'âge

n : nombre de classes d'âge

j : unité géographique

P_i : part de la classe d'âge *i* dans la population de référence

T_{ij} : taux spécifique d'incidence observé pour la classe *i* de l'unité géographique *j*

O_j : nombre total de cas observés (tous âges confondus) dans l'unité géographique *j*

E_j : nombre total de cas attendus dans l'unité géographique *j*

N_{ij} : effectif de la population correspondant à la classe d'âge *i* dans l'unité géographique *j*

ETUDE DE CORRELATIONS ECOLOGIQUES

Pour évaluer les liens entre les variations d'incidence et le niveau de défaveur sociale, une analyse de corrélation écologique a été menée. Les valeurs de l'indice de Townsend ont été discrétisées selon la méthode des quintiles pour être cartographiées et mises en relation avec les SIR lissés. La classe ayant les valeurs les plus faibles (Q1) fut prise comme référence. Un risque relatif (RR) a été ensuite calculé pour chacun des quatre autres quintiles et complété d'un calcul d'intervalle de confiance (IC) pour évaluer sa significativité. L'influence d'autres facteurs (maladie rénale initiale, prise en charge tardive, nombre de consultations) a été évaluée à l'aide du coefficient de corrélation de Pearson.

DETECTION DE CLUSTERS ATYPIQUES D'ÉVÉNEMENTS

Les clusters spatiaux et spatio-temporels d'IRCT ont été évalués à l'aide de la statistique de scan (logiciel SaTScan™ : Kulldorff, 1997 ; 2010). Les coordonnées XY des centroïdes des cantons ont été considérées comme référence géographique et les facteurs âge et sexe ont été intégrés à l'analyse, de la même manière que pour le calcul des SIR. La statistique fut testée de façon temporelle, spatiale et spatio-temporelle, de manière à détecter des clusters de surincidence et de sous-incidence. Des clusters de forme circulaire ont été considérés. La taille d'un cluster n'excède pas 25% de la population totale du Nord – Pas de Calais et 50% de la période d'étude. L'inférence statistique fut finalement évaluée à l'aide de 9999 réplifications de Monte Carlo.

Dans un second temps, une analyse spatiale a été effectuée selon la méthode isotonique de SaTScan™ (Kulldorff, 1999), afin d'estimer des risques non homogènes au sein de chaque cluster détecté. L'ensemble des cartographies ont été réalisées à l'aide du logiciel ArcGIS® v10.1.

⁶ Centre géométrique de l'aire géographique

Partie 5. Matériels et méthodes du chapitre III

Dans ce chapitre, nous nous intéressons aux facteurs de risques environnementaux des MICI. Dans ce cadre, l'objectif était d'évaluer le potentiel cartographique de la banque ADES, pour le suivi spatio-temporel de la qualité des eaux souterraines et la réalisation d'une étude écologique en lien avec les MICI.

a. Matériels

REGISTRE EPIMAD

Le Registre des maladies inflammatoires chroniques du tube digestif (EPIMAD) recense les cas de MICI depuis 1988 dans le Nord de la France, au sein de quatre départements : le Nord, le Pas de Calais, la Somme et la Seine-Maritime (Gower-Rousseau *et al.*, 1994). Les données issues du registre ont déjà fait l'objet de plusieurs analyses spatiales, démontrant non seulement des disparités d'incidence entre les cantons des quatre départements pour les patients recensés entre 1990 et 2006 (Declercq *et al.*, 2010), mais aussi la présence de clusters spatio-temporels de surincidence et de sous-incidence (Génin *et al.*, 2013). Une analyse spatio-temporelle à l'échelle de la commune et intégrant les patients incidents entre 1990 et 2012 est actuellement en cours de réalisation au sein de l'équipe du Pr. Duhamel (EA2694, Université Lille 2).

LA BANQUE D'ACCES AUX DONNEES DES EAUX SOUTERRAINES (ADES)

La banque nationale ADES est une base de données disponible en téléchargement gratuit (ADES Eau France, 2014). Elle recense l'ensemble des points d'eau de captage pour une alimentation en eau souterraine. Un point d'eau est un ouvrage souterrain dont le captage d'eau est destiné à alimenter des activités industrielles, agricoles, BTP, mais également la distribution des eaux de consommation humaine collectives et individuelles (puits). Cette base ADES permet un suivi qualitatif et quantitatif des eaux souterraines depuis les années 1980, et est alimentée par les différents acteurs de l'eau (Agence de l'eau, Agence Régionale de Santé, DREAL, collectivités, etc.). Des informations comme le mode d'utilisation de l'ouvrage souterrain, la profondeur du forage ou la hauteur de la nappe d'eau souterraine sont également disponibles. Les données utilisées dans le cadre de ces travaux de recherche concernent le suivi qualitatif des eaux (dosages de plus de 450 substances polluantes). Ces dosages sont effectués directement à la sortie du captage par des laboratoires agréés, sans traitement préalable de l'eau.

b. Méthodes

RECOLTE ET OPTIMISATION DE LA BASE DE DONNEES

Le téléchargement de la base ADES s'exécute via un formulaire de sélection des données. Il est possible de choisir l'emprise géographique, le mode d'utilisation du point d'eau et les paramètres à observer. Pour des raisons de sécurité, les coordonnées géographiques des points d'eau ne sont pas mentionnées dans la base. Ces informations ont été retrouvées et téléchargées à partir du portail InfoTerre™ du Bureau de Recherches Géologiques et Minières (BRGM : InfoTerre, 2014), via la base de données des ouvrages souterrains (Banque du Sous-Sol : BSS). Une jointure a permis d'agrèger les deux bases par un identifiant commun (le code BSS, attribué à chaque point d'eau). Nous nous sommes focalisés sur les métaux, notamment les plus recherchés au regard de contaminations du sous-sol (Al, As, Cd, Cr, Cu, Hg, Ni, Pb, Zn). Ce choix s'est fait vis à vis des hypothèses en rapport à l'étiologie des MICI. La base de données des micropolluants métalliques a été téléchargée pour la zone et la période d'étude du registre EPIMAD. Ce registre recouvre 4 départements (Nord, Pas de Calais, Somme, Seine – Maritime) et s'étend de 1990 à 2012.

Une fois téléchargée, la base de données fut mise en forme sous Excel. Les analyses critique et statistique de cette base de données (tri des variables, statistiques descriptives, etc.) ont été réalisées à l'aide du logiciel XLSTAT (Addinsoft).

ANALYSE SPATIALE ET SPATIO-TEMPORELLE DE LA CONTAMINATION DES EAUX SOUTERRAINES

L'analyse cartographique concerne les résultats de mesures de concentration en aluminium, obtenus pour des échantillons d'eau prélevés directement à la sortie des points de captage. Ces résultats ont été considérés en fonction des coordonnées spatiales des points d'eau et de la date de prélèvement. Les analyses spatiales et spatio-temporelles ont été réalisées sous ArcGIS® v10.1. Les cartographies géostatistiques sont issues de la méthode EBK (Krivoruchko, 2012).

INDICE INTEGRE DE CONTAMINATION MULTIMETALLIQUE

A l'instar des données lichéniques, un indice intégré de contamination fut généré pour évaluer la qualité globale du milieu. Le RIM a également été utilisé. Pour rappel, c'est un indice d'enrichissement du milieu, représentant le niveau d'imprégnation par rapport aux concentrations naturelles, représentées dans notre étude par des normes environnementales (tableau 7).

Tableau 7. Normes environnementales utilisées pour le calcul du RIM eaux souterraines (Arrêté du 11/01/07)

Élément	Limite de qualité des eaux brutes utilisées pour la production d'eau destinées à la consommation humaine	Valeur seuil des eaux souterraines
Aluminium		200 µg/l
Arsenic	100 µg/l	10 µg/l
Cadmium	5 µg/l	5 µg/l
Chrome	50 µg/l	50 µg/l
Cuivre		2 mg/l
Mercurure	1 µg/l	1 µg/l
Nickel		20 µg/l
Plomb	50 µg/l	10 µg/l
Zinc	5 mg/l	5 mg/l

Pour chacun des 9 éléments dosés en un point d'échantillonnage, le ratio de la concentration mesurée (C_i) sur la valeur de référence (B_i) est dans un premier temps calculé. La moyenne de ces 9 ratios représente le RIM.

$$RIM = \frac{\sum C_i/B_i}{9}$$

REFERENCES BIBLIOGRAPHIQUES

AFNOR. 2013. NF X 43-904 : Biosurveillance passive de la qualité de l'air à l'aide des lichens autochtones : de la récolte à la préparation des échantillons.

Algros E & Charissou AM. 2010. Réalisation d'un avis d'experts : « bruit de fond » des polluants dans l'environnement, analyse des données existantes et identification des lacunes. Etude RECORD N° 09-0812/1A. 163p.

Arrêté du 11/01/07 relatif aux limites et références de qualité des eaux brutes et des eaux destinées à la consommation humaine mentionnées aux articles R. 1321-2, R. 1321-3, R. 1321-7 et R. 1321-38 du code de la santé publique.

Bargagli R. Trace elements in terrestrial plants, an ecophysiological approach to biomonitoring and biorecovery, Springer, Berlin, 1998, 324 p.

Besag J, York J, Mollié A. 1991. Bayesian image restoration, with two applications in spatial statistics (with Discussion). *Annals of the Institute of Statistical Mathematics*, 43(1): 1-59.

Carstairs V, Morris R. 1990. Deprivation and health in Scotland. *Health Bull*, 48 (4): 162-175

Chen TB, Zheng YM, Lei M, Huang ZC, Wu HT, Chen H, Fan KK, Yu K, Wu X, Tian QZ. 2005. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere*, 60: 542–551.

Couchoud C, Stengel B, Landais P, Aldigier J-C, De Cornelissen F, *et al.* 2006. The renal epidemiology and information network (REIN): a new registry for end-stage renal disease in France. *Nephrology Dialysis Transplantation*, 21: 411–418.

Cuny D, Davranche L, Kempa M, Vandamme L, Pailleux N, Caous AS, Dossin M, Silvie JP, Van Haluwyn C. 2003. Comparison of lichen flora (between 1995 and 2002) in two cities in the north of France : contribution of nitrogen pollution ? In *Nitrogen in the Environment*, Netlecombe Court, Taunton, Angleterre, 24-27/02/03.

Cuny D, Davranche L, Thomas P, Kempa M, Van Haluwyn C. 2004. Spatial and temporal variations of trace element contents in *Xanthoria parietina* thalli collected in a highly industrialized area in Northern France as an element for a future epidemiological study. *Journal of Atmospheric Chemistry*, 49: 391-401.

Cuny D, Quarre S, Devred I, Jeanne B, Reynaert A, Rzepka-Cuny MA. 2010. Évolution spatiale et temporelle des concentrations en Éléments Traces Métalliques de thalles de *Xanthoria parietina* récoltés sur la zone industrialo-portuaire de Dunkerque. *Air Pur*, 77 : 43-49.

Declercq C, Gower-Rousseau C, Vernier-Massouille G, Salleron J, Baldé M *et al.* 2010. Mapping of inflammatory bowel disease in northern France: spatial variations and relation to affluence. *Inflamm. Bowel Dis.*, 16 (5): 807-12.

Decocq C. 2010. Friches industrielles et pollutions historiques. Lille Métropole Communauté Urbaine, 98p.

Delzenne-Van Haluwyn C. 1973. Contribution à l'étude de la distribution des lichens épiphytes dans le nord de la France : application au problème de la pollution atmosphérique. Thèse de Doctorat Faculté de Pharmacie Lille2. 162p.

DREAL. 2014a. L'environnement en Nord – Pas de Calais – Tome 1 Etat des lieux. 78p

- DREAL. 2014b. Plan de prévention de l'atmosphère Nord – Pas de Calais. 252p.
- Génin M, Duhamel A, Preda C, Fumery M, Savoye G *et al.* 2013. Space-time clusters of Crohn's disease in northern France. *J Public Health*, DOI 10.1007/s10389-013-0580-9.
- Gower-Rousseau C, Salomez JL, Dupas JL, Marti R, Nuttens M-C *et al.* 1994. Incidence of inflammatory bowel disease in northern France (1988-1990). *Gut.*, 35 (10): 1433-8.
- Hardouin M. 1983. Les lichens épiphytes de la région Sud-Est de Lille: généralités, distribution et estimation de la pollution atmosphérique acide. Thèse de Doctorat en Pharmacie. 137p.
- Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. 2008. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science and Medicine*, 67(12) : 2007-2016.
- INSEE. 2013. Le développement durable en Nord – Pas de Calais. 180p.
- Jany-Catrice F & Zotti R. 2008. Les régions françaises face à leur santé sociale, contribution au débat. Institut pour le développement de l'information économique et sociale.
- Krivoruchko K. 2012. Empirical Bayesian Kriging implemented in ArcGIS Geostatistical Analyst. *ArcUser*, 15 (4): 6-10.
- Kulldorff M. 1997. A spatial scan statistic. *Communication in Statistics - Theory and Methods*, 26: 1481–1496.
- Kulldorff M. 1999. An isotonic spatial scan statistic for geographical disease surveillance. *Journal of the National Institute of Public Health*, 48: 94–101.
- Kulldorff M. 2010. SaTScan User Guide for version 9.0. 109p.
- Nisse C. 2010. Imprégnation par des métaux de la population générale. Valeurs d'imprégnation par les métaux de la population générale du Nord - Pas de Calais : établissement de valeurs de base pour l'interprétation des indicateurs d'exposition en milieu professionnel. Rapport final, 109p.
- ORS Nord – Pas de Calais. 2011. Indice de défaveur sociale de Townsend par commune en 2006. Disponible à partir de www.orsnpdc.org.
- PRSE2. 2011. Plan Régional Santé Environnement Nord – Pas de Calais 2011-2014. 63p.
- Sargaonkar A & Deshpande V. 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environmental Monitoring and Assessment*, 89: 43-67.
- Scerbo R, Possenti L, Lampugnani L, Ristori T, Barale R, Barghigiani C. 1999. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Livorno Province (Tuscany, Italy). *The Science of the Total Environment*, 241: 91-106.
- Scerbo R, Ristori T, Possenti L, Lampugnani L, Barale R, Barghigiani C. 2002. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Pisa Province (Tuscany, Italy). *The Science of the Total Environment*, 286: 27-40.
- Townsend P. 1987. Deprivation. *Journal of Social Policy*, 16 (2) : 125-146.
- Van Haluwyn C. 1994. Niveaux de la qualité de l'air sur le littoral Calais-Dunkerque ; évaluation en fonction de la diversité lichénique. Rapport SPPPI du Littoral : 57-74.

Van Haluwyn C & Cottard S. 1995. Cartographie de la qualité de l'air de la communauté urbaine de Lille à partir de l'observation des lichens. Rapport AREMA LM. 14p.

Van Haluwyn C & Cuny D. 1997. Cartographie de la pollution de l'air par certains métaux lourds sur le littoral Calais-Dunkerque. Rapport SPPPI du Littoral: 25-40.

Sites internet :

ADES : www.ades.eaufrance.fr (consulté le 22.04.2014).

Agreste : www.agreste.agriculture.gouv.fr (consulté le 22.04.2014).

ATMO Nord – Pas de Calais : www.atmo-npdc.fr (consulté le 22.04.2014).

BASIAS : www.basias.brgm.fr (consulté le 22.04.2014).

BASOL : www.basol.developpement-durable.gouv.fr (consulté le 22.04.2014).

INDIQUASOL : www.indiquasol.gissol.fr (consulté le 23.04.2014).

InfoTerre: www.infoterre.brgm.fr (consulté le 27.01.2014).

INSEE : www.insee.fr (consulté le 22/04/2014).

MEDDE : www.installationsclassees.developpement-durable.gouv.fr (consulté le 22.04.2014).

ORS Nord – Pas de Calais. www.orsnpdc.org (consulté le 22.04.2014).

CHAPITRE I. INEGALITES ENVIRONNEMENTALES ET SOCIALES DE SANTE A FINE ECHELLE EN REGION NORD – PAS DE CALAIS

Partie 1. Introduction

Ce chapitre présente les résultats de l'analyse spatiale d'inégalités environnementales et sociales de santé à fine échelle, menée au niveau de trois bassins de vie de la région Nord – Pas de Calais : Dunkerque, Lille et Maubeuge. Ces trois bassins présentent des configurations différentes. Le bassin de Dunkerque est fortement industrialisé, le bassin de Lille s'inscrit dans un contexte majoritairement urbain, tandis que le bassin de Maubeuge est plutôt rural. Dans le cadre de ces travaux de recherche, nous nous intéressons au compartiment atmosphérique et plus particulièrement à la biosurveillance des ETM de l'air à l'aide des lichens épiphytes.

La première partie de ce chapitre est consacrée à l'analyse des variations spatiales des niveaux d'imprégnation de l'environnement par les ETM, mesurés chez les lichens. Elle met notamment en œuvre la cartographie géostatistique à l'aide du krigeage empirique bayésien. Cette analyse a fait l'objet d'un article paru dans la revue *Pollution Atmosphérique*, intitulé « Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques. Vers une nouvelle utilisation des données de biosurveillance lichénique » (Ocelli *et al.*, 2013). La deuxième partie s'intéresse aux inégalités environnementales, c'est-à-dire à la présence de populations défavorisées à proximité d'un environnement dégradé, estimées au niveau des 3 bassins de vie à l'échelle de l'Iris. A ce titre, les cartographies lichéniques ont été confrontées à un indicateur de défaveur sociale⁷. Un article intitulé « Inégalités environnementales à l'échelle des quartiers du bassin industriel de Dunkerque : apports de la biosurveillance lichénique de la qualité de l'air » sera prochainement soumis pour publication dans la revue *Environmental Health Perspectives*. Dans la troisième partie, nous initiions une analyse des inégalités environnementales de santé, c'est-à-dire la présence de populations présentant un état de santé dégradé et vivant dans un environnement dégradé, par la mise en relation des mesures d'imprégnation par les métaux de l'environnement avec celles des populations (dosages sanguins et urinaires)⁸.

Ainsi dans ce chapitre nous cherchons à mettre en évidence **si les populations les plus défavorisées sur un plan socioéconomique vivent dans un environnement fortement contaminé par les ETM et si ces populations sont également les plus imprégnées par ces mêmes éléments.**

Pour ce travail, nous avons développé un indice intégré, permettant de caractériser le niveau d'imprégnation multimétallique du milieu et des populations. Il s'agit du ratio d'imprégnation moyen (RIM). Le RIM est dérivé de l'indice OIP (Overall Index of Pollution) développé par Sargaonkar & Deshpande (2003) pour la surveillance des rivières en Inde, dont la méthode de calcul a été reprise par Chen *et al.* (2005) avec l'IPI (Integrated Pollution Index) pour évaluer la qualité des sols de surface en Chine. Le ratio d'imprégnation (P_i) est dans un premier temps calculé pour chaque ETM i , par le rapport de la concentration mesurée (C_i) sur la concentration de référence (B_i). Le RIM est la moyenne arithmétique des n P_i . Un RIM supérieur à 1 indique donc une contamination du milieu au regard des valeurs de références.

$$P_i = \frac{C_i}{B_i} \qquad RIM = \sum_{i=1}^n \frac{P_i}{n}$$

⁷ L'indice de Défaveur Localisé (IDL), obtenu dans le cadre d'un partenariat avec le laboratoire TVES de l'ULCO (Dr. Anne Peggy Hellequin et Rachel Bavdek, EA4477).

⁸ Données obtenues dans le cadre d'un partenariat avec le CERESTE du CHRU de Lille (Dr. Catherine Nisse et Dr. Ariane Leroyer, EA4483, Université Lille 2).

Pour la matrice environnementale, cet indice d'enrichissement est calculé sur la base des concentrations naturelles retrouvées dans les lichens. Pour le sang et l'urine, les valeurs de référence utilisées sont celles définies par l'étude régionale IMEPOGE (Nisse, 2010). L'intérêt d'utiliser de tels indices intégrés sera repris lors de la discussion générale.

Par ailleurs, les outils SIG ont été mis en œuvre pour permettre la mise en relation des données de sources et de natures diverses. Dans la première partie le krigeage empirique bayésien fut utilisé pour estimer les valeurs d'imprégnation de l'environnement en des lieux non échantillonnés. Dans la deuxième partie, les cartographies géostatistiques résultantes ont été transformées en information agrégée à l'unité géographique (l'Iris), dans le but de les confronter à l'indice de défaveur sociale. Les valeurs des concentrations dans les lichens ont pour cela été modélisées au niveau d'un maillage fin et régulier établi sur les zones d'études, puis moyennées à l'intérieur de chaque Iris. Dans la troisième partie, les géostatistiques ont permis d'estimer les valeurs environnementales au niveau des coordonnées spatiales des individus. En effet, les lieux d'habitation des sujets intégrés à l'étude ne correspondaient pas aux points d'échantillonnage des lichens.

Dans chacune de ces parties, les résultats de nos recherches sont présentés sous la forme d'un article, précédé d'une introduction détaillée. Ce chapitre s'achèvera sur une discussion générale, suivie des perspectives.

Partie 2. Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques

a. Résumé de l'article

La biosurveillance peut être employée, en complément de la surveillance physico-chimique, pour mettre en évidence les impacts biologiques de la pollution atmosphérique. Dans cette première partie, nous nous intéressons aux capacités bioaccumulatrices des lichens. Ils sont en effet capables d'accumuler sur le long terme certains polluants présents dans l'air, comme les métaux. De fait, la mesure de leur imprégnation permet de suivre l'évolution spatiale et temporelle de la contamination.

Les lichens ont été employés comme marqueurs des niveaux d'imprégnation de l'environnement atmosphérique par les ETM. Des dosages de 18 métaux ont été réalisés dans des thalles du lichen *Xanthoria parietina*, prélevés au niveau de trois bassins de vie de la région Nord – Pas de Calais. Afin d'évaluer le niveau d'imprégnation de l'environnement par ces éléments, nous avons comparé les concentrations obtenues à des concentrations naturelles appelées bruit de fond. A partir des 220 mailles d'échantillonnage (60 mailles pour Dunkerque, 120 mailles pour Lille et 40 mailles pour Maubeuge), des cartographies ponctuelles et géostatistiques des niveaux d'imprégnation ont été réalisées pour chacun des éléments mesurés. Ces cartographies ont été générées à l'aide du logiciel ArcGIS® (version 10.1). La méthode géostatistique employée est le krigeage empirique bayésien. En complément des cartographies monométalliques, le niveau de contamination globale du milieu fut évalué à l'aide d'un indice intégré : le ratio d'imprégnation moyen.

Les 18 cartographies générées montrent des variations spatiales des niveaux d'imprégnation, non seulement entre les bassins, mais aussi au sein d'un même bassin. Une structure spatiale similaire est observée sur ces cartographies, avec des charges plus importantes sur le bassin de Dunkerque, puis le bassin lillois et enfin des charges plus faibles sur le bassin de Maubeuge. La cartographie du RIM montre des niveaux d'imprégnation élevés sur l'ensemble des trois bassins. Les zones les plus contaminées sont situées à proximité des sources d'émissions industrielles, routières et urbaines. Enfin, les données de biosurveillance lichénique peuvent être représentées de manière agrégée, dans notre cas à l'Iris, afin de les rapprocher des indicateurs sanitaires et socioéconomiques.

Les résultats de ce travail ont fait l'objet d'un article récemment publié par la revue *Pollution Atmosphérique* dans son numéro 220 (Octobre – Décembre 2013) et intégré à la suite de ce résumé. Ils concernent également une communication orale réalisée lors de la conférence internationale des SIG 2013, organisée par ESRI France à Versailles en octobre 2013.

A la suite de l'article, une note apporte un éclairage complémentaire sur l'indice intégré d'imprégnation multimétallique considéré dans ce travail.

b. Article

TITRE DE L'ARTICLE

Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques. Vers une nouvelle utilisation des données de biosurveillance lichénique.

NOM DES AUTEURS ET AFFILIATIONS

Florent Occelli, E.A. 4483 « Impacts de l'environnement chimique sur la santé humaine », Laboratoire des Sciences Végétales et Fongiques (LSVF), faculté des sciences pharmaceutiques et biologiques, 3 rue du Professeur Laguesse, B.P. 83, 59006 Lille cedex.

Marie-Amélie Cuny, Association pour la Prévention de la Pollution Atmosphérique, comité régional Nord-Pas-de-Calais, 235 avenue de la Recherche, B.P. 86, 59373 Loos cedex, France.

Inès Devred, E.A. 4483 « Impacts de l'environnement chimique sur la santé humaine », Laboratoire des Sciences Végétales et Fongiques (LSVF), faculté des sciences pharmaceutiques et biologiques, 3 rue du Professeur Laguesse, B.P. 83, 59006 Lille cedex.

Annabelle Deram, E.A. 4483 « Impacts de l'environnement chimique sur la santé humaine », Laboratoire des Sciences Végétales et Fongiques (LSVF), faculté des sciences pharmaceutiques et biologiques, 3 rue du Professeur Laguesse, B.P. 83, 59006 Lille cedex. Faculté d'ingénierie et de management de la santé, 42 rue Ambroise Paré, 59120 Loos.

Stéphanie Quarré, E.A. 4483 « Impacts de l'environnement chimique sur la santé humaine », Laboratoire des Sciences Végétales et Fongiques (LSVF), faculté des sciences pharmaceutiques et biologiques, 3 rue du Professeur Laguesse, B.P. 83, 59006 Lille cedex.

Damien Cuny, E.A. 4483 « Impacts de l'environnement chimique sur la santé humaine », Laboratoire des Sciences Végétales et Fongiques (LSVF), faculté des sciences pharmaceutiques et biologiques, 3 rue du Professeur Laguesse, B.P. 83, 59006 Lille cedex.

MOTS-CLES

Biosurveillance, lichens, métaux, krigeage, santé-environnement

KEYWORDS

Biomonitoring, lichens, metals, kriging, environment and health

REFERENCE ELECTRONIQUE

Florent Occelli, Marie-Amélie Cuny, Inès Devred, Annabelle Deram, Stéphanie Quarré et Damien Cuny « Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques », Pollution atmosphérique [En ligne], N° 220, mis à jour le : 13/01/2014, URL : <http://lodel.irevues.inist.fr/pollution-atmospherique/index.php?id=2497>.

LISTE DES ILLUSTRATIONS

Figure 1. Présentation générale, répartition des centres des mailles de récoltes et caractéristiques des trois bassins de vie. *General presentation, distribution of the center of the harvest areas and characteristics of the three living areas.*

Figure 2. *Xanthoria parietina* (photo : APPA Nord-Pas-de-Calais). *Xanthoria parietina* lichen.

Figure 3. Cartographie du ratio d'imprégnation par le plomb. Dans les zones représentées en bleu, les lichens ont une imprégnation proche des références. *Map of lichens impregnation ratio for the lead. In areas shown in blue, lichens do not differ from references.*

Figure 4. Cartographie des ratios d'imprégnation moyens (RIM) sur les trois bassins étudiés. *Mapping of mean impregnation ratio (RIM) on the three studied areas.*

Figure 5. Cartographie des RIM au niveau des IRIS. *RIM mapping in IRIS areas.*

LISTE DES TABLES

Tableau I. Valeurs de référence (en mg/kg poids sec) des concentrations en éléments traces dans les lichens, et statistiques descriptives des concentrations (en mg/kg poids sec) des différents éléments traces étudiés dans les thalles de lichens récoltés sur les trois bassins de vie. Les résultats des comparaisons statistiques des concentrations sont présentés dans les dernières colonnes. « Oui » indique une différence statistiquement significative ($p < 0.05$) entre les territoires désignés. « - » indique une absence de différence. *Reference values (mg/kg dry weight) for trace elements concentrations in lichens and descriptive statistics of trace elements concentrations (mg/kg dry weight) in lichen thalli. The results of statistical comparisons of concentrations are shown in the last column. Yes indicates a statistically significant difference ($p < 0.05$) between the designated territories. “ - ” indicates no difference.*

RESUME

Du fait de leur longue persistance, les Éléments Traces Métalliques (ETM) restent une préoccupation majeure en termes de contamination de l'environnement, de risques sanitaires et environnementaux, y compris dans un contexte de diminution des émissions. Les données de contamination de l'environnement restent très partielles et le suivi des immissions insuffisamment étendu pour avoir une image précise de la situation. Dans ce travail, nous avons donc utilisé les lichens épiphytes répartis dans trois bassins de vie de la région (de typologies bien différentes), comme bioaccumulateurs dans lesquels nous avons dosé les concentrations totales en 18 ETM. Nous avons également mené plusieurs analyses des données et interpolations spatiales afin d'obtenir un indicateur non plus ponctuel mais agrégé des données de biosurveillance afin que celles-ci puissent être associées à des données socioéconomiques, sanitaires et épidémiologiques. Ce travail précurseur permettra de progresser dans la connaissance des inégalités et des points noirs environnementaux.

ABSTRACT

Because of their long persistence, trace metals remain an issue in terms of contamination of the environment, health and environmental risks, including in the context of reducing emissions. Data from environmental contamination are very partial and air-monitoring network insufficiently expanded to have a clear image of the situation. In this work we therefore used epiphytic lichens as bio-accumulators, across three catchment areas in the region (with very different types) in which we measured the total concentrations of 18 ETM. We also conducted several analyses of the data and spatial interpolation to obtain an aggregate indicator of biomonitoring data so that they can be associated with these health and epidemiological data. This pioneering work will advance the understanding of environmental inequalities and blackheads.

INTRODUCTION

Les Éléments Traces Métalliques (ETM) sont essentiellement émis lors de la mise en œuvre des procédés industriels (fonderies, usines sidérurgiques et métallurgiques...), mais aussi lors des combustions de pétrole et de charbon, ou encore des déchets ménagers. Ainsi, le secteur résidentiel tertiaire et le trafic routier (rejets et usure des véhicules et des infrastructures) contribuent également aux émissions. Enfin, des émissions naturelles (éruptions volcaniques, feu de forêt, envol de poussières des sols contaminés) viennent enrichir ce panel de sources. Dans le compartiment atmosphérique, la plupart de ces éléments sont véhiculés par les particules dont les plus fines pénètrent profondément dans l'appareil respiratoire (Marano, 2012). Même si leurs émissions ont fortement baissé ces 20 dernières années, ils se sont largement accumulés dans de nombreux compartiments de notre environnement. En outre, les ETM peuvent être transférés au travers des réseaux trophiques. Ils constituent ainsi encore un risque pour l'homme et pour l'environnement (Cuny, 2009 ; Davranche *et al.*, 2013, Liu *et al.*, 2013). De nombreuses études ont montré l'hétérogénéité des cibles des ETM dans l'organisme humain, ainsi que leur toxicité (atteintes neurologiques, cardio-vasculaires, respiratoires, rénales...) voire leur cancérogénicité (cadmium, chrome VI, arsenic) (IARC, 1993, 1997 ; Chen et Lippmann, 2009 ; Cooksey, 2012).

L'ensemble des données montre que le Nord-Pas-de-Calais se classe fréquemment en tête des régions concernant les émissions de métaux, exprimées en grammes par habitants comme en grammes par hectares (CITEPA, 2012 ; DREAL 2012). De plus, il persiste de nombreuses zones où les émissions, les immissions (mesurées sur l'ensemble de la région grâce à 8 stations, principalement urbaines ou de proximité industrielle, de l'Association Agréée de Surveillance de la Qualité de l'Air ATMO Nord-Pas-de-Calais) et les concentrations résiduelles dans l'environnement ne sont que peu voire pas connues. Enfin, il existe une grande disparité régionale dans la répartition des sources, avec comme bassin industriel le plus important, la zone industrialo-portuaire de Dunkerque. Cette hétérogénéité territoriale dans les émissions et les immissions génère des différences d'exposition des populations aux ETM. Compte tenu de cette situation, nous avons décidé de mener des investigations sur la contamination de l'environnement par les ETM en utilisant des lichens en tant que bioaccumulateurs. Dans son article 4 au paragraphe 10, la directive européenne 2004/7 (relative à la surveillance de l'arsenic, du cadmium, du mercure, du nickel et des HAP) stipule que « l'utilisation des bioindicateurs peut être considérée là où les impacts sur les écosystèmes doivent être évalués ». En effet, les bioaccumulateurs permettent de caractériser l'imprégnation de l'environnement, résultant de l'équilibre dynamique entre l'air et l'organisme utilisé, intégrant en permanence de nombreux facteurs, comme les paramètres climatiques, l'augmentation de la biomasse ou l'éventuelle métabolisation du polluant... Ainsi, les concentrations résultantes mesurées sont « les traces » laissées dans l'environnement et non directement les concentrations atmosphériques. C'est en cela que la biosurveillance est une méthode complémentaire aux analyses physico-chimiques atmosphériques, qui permet de mettre en évidence l'imprégnation de l'environnement par les polluants tels que les ETM ainsi que leurs effets (Cuny *et al.*, 2008). Les lichens, en tant que symbiose entre une algue et un champignon sont particulièrement indiqués pour le suivi des ETM, comme le témoigne la littérature scientifique très abondante à ce sujet (Loppi *et al.*, 2004 ; Rzepka et Cuny, 2009 ; Kar *et al.*, 2013 ; Majumder *et al.*, 2013).

Les résultats issus de la biosurveillance permettent, d'une manière large, de caractériser la contamination de l'environnement extérieur dans lequel des populations vivent. Néanmoins, il est possible d'affiner les analyses en associant ces résultats avec des données sanitaires ou socio-économiques. Cependant, ces dernières sont très fréquemment rapportées pour une aire géographique donnée alors que les données de biosurveillance sont ponctuelles (résultats obtenus sur chaque station). Il en résulte donc une adaptation nécessaire des données « de terrain » dans le but de progresser dans la connaissance des liens complexes entre environnement et santé. L'IRIS (Ilots Regroupés pour l'Information Statistique) étant l'échelle la plus fréquemment utilisée pour les paramètres socio-économiques et sanitaires, il convient donc d'analyser les données environnementales sur cette base.

Les objectifs de ce travail étaient donc (1) d'obtenir des données de contamination de l'environnement par les métaux dans trois bassins de vie caractérisés par des activités humaines différentes et (2) de traiter ces données afin qu'elles puissent, par la suite, être rapprochées de données sanitaires et socio-économiques à l'échelle de l'IRIS.

MATERIELS ET METHODES

Les 3 bassins de vie choisis dans la Région Nord-Pas-de-Calais sont : la Communauté Urbaine de Dunkerque (CUD), la Communauté Urbaine De Lille (CUDL) et la Communauté de Commune du Val de Sambre (CCVS). Les principales caractéristiques de ces trois bassins sont reprises dans la figure ci-dessous (Figure 1). Pour des raisons de calculs géostatistiques et de représentation cartographique, nous avons étendu la zone de la CCVS au-delà des 67 communes incluses. Néanmoins, pour des raisons d'homogénéité de présentation, nous conserverons cette dénomination.

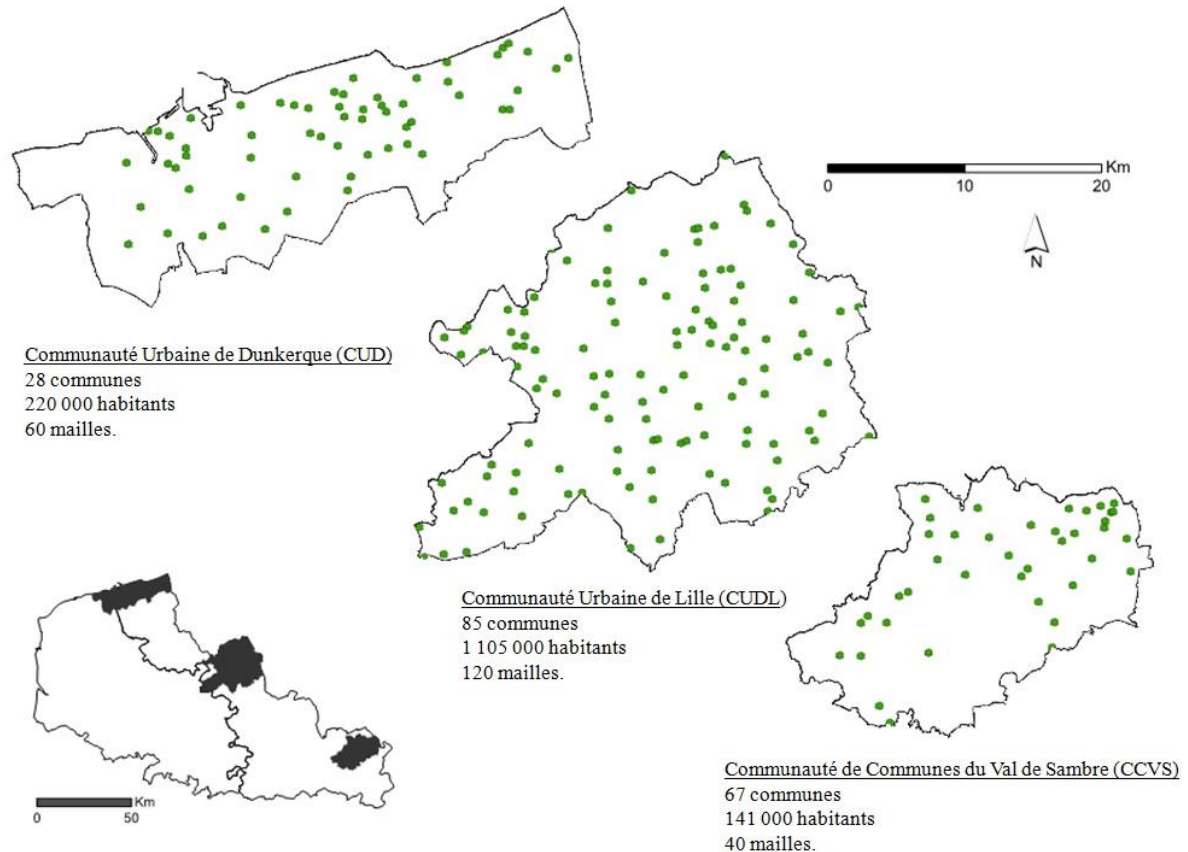


Figure 1. Présentation générale, répartition des centres des mailles de récoltes et caractéristiques des trois bassins de vie.

General presentation, distribution of the center of the harvest areas and characteristics of the three living areas.

Nous basons notre travail sur l'hypothèse qu'il existe un gradient de contamination ouest-est avec des concentrations plus importantes en ETM dans la CUD possédant une importante activité industrielle. Celle-ci est moins présente sur l'agglomération lilloise qui se caractérise d'avantage par un habitat dense et de très nombreuses voies de circulation importantes. La CCVS, où le milieu rural est dominant, présenterait une imprégnation métallique moins importante. Ces hypothèses sont fondées sur la présence des sources. Notre travail consiste à confirmer que les imprégnations à long terme de l'environnement de ces trois bassins correspondent à ces hypothèses.

Nous avons utilisé des thalles du lichen épiphyte *Xanthoria parietina*. C'est une espèce nitrophile très courante dans la région. Son thalle est foliacé, de couleur orange assez vif. Cette espèce a un rythme de croissance d'environ 1 cm par an. Nous avons récolté des thalles de 6 cm de diamètre (ce qui correspond donc à une exposition sur les 6 dernières années environ).



Figure 2. *Xanthoria parietina* (photo : APPA Nord-Pas-de-Calais).
Xanthoria parietina lichen.

Dans un même bassin, les lichens ont été récoltés le même jour dans chacune des mailles à l'aide d'outils en céramique (60 mailles pour la CUD, 120 mailles dans la CUDL et 40 mailles dans la CCVS). Pour pouvoir réaliser cela, chaque maille de récolte (1 km x 1 km) a fait l'objet d'un repérage préliminaire précis.

Au sein de chacune des mailles, 6 à 10 thalles ont été récoltés, en 3 lieux différents géoréférencés, par maille, sur des troncs à au moins 1,5 m du sol et placés dans des piluliers numérotés.

Au laboratoire, les thalles ont été séchés à l'étuve (35 °C), soigneusement débarrassés des débris d'écorce et ensuite broyés dans un mortier en céramique. Tous les thalles d'une même maille ont par la suite été rassemblés en 1 seul échantillon. Il y a donc eu au total 220 échantillons dosés.

Les éléments traces recherchés dans les thalles de lichens étaient les suivants : l'aluminium (Al), l'antimoine (Sb), l'arsenic (As), le béryllium (Be), le cadmium (Cd), le cobalt (Co), le chrome (Cr), le cuivre (Cu), le manganèse (Mn), le mercure (Hg), le nickel (Ni), le plomb (Pb), le palladium (Pd), le platine (Pt), le rhodium (Rh), le titane (Ti), le vanadium (V), le zinc (Zn).

Ce sont les concentrations totales en ces éléments qui ont été déterminées, c'est-à-dire sans tenir compte de leurs différentes spéciations au sein des thalles.

L'aluminium, le manganèse, le titane et le zinc ont été dosés par ICP-AES (spectroscopie d'émission optique avec plasma induit par haute fréquence). Les autres éléments ont été dosés par ICP-MS (spectrométrie de masse avec plasma à couplage inductif). L'échantillon européen de référence (numéro d'identification 482 – lichen *Pseudevernia furfuracea* contaminé par des métaux) a été utilisé dans le contrôle qualité des dosages.

Les statistiques descriptives des résultats et les différentes analyses ont été réalisées à l'aide du logiciel XLSTAT (Addinsoft – Paris, France). Les comparaisons entre les différents territoires ont été réalisées à l'aide d'un test non paramétrique de Dunn ($p < 0.005$). Les coefficients de corrélation entre les ETM ont été calculés, et la significativité des différences évaluée grâce à un test de Spearman ($p < 0,05$).

Les cartographies et les différentes interpolations géostatistiques (Krigage empirique Bayésien - EBK) ont été réalisées grâce au logiciel Arcview (Esri – Meudon, France).

Les résultats obtenus ont permis de constituer une base de données de contamination lichénique et de réaliser 18 cartes de la contamination métallique des environnements étudiés. En plus de ces données monométalliques, nous avons développé un indicateur composite de contamination des milieux à partir des travaux de Tong-Bin *et al.* (2005). Pour chaque élément dosé en un point, nous avons calculé le ratio de la concentration mesurée sur la valeur de référence. Pour chaque point, nous avons calculé la moyenne de ces ratios, obtenant ainsi un Ratio d'Imprégnation Moyen (RIM).

La réalisation de ce RIM nous a également permis de calculer les valeurs pour chaque IRIS, unité géographique prise comme référence pour les données socio-économiques. Ce traitement a, pour chaque bassin, nécessité la constitution d'un maillage très fin (100 m x 100 m). Une valeur de RIM a été calculée pour chacune de ces petites mailles à partir des résultats. Les limites des IRIS sont ensuite intégrées (source : IGN). Les RIM des petites mailles appartenant à un même IRIS ont ensuite permis le calcul d'un RIM spécifique à cet IRIS.

RESULTATS

Il n'existe pas de normes établies de concentration en éléments traces dans les lichens. Pour interpréter les résultats obtenus pour chaque élément trace, nous utilisons comme valeurs de référence, des gammes de concentration retrouvées dans la bibliographie, notamment les travaux de Van Haluwyn et Cuny (1997), Bargagli (1998), et Scerbo *et al.* (1999, 2002), Cuny *et al.* (2004, 2010). Ces valeurs sont présentées dans le tableau I ci-après. Ce tableau présente également les statistiques descriptives réalisées sur les valeurs des concentrations des différents éléments traces dosés dans les thalles ainsi que les résultats des tests de comparaison. Ces résultats sont globalisés pour chaque bassin de vie afin de pouvoir établir les comparaisons. Nous observons pour la quasi-totalité des éléments des coefficients de variation importants, ce qui traduit des grands écarts de concentrations et donc l'hétérogénéité de la contamination de l'environnement.

Statistiques descriptives :		Territoires	Minimum	Maximum	Moyenne	Ecart-type	Résultat des tests de Dunn	
Eléments	Ref. (mg.lg poids sec)		(Concentrations en mg/kg poids sec)				CCVS	CUDL
Al	100 à 500	CCVS	161,0	1933,0	730,7	424,8	CCVS	-
		CUDL	221,0	8745,7	1092,9	1175,6	CUDL	-
		CUD	349,0	6681,7	1504,0	1227,8	CUD	Oui
As	0,1 à 0,5	CCVS	0,2	2,4	0,8	0,5	CCVS	-
		CUDL	0,2	5,1	1,1	0,8	CUDL	-
		CUD	0,4	9,0	2,1	1,7	CUD	Oui
Be	0,1 à 0,3	CCVS	0,3	0,3	0,3	0,0	CCVS	-
		CUDL	0,0	0,3	0,2	0,1	CUDL	-
		CUD	0,0	0,4	0,2	0,1	CUD	-
Cd	0,1 à 0,2	CCVS	0,1	0,8	0,3	0,2	CCVS	-
		CUDL	0,1	10,6	0,7	1,0	CUDL	Oui
		CUD	0,2	30,1	1,4	3,9	CUD	Oui
Co	3 à 7	CCVS	0,3	1,4	0,6	0,3	CCVS	-
		CUDL	0,2	9,6	0,7	0,9	CUDL	-
		CUD	0,2	13,1	1,5	2,1	CUD	Oui
Cr	1 à 4	CCVS	0,9	11,0	3,8	2,4	CCVS	-
		CUDL	1,0	45,1	4,6	4,7	CUDL	-
		CUD	2,6	1172,4	38,3	153,0	CUD	Oui
Cu	5 à 15	CCVS	3,9	22,1	9,5	4,5	CCVS	-
		CUDL	4,7	259,2	19,1	26,2	CUDL	Oui
		CUD	5,1	95,5	19,0	17,5	CUD	Oui
Hg	0,1 à 0,2	CCVS	0,1	0,1	0,1	0,0	CCVS	-
		CUDL	0,0	0,8	0,1	0,1	CUDL	Oui
		CUD	0,0	1,1	0,2	0,2	CUD	Oui
Mn	5 à 30	CCVS	20,0	165,0	57,3	31,9	CCVS	-
		CUDL	16,0	309,0	54,0	42,3	CUDL	-
		CUD	45,4	6305,0	556,9	912,5	CUD	Oui
Ni	3 à 10	CCVS	0,8	8,7	2,6	1,5	CCVS	-
		CUDL	0,7	72,9	3,2	6,4	CUDL	-
		CUD	1,2	794,5	24,2	102,3	CUD	Oui
Pb	8 à 15	CCVS	2,4	55,0	11,5	9,0	CCVS	-
		CUDL	1,3	873,0	31,1	77,9	CUDL	Oui
		CUD	3,0	371,9	32,7	50,6	CUD	Oui
Pd	15	CCVS	15,1	15,8	15,4	0,2	CCVS	-
		CUDL	15,1	231,9	21,7	21,6	CUDL	-
		CUD	15,2	61,2	21,0	11,3	CUD	-
Pt	15	CCVS	15,1	15,8	15,4	0,2	CCVS	-
		CUDL	15,0	15,9	15,4	0,2	CUDL	-
		CUD	15,2	15,7	15,4	0,1	CUD	-
Rh	15	CCVS	15,1	15,8	15,4	0,2	CCVS	-
		CUDL	15,0	58,2	15,8	4,0	CUDL	-
		CUD	15,2	15,7	15,4	0,1	CUD	-
Sb	0,3	CCVS	0,3	2,1	0,6	0,4	CCVS	-
		CUDL	0,3	14,0	1,7	1,8	CUDL	Oui
		CUD	0,3	25,2	1,8	3,3	CUD	Oui
Ti	5	CCVS	5,1	38,0	15,5	6,6	CCVS	-
		CUDL	0,0	76,0	15,4	10,9	CUDL	-
		CUD	0,0	156,0	30,2	34,9	CUD	Oui
V	3 à 12	CCVS	0,7	8,2	3,0	1,6	CCVS	-
		CUDL	0,3	18,0	3,6	2,8	CUDL	-
		CUD	2,6	102,0	13,2	15,1	CUD	Oui
Zn	30 à 50	CCVS	23,3	267,4	74,4	41,5	CCVS	-
		CUDL	27,9	1582,7	112,8	141,5	CUDL	Oui
		CUD	40,4	2407,9	178,9	310,2	CUD	Oui

Tableau I. Valeurs de référence (en mg/kg poids sec) des concentrations en éléments traces dans les lichens, et statistiques descriptives des concentrations (en mg/kg poids sec) des différents éléments traces étudiés dans les thalles de lichens récoltés sur les trois bassins de vie. Les résultats des comparaisons statistiques des concentrations sont présentés dans les dernières colonnes. « Oui »

indique une différence statistiquement significative ($p < 0.05$) entre les territoires désignés. « - » indique une absence de différence.

Reference values (mg/kg dry weight) for trace elements concentrations in lichens and descriptive statistics of trace elements concentrations (mg/kg dry weight) in lichen thalli. The results of statistical comparisons of concentrations are shown in the last column. Yes indicates a statistically significant difference ($p < 0.05$) between the designated territories. " - " indicates no difference.

Notre hypothèse de départ est confirmée dans la grande majorité des cas. Nous observons des imprégnations plus importantes sur la CUD puis la CUDL et enfin sur la CCVS. Les variations en termes de sources (nature, activité, importance...) entre les territoires se traduisent par des différences significatives en ce qui concerne les imprégnations de l'environnement.

Il existe quelques exceptions, notamment entre la CUDL et la CCVS malgré des tendances marquées. La première concerne le béryllium pour lequel les résultats obtenus sont très homogènes. Cet élément est principalement utilisé en plasturgie (fabrication de moules pour matières plastiques), en métallurgie et en électronique sous forme d'oxyde de béryllium. En raison de sa forte toxicité, il a, dans de très nombreux cas, été substitué. Ainsi, sur les trois territoires que nous avons étudiés, le béryllium n'est pas présent d'une manière importante et il n'existe pas de différence d'imprégnation environnementale notable.

Les imprégnations en cuivre sur la CUDL et la CUD ne diffèrent pas significativement. Nous pouvons remarquer qu'elles sont plus importantes que les niveaux de référence. Ceci, compte tenu de l'activité métallurgique présente sur la CUD, peut paraître surprenant. En réalité, en France, le cuivre provient majoritairement de l'usure des caténaires induite par le trafic ferroviaire et de l'usure des plaquettes de frein des véhicules routiers (CITEPA, 2012). Ce sont donc principalement les transports qui sont émetteurs. À noter que l'utilisation de pesticides peut générer également des émissions de Cu. Ce dernier point pourrait contribuer à expliquer le fait que, sur la CCVS, les concentrations sont également bien au-dessus des valeurs de référence.

Un autre élément attire l'attention, il s'agit du plomb. En effet, celui-ci était au niveau français principalement émis par le trafic routier et cela jusque dans les années 2000 (le plomb dans l'essence a été interdit le 1^{er} janvier 2000 mais cette mesure n'a pas eu immédiatement des effets significatifs). Actuellement, les principales sources de cet élément sont la première et la seconde fusion du plomb, la fabrication de batteries électriques, la fabrication de certains verres (cristal). Il ne faut pas négliger les émissions du secteur résidentiel tertiaire qui représentent 16 % en France (CITEPA, 2012) et enfin, même si cela est plus difficilement quantifiable, la mise en suspension de particules provenant de sols contaminés (présents dans notre région). Sur la figure 3 ci-dessous, nous remarquons que les zones urbanisées sont systématiquement plus imprégnées (bien entendu cela est plus net sur la CUDL et la CUD). Dans la CUD, nous retrouvons également la zone industrialo-portuaire nettement marquée, ce qui correspond à des résultats précédemment décrits (Cuny *et al.*, 2004, 2010).

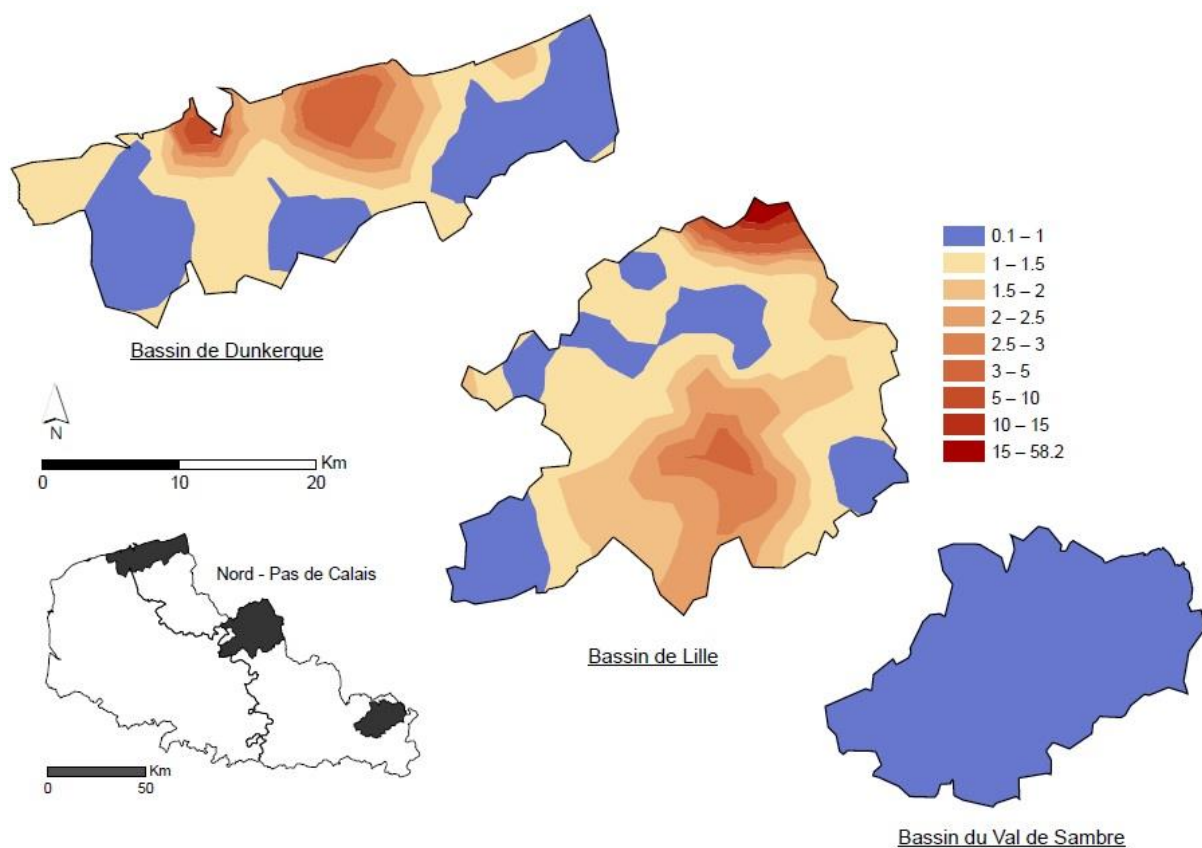


Figure 3. Cartographie du ratio d'imprégnation par le plomb. Dans les zones représentées en bleu, les lichens ont une imprégnation proche des références.

Map of lichens impregnation ratio for the lead. In areas shown in blue, lichens do not differ from references.

Dans nos travaux, l'imprégnation en antimoine n'est pas différente entre la CUD et la CUDL. Cet élément est très lié au plomb dans de nombreuses utilisations industrielles car il en augmente la dureté. Ainsi, dans des produits tels que les accumulateurs, on retrouve de l'antimoine lié au plomb. Dans nos résultats, ces deux éléments sont significativement corrélés sur ces deux zones, ce qui pourrait confirmer leurs origines communes.

Les résultats obtenus pour le titane sont parmi ceux qui soulèvent le plus de questions quant à leur interprétation. En effet, cet élément fait l'objet de très nombreuses applications mais il est aussi un élément naturel que l'on retrouve grâce à la mise en suspension des poussières des sols. Ainsi, il est possible que sur les différents territoires nous ayons à la fois de grandes disparités des concentrations mais aussi des sources qui font, qu'*in fine*, les résultats ne puissent se distinguer d'une manière significative.

Enfin, les imprégnations de l'environnement des trois bassins de vie par les platinoïdes (rhodium, palladium et platine) ne sont pas significatives hormis pour quelques points particuliers de la CUDL. Nous avons déjà obtenu des résultats comparables à l'aide d'une méthode de biosurveillance différente (Rzepka et Cuny, 2011). L'un des objectifs suivis ici était de mettre en évidence une éventuelle contribution du trafic routier car ces éléments sont susceptibles d'être émis dans l'environnement suite à l'usure des pots catalytiques. Les émissions de ces éléments sont assez faibles (de l'ordre du ng/km) ainsi, comme limite à ce présent travail, nous pourrions simplement évoquer des seuils de détection et de quantification des méthodes analytiques insuffisamment bas. Compte tenu des enjeux de ces éléments, leur biosurveillance environnementale nécessitera d'être développée.

Nous proposons deux approches agrégées. La première est la cartographie des RIM pour les trois bassins de vie (présentée figure 4), et la seconde la cartographie des RIM calculés et représentés pour chaque IRIS (figure 5).

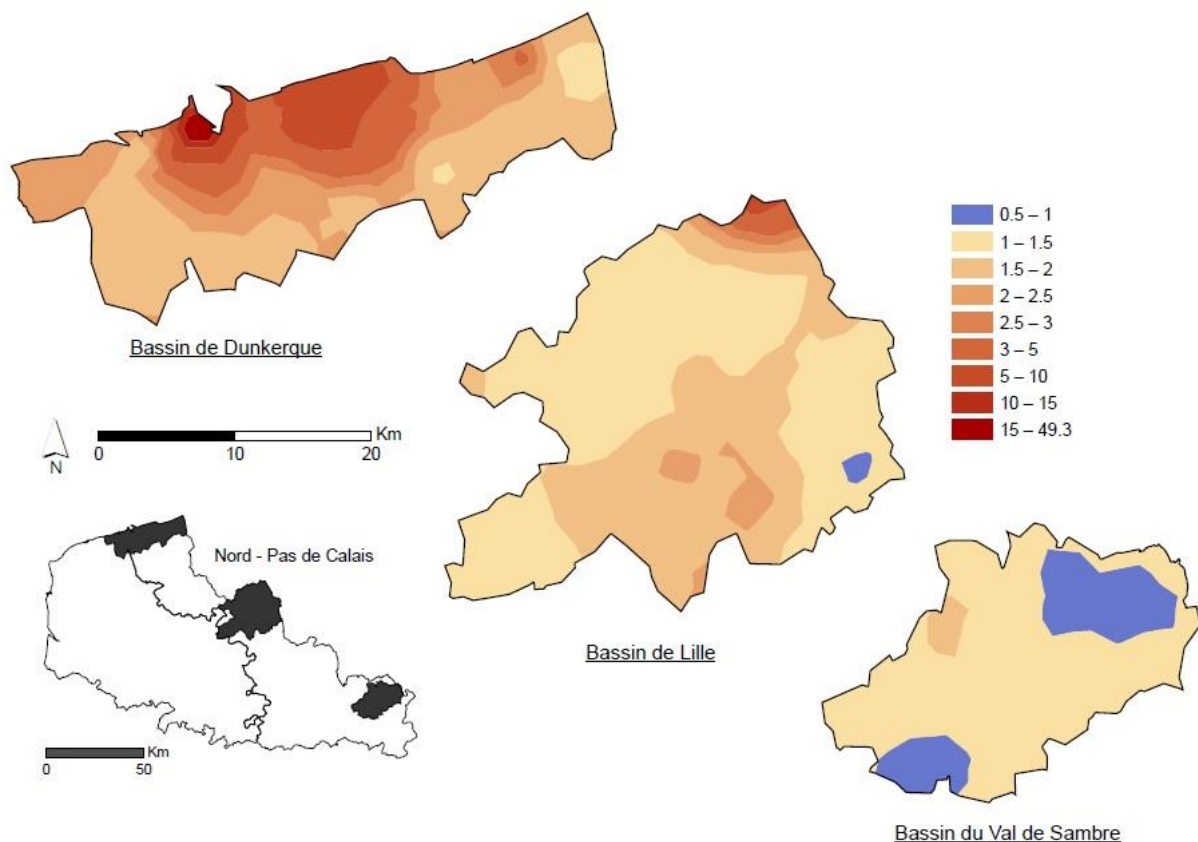


Figure 4. Cartographie des ratios d'imprégnation moyens (RIM) sur les trois bassins étudiés.
Mapping of mean impregnation ratio (RIM) on the three studied areas.

Nous y retrouvons la répartition décroissante précédemment décrite entre les bassins, de la CUD vers la CCVS. Il n'y a que de rares données ponctuelles qui présentent un RIM inférieur à 1, ce qui amène à une modélisation de l'ensemble de la zone avec des valeurs de RIM supérieures à 1. La déformation des zones de contamination vers le sud de la CUD, préalablement décrite par Cuny *et al.* (2004) et attribuée au transport des ETM par les vents (et principalement les brises de mer) perpendiculaires au littoral, se confirme. Les zones les plus contaminées correspondent aux deux zones industrialo-portuaires du secteur. Néanmoins, compte tenu de la proximité des zones urbaines, celles-ci intègrent également les secteurs pour lesquels les RIM sont les plus importants. Sur la CUDL, là aussi, la contamination est assez générale même si deux zones se distinguent plus particulièrement. La première au sud correspond à la fois à un axe majeur de transit et d'activités (l'autoroute A1 et toutes les zones d'activités aux alentours). On ne peut toutefois pas exclure que cette zone soit aussi sous l'influence des retombées du bassin minier situé juste au sud. De la même manière, à l'extrémité nord, il existe quelques activités émettrices d'ETM (industries et incinération notamment) mais nous ne pouvons exclure la contribution des retombées issues d'activités localisées de l'autre côté de la frontière belge.

Enfin, sur la CCVS, l'imprégnation là aussi est assez générale mais plus faible que sur les bassins précédents. Il est à noter que nous n'avons pas mis en évidence de contribution significative des zones les plus habitées et/ou industrielles telles que Jeumont ou Hautmont.

Nous remarquons au final qu'il existe une contamination de fond présente sur l'ensemble de la région car même dans la CCVS les zones avec des RIM supérieurs à 1 sont très largement répandues.

Ainsi, nous sommes face à une contamination assez générale de l'environnement par les éléments traces. Bien entendu, il est nécessaire de nuancer ce résultat car nos dosages intègrent la contribution du fond géochimique (fond naturel).

La figure 5 présente les RIM calculés et représentés à l'échelle des IRIS.

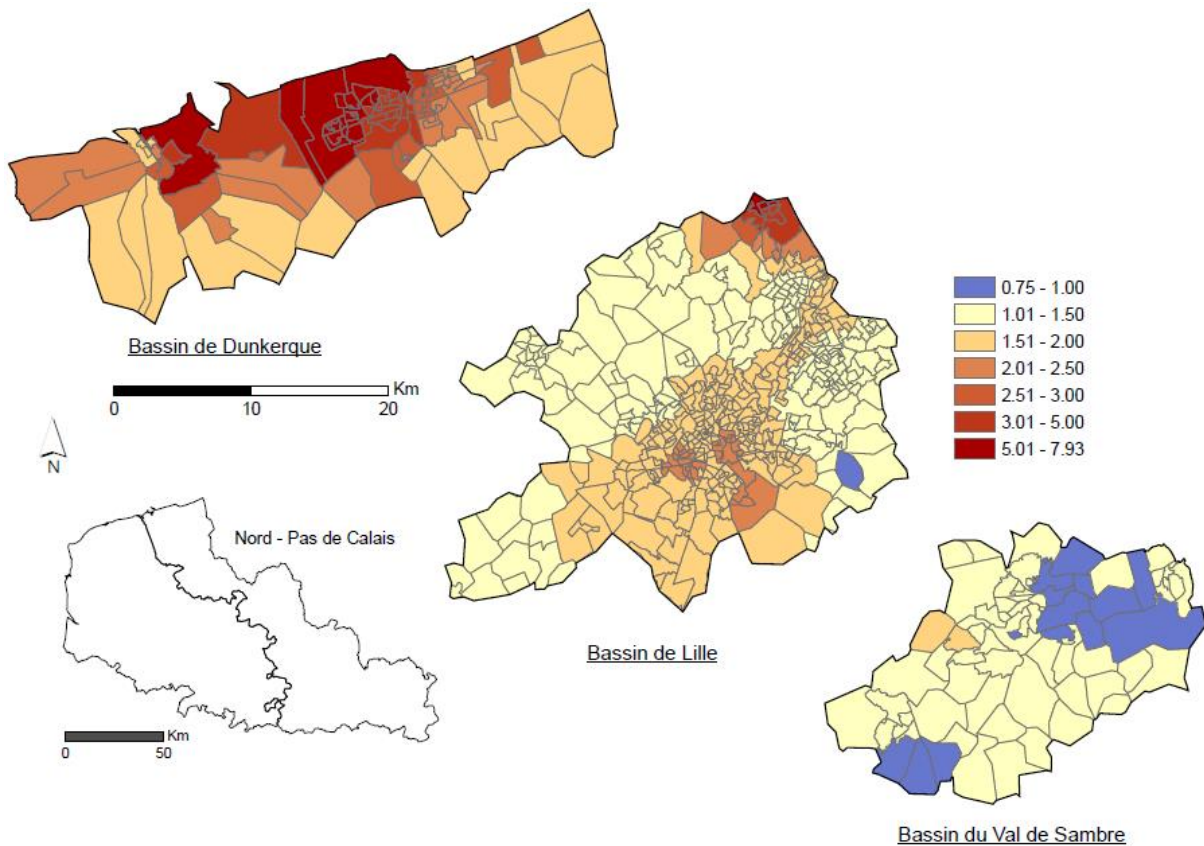


Figure 5. Cartographie des RIM au niveau des IRIS.

RIM mapping in IRIS areas.

Cette cartographie confirme et affine un certain nombre d'observations précédentes. Cela est principalement dû au calcul de krigeage qui s'étend sur des zones moins importantes et donc plus homogènes. Elle permet notamment de distinguer des IRIS dans lesquels les RIM sont inférieurs à 1 sur la CCVS mais aussi de mettre en évidence des IRIS pour lesquels les RIM sont très importants sur la CUD et qui correspondent à des IRIS densément peuplés. C'est, à notre connaissance, la première fois que des données de biosurveillance lichénique sont agrégées à l'IRIS. L'étape suivante de ce travail sera, dans un très proche avenir, d'associer ces résultats à ceux des mesures d'imprégnation des populations ainsi qu'à diverses données socio-économiques afin d'affiner les liens santé-environnement, en intégrant à partir de questionnaires les autres sources possibles d'exposition aux ETM.

DISCUSSION

Ce travail est le premier de cette ampleur dans la région Nord-Pas-de-Calais et, à notre connaissance, n'a que peu d'équivalent en France et en Europe. Les dosages atmosphériques des 4 métaux réglementés montrent que les concentrations respectent les valeurs limites imposées mais cela n'intègre pas le caractère cumulatif des polluants inorganiques. Or nous observons qu'en ce qui concerne l'imprégnation de l'environnement, les valeurs de référence (qui n'ont pas de caractère réglementaire) sont bien souvent largement dépassées. Cette cartographie nous a permis de mettre en évidence la contribution des sources anthropiques principales dans l'imprégnation de l'environnement par les métaux. Le but de notre travail n'était pas d'identifier d'une manière très précise ces sources. Dans ce cas, nous aurions utilisé d'autres techniques comme le marquage isotopique, par exemple. Notre principal but était d'évaluer l'imprégnation de l'environnement et de pouvoir ensuite utiliser ces résultats en association avec d'autres paramètres. Il s'agit là d'une méthode très novatrice dans le domaine de la biosurveillance environnementale. Les objectifs de ces associations sont nombreux et visent notamment à réaliser d'une manière objective des cartes où les points noirs environnementaux et sanitaires pourront être facilement mis en évidence. Ces cartographies pourront être utilisées comme outils de décision dans la gestion des territoires soumis à des inégalités environnementales importantes.

Ce travail a bénéficié des financements de la DRRT, des fonds FEDER et de cofinancements par le Plan Régional de Santé Publique (PRSP), la Communauté Urbaine de Dunkerque et le Conseil Régional Nord-Pas-de-Calais. Il a été réalisé dans le cadre de l'IRENI (Institut de Recherche en Environnement Industriel).

RÉFÉRENCES

- Bargagli R. Trace elements in terrestrial plants, an ecophysiological approach to biomonitoring and biorecovery, Springer, Berlin, 1998, 324 p.
- Chen L.C., Lippmann M. (2009). Effects of metals within ambient air particulate matter (PM) on human health. *Inhal. Toxicol.*, n° 21(1), p. 1-31.
- CITEPA. Inventaire des émissions de polluants atmosphériques et de gaz à effet de serre en France. Séries sectorielles et Analyses étendues (format SECTEN). Rapport National d'Inventaire, 2012, 336 p.
- Cooksey C. (2012). Health concerns of heavy metals and metalloids. *Sci. Prog.*, 95 (Pt 1), p. 73-88.
- Cuny D. (2009). Aperçu bibliographique sur la contamination du lait maternel par les polluants environnementaux. I. : les éléments traces métalliques. *Air Pur*, n° 75, p. 59-65.
- Cuny D., Davranche L., Thomas P. *et al.* (2004). Spatial and temporal variations of trace element contents in *Xanthoria parietina* thalli collected in a highly industrialized area in Northern France as an element for a future epidemiological study. *Journal of Atmospheric Chemistry*, n°49, p. 391-401.
- Cuny D., Quarre S., Devred I. *et al.* (2010). Évolution spatiale et temporelle des concentrations en Éléments Traces Métalliques de thalles de *Xanthoria parietina* récoltés sur la zone industrialo-portuaire de Dunkerque. *Air Pur*, n° 77, p. 43-49.
- Cuny D., Rzepka M.A., Van Haluwyn C. (2008). Introduction à la biosurveillance végétale et fongique de la qualité de l'air – concept et intérêt. *Spectra Analyse*, n° 264, p. 31-35.
- Davranche L., Van Haluwyn C., Cuny D. (2013). Approche du risque sanitaire lié à la consommation de champignons contaminés par les Éléments Traces Métalliques – Documents Mycologiques, 35, p. 311-323.
- Direction Régionale de l'Environnement, de l'Aménagement et du Logement (DREAL) Nord-Pas de Calais. L'Industrie au Regard de l'Environnement en 2011 - IRE 2012, 151 p.
- International Agency for Research on Cancer (IARC). (1993). IARC monographs on the evaluation of carcinogenic risks to humans, beryllium, cadmium, mercury and exposures in the glass, Manufacturing industry, IARC edition, Lyon, n° 58, 453 p.
- International Agency for Research on Cancer (IARC). (1997). IARC monographs on the evaluation of carcinogenic risks to humans, chromium, nickel and welding, IARC edition, Lyon, n° 49, 36 p.
- Kar S., Samal A.C., Maity J.P., Santra S.C., (2013). Diversity of epiphytic lichens and their role in sequestration of atmospheric metals. *Int. J. Env. Sci. Tech.*, DOI 10.1007/s13762-013-0270-8.
- Liu X., Song Q., Tang Y. *et al.* (2013). Human health risk assessment of heavy metals in soil-vegetable system: A multi-medium analysis. *Sci. Total Environ.*, n°4, p. 530-540.
- Loppi S., Frati L., Paoli L. *et al.* (2004). Biodiversity of epiphytic lichens and heavy metals contents of *Flavoparmelia caperata* thalli as indicators of temporal variations of air pollution in the town of Montecatini Terme (central Italy). *Sci. Total Env.*, n° 326, p. 113-122.
- Majumder S., Mishra D., Ram S.S. (2013). Physiological and chemical response of the lichen, *Flavoparmelia caperata* (L.) Hale, to the urban environment of Kolkata. *India Environmental Science and Pollution Research*, n° 20 (5), p. 3077-3085.
- Marano F. (2012). Mécanismes d'action des particules atmosphériques fines et ultrafines. *Pollution Atmosphérique*, N° spécial Particules, p. 20-23.
- Rzepka M.A., Cuny D. (2009). Biosurveillance végétale et fongique des éléments traces métalliques atmosphériques. *Air Pur*, n° 75, p. 66-77.

Rzepka-Cuny M.A. Cuny D. (2011). Preliminary researches on the use of an air biomonitoring network using plants in a complex industrial zone: application of ray-grass in Dunkerque. *Acta Bot. Gal.*, 158, p. 387-400.

Scerbo R., Possenti L. Lampugnani L. *et al.* (1999). Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Livorno Province (Tuscany, Italy). *Sci. of the Tot. Env.*, n°241, p. 91-106.

Scerbo R., Ristori T., Possenti L. *et al.* (2002). Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Pisa Province (Tuscany, Italy). *Sci. of the Tot. Env.*, n° 286, p. 91-106.

Tong-Bin C., Yuan-Ming Z., Mei L., *et al.* (2005). Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere*, n° 60 (4), p. 542-551.

Van Haluwyn C. et Cuny D. (1997). Cartographie de la pollution de l'air par certains métaux lourds sur le littoral Calais-Dunkerque. Rapport SPPPI du Littoral, p. 25-40.

c. Discussion à propos de l'indice multimétallique

Dans la première partie de ce chapitre, les 18 cartographies monométalliques générées montrent que les populations du Nord – Pas de Calais vivent dans un environnement multi-contaminé par les métaux. L'Homme étant exposé à l'ensemble des contaminants présents dans son environnement, il est nécessaire de réfléchir en termes de pollution globale et intégrée. Ce travail fut donc l'occasion de développer un indice intégré de qualité des milieux. Cet indice a pour objectif de synthétiser en un indicateur le niveau d'imprégnation multimétallique de l'environnement au niveau d'un point de mesure. Cela permettra au même titre de confronter les données de biosurveillance lichénique à des indicateurs sanitaires et socio-économiques.

Suite à l'étude bibliographique réalisée dans le cadre de la thèse (voir Introduction générale), nous avons pu relever un certain nombre de critères permettant d'évaluer la pertinence des indices intégrés utilisés pour représenter la qualité globale des milieux (Caeiro *et al.*, 2005 ; Cairncross *et al.*, 2007 ; Singh *et al.*, 2008 ; Bharti & Katyal, 2011). De fait, un tel indice se doit :

- de ne pas surestimer (ambiguïté) ou sous-estimer (éclipse) la contamination par rapport aux valeurs de concentration des paramètres ;
- d'être simple, basique, crédible et transparent en termes de construction ;
- d'être facile à interpréter (comparable à un niveau de seuil, traduction aisée de la valeur brute de l'indice) ;
- d'être à la fois sensible et robuste aux variations d'un paramètre ;
- d'être représentatif (capacité de représenter une image de l'état du milieu) ;
- de justifier d'un niveau d'incertitude acceptable ;
- d'être capable de mettre en évidence des points noirs environnementaux.

Le Ratio d'Imprégnation Moyen (RIM) développé dans notre étude est adapté de l'indice de qualité des eaux Overall Index of Pollution, développé par Sargaonkar & Deshpande (2003), repris plus tard par Chen *et al.* (2005) en tant qu'indice de qualité des sols : l'Integrated Pollution Index. Cet indice remplit plusieurs de ces critères. Il est simple de calcul et transparent, puisqu'il représente une moyenne de ratios de pollution. Il est par nature très communicant et son interprétation est relativement aisée, car il mesure de combien de fois le milieu est enrichi au regard de la valeur de référence.

Néanmoins, nous avons souhaité valider notre indice en termes de représentativité de la qualité globale du milieu, de sensibilité et de robustesse, par le biais de tests statistiques. Le premier test consiste à étudier la représentativité de l'indice, en évaluant le niveau de corrélation entre les 18 métaux et le RIM. Une matrice de corrélation a pour cela été établie sur la base des résultats présentés précédemment. La sensibilité de l'indice est ici considérée comme sa faculté à s'adapter au contexte de pollution. Elle a été évaluée en renouvelant la matrice de corrélation pour chacun des bassins de vie. Enfin, la robustesse représente dans notre étude la stabilité de performance dans des conditions différentes. Celle-ci a été évaluée en comparant les résultats de RIM calculés à partir de 18, 15, 10 et 5 ETM.

REPRESENTATIVITE DE L'INDICE MULTIMETALLIQUE

Afin d'évaluer le niveau de représentativité globale du RIM, nous avons étudié les corrélations entre ce dernier et les éléments qui le composent. Une matrice de corrélation de Spearman a été établie à partir des résultats des 220 points échantillonnés sur les 3 bassins (tableau 8). Le tableau 9 complète les résultats de la matrice de corrélation. Ce dernier classe les éléments en fonction de leur degré de corrélation au RIM. Il fournit ensuite la moyenne et l'écart-type des valeurs de concentrations et du ratio d'imprégnation monométallique (P_i).

Tableau 8. Matrice de corrélation de Spearman entre les 18 métaux et le RIM pour les trois bassins – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 5%. Les valeurs en vert sont ≥ 0.75

Variables	Al	Cr	Cu	As	Hg	Cd	Pd	Mn	Be	Rh	Co	Sb	Ni	Pt	V	Pb	Ti	Zn	RIM
Al	1																		
Cr	0.858	1																	
Cu	0.628	0.637	1																
As	0.765	0.795	0.646	1															
Hg	0.431	0.532	0.492	0.623	1														
Cd	0.413	0.453	0.482	0.369	0.583	1													
Pd	0.284	0.286	0.393	0.362	0.321	0.546	1												
Mn	0.708	0.855	0.497	0.757	0.496	0.384	0.214	1											
Be	-0.176	-0.131	0.049	0.105	-0.049	-0.085	0.409	0.013	1										
Rh	0.175	0.133	0.049	0.206	-0.002	0.077	0.587	0.137	0.676	1									
Co	0.857	0.826	0.699	0.901	0.543	0.445	0.403	0.761	0.052	0.243	1								
Sb	0.576	0.573	0.852	0.538	0.527	0.511	0.363	0.384	-0.114	0.036	0.604	1							
Ni	0.802	0.918	0.644	0.831	0.562	0.484	0.331	0.879	0.022	0.169	0.866	0.518	1						
Pt	0.144	0.099	0.016	0.176	-0.034	0.044	0.556	0.109	0.707	0.972	0.210	0.002	0.135	1					
V	0.836	0.910	0.584	0.885	0.612	0.431	0.318	0.854	-0.053	0.198	0.883	0.520	0.907	0.168	1				
Pb	0.580	0.576	0.813	0.605	0.566	0.635	0.473	0.511	0.119	0.132	0.701	0.755	0.640	0.096	0.574	1			
Ti	0.430	0.479	0.491	0.647	0.315	0.169	0.280	0.555	0.436	0.143	0.609	0.263	0.575	0.110	0.543	0.516	1		
Zn	0.566	0.612	0.715	0.586	0.559	0.742	0.508	0.562	0.116	0.137	0.687	0.696	0.651	0.102	0.604	0.802	0.478	1	
RIM	0.801	0.859	0.809	0.829	0.679	0.683	0.477	0.777	-0.022	0.138	0.882	0.768	0.870	0.103	0.860	0.819	0.541	0.837	1

Tableau 9. Classement des éléments en fonction de leur degré de corrélation (coefficient de Spearman) avec le RIM, concentration moyenne de chaque élément dans les lichens (en mg/kg) et ratio d'imprégnation monométallique

Élément	Coefficient Spearman ETM-RIM	Concentration moyenne (écart-type) en mg/kg	Ratio (Pi) moyen (écart-type)
Co	0.88	0.89 (1.31)	0.13 (0.19)
Ni	0.87	8.61 (53.11)	0.86 (5.32)
V	0.86	6.01 (9.03)	0.50 (0.75)
Cr	0.86	13.19 (78.84)	3.30 (19.75)
Zn	0.84	123.42 (194.03)	2.47 (3.89)
As	0.83	1.30 (1.17)	2.60 (2.35)
Pb	0.82	28.07 (64.15)	1.87 (4.29)
Cu	0.81	17.39 (21.90)	1.16 (1.46)
Al	0.80	1135.92 (1120.06)	2.27 (2.25)
Mn	0.78	185.29 (512.26)	6.18 (17.11)
Sb	0.77	1.52 (2.20)	5.06 (7.35)
Cd	0.68	0.82 (2.16)	4.09 (10.84)
Hg	0.68	0.14 (0.12)	0.69 (0.62)
Ti	0.54	19.25 (20.66)	3.85 (4.14)
Pd	0.48	20.31 (17.29)	1.35 (1.16)
Rh	0.14	15.62 (3.02)	1.04 (0.20)
Pt	0.10	15.40 (0.15)	1.03 (0.01)
Be	-0.02	0.24 (0.06)	0.08 (0.02)

La première observation importante fournie par la matrice est la corrélation significative entre la plupart des métaux, comme l'attestent les valeurs en gras. Be, Pt et Rh figurent parmi les moins corrélés aux autres ETM. Cela peut s'expliquer par la faible dispersion de leurs valeurs, toutes proches ou inférieures aux concentrations naturelles. Ces trois éléments présentent respectivement une concentration moyenne de 0.24 ± 0.06 mg/kg, 15.4 ± 0.15 mg/kg et 15.62 ± 3.02 mg/kg. Nous remarquons ensuite qu'en dehors de ces 3 éléments, le RIM est significativement corrélé aux 15 autres ETM. Parmi ces derniers, le manganèse ($P_i : 6.18 \pm 17.11$), l'antimoine ($P_i : 5.06 \pm 7.35$) et le cadmium ($P_i : 4.09 \pm 10.84$) sont les éléments présentant les valeurs d'enrichissement les plus élevées au regard des valeurs naturelles.

Lorsqu'on observe les coefficients de Spearman, ils varient de 0.48 pour le palladium à 0.88 pour le cobalt. Parmi ces 15 éléments, 11 sont liés à plus de 75% à l'indice : Al, Cr, Cu, As, Mn, Co, Sb, Ni, V, Pb, Zn. Les plus corrélés sont Co (0.88), Ni (0.87), V et Cr (0.86) et les 4 éléments les moins corrélés sont Pd (0.48), Ti (0.54), Hg (0.68) et Cd (0.68). En observant l'ordre de grandeur des concentrations et des P_i , nous pouvons finalement constater que les éléments présentant des valeurs fortes ne sont pas les plus corrélés à l'indice, ce qui témoigne de sa faible sensibilité aux gammes élevées de concentrations.

La forte corrélation des éléments entre eux, ainsi que la corrélation significative du RIM avec la majorité de ces paramètres montre que notre indice est représentatif de la contamination multimétallique du milieu.

SENSIBILITE DE L'INDICE MULTIMETALLIQUE

Afin d'évaluer la sensibilité de notre indice au contexte de pollution, nous avons calculé le coefficient de corrélation de Spearman en isolant les échantillons issus d'un même bassin ($n = 60$ pour Dunkerque ; $n = 120$ pour Lille ; $n = 40$ pour Maubeuge). La comparaison des éléments les plus corrélés en fonction des trois bassins peut illustrer la capacité du RIM à s'adapter au contexte de pollution. Le tableau 10 synthétise les niveaux de corrélation des éléments avec le RIM en fonction des bassins. Les éléments sont à chaque fois classés du plus corrélé au moins corrélé.

Ce tableau nous montre dans un premier temps que les niveaux de corrélation entre le RIM et les ETM varient en fonction des bassins. Ils sont en moyenne plus élevés sur le bassin de Dunkerque (0.69) et plus faibles sur le bassin de Maubeuge (0.63, contre 0.65 pour les 3 bassins et 0.66 pour Lille). Les éléments significativement corrélés au RIM sont les mêmes pour le contexte global et le dunkerquois. Pour le bassin de Lille, les mêmes éléments sont retrouvés, ainsi que le rhodium qui devient significativement corrélé au RIM. En revanche, sur le secteur rural de Maubeuge, 5 éléments ne sont pas corrélés à l'indice (Be, Hg, Pd, Pt, et Rh), notamment le mercure qui est principalement émis dans l'atmosphère par le secteur industriel (96% des émissions atmosphériques totales : CITEPA, 2014).

Lorsque l'on se focalise sur le classement établi pour chaque contexte, nous remarquons que les éléments les plus liés au RIM ne sont pas les mêmes. Même si le cobalt occupe la première place pour deux des contextes (les 3 bassins confondus et Maubeuge), il est classé 3^e et 6^e sur Dunkerque et Lille. Pour ces deux bassins, ce sont respectivement le manganèse et l'antimoine qui dominent le classement. Parmi les 5 éléments les plus corrélés au RIM, aucun n'est retrouvé pour les 4 contextes. Si l'on s'intéresse maintenant aux valeurs des coefficients par élément, on constate de grands écarts entre les bassins. Par exemple, le coefficient de corrélation du plomb est de 0.82 pour les 3 bassins, 0.93 pour Dunkerque, 0.86 pour Lille et 0.75 pour Maubeuge. Majoritairement émis par le secteur industriel, le mercure est fortement corrélé au RIM sur le bassin de Dunkerque (0.86) et l'est beaucoup moins sur les deux autres bassins (0.49 pour Lille et 0.30 pour Maubeuge). D'autres comme l'aluminium sont en revanche plus stables (0.80 pour les 3 bassins, 0.86 pour Dunkerque, 0.76 pour Lille et 0.79 pour Maubeuge). L'un des plus gros écarts concerne le titane. Il figure parmi les moins corrélés pour 3 contextes (coefficients de 0.54 pour les 3 bassins, 0.51 pour Dunkerque et 0.58 pour Lille) et occupe la 2^e place sur le bassin de Maubeuge (0.86). Le RIM peut donc être considéré différemment sur chacun des bassins. Le fait qu'il s'adapte au cocktail de pollution en fonction de la typologie du territoire confirme sa sensibilité.

Tableau 10. Classement des éléments en fonction de leur niveau de corrélation (coefficient de Spearman) avec le RIM – analyse par bassin – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 5%

3 Bassins		Dunkerque		Lille		Maubeuge	
Co	0.88	Mn	0.95	Sb	0.88	Co	0.87
Ni	0.87	V	0.95	Ni	0.88	Ti	0.86
V	0.86	Co	0.94	Zn	0.87	V	0.84
Cr	0.86	Pb	0.93	Cr	0.87	Cr	0.83
Zn	0.84	Zn	0.89	Pb	0.86	Mn	0.83
As	0.83	Ni	0.89	Co	0.86	Ni	0.82
Pb	0.82	Cu	0.87	Cu	0.85	As	0.80
Cu	0.81	Al	0.86	V	0.83	Al	0.79
Al	0.80	Hg	0.86	Mn	0.80	Cu	0.75
Mn	0.78	Cr	0.85	As	0.76	Pb	0.75
Sb	0.77	As	0.84	Al	0.76	Zn	0.69
Cd	0.68	Cd	0.83	Cd	0.61	Sb	0.57
Hg	0.68	Sb	0.78	Ti	0.58	Cd	0.51
Ti	0.54	Pd	0.51	Pd	0.55	Pd	0.31
Pd	0.48	Ti	0.51	Hg	0.49	Rh	0.31
Rh	0.14	Be	0.02	Rh	0.21	Pt	0.31
Pt	0.10	Rh	-0.04	Pt	0.16	Hg	0.30
Be	-0.02	Pt	-0.04	Be	0.05	Be	0.20

En dehors de certains éléments traceurs, comme le cuivre pour les transports routiers et le mercure pour l'industrie, il est difficile d'émettre l'hypothèse que le RIM symbolise une contamination de type industrielle sur le bassin de Dunkerque, urbaine sur le bassin de Lille ou rurale sur le bassin de Maubeuge. La majorité des éléments se retrouvent en effet dans l'environnement du fait d'activités anthropiques multiples (CITEPA, 2014). Néanmoins, sur le bassin de Maubeuge, les éléments majoritairement émis par des activités anthropiques (Cd, Cu, Hg, Pb, Zn) se retrouvent dans le bas du classement.

ROBUSTESSE DE L'INDICE MULTIMETALLIQUE

La robustesse de notre indice aux différents éléments qui le composent a ensuite été testée. Pour chacun des 220 points d'échantillonnage, l'indice a pour cela été recalculé à partir de 15, 10 et 5 des 18 éléments. Le choix des éléments intégrés à ces calculs est basé sur leur degré de corrélation au RIM initial (RIM18) dans le contexte des trois bassins (tableau 10). Ainsi, le « RIM15a » est calculé à partir des 15 éléments les plus corrélés au RIM18, alors que le « RIM15b » est calculé à partir des 15 éléments les moins corrélés au RIM18. Les RIM10a, RIM10b, RIM5a et RIM5b ont été calculés de la même manière à partir de 10 et 5 éléments. La figure 25 montre la distribution de ces nouveaux indices. Les différences statistiques entre les 7 variables ont été évaluées à l'aide du test multiple de Dunn. Les résultats montrent que les RIM15a, RIM15b et RIM10a ne sont pas différents du RIM18. Au contraire, RIM10b, RIM5a et RIM5b sont statistiquement différents de RIM18 ($p < 0.001$).

Les RIM calculés à partir de 15 éléments sont similaires au RIM18 (moyenne de 2.47 ± 4.18 pour RIM15a, de 2.57 ± 4.05 pour RIM15b et de 2.23 ± 3.78 pour RIM18). Cela traduit une robustesse de l'indicateur par rapport aux facteurs intégrés au calcul. En revanche, lorsque le RIM est calculé à partir d'un nombre plus faible d'éléments, le résultat varie plus fortement. Bien que le RIM10a,

calculé à partir des 10 métaux les plus corrélés au RIM initial, montre un niveau d'imprégnation similaire (2.16±4.54), le résultat est différent pour les RIM10b (2.72±3.55). Les indices calculés à partir de 5 éléments sous-estiment le niveau d'imprégnation globale (1.55±7.05 pour RIM5a et 1.39±1.04 pour RIM5b).

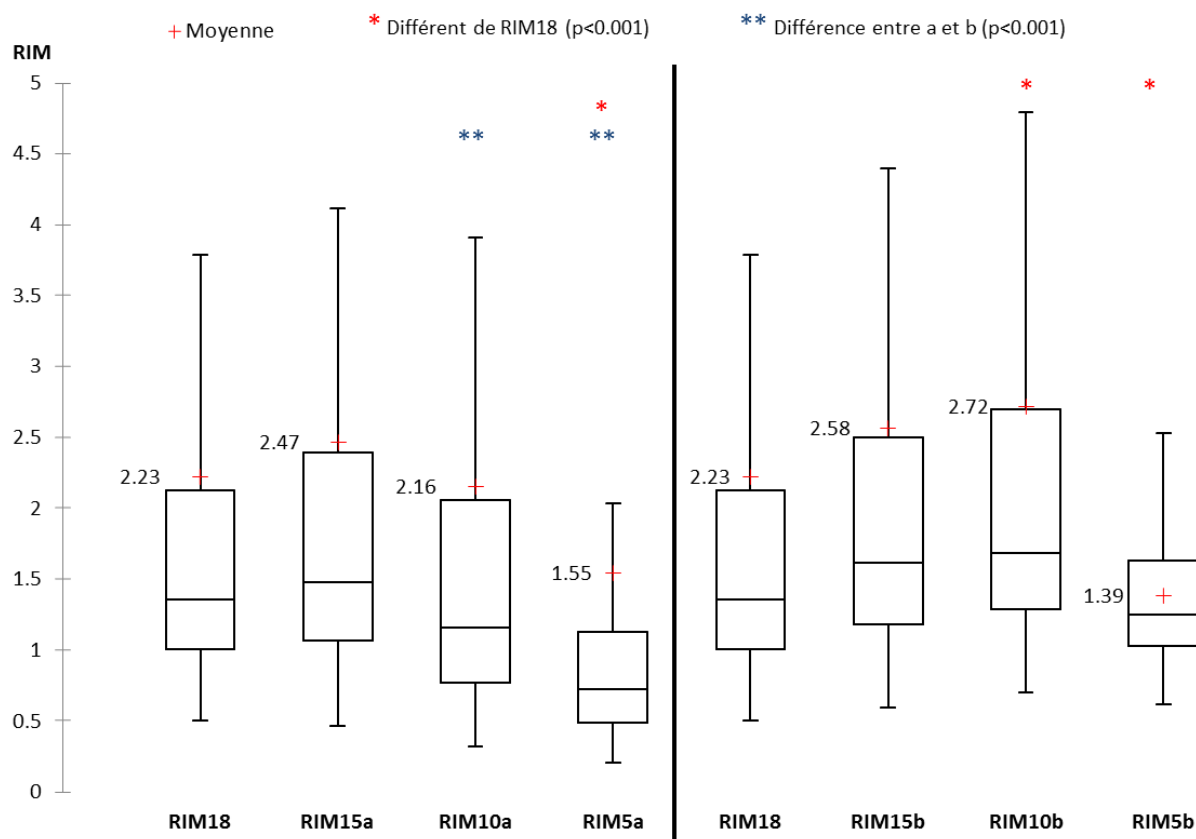


Figure 25. Variabilité des RIM en fonction du nombre d'éléments traces métalliques intégrés au calcul, sous forme de box plot – les indices sont calculés à partir des 15, 10 et 5 éléments les plus corrélés (a) ou les moins corrélés (b) au RIM18 – les différences entre les RIM sont évaluées à l'aide du test de Dunn (p<0.001)

CONCLUSION SUR LA PERTINENCE DE L'INDICE MULTIMÉTALLIQUE

En dernier lieu, le RIM a été comparé à un autre indice composite de qualité des milieux. Il s'agit du PLI (Pollution Load Index), développé par Tomlinson *et al.* (1980) pour caractériser le niveau de charge globale en contaminants chez plusieurs organismes vivant dans les sédiments (algues crustacés). Cet indice a notamment été utilisé à deux reprises pour évaluer le niveau de charge multimétallique accumulé dans les lichens (Boamponsem *et al.*, 2010 ; Salo *et al.*, 2012). Pour le calcul du PLI, le ratio d'imprégnation (P) est dans un premier temps calculé pour chaque ETM i , par le rapport de la concentration mesurée (C_i) sur la concentration de référence (B_i). Le PLI est la $n^{\text{ème}}$ racine du produit des P_i . A l'instar du RIM, un PLI supérieur à 1 indique une contamination du milieu au regard des valeurs de référence.

$$P_i = \frac{C_i}{B_i}$$

$$PLI = \sqrt[n]{P_1 \times P_2 \times \dots \times P_n}$$

La figure 26 montre le nuage de corrélation entre le RIM et le PLI selon les 220 points d'échantillonnage. On observe un degré de corrélation très élevé, puisque le coefficient de Spearman est supérieur à 0.93. Caeiro *et al.*, 2005 classent le PLI parmi les indices les plus pertinents en termes de simplicité, représentativité, crédibilité, comparabilité, sensibilité/robustesse et incertitudes. La forte corrélation de notre indice avec celui-ci conforte donc notre choix.

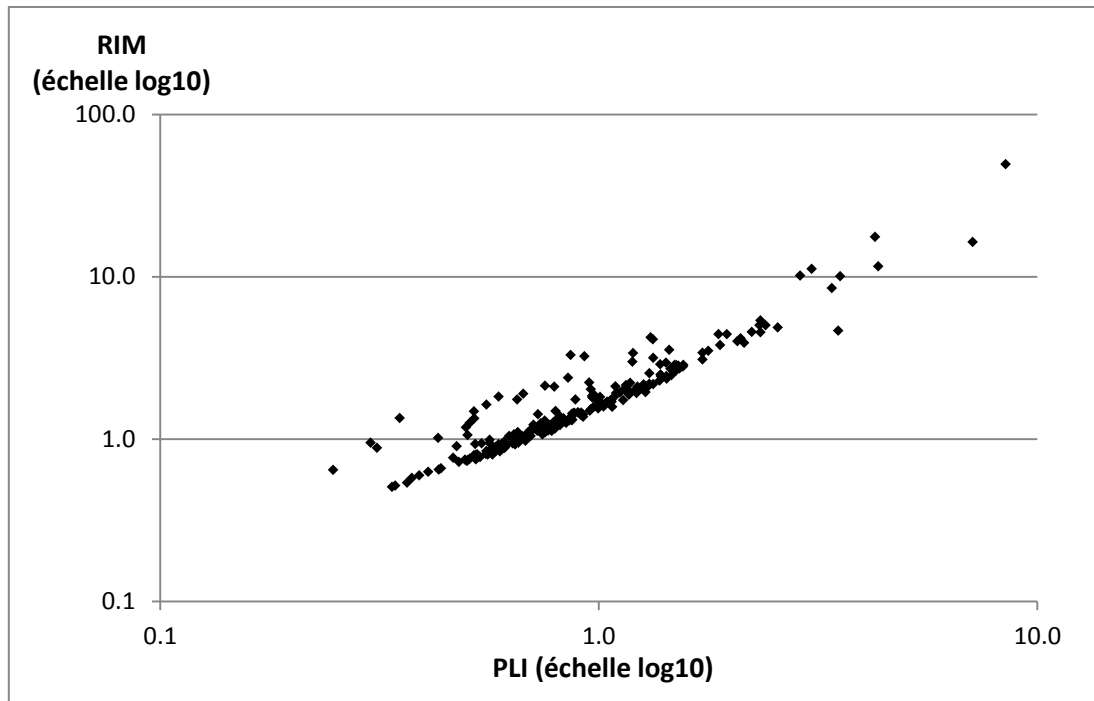


Figure 26. Corrélation entre le RIM et le PLI pour les trois bassins (échelle log10)

Par ailleurs, le PLI emploie une fonction racine qui a pour objectif de réduire l'influence des éléments présentant de fortes concentrations. Il est de fait moins sensible aux valeurs extrêmes et sous-estime la contamination globale lorsqu'un des éléments se retrouve à des niveaux très élevés de concentrations (ambiguïté). Le RIM est quant à lui plus sensible à ce type de phénomènes et s'adapte, nous l'avons vu, au contexte de contamination. De plus, dans notre étude la plupart des éléments traces sont corrélés entre eux. En d'autres termes, lorsqu'une contamination importante est observée pour l'un des éléments, elle l'est également pour les autres. Dans notre cas d'étude, le RIM ne surestime (ambiguïté) et ne sous-estime (éclipse) donc pas la contamination globale du milieu.

Plusieurs arguments bibliographiques et statistiques permettent ainsi de justifier l'utilisation du RIM pour évaluer le niveau d'imprégnation globale du milieu. Dans le cadre de notre étude, le RIM représente une image fiable de l'état global de contamination par l'ensemble des ETM. Il est de plus sensible au contexte de pollution. Sa robustesse aux faibles modifications des paramètres intégrés au calcul est démontrée, mais elle n'est plus valable lorsque les conditions sont très différentes.

Partie 3. Inégalités environnementales

a. Introduction

Le concept des inégalités environnementales est basé sur la proximité de populations défavorisées sur un plan socio-économique aux sources de nuisance (Glickmann, 1994 ; Brulle & Pellow, 2006) et s'est récemment étendu à la présence de ces populations sur des territoires imprégnés par les contaminants environnementaux (McLeod *et al.*, 2000 ; Chaix *et al.*, 2006 ; Havard *et al.*, 2008). Ces inégalités ont été évaluées à l'échelle des Iris du bassin industriel de Dunkerque.

Un indice de défaveur sociale (Indice de Défaveur Localisé : IDL) fut pour cela développé par le laboratoire TVES ULCO selon une approche éco-sociologique, au même titre que l'Indicateur de Santé Sociale des régions (Jany-Catrice & Zotti, 2008). Il caractérise le statut socio-économique des populations à l'échelle des Iris à partir de 14 variables, regroupées en six grands thèmes : les formes d'emploi, l'exclusion de l'emploi, l'éducation, le lien social, le revenu et enfin le logement. En parallèle, les cartographies géostatistiques, générées à partir des mesures de 18 éléments traces dans les lichens ainsi que celle du RIM, ont été agrégées à l'Iris par le calcul de la valeur moyenne observée dans chaque unité spatiale.

La corrélation entre les indicateurs environnementaux et l'IDL a dans un premier temps été évaluée à l'aide de corrélations de Spearman. Afin de mettre en évidence les inégalités environnementales, les Iris ont ensuite été catégorisés selon des quintiles en fonction des valeurs d'IDL. Le test de Dunn a enfin permis de tester les différences de contamination métallique entre les différentes catégories de défaveur.

La cartographie de l'IDL révèle, des disparités de statut socio-économique entre les populations. Les Iris des villes-centre, situés à proximité des usines du dunkerquois sont les plus défavorisés, alors que les Iris situés en périphérie sont les plus favorisés. L'IDL est significativement corrélé à l'indice multimétallique et à 16 des 18 ETM. Des différences significatives de niveau de contamination sont observées, notamment entre les groupes extrêmes de défaveur sociale. Ces résultats témoignent ainsi de la présence d'inégalités environnementales à l'échelle des quartiers du bassin industriel de Dunkerque. Les populations les plus défavorisées vivent dans un environnement fortement imprégné par les éléments traces. L'utilisation de données de biosurveillance environnementale et d'un indice intégré de contamination est pionnier dans le domaine des inégalités environnementales.

L'intégralité de ce travail a fait l'objet d'un article à soumettre prochainement à la revue *Environmental Health Perspectives* et intégré à la suite de ce résumé. Une partie de ces résultats a par ailleurs été présentée sous la forme d'une communication orale lors des Journées Interdisciplinaires de la Qualité de l'Air (JIQA) qui se sont tenues à Lille en février 2014.

b. Article

TITRE

Inégalités environnementales à l'échelle des quartiers du bassin industriel de Dunkerque : apports de la biosurveillance lichénique de la qualité de l'air.

AUTEURS ET AFFILIATIONS

Florent Occelli, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France.

Rachel Bavdek, EA4477, Université Littoral Côte d'Opale (ULCO), laboratoire Territoires Villes Environnement et Société (TVES), Dunkerque, France.

Annabelle Deram, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France. Faculté Ingénierie et Management de la Santé (ILIS), Loos, France.

Anne Peggy Hellequin, laboratoire Territoires Villes Environnement et Société (TVES) de l'Université Littoral Côte d'Opale (Dr. Anne Peggy Hellequin et Rachel Bavdek, EA4477, ULCO).

Zwarterook I, laboratoire Territoires Villes Environnement et Société (TVES) de l'Université Littoral Côte d'Opale (Dr. Anne Peggy Hellequin et Rachel Bavdek, EA4477, ULCO).

Damien Cuny, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France.

MOTS CLES

Indice de défaveur sociale, biosurveillance environnementale, lichens, éléments traces métalliques, indice intégré, inégalités environnementales

LISTE DES FIGURES

Figure 1. Défaveur sociale par Iris, représentée par l'Indice de Défaveur Localisé (IDL)

Figure 2. Charges métalliques dans les lichens par Iris, représenté par le Ratio d'Imprégnation Moyen (RIM)

Figure 3. Distribution du RIM en fonction des quintiles d'IDL

Figure 4. Inégalités environnementales à l'échelle des Iris du bassin de Dunkerque

LISTE DES TABLES

Tableau 1. Composition de l'Indice de Défaveur Localisé (IDL)

Tableau 2. Résultats de l'analyse écologique entre l'IDL et les charges métalliques dans les lichens (18 ETM et RIM) – pour le coefficient de Spearman, * = $p < 0.05$ et ** = $p < 0.01$ – pour le test de Dunn, les valeurs en gras sont significatives ($p < 0.005$)

RESUME

Le concept des inégalités environnementales est basé sur la proximité de populations défavorisées sur un plan socio-économique ou de certaines minorités ethniques à des sources de nuisances et s'est récemment étendu à la contamination des milieux environnementaux. Dans ce travail, ces inégalités ont été évaluées d'une manière novatrice à l'échelle des quartiers d'un bassin industriel de Dunkerque, en associant un indice de défaveur sociale et des mesures de biosurveillance environnementale.

Un Indicateur de Défaveur Localisé (IDL) a été développé selon une approche éco-sociologique. Il caractérise le statut socio-économique des populations à l'échelle des quartiers. Des dosages de 18 métaux ont été effectués dans des échantillons de lichens épiphytes, récoltés dans la zone d'étude. Les cartographies géostatistiques générées à partir de ces mesures ont permis d'attribuer une valeur moyenne des concentrations métalliques à chaque quartier. En complément des 18 éléments, un indice intégré (Ratio d'Imprégnation Moyen : RIM) a été généré pour évaluer le niveau de contamination multimétallique. Les inégalités environnementales possibles ont été évaluées à l'aide de corrélations de Spearman entre les indicateurs environnementaux et l'IDL. Le test de Dunn a ensuite permis de tester les différences de contamination métallique entre les catégories de défaveur.

La cartographie de l'IDL révèle des disparités socio-économiques entre les populations. Les quartiers situés à proximité des usines sont les plus défavorisés, alors que les quartiers situés en périphérie sont les plus favorisés. Les cartographies métalliques et multimétallique montrent des variations spatiales de contamination de l'environnement. Les concentrations les plus élevées sont retrouvées à proximité des zones industrielles. L'IDL est significativement corrélé au RIM et à 16 des 18 métaux. Des différences significatives de niveau de contamination sont observées entre les groupes extrêmes de défaveur sociale.

Nos résultats témoignent de la présence d'inégalités environnementales à l'échelle des quartiers de l'agglomération de Dunkerque. Les populations les plus défavorisées vivent dans un environnement fortement imprégné par les métaux. L'utilisation de données de biosurveillance et d'un indice intégré de contamination est pionnier dans le domaine de la prospection sur les inégalités environnementales.

INTRODUCTION

Le concept de justice environnementale trouve ses racines au début des années 1970 aux Etats-Unis (Bullard, 2005 ; Brulle & Pellow, 2006). Né du constat d'une sur-représentation des minorités ethniques à proximité des sites pollués et/ou dangereux, cette problématique va peu à peu trouver écho auprès des communautés scientifique et politique. Aussi, en 1994, la définition de la justice environnementale sera institutionnalisée par l'Environmental Protection Agency (EPA) et définie comme "le traitement juste et l'implication commune de tous indépendamment de son origine ethnique, sa nationalité ou son revenu dans le développement, la mise en œuvre et l'exécution des lois, règlements et politiques relative à l'environnement" (EPA, 2010).

Ce concept sera longtemps étudié en termes de proximité aux sources possibles de nuisances pour les individus (Glickmann, 1994 ; Chakraborty & Armstrong, 1997 ; Sheppard *et al.*, 1999 ; Walker *et al.*, 2005). Néanmoins, depuis l'expansion de la santé environnementale et le développement de puissants outils géostatistiques (Cressie, 1993), le concept se base de plus en plus sur l'évaluation de l'exposition aux pollutions de l'environnement, notamment les substances réglementées de la qualité de l'air (McLeod *et al.*, 2000 ; Mitchell, 2005 ; Chaix *et al.*, 2006). Ces substances sont alors étudiées indépendamment les unes des autres.

En France, les premiers travaux consacrés à la justice environnementale datent du début des années 2000. En 2001, Theys et Emilianoff abordent la ville durable, sans toutefois parler d'« inégalités environnementales », pour traiter du lien entre exposition aux nuisances environnementales et conditions socio-économiques. Peu de temps après, le ministère de l'Environnement (2003) a lancé un appel à proposition de recherche contenant un axe sur les inégalités écologiques. Enfin, en 2004 l'article 1^{er} de la Charte Française de l'Environnement stipule que « Chacun a le droit de vivre dans un environnement équilibré et respectueux de la santé » (Charte Française de l'Environnement, 2004). Depuis, les études scientifiques se sont multipliées, que ce soit sur les nuisances comme le bruit (Faburel & Malevre, 2007 ; Bocquier *et al.*, 2011 ; BruitParif, 2013), la proximité aux établissements polluants et/ou à risques (Laurian, 2008 ; Viel *et al.*, 2010), l'exposition à des polluants atmosphériques (Havard *et al.*, 2008) ou encore l'utilisation d'un indice global de qualité de l'environnement alliant les ressources (espaces verts, plans d'eau) et les handicaps (pollution de l'air, proximité aux industries, bruit, inondations) (Faburel & Gueymard, 2008). Certaines de ces études indiquent que la proximité aux industries n'est pas toujours synonyme de défaveur (Flanquart *et al.*, 2013).

Dans cette étude, la biosurveillance environnementale a été employée pour décrire la qualité de l'air d'un territoire industrialo-portuaire du Nord de la France, notamment en termes de contaminations métalliques. La biosurveillance est « l'utilisation des réponses à tous les niveaux d'organisation biologique (moléculaire, biochimique, cellulaire, physiologique, tissulaire, morphologique, écologique) d'un organisme ou d'un ensemble d'organismes pour prévoir et/ou révéler une altération de l'environnement et pour en suivre son évolution » (Garrec & Van Haluwyn, 2002). Cette approche biologique n'est pas substitutive, mais complémentaire des méthodes de surveillance physico-chimiques (Cuny *et al.*, 2008). Dans le contexte de la qualité de l'air, la biosurveillance végétale et fongique met en évidence les impacts biologiques de la qualité de l'air. Parmi les organismes utilisés, les lichens épiphytes sont les représentants du règne fongique les plus couramment employés. Ils sont exclusivement dépendants de l'atmosphère pour leur nutrition et ils absorbent les polluants présents dans l'air en même temps que les éléments qui leur sont nécessaires (Loppi *et al.*, 1997; Loppi & Nascimbene, 2010). Les ETM retrouvés dans l'environnement

sont principalement issus d'activités anthropiques. Ils sont classés parmi les groupes de polluants les plus dangereux en raison de leur toxicité et de leur persistance dans l'environnement. La capacité bioaccumulatrice des lichens a largement été employée dans le cadre de la recherche en biosurveillance (Garty, 1993; Rzepka & Cuny, 2008). Ces organismes constituent d'excellents indicateurs de la pollution atmosphérique de fond (Bargagli, 1998; Cuny *et al.*, 2001; Garty, 2001; Boamponsem *et al.*, 2010) et sont très utiles pour le suivi spatial et temporel des contaminations métalliques (Scerbo *et al.*, 1999, 2002; Loppi *et al.*, 2003; Cuny *et al.*, 2004; Cloquet *et al.*, 2009; Balabanova *et al.*, 2012; Ocelli *et al.*, 2013).

Dans ce contexte de la biosurveillance environnementale, les indices intégrés de pollution sont de plus en plus utilisés pour décrire les contaminations multimétalliques (Madejon *et al.*, 2006; Boamponsem *et al.*, 2010; Guéguen *et al.*, 2012; Salo *et al.*, 2012). Ils agrègent les concentrations de plusieurs substances en un indicateur qui caractérise à lui seul le statut général de qualité environnementale (Bharti & Katyal, 2011; Zhao *et al.*, 2012; Golge *et al.*, 2013). Bien que les indices intégrés représentent une vulgarisation de la qualité environnementale, ils peuvent être considérés comme de puissants outils de traitement et d'analyse globale, à partir d'informations environnementales brutes (Caeiro *et al.*, 2005 ; Golge *et al.*, 2013). Les populations sont exposées à l'environnement de manière globale, c'est-à-dire à l'ensemble des contaminants. Cela nous amène à considérer la qualité du milieu de façon globale et intégrée.

Traiter des inégalités environnementales suppose également de définir l'aspect sociétal étudié. De nombreuses études, notamment américaines, s'intéressent à un aspect particulier des caractéristiques des individus, comme l'origine ethnique ou le revenu (Bullard, 1993 ; McLaren, 1999 ; Chaix, 2006). En parallèle de ces approches, une démarche de caractérisation socio-économique synthétique des individus s'est développée. Ces indicateurs reposent majoritairement sur les recensements nationaux de populations, rendant compte d'une information précise, ponctuelle dans le temps et systématisée. Nous pouvons citer par exemple l'Index of Multiple Deprivation (DETR, 2000), ou encore le Scottish Index of Multiple Deprivation (Scottish Executive Order, 2004). Parmi ces indicateurs, peu d'entre eux sont transposables à d'autres territoires au contraire de l'indicateur de Townsend (Townsend, 1987). Ce dernier mobilise quatre variables classiquement présentes dans les recensements, mais a pour caractéristique de fortement distinguer le milieu urbain du milieu rural. Afin d'éviter ce phénomène, d'autres indicateurs ont récemment été développés, tels que l'indicateur de défaveur de Havard, basé sur une analyse statistique (Havard *et al.*, 2008). Il permet à partir d'une Analyse en Composante Principale de déterminer les variables les plus significatives dans l'estimation de la défaveur d'un territoire. Néanmoins, l'inconvénient principal de cet indicateur est de ne pas être transposable à d'autres territoires que celui de référence. Cet article propose de pallier cette difficulté, à l'aide d'un indicateur de défaveur généré à partir d'une approche éco-sociologique.

Cet article s'intéresse à l'élaboration d'un indice de défaveur sociale à l'échelle des quartiers, pouvant être généralisé à d'autres territoires. Nous caractérisons ensuite les niveaux d'imprégnation par les ETM de l'environnement de vie des populations grâce à la biosurveillance lichénique. Finalement, notre approche consiste à évaluer si les populations les plus vulnérables sur un plan socioéconomique, et par conséquent plus vulnérables sur un plan sanitaire (OMS, 2012 ; Besse *et al.*, 2014), vivent dans un environnement contaminé par les métaux.

MATERIEL ET METHODES

Aire d'étude et unité spatiale

Ce travail a été réalisé sur le littoral dunkerquois, situé au nord de la France. Ce bassin géographique est composé d'une importante zone industrialo-portuaire, entourée d'une aire urbaine densément peuplée (≈220 000 habitants). Le complexe industriel s'étend sur une dizaine de kilomètres et comprend des activités émettrices de polluants atmosphériques, telles que la chimie, la pétrochimie, la métallurgie et la production d'énergies (Cuny *et al.*, 2002). Afin de rendre compte de la présence ou non d'inégalités environnementales à un échelon territorial le plus finement possible, l'unité de recherche la plus pertinente est l'îlot regroupé pour l'information statistique (Iris). L'Iris est l'unité administrative privilégiée de diffusion de données infra-communales par l'INSEE et comprend environ 2000 habitants. Elle garantit la fiabilité et l'homogénéité des données, ainsi que leur pérennité dans le temps. L'agglomération de Dunkerque est ici composée de 28 communes, divisées en 102 Iris. Deux de ces Iris sont dévoués aux activités industrielles et ne comportent pas de population. Ils ne seront pas pris en compte dans ce travail.

Indice de défaveur localisé

Un indice de défaveur sociale nécessite une restitution la plus complète possible de la situation socio-économique des habitants d'un territoire, à travers une seule information claire, simple, disponible et flexible (Booyesen, 2002). Cette information doit refléter « les faits et tendances les plus significatifs dans un contexte social et politique donné » (Perret, 2003). Dans cet objectif, nous avons opté pour une approche éco-sociologique des indicateurs, telle que celle de l'Indicateur de Santé Sociale des régions (ISS : Jany-Catrice & Zotti, 2008). Cet indicateur est reconnu, institutionnalisé et facilement appropriable. Il vise à identifier les grands enjeux du bien-être des individus, à travers une démarche participative des spécialistes et des citoyens. Huit grandes thématiques y sont identifiées : le revenu, le travail et l'emploi, l'éducation, la santé, le logement, la justice, le lien social et le lien interindividuel. Ainsi, nous avons choisi de créer une sorte "d'idéal type" de cet indicateur à l'échelle infra-communale.

Les grands thèmes de l'ISS ont été repris, afin de rechercher une concordance assez proche et pertinente au quartier. Deux thématiques ont, dès lors, été évincées de l'indicateur pour des raisons de confidentialités de la donnée : la justice et la santé. Six grands thèmes assimilables à ceux de l'ISS ont finalement été retenus : les formes d'emploi, l'exclusion de l'emploi, l'éducation, le lien social, le revenu et enfin le logement. Ils sont eux-mêmes composés de 14 variables traitant des différentes facettes de la défaveur (Tableau 1). Suite à cette sélection, chacune des variables a été normalisée de 0 à 1 afin de rendre comparables des unités de mesures différentes et rendre compte de l'hétérogénéité du territoire étudié. Les variables d'un même thème sont ensuite sommées, puis normalisées de nouveau. Ce traitement est renouvelé à chaque étape de construction pour obtenir au final 6 scores. Un poids similaire est affecté à chacun de ces scores car ils représentent chacun une dimension à part entière de la défaveur socio-économique. Ces derniers sont ensuite sommés et normalisés pour obtenir un indicateur final variant de 0 à 10. Plus l'indice est faible comparativement aux autres, plus le territoire est considéré comme défavorisé.

Tableau 1. Composition de l'Indice de Défavor Localisé (IDL)

Forme d'emploi	Catégorie socio-professionnelle	Part d'ouvriers actifs de 15 à 64 ans
	Emplois atypiques	Part de salariés de 15 ans ou plus en CDD ou intérim Part de salariés de 15 ans ou plus à temps partiel
Exclusion de l'emploi	Chômage	Taux de chômage* des 15-64 ans Ratio (Taux de chômage des femmes de 15 à 64 ans / Taux de chômage des hommes de 15 à 64 ans) Ratio (Taux de chômage des 15-24 ans / Taux de chômage des 25-54 ans)
		Inactivité
	Education	Scolarité
Diplôme		Taux de 15 ans ou plus non scolarisés, titulaires, au maximum d'un BEPC
Lien social	Monoparentalité	Part de familles monoparentales
	Solitude	Part de ménages composés d'hommes seuls de moins de 60 ans
Revenu	Revenu Fiscal Médian par UC**	
Logement	Surpeuplement	Part de logements occupés par plus d'une personne par pièce***

*Taux de chômage au sens du recensement

**Données 2009. Pour quatre communes du bassin le revenu des IRIS est considéré comme celui de la commune (données indisponibles à cette échelle).

*** Définition donnée par Townsend (1987)

Source : INSEE – Recensement de la population 2009 et 2010

Mesures de métaux dans les lichens et calcul de l'indice intégré

Durant l'année 2009, des thalles du lichen *Xanthoria parietina* ont été récoltés au sein de 60 stations d'échantillonnage, tirées au sort selon une grille kilométrique, pour la mesure des concentrations de 18 éléments traces (Ocelli *et al.*, 2013). Au sein de chacune des mailles, 6 à 10 thalles ont été récoltés sur des troncs à au moins 1,5 m du sol et placés dans des piluliers numérotés. Au laboratoire, les thalles ont été séchés à l'étuve (35 °C), soigneusement débarrassés des débris d'écorce et ensuite broyés dans un mortier en céramique. Tous les thalles d'une même maille ont par la suite été rassemblés en 1 seul échantillon. Les concentrations en manganèse (Mn), titane (Ti) et zinc (Zn) ont été dosées par ICP-AES (spectroscopie d'émission optique avec plasma induit par haute fréquence). Celles de l'aluminium (Al), de l'antimoine (Sb), de l'arsenic (As), du béryllium (Be), du cadmium (Cd), du cobalt (Co), du chrome (Cr), du cuivre (Cu), du mercure (Hg), du nickel (Ni), du plomb (Pb), du palladium (Pd), du platine (Pt), du rhodium (Rh) et du vanadium (V) ont été dosées par ICP-MS (spectrométrie de masse avec plasma à couplage inductif). Ce sont les concentrations totales en ces éléments qui ont été déterminées, c'est-à-dire sans tenir compte de leurs différentes spéciations au sein des thalles. Avec un rythme de croissance d'environ 1 cm par an, la récolte de thalles de 6 cm de diamètre correspond à une exposition sur les 6 dernières années environ.

Plusieurs modèles d'indices intégrés de pollution existent dans la littérature. Dérivé du calcul de l'Overall Index of Pollution (OIP : SarGaonkar & Deshpande, 2003) et de l'Integrated Pollution Index (IPI : Chen *et al.*, 2005), nous avons dans un premier temps calculé un sous-indice pour chacun des 18 éléments, représenté par le ratio de la concentration mesurées sur la valeur environnementale de fond retrouvée dans la littérature (Van Haluwyn & Cuny, 1997; Bargagli, 1998; Scerbo *et al.*, 1999, 2002; Cuny *et al.*, 2004, 2010). L'indice intégré (le Ratio d'Imprégnation Moyen : RIM) est alors la moyenne de ces 18 sous-indices. A l'inverse de l'IDL, plus le RIM est faible, moins le milieu est imprégné par les métaux. Lorsque la valeur de cet indice est supérieure à 1, le milieu est considéré comme globalement contaminé au regard des valeurs de références.

Cartographie des concentrations métalliques

Ces données de biosurveillance ont ensuite été saisies dans un Système d'Information Géographique (ArcGIS 10.1 ESRI). Une méthode géostatistique d'interpolation spatiale a été employée pour estimer les valeurs de concentrations en tous points de la zone d'étude. Nous avons pour cela utilisé le krigeage empirique bayésien (EBK). EBK diffère des méthodes classiques de krigeage par la prise en compte de l'erreur liée à l'estimation du modèle de variogramme, à l'aide d'un processus d'échantillonnage et de simulations (Krivoruchko, 2012). La couche géostatistique en sortie est une surface pixelisée. Afin d'évaluer les inégalités environnementales dans une étude écologique, les 19 cartographies géostatistiques (les 18 ETM et le RIM) ont été agrégées à l'échelle de l'Iris, par le calcul d'une valeur moyenne à partir des pixels de 100m*100m inclus dans chaque unité spatiale.

Analyse écologique

Une analyse écologique a finalement été menée pour évaluer les corrélations entre l'IDL et les charges métalliques dans les lichens. Après une première analyse statistique de nos indicateurs, des tests non-paramétriques ont été utilisés. Nous avons dans un premier temps étudié les corrélations entre l'indice de défaveur sociale, l'indice multimétallique et chacun des 18 ETM par une matrice de corrélations de Spearman. L'IDL étant une variable ordinale, une corrélation par rang peut être utilisée. Dans une seconde analyse, nous avons considéré les valeurs d'IDL des Iris selon des quintiles pour tester les différences de contaminations entre chaque catégorie de défaveur. Chaque quintile contient 20 Iris, Q1 représente les plus défavorisés et Q5 les moins défavorisés. Un test de Kurskal-Wallis, avec comparaisons multiples de Dunn a permis de comparer les couples de quintiles. Pour chacun de ces quintiles, la médiane des valeurs environnementales a en parallèle été calculée.

RESULTATS

Analyse spatiale de la défaveur sociale

La figure 1 représente la cartographie des IDL dans les différents IRIS en fonction des quintiles. Nous observons une concentration de la défaveur dans quelques territoires restreints. Effectivement, les 20 IRIS identifiées comme les plus défavorisés (cinquième quintile, en rouge sur la carte) sont regroupés sur 4 communes (Dunkerque, Fort-Mardyck, Grande-Synthe et Saint-Pol-sur-Mer). Les communes de Grande-Synthe et de Saint-Pol-sur-Mer sont particulièrement concernées. Les communes les plus favorisées se situent globalement à la marge du territoire, notamment dans les territoires de banlieue proches de la ville de Dunkerque, mais également sur le littoral touristique de Malo les Bains.

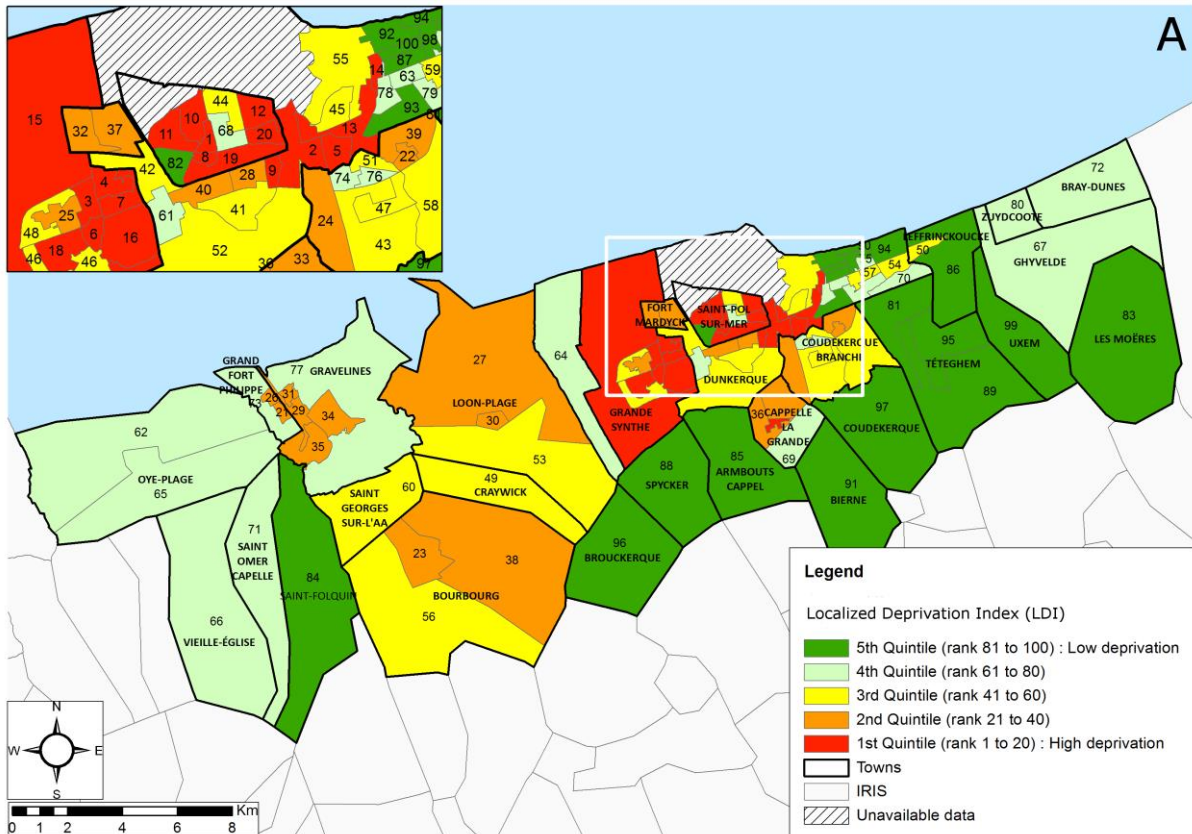


Figure 1. Défaveur sociale par Iris, représentée par l'Indice de Défavorisation Localisé (IDL)

Analyse spatiale des charges métalliques dans les lichens

Parmi les 18 éléments mesurés dans les lichens, 16 présentent des variations spatiales de concentrations. Les valeurs de Pt et Rh sont en majorité en dessous des limites de détection et des valeurs de référence, ce qui limite la variabilité et engendre une représentation spatiale uniforme. Grâce aux 16 cartographies monométalliques et à la cartographie multimétallique, nous constatons une contamination globale importante sur l'ensemble du bassin, avec des concentrations majoritairement supérieures aux valeurs régionales de fond (Occelli *et al.*, 2013). Il existe une forte variabilité spatiale de l'imprégnation de l'environnement par ces métaux. Les cartographies présentent une structure spatiale similaire représentée sur la figure 2 par l'indice intégré. Comme pour l'IDL, les Iris ont été catégorisés selon les quintiles. Les zones les plus contaminées (cinquième quintile, en rouge sur la carte) correspondent aux deux zones industrialo-portuaires du secteur. Nous y observons des niveaux d'imprégnation jusqu'à 13 fois plus élevés que les valeurs de fond observées dans les lichens. A cause de la proximité des zones urbaines, des Iris densément peuplés sont également concernés par de fortes imprégnations multimétalliques. Les concentrations métalliques dans les zones plutôt rurales en périphérie sont plus faibles (quintiles 1 et 2). Néanmoins, des valeurs

d'indice supérieures à 1 sont catégorisées dans le premier quintile (RIM entre 0.88 et 1.73), indiquant une contamination multimétallique importante sur l'ensemble du bassin. Cela justifie le choix de ce bassin pour l'étude des inégalités environnementales.

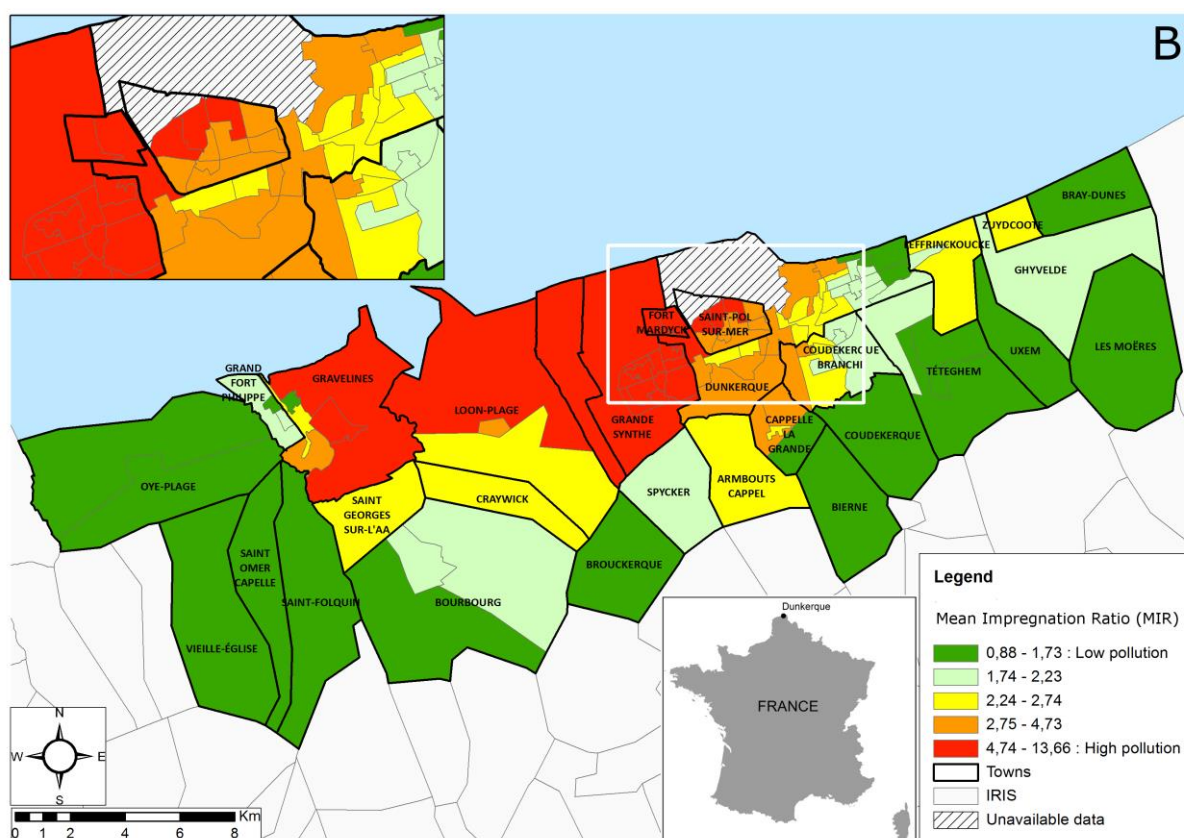


Figure 2. Charges métalliques dans les lichens par Iris, représenté par le Ratio d'Imprégnation Moyen (RIM)

Analyse écologique

Une analyse écologique a enfin été menée pour les 100 Iris habités du territoire, afin d'associer les mesures d'imprégnation environnementale et l'indice de défaveur sociale. Le tableau 2 présente pour chaque variable environnementale (18 ETM et RIM) les médianes des valeurs dans les territoires les plus défavorisés (Med-Q1) et les plus favorisés (Med-Q5), ainsi que le coefficient de corrélation de Spearman avec l'IDL. La partie droite du tableau présente les résultats du test de comparaisons multiples de Dunn. Celui-ci teste la différence entre les valeurs lichéniques, pour chaque couple de quintile d'IDL.

Tableau 2. Résultats de l'analyse écologique entre l'IDL et les charges métalliques dans les lichens (18 ETM et RIM) – pour le coefficient de Spearman, * = $p < 0.05$ et ** = $p < 0.01$ – pour le test de Dunn, les valeurs en gras sont significatives ($p < 0.005$)

	Med-Q1 lichens (mg/kg)	Med-Q5 lichens (mg/kg)	Spearman coefficient (lichens-IDL)	Test de Dunn pour chaque couple de quintiles d'IDL (valeur de p)									
				Q1 ≠ Q2	Q2 ≠ Q3	Q3 ≠ Q4	Q4 ≠ Q5	Q1 ≠ Q3	Q2 ≠ Q4	Q3 ≠ Q5	Q1 ≠ Q5	Q2 ≠ Q5	
				Q2	Q3	Q4	Q5	Q3	Q4	Q5	Q4	Q5	
Al	1 931.93	1 242.32	-0.389**	0.007	0.806	0.065	0.401	0.014	0.11	0.316	<0.001	0.449	0.001
As	2.63	1.17	-0.620**	0.068	0.97	0.013	0.295	0.062	0.012	<0.001	<0.001	<0.001	<0.001
Be	0.24	0.18	-0.375**	0.351	0.133	0.711	0.233	0.015	0.257	0.411	0.039	0.02	0.001
Cd	0.91	0.48	-0.494**	0.316	0.191	0.288	0.517	0.021	0.018	0.087	0.001	0.003	<0.001
Co	1.73	0.65	-0.601**	0.153	0.231	0.115	0.343	0.009	0.006	0.012	<0.001	<0.001	<0.001
Cr	15.31	8.42	-0.329*	0.931	0.624	0.299	0.185	0.687	0.126	0.018	0.149	0.004	0.006
Cu	23.27	14.21	-0.509**	0.026	0.744	0.047	0.386	0.057	0.096	0.004	<0.001	0.011	<0.001
Hg	0.22	0.13	-0.625**	0.224	0.389	0.064	0.15	0.038	0.007	0.001	<0.001	<0.0001	<0.001
Mn	697.9	243.46	-0.546**	0.026	0.647	0.088	0.605	0.007	0.03	0.026	<0.001	0.007	<0.001
Ni	10.27	4.35	-0.595**	0.862	0.026	0.371	0.146	0.016	0.002	0.019	0.001	<0.0001	<0.001
Pb	39.42	13.08	-0.581**	0.032	0.482	0.22	0.228	0.004	0.054	0.015	<0.001	0.002	<0.001
Pd	24.56	16.34	-0.618**	0.134	0.719	0.023	0.228	0.063	0.008	<0.001	<0.001	<0.001	<0.001
Pt	15.39	15.4	0.18	Not Significant									
Rh	15.44	15.41	0.177	Not Significant									
Sb	2.14	1.49	-0.449**	0.003	0.62	0.043	0.655	0.012	0.136	0.114	<0.001	0.278	<0.001
Ti	44.59	20.63	-0.553**	0.003	0.926	0.127	0.407	0.004	0.152	0.019	<0.001	0.024	<0.001
V	16.02	7.11	-0.595**	0.03	0.715	0.044	0.449	0.011	0.017	0.006	<0.001	0.002	<0.001
Zn	143.47	82.79	-0.519**	0.246	0.131	0.049	0.166	0.008	0.049	0.065	0.002	0.001	<0.001
RIM	4.19	1.83	-0.563**	0.062	0.496	0.155	0.273	0.011	0.035	0.012	<0.001	0.001	<0.001

Plusieurs éléments indiquent une relation entre l'IDL et les indicateurs d'imprégnation métallique. Les médianes des concentrations métalliques mesurées dans Q1 et Q5 mettent en évidence que les Iris les plus défavorisés ont les niveaux d'imprégnation métallique les plus élevés. Il existe un rapport de 3 entre ces deux valeurs concernant le plomb, de 2.9 pour le manganèse et de 2.7 pour le cobalt. Pour l'indice multimétallique, le rapport entre Q1 et Q5 est de 2.3.

L'hypothèse d'un lien entre l'IDL et la contamination métallique du milieu se vérifie à l'étude des corrélations. En effet, hormis Rh et Pt, les coefficients de Spearman sont tous significatifs ($p < 0.05$ pour Cr et $p < 0.01$ pour les autres variables). Ils indiquent une corrélation inverse, c'est-à-dire que les populations les moins favorisées se retrouvent dans les territoires les plus contaminés par les métaux. Le degré de corrélation n'est pas la même selon les éléments. Les variables Cr, Be et Al sont les moins corrélées à l'IDL, puisque les coefficients sont respectivement de -0.329, -0.375 et -0.389. En revanche, l'IDL est fortement corrélé à Hg (-0.625), As (-0.620), Pd (-0.618) et Co (-0.601). Le coefficient de corrélation se situe entre -0.5 et -0.6 pour Ni, V, Pb, Ti, Mn, Zn et Cu. Il est inférieur à -0.5 pour Cd et Sb. Le coefficient de corrélation entre les deux indicateurs (RIM et IDL) est également important (-0.563).

Ces premiers éléments nous éclairent sur le lien global pouvant relier l'IDL et les contaminants métalliques. Par ailleurs, les tests de Kurskal-Wallis et de Dunn cherchent à comparer les différentes catégories de défaveur de notre indice (de Q1 à Q5) en fonction de la contamination. Le test de Kurskal-Wallis confirme qu'il n'y a pas de différence significative entre les groupes d'IDL pour Pt et Rh. En revanche, des différences sont observées pour Be et Cr ($p < 0.05$) et pour les 15 autres variables ($p < 0.01$). Les résultats du test de Dunn indiquent ensuite une différence croissante en fonction de

l'écart entre les groupes d'IDL. De fait, les 4 premières colonnes montrent très peu de différences entre les catégories directement voisines. Seuls les territoires défavorisés Q1 et Q2 sont significativement différents pour Ti ($p=0.003$) et Sb ($p=0.003$). Dans les 3 colonnes suivantes, nous constatons des différences entre Q1 et Q3 pour Pb ($p=0.004$) et Ti ($p=0.004$), entre Q2 et Q4 pour Ni ($p=0.002$) et entre Q3 et Q5 pour As, Cu, Hg et Pd ($p<0.005$). Alors que des différences significatives sont observées pour 10 éléments entre Q2 et Q5 ($p<0.005$), cela concerne 14 éléments entre Q1 et Q4. Finalement, pour l'écart le plus important (Q1 – Q5), seule la teneur en chrome n'est pas significativement différente entre ces deux groupes. Notons à ce titre que le groupe d'Iris les plus défavorisés (Q1) est différent de chacun des autres groupes pour au moins un élément. Au contraire, la situation considérée comme moyenne (Q3) se distingue très peu des autres catégories de défaveur.

Si l'on s'intéresse plus particulièrement à l'indice multimétallique, nous remarquons que les valeurs de RIM attribuées aux territoires les plus défavorisés (Q1) sont statistiquement différentes de celles des territoires les plus favorisés Q4 ($p<0.001$) et Q5 ($p<0.001$). Cela concerne également les groupes Q2 et Q5, mais à un niveau de significativité différent ($p=0.001$). La répartition statistique des RIM dans chaque catégorie d'IDL est ensuite représentée sous la forme de boxplot (figure 3). Pour chaque catégorie d'IDL, la moyenne des RIM est supérieure à la médiane. L'écart interquartiles est également plus important entre les deuxième et troisième quartiles de RIM, qu'entre les premiers et deuxièmes quartiles. Cela nous indique la présence de valeurs élevées d'indice, donc une contamination importante de l'environnement, même pour certains Iris les plus favorisés. Nous remarquons ensuite une décroissance des valeurs de médiane et de moyenne des RIM en fonction de la favorisation des populations. Les valeurs de RIM observées pour les Iris défavorisés (Q4 et Q5) sont de manière homogène très élevées (RIM toujours 2 pour Q5), alors que celles des Iris favorisés (Q1 et Q2) sont plus faibles, avec la présence ponctuelle de valeurs extrêmes.

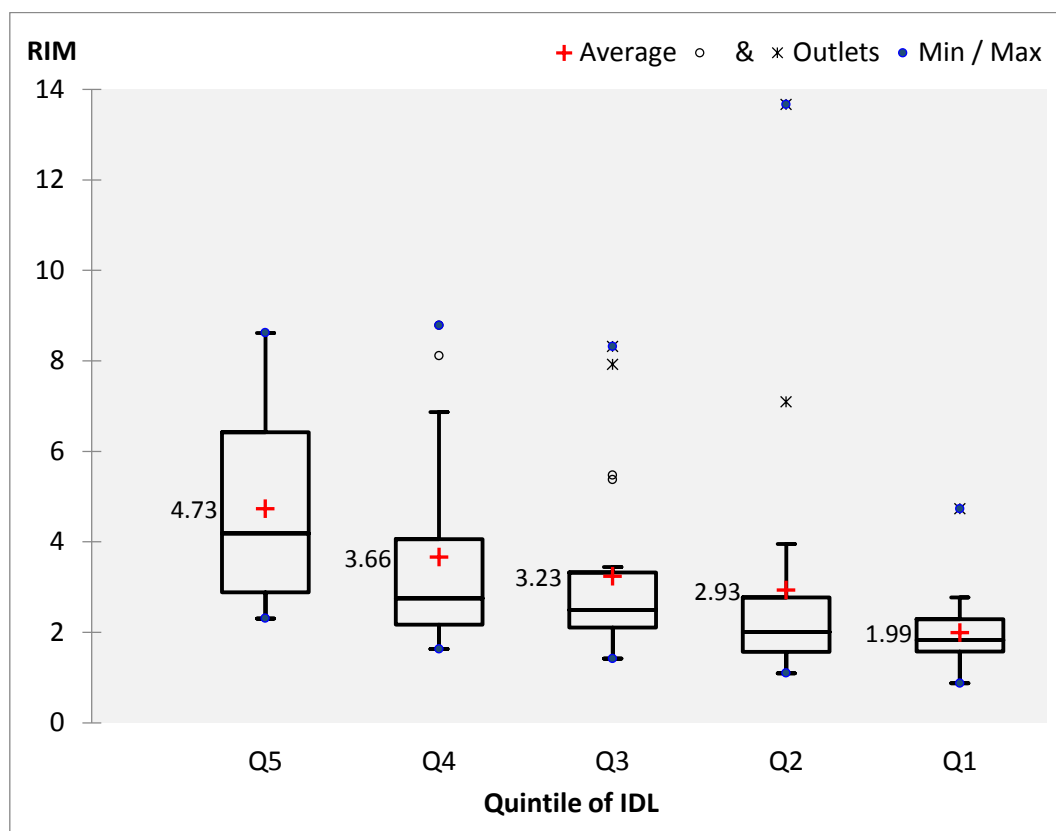


Figure 3. Distribution du RIM en fonction des quintiles d'IDL

DISCUSSION

Lors de ce travail, deux nouveaux outils ont été développés. L'IDL est un indice synthétique des conditions socio-économiques des populations à l'échelle des quartiers. Il permet de mettre en évidence des disparités spatiales de défaveur sociale. Le RIM est un indice intégré construit à partir de mesures de biosurveillance environnementale, utilisé pour décrire les variations spatiales de l'imprégnation multimétallique du milieu. Ces deux outils ont été étudiés conjointement afin d'évaluer les inégalités environnementales à l'échelle des quartiers d'un bassin industriel du Nord de la France.

L'Indicateur de Défaveur Localisé développé peut être considéré comme un indicateur composite représentant la « mesure scientifique d'un phénomène, d'un processus social, entendu au sens large, qui va au-delà de la sphère monétaire et concerne plusieurs des dimensions de la vie sociale » (Lebaron, 2011). Il vise ainsi à décrire « la vie en société » (Méda, 2002) et permet un nouveau traitement de l'analyse des inégalités à fine échelle, et de ce fait, il se veut appropriable et transposable à l'échelle de l'Iris. Cependant, dans l'objectif de tester à la fois de la robustesse de l'IDL, mais également de le critiquer, des corrélations de Spearman ont été mises en place à partir des indicateurs existants : Townsend (Townsend, 1987), Carstairs (Carstairs & Morris, 1991), Havard (Havard *et al.*, 2008), ISS, IDH-2 et PIB/hab (Région Nord-Pas-de-Calais, 2008). Cette corrélation a été testée à une échelle fine (l'Iris) et à une échelle plus large (la région). L'IDL est significativement corrélé avec chacun des six indicateurs testés. Par exemple, la corrélation atteint -0.891 pour Townsend et 0.612 pour l'ISS ($p < 0.05$). Par ailleurs, les corrélations importantes entre les différentes variables mobilisées dans notre indicateur démontrent la dimension cumulative de la défaveur. Cela nous permet ainsi de créer de réels profils de populations et d'identifier des caractéristiques particulières à chacun des territoires.

La spatialisation des conditions socio-économiques semble trouver écho, en partie, dans la construction historique du territoire de Dunkerque et plus précisément dans son héritage industriel. Cela concerne notamment la commune de Grande-Synthe. Localisée à proximité des grandes industries historiques qui ont engendré des flux massifs de population ouvrière à la fin des années 1960. Ces populations étaient alors logées dans du parc social proche (Castells & Godard, 1974). Néanmoins, un double processus va peu à peu paupériser la population de ces quartiers : un déplacement des ouvriers vers le péri-urbain et le déclin de l'industrie. Ces phénomènes vont entraîner un délaissement du territoire et un non renouvellement de la main d'œuvre. Dans l'objectif d'éviter la vacance des logements sociaux, des populations davantage en difficultés socio-économiques y ont été installées (Bondue, 1986). Ces populations très défavorisées sont encore aujourd'hui présentes. Bien sûr, ces données ne sauraient, à elles seules, constituer un argumentaire complet quant à la spatialisation de la défaveur, mais il est important de tenir compte des liens pouvant exister entre histoire industrielle et caractérisation socio-économique des populations.

Les populations les plus favorisées se situent essentiellement dans les banlieues proches de la ville de Dunkerque et de ses aménités, permettant de concilier plus aisément cadre de vie plus agréable et vie professionnelle. Cependant, si ces premiers constats laissent à penser une distinction nette entre urbain et banlieue, le quartier de Malo-les-Bain à Dunkerque fait figure de contre-exemple. Quartier très urbanisé parmi les plus favorisés, sa condition tient surtout dans l'histoire de son aménagement qui en a fait une station balnéaire reconnue le long du littoral, attirant par là même des populations relativement avantagées socio-économiquement.

En parallèle, nous avons pu mettre en évidence des disparités spatiales d'imprégnation de l'environnement par les éléments traces accumulés au sein des lichens (Ocelli *et al.*, 2013). En complément de la surveillance physico-chimique de l'air, les concentrations mesurées chez les lichens représentent la fraction biodisponible, accumulée de la contamination atmosphérique sur le long terme. Sur le bassin de Dunkerque, les secteurs industrialo-portuaires, situés à proximité des zones densément peuplées, correspondent aux deux zones les plus contaminées, contrairement à la périphérie. Ces résultats confirment ceux précédemment obtenus sur ce territoire (Cuny *et al.*, 2004, 2010).

Dans ce contexte de contamination multimétallique, un indice intégré a été utilisé pour représenter le niveau global de contamination pour l'ensemble des 18 éléments, en rapport aux concentrations naturelles retrouvées dans les lichens. Afin d'évaluer le niveau de fiabilité de cet indice, plusieurs critères ont été recensés dans la littérature (Caeiro *et al.*, 2005 ; Cairncross *et al.*, 2007 ; Singh *et al.*, 2008 ; Bharti & Katyal, 2011). De fait, un indice composite est considéré comme pertinent lorsqu'il :

- minimise la surestimation (ambiguïté) ou la sous-estimation (éclipse) de la contamination ;
- est simple de calcul, crédible et transparent ;
- est relativement facile à utiliser et à comprendre ;
- est à la fois sensible et robuste à la variabilité des paramètres ;
- est une image représentative d'un bon ou mauvais état qualitatif du milieu ;
- apporte un faible niveau d'incertitudes.

L'indice développé dans cette étude est dérivé de l'integrated pollution index (IPI: Chen *et al.*, 2005) et de l'overall index of pollution (OIP: SarGaonkar & Deshpande, 2003). Récemment, Bharti & Katyal (2011) classaient l'OIP parmi les 8 indicateurs les plus simples, basiques et performants en termes de calcul. La traduction de notre indice est relativement simple du fait de sa comparabilité aux valeurs de fond (représentées par la valeur 1). Avec un coefficient de corrélation de -0.563, le RIM est la 8^e variable environnementale la plus corrélée à l'IDL (tableau 2). Cette position centrale au sein des 18 ETM confirme sa représentativité en termes de pollution globale sur le territoire. La structure spatiale des cartographies monométalliques montre que les éléments traces sont fortement corrélés entre eux. Cette corrélation indique une certaine robustesse de l'indice (pas d'ambiguïté ou d'éclipse).

Nos analyses montrent la réalité de l'existence d'inégalités environnementales sur le bassin dunkerquois. Celles-ci sont davantage marquées pour les populations antagonistes, même si de manière globale, plus la faveur d'un territoire augmente, moins son imprégnation par l'ensemble des éléments est importante. Cette répartition laisse apparaître des quartiers défavorisés, concentrés dans des périmètres restreints et urbains, hérités de l'histoire industrielle du territoire. L'imprégnation de l'environnement de ces derniers est majoritairement le fruit de cette activité industrielle, notamment en bordure du littoral.

Ce lien fort entre population défavorisée et industrie affecte sur notre démarche en rendant délicate la notion d'injustice environnementale. En effet, la relation définie aux Etats-Unis repose davantage sur les « forces du marché » (Laurian, 2008), avec des industries qui s'installent sur des territoires au foncier avantageux où la main d'œuvre est facilement accessible. Or, cette situation ne se vérifie pas dans le bassin de Dunkerque, où au contraire, les industries sont primo-arrivantes, de part notamment la localisation en bord de mer et la facilité des échanges internationaux. Le territoire et son urbanisation se sont en partie construits en réponse à cette situation, avec notamment la

création de logements sociaux. Comme évoqué par Beck (1986), nous constatons bien que « les problèmes environnementaux sont fondamentalement basés sur la façon dont la société est organisée ». Lorsqu'une industrie s'installe, elle nécessite beaucoup de main d'œuvre et engendre un flux de population. Il est donc nécessaire de « situer ce phénomène dans une vision plus large de la production sociale d'inégalités et de la dégradation de l'environnement » (Brulle & Pellow, 2006). Si cette dynamique ouvrière n'est plus d'actualité, les phénomènes de ségrégation sont tels que la mobilité reste quasi-inexistante pour les populations en difficulté, vivant à proximité des usines et donc plus exposées à la pollution atmosphérique métallique (figure 4). Face à ces populations cumulant défaveur et exposition, les populations les plus favorisées se situent quant à elles davantage en marge de la zone portuaire et sont par ailleurs nettement moins exposées à un environnement extérieur contaminé par les métaux (figure 4).

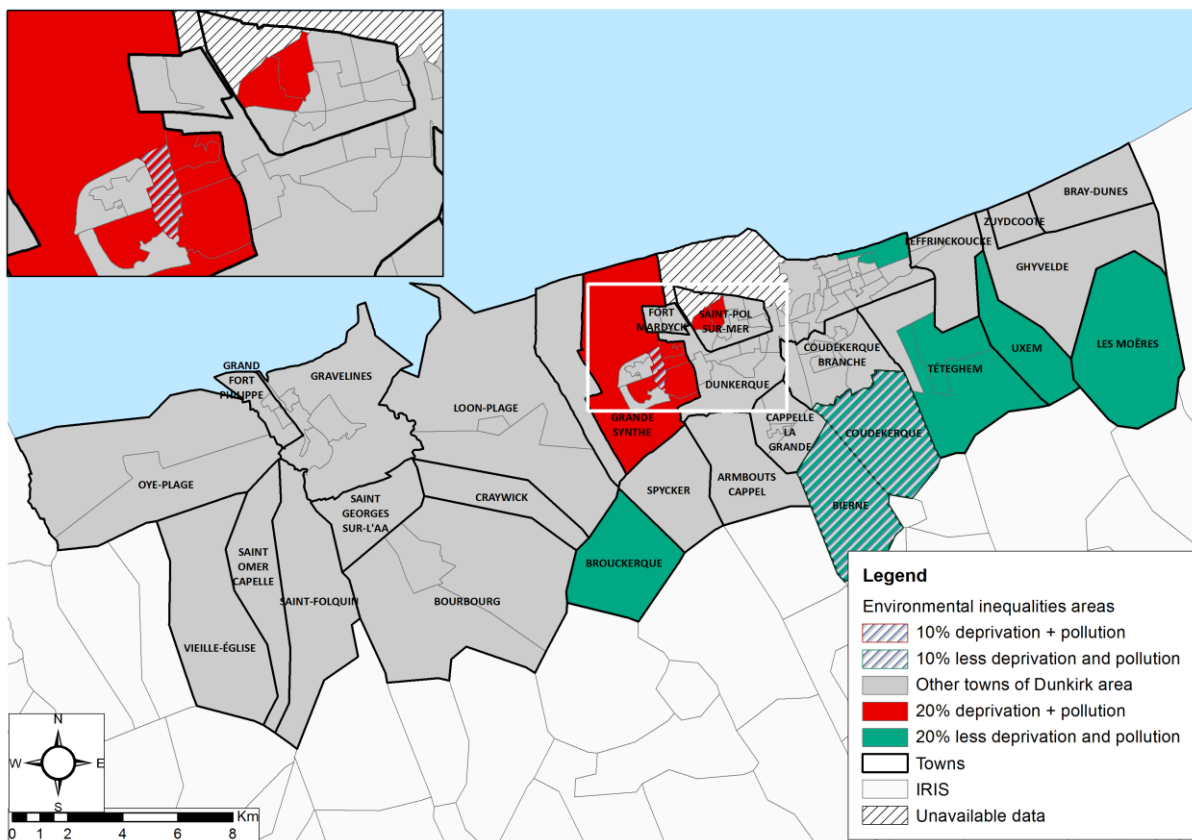


Figure 4. Inégalités environnementales à l'échelle des Irises du bassin de Dunkerque

CONCLUSION

Deux indices ont été construits pour cette étude. L'IDL est un indice de défaveur sociale construit sur la base d'une approche éco-sociologique, permettant de catégoriser les populations à l'échelle des quartiers selon leur situation socio-économique et transposable à d'autres territoires. En parallèle, nous avons calculé un indice intégré de pollution multimétallique à partir des concentrations mesurées dans les lichens. Du fait de l'exposition des populations à l'ensemble des substances présentes dans leur environnement, l'utilisation de tels indices s'avère indispensable pour réfléchir en termes de pollution globale. Le RIM satisfait plusieurs critères répertoriés dans la littérature, ce qui lui confère une bonne représentativité de la pollution globale.

Notre étude révèle la présence d'inégalités environnementales à l'échelle des quartiers du bassin industriel de Dunkerque. Les populations les plus défavorisées vivent dans un environnement fortement imprégné par les éléments traces. La structure du bassin dunkerquois fait que ces populations logent dans la ville-centre, située à proximité des zones industrielles.

L'utilisation de modèles biologiques de surveillance des milieux dans le cadre des inégalités environnementales est extrêmement rare dans la littérature. La biosurveillance environnementale, telle que la mesure de bioaccumulation des métaux peut aider à l'interprétation de données de santé, notamment pour l'identification de facteurs étiologiques des maladies dans le cadre d'études épidémiologiques. Dans un futur proche, cette analyse sera réitérée pour d'autres bassins de vie, afin d'évaluer la variabilité des résultats en fonction de la typologie et de la structure de la zone étudiée. Il s'agit en l'occurrence du bassin urbain de Lille et du bassin rural de Maubeuge, tous deux situés dans la région Nord de la France (Ocelli *et al.*, 2013).

REFERENCES

- Balabanova B, Stafilov T, Šajn R, Baèeva K. 2012. Characterisation of heavy metals in lichen species hypogymnia physodes and evernia prunastri due to biomonitoring of air pollution in the vicinity of copper mine. *International Journal of Environmental Research*, 6(3): 779-792.
- Bargagli R. 1998. *Trace Elements in Terrestrial Plants, an Ecophysiological Approach to Biomonitoring and Biorecovery*. Springer, Berlin: 324p.
- Beck U. 1986. *La société du risque : sur la voie d'une autre modernité*, trad. de l'allemand par Laure Bernardi, Paris : Flammarion, ed. 2003 : 521p.
- Bharti N, Katyal D. 2011. Water quality indices used for surface water vulnerability assessment. *International Journal of Environmental Sciences*, 2(1): 154-173.
- Boamponsem LK, Adam JI, Dampare SB, Nyarko BJB, Essumang DK. 2010. Assessment of atmospheric heavy metal deposition in the tarkwa gold mining area of ghana using epiphytic lichens. *Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms*, 268(9): 1492-1501.
- Bocquier A, Cortaredona S, Boutin C, Bigot A, David A, Verger P. 2011. Bruit routier et statut socio-économique : étude des inégalités environnementales à un niveau géographique fin au sein de la commune de Marseille. *Environnement, Risques & Santé*, 10 (3): 225-229.
- Bondue J.-P., 1986, « Vacances du logement et différenciations sociales : l'exemple de Grande-Synthe dans l'agglomération dunkerquoise », *Espace, Population et Sociétés*, n° 1, p. 73-83.
- Booyesen F. 2002. An overview and Evaluation of Composite Indices of Developpement. *Social Indicators Research*, 59 (2): 115-151.
- BruitParif. 2013. *Système d'Information Géographique développé en vue du croisement des caractéristiques socio-économiques des populations set des données environnementales (air, bruit) autour des grandes plateformes aéroportuaires franciliennes*. Note de synthèse, 13p.
- Brulle RJ & Pellow DN. 2006. Environmental justice: Human health and environmental inequalities. *Annual Review of Public Health*, 27: 103-124.
- Bullard RD. 2005. Environmental Justice in the Twenty-first Century. In RD Bullard (Ed.). *The Quest of Environmental Justice*, p 19-42, Sierra Club Books.
- Bullard RD. 1993. Race and Environmental Justice in the United States. *Yale Journal International Law*, 18 (1): 319-335.
- Caeiro S, Costa MH, Ramos TB, Fernandes F, Silveira N, Coimbra A, Medeiros G, Painho M. 2005. Assessing heavy metal contamination in Sado Estuary sediment: An index analysis approach. *Ecological Indicators*, 5: 151-169.
- Cairncross EK, John J, Zunckel M. 2007. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. *Atmos. Environ*, 41: 8442-8454.
- Carstairs V & Morris R. 1990. Deprivation and health in Scotland. *Health Bull*, 48 (4): 162-175.
- Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, Merlo J. 2006. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *Journal of Epidemiology and Community Health*, 60 (3): 234-241.
- Chakraborty J & Armstrong MP. 1997. Exploring the use of buffer abalysis for the identification of impacted areas in environmental equity assessment. *Cartography and Geographic Information Systems*, 24 (3): 145-157.

- Chen TB, Zheng YM, Lei M, Huang ZC, Wu HT, Chen H, Fan KK, Yu K, Wu X, Tian QZ. 2005. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere*, 60: 542–551.
- Cloquet C, De Muynck D, Signoret J, Vanhaecke F. 2009. Urban/peri-urban aerosol survey by determination of the concentration and isotopic composition of Pb collected by transplanted lichen. *Hypogymnia physodes*. *Environmental Science and Technology*, 43 (3): 623-9.
- Cressie NAC. 1993. *Statistics for spatial data*. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. John Wiley & Sons Inc., New York. Revised reprint of the 1991 edition, A Wiley-Interscience Publication.
- Cuny D, Davranche L, Thomas P, Kempa M, Van Haluwyn C. 2004. Spatial and Temporal Variations of Trace Element Contents in *Xanthoria Parietina* Thalli Collected in a Highly Industrialized Area in Northern France as an Element for a Future Epidemiological Study. *Journal of Atmospheric Chemistry*, 49: 391–401.
- Cuny D, Quarre S, Devred I, Jeanne B, Reynaert A, Rzepka-Cuny MA. 2010. Évolution spatiale et temporelle des concentrations en Éléments Traces Métalliques de thalles de *Xanthoria parietina* récoltés sur la zone industrialo-portuaire de Dunkerque. *Air Pur*, 77 : 43-49.
- Cuny D, Rzepka MA, Van Haluwyn C. 2008. Introduction à la biosurveillance végétale et fongique de la qualité de l'air, concept et intérêt. *Spectra Analyse*, 264: 31-35.
- Cuny D, Davranche L, Van Haluwyn C, Plaisance H, Caron B, Malrieu V. 2002. Monitoring ozone by using tobacco, automatic network and passive samplers in an industrial area in France. In A. Klump, W. Ansel and G. Klump (eds.). *Urban Air Pollution Bioindication and Environmental Awareness*, Cuvillier Verlag, Göttingen: 97–108.
- Cuny D, Van Haluwyn C, Pesch R. 2001. Biomonitoring of trace elements in air and soil compartments along the major motorway in France. *Water, Air, and Soil Pollution*, 125: 273-289.
- DETR. 2000. *Indice of deprivation 2000*. London: Department of the Environment, Transport and the Regions.
- Environmental Protection Agency : www.epa.gov/environmentaljustice/ (12.06.2014).
- Faburel G & Maleyre I. 2007. Le bruit des avions comme facteur de dépréciations immobilières, de polarisation sociale et d'inégalités environnementales. Le cas d'Orly. *Développement durable et territoires* [En ligne], Dossier 9 mis en ligne le 10 janvier 2007, consulté le 26 juin 2014. URL : <http://developpementdurable.revues.org/2775> ; DOI : 10.4000/developpementdurable.2775.
- Faburel G & Gueymard S. 2008. Inégalités environnementales en région Ile-de-France : le rôle structurant des facteurs négatifs de l'environnement et des choix politiques afférents. *Revue Espace-Populations-Sociétés*, 1: 159-172.
- Flanquart H, Hellequin A-P, Vallet P. 2013. Living alongside hazardous factories: risk, choice and necessity. *Health, Risk & Society*, 15 (8) : 663-680.
- Garrec JP & Van Haluwyn C. 2002. *Biosurveillance végétale de la qualité de l'air*, Editions Tec &Doc, Paris, 117p.
- Garty J. 1993. Lichens as biomonitors for heavy metal pollution. In B. Market (ed.). *Plants as Biomonitors, Indicators for Heavy Metals in the Terrestrial Environment*, VCH, Weinheim: 193–263.
- Garty J. 2001. Biomonitoring atmospheric heavy metals with lichens: Theory and application. *Crit. Rev. Plant Sci.*, 20(4): 309–371.
- Glickman TS. 1994. Measuring environmental equity with geographic information systems. *Renewable Resources Journal*, 12 (3): 17-21.

- Golge M, Yenilmez F, Aksoy A. 2013. Development of pollution indices for the middle section of the Lower Seyhan Basin (Turkey). *Ecological Indicators*, 29: 6–17.
- Guéguen F, Stille P, Geagea ML, Boutin R. 2012. Atmospheric pollution in an urban environment by tree bark biomonitoring – Part I: Trace element analysis. *Chemosphere*, 86: 1013–1019.
- Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. 2008. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science and Medicine*, 67(12): 2007–2016. INSEE, Recensement de la population, 2010.
- Jany-Catrice F & Zotti R. 2008. Les régions françaises face à leur santé sociale, contribution au débat. Institut pour le développement de l'information économique et sociale.
- Krivoruchko K. 2012. Empirical Bayesian Kriging implemented in ArcGIS Geostatistical Analyst. *ArcUser*, 15 (4): 6-10.
- Laurian L. 2008. Environmental Injustice in France. *Journal of Environmental Planning and Management*, 51 (1): 55-79.
- Lebaron F. 2011. Les indicateurs sociaux au XXI^e siècle, coll. « Les Topos », Dunod, 120p
- Loppi S, Nelli L, Ancora S, Bargagli R. 1997. Accumulation of trace elements in the peripheral and central parts of a foliose lichen thallus. *Bryologist*, 100: 251–253.
- Loppi S & Nascimbene J. 2010. Monitoring H₂S air pollution caused by the industrial exploitation of geothermal energy: The pitfall of using lichens as bioindicators. *Environmental Pollution*, 158: 2635-2639.
- Loppi S, Frati L, Paoli L, Bibabli V, Rossetti C, Bruscoli C, Corsini A. 2004. Biodiversity of epiphytic lichens and heavy metal contents of *Flavoparmelia caperata* thalli as indicators of temporal variations of air pollution in the town of Montecatini Terme (central Italy). *Science of the Total Environment*, 326: 113–122.
- Madejon P, Maranon T, Murillo JM. 2006. Biomonitoring of trace elements in the leaves and fruits of wild olive and holm oak trees. *Science of the Total Environment*, 355: 187–203.
- McLaren D, Cottray O, Taylor M, Pipes S, Bullock S. 1999. The geographic relation between household income and polluting factories. Friends of the Earth, London.
- McLeod H, Langford IH, Jones AP, Stedman JR, Day RJ, Lorenzoni I, Batement IJ. 2000. The relationship between socio-economic indicators and air pollution in England and Wales: implications for environmental justice. *Regional Environmental Change*, 1 (2): 78-85.
- Meda D. 2002. Le capital social : un point de vue critique. *L'économie politique*, 14 : 36-47.
- Ministère de l'écologie et du développement durable et ministère des transports de l'équipement du tourisme et de la mer. 2003. Appel à Proposition de Recherche "Politiques territoriales et développement Durable".
- Mitchell G. 2005. Forecasting environmental equity: air quality responses to road user charging in Leeds, UK. *Journal of Environmental Management*, 77 (3): 212-226.
- Nieboer E, Richardson DHS, Tomassini FD. 1978. Mineral uptake and release by lichens - overview. *Bryologist*, 81: 226-246.
- Ocelli F, Cuny M-A, Devred I, Deram A, Quarré S, Cuny D. 2013. Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques. *Pollution atmosphérique [En ligne]*, 220 mis à jour le : 13/01/2014, URL : <http://lodel.irevues.inist.fr/pollution-atmospherique/index.php?id=2497>.
- Perret B. 2003. Indicateurs sociaux et théorie sociale. *Revue Mauss*, 21 : 261-275.

- Région Nord-Pas-de-Calais, coord. Grégory Marlier. 2008. Un indicateur de santé sociale pour les régions françaises. Note D2PE n°20 : 20p.
- République française. 2005. Charte de l'environnement, Loi constitutionnelle n° 2005-205 du 1^{er} mars 2005.
- Rzepka MA, Cuny D. 2008. Biosurveillance végétale et fongique des éléments traces métalliques atmosphériques. *Air Pur*, 75: 66—77.
- Salo H, Bucko MS, Vaahtovuori E, Limo, J, Mäkinen J, Pesonen LJ. 2012. Biomonitoring of air pollution in SW Finland by magnetic and chemical measurements of moss bags and lichens. *Journal of Geochemical Exploration*, 115: 69–81.
- Sargaonkar A & Deshpande V. 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environmental Monitoring and Assessment*, 89: 43-67.
- Scerbo R, Possenti L, Lampugnani L, Ristori T, Barale R, Barghigiani C. 1999. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Livorno Province (Tuscany, Italy). *The Science of the Total Environment*, 241: 91-106.
- Scerbo R, Ristori T, Possenti L, Lampugnani L, Barale R, Barghigiani C. 2002. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Pisa Province (Tuscany, Italy). *The Science of the Total Environment*, 286: 27-40.
- Scottish Executive. 2004. Scottish index of Multiple Deprivation 2004, Scottish Executive Edinburgh.
- Sheppard E, Leitner H, McMaster RB, Hongguo T. 1999. GIS based measures of environmental equity: exploring their sensitivity and significance. *Journal of Exposure Analysis and Environmental Epidemiology*, 9: 18-28.
- Singh RP, Nath S, Prasad SC, Nema AK. 2008. Selection of Suitable Aggregation Function for Estimation of Aggregate Pollution Index for River Ganges in India. *Journal of Environmental Engineering*, 134 (8): 689-701.
- Theys J & Emelianoff C. 2001. Les contradictions de la ville durable. *Le Débat*, 113: 122-135.
- Townsend P. 1987. Deprivation. *Journal of Social Policy*, 16 (2): 125-146.
- Van Haluwyn C & Cuny D. 1997. Cartographie de la pollution de l'air par certains métaux lourds sur le littoral Calais-Dunkerque. *Rapport SPPPI du Littoral*: 25-40.
- Viel JF, Hagi M, Upegui E, Laurian L. 2011. Environmental injustice in a French industrial region: are polluting facilities equally distributed? *Health & Place*, 17 (11): 257-262.
- Walker G, Mitchell G, Fairburn J, Smith G. 2005. Industrial pollution and social deprivation: evidence and complexity in evaluating and responding to environmental inequality. *Local Environment*, 10 (4): 361-377.
- World Health Organization. 2012. Environmental health inequalities in Europe. Copenhagen, OMS Europe. 190p.

Partie 4. Inégalités environnementales de santé

a. Introduction

La mesure de substances chimiques ou de leurs métabolites dans les liquides biologiques, appelée biométrie, a pris une place privilégiée dans la mesure de l'exposition des populations. Elle fournit des informations sur l'exposition récente ou la dose accumulée dans l'organisme, en tenant compte de l'ensemble des voies d'exposition, qu'elles soient professionnelles, alimentaires, domestiques ou environnementales et en intégrant toutes les voies d'absorption (Nisse, 2010). De fait, la biométrie est devenu un outil indispensable à l'évaluation et la gestion des risques (Albertini *et al.*, 2006), permettant par exemple d'identifier les sources responsables des hauts niveaux d'exposition (Angerer, 2007) et d'étudier l'évolution de l'imprégnation des populations dans le temps (Paustenbach & Galbraith, 2006a ; 2006b).

L'étude IMEPOGE (Imprégnation par les Métaux de la Population Générale) menée par l'équipe du Dr. Catherine Nisse (EA4483 Université Lille 2 - Centre de Recherches en Santé - Travail - Environnement – CHRU de Lille) vise à évaluer l'imprégnation de populations par différents métaux, ainsi que les variations éventuelles en fonction des caractéristiques individuelles : âge, sexe, expositions d'origine professionnelle, domestique, ou environnementale (Nisse, 2010). Cette étude est basée sur la mesure des concentrations sanguine et urinaire dans un échantillon de population générale fréquentant les Centres d'exams de Santé du Nord Pas de Calais et l'administration d'un auto-questionnaire auprès des volontaires. Le recueil des données fut réalisé entre 2008 et 2010.

Dans ce premier chapitre de thèse, nous avons mis en évidence des variations spatiales de l'imprégnation des lichens par différents éléments au sein de trois territoires de la région Nord – Pas de Calais (Ocelli *et al.*, 2013). En se focalisant sur le bassin de Dunkerque, nous avons ensuite souligné la présence de populations défavorisées sur les territoires les plus imprégnés par les ETM (Ocelli *et al.*, à soumettre). L'Homme étant exposé à son environnement d'une manière globale, nous avons également développé un indice d'imprégnation multimétallique des lichens appelé RIM. Celui-ci peut être appliqué à différentes matrices biologiques, telles que le sang et les urines.

Cette étude s'intéresse aux inégalités environnementales de santé et propose d'évaluer les niveaux d'imprégnation d'un échantillon de la population générale et de lichens épiphytes, pour différents métaux sur un territoire commun. L'objectif est de mettre en relation ces données afin d'évaluer si les populations les plus imprégnées par les ETM vivent dans un environnement fortement contaminé par ces éléments. Dans cette partie, nous travaillons au niveau individuel, car les indicateurs environnementaux ont été estimés pour chacun des sujets grâce aux outils SIG. Les corrélations entre les concentrations mesurées dans les trois matrices (sang, urines, lichens) sont étudiées en fonction des différents éléments (11) et de l'indice multimétallique (RIM). Dans la continuité de la partie précédente, cette analyse se focalise sur le bassin dunkerquois. Elle intègre à ce jour deux facteurs confondants (l'âge et le tabagisme).

b. Matériel et méthodes

AIRES D'ETUDE ET SOURCES DES DONNEES

Nous avons travaillé à partir des données lichéniques présentées dans les parties précédentes. Les données populationnelles sont issues d'une étude descriptive transversale réalisée entre 2008 et 2010. Celle-ci vise à décrire l'imprégnation par certains métaux, par le biais de l'analyse de ceux-ci dans le sang et dans les urines sur un échantillon de 2000 personnes issues de la population générale du Nord – Pas de Calais se présentant dans les Centres d'exams de santé (Nisse, 2010). Afin de disposer de données issues des trois matrices sur un même territoire, les individus intégrés à cette étude ont été sélectionnés en fonction du bassin d'échantillonnage des lichens et de leur lieu d'habitation. Parmi les 2000 sujets de l'étude IMEPOGE, 292 vivaient dans le bassin de Dunkerque au moment des prélèvements.

Parmi les éléments recherchés dans les trois matrices, 11 ont été mesurés à la fois chez l'Homme et dans les lichens. Il s'agit de l'aluminium (Al), de l'arsenic (As), du cadmium (Cd), du cobalt (Co), du chrome (Cr), du mercure (Hg), du manganèse (Mn), du nickel (Ni), du plomb (Pb), du vanadium (V) et du zinc (Zn). Les autres informations disponibles pour caractériser les individus intégrés à l'analyse sont la classe d'âge (moins de 30 ans ; entre 30 et 39 ans ; entre 40 et 49 ans ; 50 ans et plus) et le statut tabagique (fumeur / non-fumeur).

ANALYSE SPATIALE ET CREATION DE LA BASE DE DONNEES

A partir des 11 ETM dosés, le RIM fut généré pour chacune des trois matrices (lichens : RIML ; sang : RIMs ; urines : RIMu). Pour le calcul des RIM sanguin et urinaire, les valeurs de références utilisées sont celles définies dans l'étude IMEPOGE et correspondent aux percentiles 95 des résultats globaux de l'enquête régionale (Nisse, 2010). Ces valeurs de référence sont répertoriées dans le tableau 11.

Tableau 11. Valeurs de référence utilisées pour le calcul des indices multimétalliques d'imprégnation

Matrice	Al	As	Cd	Co	Cr	Hg	Mn	Ni	Pb	V	Zn
Sang (µg/L)	11.24	6.72	1.67	0.54	1.26	5.06	12.94	2.67	49.35	0.14	7272
Urines (µg/L)	11.54	130.83	1.33	1.89	1.6	6.6	1.07	5.99	3.76	1.21	1039

Les cartographies lichéniques ont été générées à l'aide d'une méthode géostatistique d'interpolation spatiale du logiciel ArcGIS® : le krigeage empirique bayésien (voir chapitre I, Partie 2). A partir de ces cartographies, les valeurs environnementales de concentration en ETM et de RIM ont été attribuées à chacun des 292 individus, en fonction des coordonnées spatiales du lieu d'habitation. Cette opération fut également réalisée à l'aide de l'outil Geostatistical Analyst d'ArcGIS®. Ainsi, pour chaque individu, nous disposons de 38 variables reprenant :

- les mesures de concentrations en 11 éléments traces pour chacune des trois matrices ;
- le RIM pour chacune des trois matrices (RIML, RIMs, RIMu) ;
- la classe d'âge ;
- le statut tabagique (fumeur ou non).

ANALYSE STATISTIQUE

Avant d'étudier les corrélations entre les mesures sanguines, urinaires et lichéniques, une analyse critique des données populationnelles fut réalisée, notamment en vue d'évaluer l'influence des facteurs que sont l'âge et le tabac sur les niveaux d'imprégnation. Pour chacune de ces variables, nous avons utilisé un test de comparaison de moyennes afin d'identifier une différence significative entre les groupes d'individus. Les données populationnelles n'étant pas distribuées sous forme gaussienne, même après transformation logarithmique, le test non-paramétrique de Mann et Whitney a été utilisé, avec un risque alpha de 5%. Concernant le facteur âge, cette comparaison fut réalisée entre les deux classes d'âge extrêmes, composées de 74 individus de moins de 30 ans et de 55 individus de plus de 50 ans.

Une analyse de corrélations entre les trois matrices a ensuite été réalisée, en intégrant les facteurs confondants précédemment évoqués. La matrice de corrélation de Spearman a pour cela été générée pour mesurer le niveau de corrélation entre les 36 variables métalliques. Le coefficient de Spearman (CS) ne mesure pas la corrélation entre les valeurs brutes de concentrations des individus, mais entre les rangs qu'elles occupent au sein de la distribution. Les données n'étant pas de forme gaussienne, ce dernier représente l'avantage d'être moins influencé par les valeurs extrêmes que le coefficient de Pearson. Les analyses statistiques ont été réalisées à l'aide du logiciel XLSTAT.

c. Résultats

STRUCTURES DES DONNEES POPULATIONELLES

Une première phase consistait à étudier la structure et le contenu des données populationnelles, afin de valider la base utilisée pour la suite de l'étude. Nous avons mis en évidence plusieurs critères de traitement de la base. Le premier concerne son renseignement. Alors que des mesures sanguines ont été effectuées pour l'ensemble des individus, les analyses urinaires ne sont pas recensées pour 12 d'entre eux (4%). Ces derniers ont été supprimés de la base.

Le deuxième critère est basé sur la distribution statistique des variables sous la forme d'un nuage de points. Des valeurs extrêmes peuvent être observées, comme le montre la figure 27. Les individus inclus dans notre étude sont issus d'un échantillon de population générale. Il n'est donc pas exclu que des concentrations très élevées pour certains métaux dans les fluides biologiques soient directement issues de traitements médicaux (le cobalt présent dans les prothèses par exemple) ou d'une exposition professionnelle récente à forte dose. De telles observations risquent de masquer l'influence de l'environnement dans notre étude de corrélations. Un test de valeurs extrêmes a donc été mis en œuvre pour identifier les valeurs statistiquement aberrantes. Le test de Grubbs unilatéral à droite (à un niveau de significativité de 1%), confirme la présence de valeurs statistiquement aberrantes pour plusieurs éléments (figure 27). Les individus correspondants sont différents en fonction des métaux. Par souci de conserver un échantillon d'individus suffisamment grand pour la suite de l'analyse, seules les valeurs extrêmes ont été supprimées de la base. Les individus en question sont conservés pour les autres éléments. Au final, 280 sujets sur les 292 sont conservés pour la suite de l'étude.

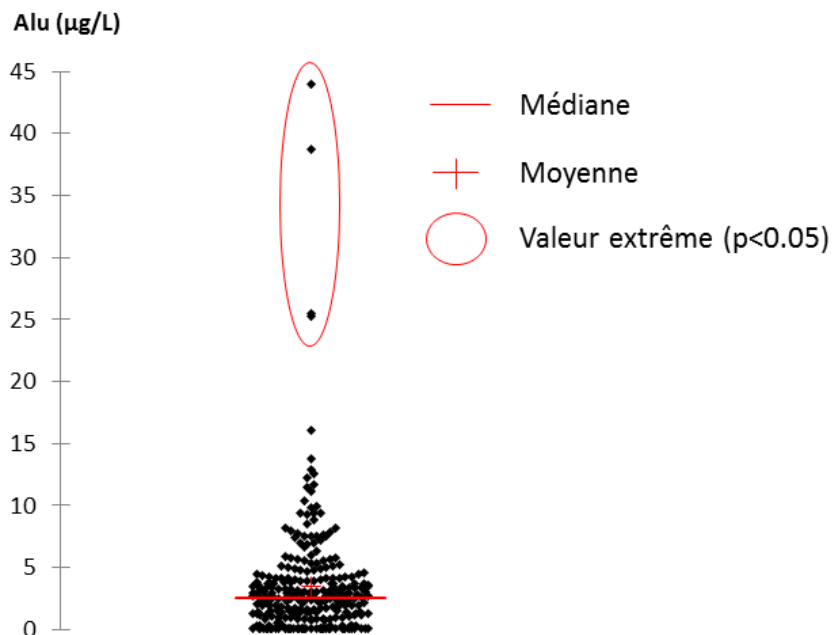


Figure 27. Exemple de distribution statistique des concentrations (en µg/L) dans les urines pour l'aluminium (Alu)

ETUDE DES FACTEURS CONFONDANTS

Avant d'étudier les corrélations entre l'imprégnation métallique des populations et de l'environnement, nous nous sommes intéressés à certains facteurs confondants. Le premier concerne le statut tabagique des individus. La fumée de cigarette contenant des métaux, les fumeurs présentent généralement des concentrations biologiques plus élevées que les non-fumeurs. Nous avons donc effectué un test de comparaison de moyennes entre ces deux groupes de population. Le test bilatéral de Mann et Whitney (tableau 12) montre des différences significatives de concentrations entre les fumeurs (n = 97) et les non-fumeurs (n = 183) pour certains éléments.

Tableau 12. Résultats du test Mann et Whitney de comparaison de moyennes pour les ETM sanguins (s) et urinaires (u) entre les fumeurs et les non-fumeurs (NS : non-significatif ; *** : significatif) – les valeurs en gras sont significativement supérieures, à un risque alpha de 0.05

Variable	Différence (p<0.05)	Fumeurs (µg/L) Moyenne ± écart-type	Non-Fumeurs (µg/L) Moyenne ± écart-type
Als	NS	3.38 ± 2.60	2.97 ± 2.37
Ass	NS	2.45 ± 2.77	2.63 ± 2.32
Cds	***	1.06 ± 0.64	0.32 ± 0.20
Cos	NS	0.27 ± 0.09	0.29 ± 0.09
Crs	NS	0.63 ± 0.33	0.68 ± 0.36
Hgs	***	1.68 ± 1.31	2.05 ± 1.51
Mns	NS	7.56 ± 2.59	7.97 ± 2.69
Nis	***	1.13 ± 0.44	0.97 ± 0.33
Pbs	***	23.78 ± 19.58	18.86 ± 11.48
Vs	NS	0.15 ± 0.15	0.14 ± 0.16
Zns	NS	6045.78 ± 798.99	5956.91 ± 861.05
RIMs	***	0.45 ± 0.11	0.41 ± 0.11
Alu	NS	4.16 ± 3.73	4.11 ± 4.38
Asu	NS	24.20 ± 29.43	36.66 ± 56.22
Cdu	***	0.72 ± 0.54	0.51 ± 0.37
Cou	NS	0.64 ± 0.65	0.63 ± 0.50
Cru	NS	0.71 ± 0.36	0.67 ± 0.32
Hgu	NS	1.51 ± 1.23	1.54 ± 1.19
Mnu	NS	0.36 ± 0.19	0.36 ± 0.24
Niu	NS	1.98 ± 1.10	2.02 ± 1.40
Pbu	***	1.82 ± 2.24	1.24 ± 1.03
Vu	NS	0.61 ± 0.35	0.54 ± 0.32
Znu	***	538.90 ± 357.78	416.47 ± 302.80
RIMu	***	0.35 ± 0.14	0.31 ± 0.14

Concernant le sang, il n'y a pas de différence significative entre les deux groupes pour Al, As, Co, Cr, Mn, V et Zn. Pour le cadmium, le plomb et le nickel, le test montre une différence significative, avec des concentrations en moyenne plus élevées chez les fumeurs. Pour le cadmium, la concentration moyenne chez les fumeurs est plus de trois fois supérieure à celle des non-fumeurs (1.06 µg/L contre 0.32 µg/L). Pour les deux autres éléments, la différence est de 1.03 µg/L contre 0.97 µg/L pour Ni et 23.78 µg/L contre 18.86 µg/L pour Pb. Une différence est également observée pour Hg, mais cette fois-ci les non-fumeurs sont les plus imprégnés (2.05 µg/L contre 1.68 µg/L). Concernant l'indice

multimétallique, le niveau d'imprégnation global est statistiquement plus important chez les fumeurs (0.45 contre 0.41).

Dans les urines, ce sont le cadmium, le plomb et le zinc qui présentent des valeurs significativement plus élevées chez les fumeurs. Les concentrations moyennes sont respectivement de 0.72 µg/L, 1.82 µg/L et 539 µg/L chez les fumeurs et de 0.51 µg/L, 1.24 µg/L et 416 µg/L chez les non-fumeurs. Les autres éléments ne montrent au contraire aucune différence significative entre les deux groupes. Le RIM urinaire est, comme le RIM sanguin, statistiquement plus élevé chez les fumeurs (0.35µg/L contre 0.31 µg/L).

Selon nos résultats, les concentrations sanguines en Al, As, Co, Cr, Mn, V et Zn et urinaires en Al, As, Co, Cr, Hg, Mn, Ni et V sont indépendantes du statut tabagique des individus. En revanche, le tabagisme constitue un facteur confondant pour Cd, Hg, Ni et Pb dans le sang, pour Cd, Pb et Zn dans les urines et pour les deux RIM. Les corrélations de ces variables avec les concentrations lichéniques seront donc uniquement étudiées en fonction du statut tabagique et particulièrement chez les non-fumeurs, car non impactés par ce facteur confondant.

Le second facteur étudié est l'âge des individus. Les métaux étant des éléments non dégradés par l'organisme et cumulatifs, les personnes plus âgées sont susceptibles de présenter des niveaux de concentrations plus élevés que les personnes jeunes. Nous avons donc effectué un test de comparaison de moyennes entre deux groupes de population extrêmes : les moins de 30 ans (n= 74) et les plus de 50 ans (n = 55), pour connaître la sensibilité des résultats face à ce facteur. Le test bilatéral de Mann et Whitney (tableau 13) montre des moyennes significativement différentes entre les deux groupes d'âge pour plusieurs éléments.

Dans le sang, des concentrations significativement plus élevées en arsenic, chrome, mercure et plomb sont mesurées chez les plus de 50 ans. Cette différence est notamment importante pour le plomb, avec des valeurs pratiquement doublées (27.99 µg/L contre 17.65 µg/L). Nous remarquons également une différence significative pour le RIM sanguin (0.45 µg/L pour les plus de 50 ans, contre 0.41 µg/L pour les moins de 30 ans). En revanche, la concentration en nickel est significativement plus élevée chez les plus jeunes (1.08 µg/L contre 0.87 µg/L). Pour Al, Cd, Co, Mn, V et Zn, les concentrations ne sont pas statistiquement différentes et n'indiquent pas une influence de l'âge.

Dans les urines, le cadmium et le plomb sont présents à des niveaux plus élevés chez les plus âgés (respectivement 0.74 µg/L et 1.77 µg/L contre 0.48 µg/L et 1.51 µg/L). A l'inverse de ces deux éléments, Co et Ni sont retrouvés dans des concentrations plus importantes chez les sujets jeunes (respectivement 0.75 µg/L et 2.12 µg/L contre 0.62 µg/L et 1.69 µg/L). Le RIM urinaire, ainsi que Al, As, Cr, Hg, Mn, V et Zn ne diffèrent pas selon le groupe d'âge.

Selon nos résultats, les concentrations sanguines en Al, Cd, Co, Mn, V et Zn et urinaires en Al, As, Cr, Hg, Mn, V et Zn, ainsi que le RIM urinaire sont indépendants de l'âge des individus. En revanche, l'âge constitue un facteur confondant pour As, Cr, Hg et Pb dans le sang, pour Cd, Co, Ni et Pb dans les urines, ainsi que pour le RIM sanguin. Les corrélations avec les concentrations lichéniques seront étudiées en fonction de groupes d'âge.

Tableau 13. Résultats du test Mann et Whitney de comparaison de moyennes pour les ETM sanguins (s) et urinaires (u) entre les moins de 30 ans et les plus de 50 ans (NS : non-significatif ; *** : significatif) – les valeurs en gras sont significativement supérieures, à un risque alpha de 0.05

Variable	Différence (p<0.05)	< 30 ans (µg/L) Moyenne ± écart-type	> 50 ans (µg/L) Moyenne ± écart-type
Als	NS	2.78 ± 2.10	3.30 ± 2.73
Ass	***	2.10 ± 1.97	3.31 ± 2.97
Cds	NS	0.64 ± 0.63	0.48 ± 0.42
Cos	NS	0.29 ± 0.08	0.29 ± 0.12
Crs	***	0.60 ± 0.30	0.80 ± 0.39
Hgs	***	1.54 ± 1.16	2.36 ± 1.64
Mns	NS	8.44 ± 2.87	8.12 ± 2.78
Nis	***	1.08 ± 0.48	0.87 ± 0.30
Pbs	***	17.65 ± 20.34	27.99 ± 12.92
Vs	NS	0.17 ± 0.15	0.12 ± 0.07
Zns	NS	5900.18 ± 831.41	6099.13 ± 816.06
RIMs	***	0.41 ± 0.10	0.45 ± 0.11
Alu	NS	4.72 ± 3.45	3.86 ± 4.11
Asu	NS	29.54 ± 43.47	35.36 ± 51.38
Cdu	***	0.48 ± 0.42	0.74 ± 0.57
Cou	***	0.75 ± 0.48	0.62 ± 0.78
Cru	NS	0.62 ± 0.29	0.64 ± 0.28
Hgu	NS	1.68 ± 1.49	1.55 ± 1.15
Mnu	NS	0.35 ± 0.21	0.32 ± 0.23
Niu	***	2.12 ± 1.18	1.69 ± 1.05
Pbu	***	1.51 ± 2.52	1.77 ± 1.34
Vu	NS	0.63 ± 0.41	0.58 ± 0.30
Znu	NS	495.15 ± 365.94	431.38 ± 280.84
RIMu	NS	0.33 ± 0.15	0.34 ± 0.13

Pour résumer cette étude des facteurs confondants, le tableau 14 synthétise l'influence significative supérieure (+) ou significative inférieure (-) de l'âge et du tabac sur les concentrations sanguines et urinaires. Lors de l'analyse des corrélations, l'ensemble des individus ont été pris en considération pour Als, Cos, Mns, Vs, Zns, Alu, Asu, Cru, Hgu, Mnu et Vu.

Tableau 14. Influence de l'âge et du tabac sur les concentrations sanguines et urinaires en métaux (NS : non significatif ; + : augmentation significative des concentrations ; - : diminution significative des concentrations)

Variable	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMs	RIMu
Effet tabac	NS	NS	+	NS	NS	-	NS	+	+	NS	NS	NS	NS	+	NS	NS	NS	NS	NS	+	NS	+	+	+
Effet âge	NS	+	NS	NS	+	+	NS	-	+	NS	NS	NS	NS	+	-	NS	NS	NS	-	+	NS	NS	+	NS

Une analyse a ensuite été essentiellement basée sur les non-fumeurs pour Cds, Nis, Pbs, Cdu, Pbu, Znu, RIMs et RIMu, car ces variables sont influencées par le statut tabagique. Au sein de la base de données d'imprégnation de l'Homme, l'âge est renseigné selon 4 classes (<30 ans ; 30-39 ans ; 40-49 ans ; ≥50 ans). Afin de conserver un nombre conséquent de sujets, nous avons considéré les individus

selon deux groupes (les moins de 40 ans, n = 151 et les 40 ans et plus, n = 129) pour l'analyse de corrélations concernant Ass, Crs, Hgs, Nis, Pbs, Cdu, Cou, Niu, Pbu et RIMs. Finalement, certaines variables sont influencées par les deux facteurs. Les analyses de corrélations pour Hgs, Nis, Pbs, Cdu, Pbu et RIMs ont été réalisées pour 2 groupes de population intégrant ces deux facteurs : les sujets non-fumeurs de moins de 40 ans (groupe A, n = 86) et les sujets non-fumeurs de plus de 40 ans (groupe B, n = 64).

ANALYSE DE CORRELATIONS

Nous avons dans un premier temps étudié la relation entre les deux matrices humaines. Le tableau 15 représente la matrice de corrélation entre les concentrations métalliques mesurées dans le sang et les urines pour chacun des individus du bassin de Dunkerque. Sur les tableaux 15 à 19, l'indicateur statistique renseigné est le coefficient de Spearman (CS). Les cellules colorées en vert sont mises en évidence pour faciliter la lecture des tableaux. Elles correspondent à la corrélation entre un même élément dosé dans des matrices différentes. Les valeurs en gras indiquent une corrélation significative à un risque alpha de 5%. Toujours dans l'optique de faciliter la lecture, seule une partie des matrices est à chaque fois présentée. Ces matrices de corrélations sont disponibles en intégralité en annexe I.

Tableau 15. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) et urinaires (u) pour l'ensemble des individus – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.a

	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	RIMu
Alu	0.08	-0.05	0.08	0.11	0.15	-0.09	0.07	0.02	-0.01	0.21	0.04	0.10	0.47
Asu	-0.05	0.59	-0.03	0.07	0.15	0.35	-0.03	-0.09	0.09	0.12	-0.02	0.29	0.33
Cdu	0.01	0.07	0.36	0.09	0.11	0.03	0.03	0.00	0.20	0.37	0.06	0.23	0.64
Cou	0.00	0.04	0.21	0.42	0.00	-0.06	0.22	0.06	-0.11	0.06	-0.20	0.20	0.47
Cru	0.20	0.00	0.02	0.00	-0.13	-0.14	0.03	0.03	0.01	0.49	0.18	0.02	0.42
Hgu	-0.02	-0.02	-0.04	0.08	-0.01	0.18	0.04	-0.14	-0.07	0.17	-0.06	0.05	0.37
Mnu	0.01	0.02	-0.02	0.07	-0.08	0.00	0.04	0.13	0.00	-0.07	0.26	0.02	0.35
Niu	0.01	0.04	0.23	0.30	0.11	-0.17	0.07	0.28	-0.01	0.12	0.02	0.17	0.50
Pbu	-0.04	0.02	0.26	-0.13	0.02	0.04	-0.08	-0.01	0.63	0.01	0.06	0.26	0.62
Vu	0.04	0.05	0.11	-0.04	0.15	-0.01	0.03	-0.14	0.10	0.12	-0.06	0.15	0.60
Znu	-0.06	-0.09	0.16	-0.14	-0.02	-0.03	-0.05	0.02	0.14	0.07	0.19	0.04	0.56
RIMu	0.04	0.14	0.24	0.11	0.09	0.07	0.03	-0.01	0.23	0.21	0.10	0.33	1.00
RIMs	0.33	0.55	0.35	0.38	0.52	0.36	0.22	0.10	0.40	0.61	0.16	1.00	

Le tableau 15 montre une corrélation significative entre les niveaux sanguins et urinaires pour 7 éléments, cités par ordre décroissant du CS : plomb (0.63), arsenic (0.59), cobalt (0.42), cadmium (0.36), nickel (0.28), zinc (0.19) et mercure (0.18). On observe également une corrélation significative entre les deux indices multimétalliques (CS : 0.33). En revanche, il n'y a pas de relation significative entre les concentrations sanguine et urinaire pour l'aluminium (CS : 0.08), le chrome (CS : -0.13), le manganèse (CS : 0.04) et le vanadium (CS : 0.12).

Nous pouvons ensuite remarquer que les métaux sont fortement corrélés aux RIM correspondants, ce qui valide le caractère synthétique de cet indice. Le RIM urinaire est corrélé à l'ensemble des éléments qui le composent, le CS varie de 0.33 pour l'arsenic à 0.63 pour le cadmium. Pour le RIM sanguin, seul le nickel n'est pas significativement corrélé. Les CS sont globalement moins élevés et varient de 0.16 pour le zinc à 0.61 pour le vanadium.

Les inégalités environnementales de santé ont ensuite été caractérisées par la relation entre les concentrations biologiques mesurées chez la population et le niveau d'imprégnation de l'environnement, représenté par les concentrations lichéniques. La première analyse concerne les variables non influencées par les facteurs âge et tabac (tableau 16). Un seul élément dosé dans le sang montre une corrélation significative avec les concentrations environnementales. Il s'agit du vanadium (CS : 0.33). Les autres éléments, ainsi que la totalité de ceux mesurés dans les urines ne présentent pas de corrélation significative entre les matrices humaines et lichéniques, les coefficients étant globalement très faibles (moins de 0.15).

Tableau 16. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour l'ensemble des individus – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.b

	Als	Cos	Mns	Vs	Zns	Alu	Asu	Cru	Hgu	Mnu	Vu
AlL	0.12	0.04	0.07	0.15	0.11	-0.04	-0.03	-0.01	0.09	0.07	-0.02
AsL	0.09	0.01	-0.04	0.27	0.08	-0.01	-0.07	0.08	0.11	0.03	0.02
CoL	0.16	0.08	0.05	0.32	0.09	0.03	-0.07	0.04	0.10	0.05	0.02
CrL	0.13	-0.01	-0.06	-0.12	-0.09	-0.06	-0.11	0.00	0.04	0.03	-0.05
HgL	0.17	0.08	0.05	0.31	0.08	0.01	-0.13	0.06	0.08	0.08	-0.02
MnL	0.20	0.09	0.07	0.32	0.14	0.00	-0.07	0.09	0.12	0.09	0.00
VL	0.18	0.07	0.03	0.33	0.13	0.01	-0.05	0.08	0.11	0.06	0.00
ZnL	0.16	0.10	0.01	0.06	0.05	-0.07	-0.09	0.00	0.04	0.09	-0.02

En ce qui concerne les corrélations pour les éléments influencés par le statut tabagique, les sujets non-fumeurs ont été isolés (tableau 17). Aucune corrélation significative entre les dosages effectués chez les populations et les concentrations lichéniques n'est observée. Les CS sont tous inférieurs à 0.15. Ainsi, en tenant compte du facteur tabac, l'imprégnation multimétallique des populations ne semble pas statistiquement liée à l'imprégnation multimétallique de l'environnement.

Tableau 17. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus non-fumeurs – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.c

	Cds	Nis	Pbs	RIMs	Cdu	Pbu	Znu	RIMu
CdL	0.09	0.00	0.16	0.11	0.08	0.13	0.07	0.12
NiL	0.08	0.08	0.06	0.00	0.02	0.02	0.06	0.01
PbL	0.15	0.10	0.13	0.10	0.08	0.11	0.10	0.17
ZnL	0.12	0.12	0.19	0.10	0.05	0.08	0.09	0.11
RIML	0.15	0.09	0.14	0.09	0.05	0.09	0.08	0.12

Concernant les éléments influencés par l'âge des individus, l'analyse de corrélations concerne le groupe des moins de 40 ans et celui des 40 ans et plus (tableau 18). Chez les individus de plus de 40 ans (B), aucune corrélation significative n'est mise en évidence, les CS étant même proches de 0. Les concentrations biologiques des individus âgés de moins de 40 ans (A) ne sont pas liées aux concentrations environnementales pour As, Crs, Hgs, Nis et Cou. Les indices multimétalliques ne sont pas non plus corrélés. En revanche, les concentrations environnementales sont significativement corrélées aux concentrations mesurées chez les sujets pour 4 éléments. Pour le sang, cela concerne

le plomb (CS : 0.26). Pour l'urine, il s'agit du cadmium (CS : 0.24), du nickel (CS : 0.21) et du plomb (CS : 0.19).

Tableau 18. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus de moins de 40 ans (A) et de 40 ans et plus (B) – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.d

A	Ass	CrS	Hgs	Nis	Pbs	RIMs	Cdu	Cou	Niu	Pbu
AsL	0.02	-0.01	0.04	-0.06	0.23	0.10	0.19	0.06	0.15	0.18
CdL	-0.12	0.02	-0.02	0.04	0.11	0.05	0.24	0.04	0.22	0.17
CoL	0.00	0.03	0.08	-0.03	0.18	0.14	0.23	0.15	0.24	0.20
CrL	0.01	0.06	0.01	0.06	0.07	0.01	0.01	-0.14	0.07	0.07
HgL	-0.07	0.01	0.04	-0.03	0.17	0.07	0.20	0.11	0.21	0.17
NiL	-0.18	-0.01	-0.02	0.00	0.05	0.00	0.15	0.02	0.21	0.13
PbL	-0.02	0.01	0.07	0.01	0.26	0.12	0.19	0.05	0.18	0.19
RIML	-0.03	0.02	0.03	0.01	0.17	0.10	0.18	0.05	0.20	0.16
B	Ass	CrS	Hgs	Nis	Pbs	RIMs	Cdu	Cou	Niu	Pbu
AsL	-0.08	-0.18	-0.20	0.12	0.03	0.00	0.05	0.07	0.08	0.14
CdL	-0.02	0.00	-0.09	0.05	0.11	0.07	0.05	-0.11	-0.04	0.09
CoL	-0.08	-0.03	-0.16	0.08	0.02	0.07	0.04	0.02	0.00	0.07
CrL	-0.04	-0.07	-0.23	0.17	0.05	-0.14	-0.03	-0.13	0.01	-0.07
HgL	-0.10	-0.09	-0.17	0.10	0.04	0.08	0.06	0.02	-0.01	0.08
NiL	-0.16	-0.01	-0.26	0.11	0.05	0.01	0.07	-0.06	0.02	0.00
PbL	-0.03	-0.01	-0.13	0.08	0.02	0.09	0.05	0.04	0.04	0.06
RIML	-0.06	-0.05	-0.16	0.13	0.04	0.04	0.01	-0.05	0.00	0.01

Les analyses de corrélations ont été réalisées pour 2 groupes de population intégrant ces deux facteurs confondants (tableau 19). Concernant le groupe B (sujets non-fumeurs de plus de 40 ans), aucun élément mesuré chez l'Homme ne montre de lien statistique avec les mesures environnementales. Pour le groupe A (sujets non-fumeurs de moins de 40 ans), Hgs et Nis présentent des degrés de corrélation très faibles avec les valeurs lichéniques (respectivement 0.08 et 0.06). Les CS sont de 0.18 entre les indices multimétalliques et de 0.19 entre le cadmium sanguin et lichénique, la relation n'est suffisamment forte pour être significative. Une corrélation significative avec les concentrations lichéniques est mise en évidence concernant le plomb sanguin (CS : 0.31) et urinaire (CS : 0.28).

Tableau 19. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus en fonction des quatre groupes de risque (A : non-fumeur et moins de 40 ans et B : non-fumeur et plus de 40 ans) – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05 – la matrice est disponible dans son intégralité en annexe I.e

A	Hgs	Nis	Pbs	RIMs	Cdu	Pbu	B	Hgs	Nis	Pbs	RIMs	Cdu	Pbu
CdL	-0.03	0.04	0.22	0.14	0.19	0.29	CdL	-0.05	-0.04	0.14	0.04	-0.06	-0.03
HgL	0.08	0.02	0.28	0.12	0.16	0.29	HgL	-0.18	0.06	0.01	0.04	-0.03	-0.06
NiL	-0.03	0.06	0.15	0.04	0.10	0.20	NiL	-0.21	0.08	0.06	-0.02	-0.05	-0.13
PbL	0.06	0.09	0.31	0.20	0.14	0.28	PbL	-0.17	0.10	0.02	0.06	0.04	-0.03
RIML	0.04	0.06	0.25	0.18	0.15	0.32	RIML	-0.18	0.12	0.04	0.01	-0.08	-0.11

d. Discussion

L'objectif de cette étude était d'évaluer les liens entre les niveaux d'imprégnation des populations et de l'environnement par les ETM, afin d'identifier des inégalités environnementales de santé. Les données populationnelles issues de l'étude IMEPOGE (Nisse, 2010) ont été sélectionnées sur le bassin de Dunkerque, puis mises en forme pour l'analyse des corrélations. Les données environnementales proviennent du suivi spatio-temporel de l'imprégnation des lichens par les ETM sur la région Nord – Pas de Calais.

Nous apercevons à travers cette étude les atouts que peuvent représenter les outils SIG, notamment les géostatistiques, dans le cadre de la recherche en santé – environnement. En effet, la mise en relation de ces deux bases de données n'est pas possible avec les jeux de données initiaux, du fait de la non-concordance géographique des points d'échantillonnage. La réalisation de mesures environnementales au niveau du lieu d'habitation de chaque sujet n'était pas non plus envisageable. Grâce aux géostatistiques, les valeurs de concentrations environnementales des 11 ETM et du RIM ont pu être estimées au niveau des coordonnées de ces lieux d'habitation, à partir de données existantes.

Une fois la base de données générée et mise en forme, nous avons étudié l'influence de plusieurs facteurs sur les niveaux d'imprégnation des populations. Au niveau du bassin de Dunkerque, l'âge et le statut tabagique se révèlent être d'importants facteurs favorisant l'excès de concentrations sanguines et urinaires en métaux. Ces résultats concordent avec conclusions de l'étude régionale, notamment pour le cadmium et le plomb (Nisse, 2010), ainsi qu'avec tendances observées à l'échelle mondiale (Apostoli *et al.*, 2002 ; Batariova *et al.*, 2006 ; Heitland & Koster, 2006). L'ensemble des métaux considérés ici est retrouvé dans la fumée de cigarette (Chang *et al.*, 2003 ; Behera *et al.*, 2014). Leur présence est notamment due à l'absorption de ces éléments présents dans le sol par la plante de tabac, mais également par l'application de fertilisants et pesticides sur les cultures (Chiba & Masironi, 1992). Cependant, certaines études démontrent que le tabagisme affecte les charges retrouvées dans les fluides biologiques uniquement pour certains d'entre eux, comme le cadmium, le chrome, le plomb (Bernhard *et al.*, 2005).

Cette analyse a également permis d'identifier les variables influencées par l'âge de l'individu, son statut tabagique ou ces deux facteurs cumulés. Les corrélations entre les concentrations métalliques des trois matrices ont donc été évaluées, selon les variables, en fonction de ces deux facteurs. Le tableau 20 reprend l'ensemble des résultats des différentes matrices de corrélations. Les valeurs en vert et en gras sont significativement corrélées aux mesures effectuées dans les lichens ($p < 0.05$).

Tableau 20. Synthèse des résultats d'analyse de corrélation entre les dosages sanguins (s) ou urinaires (u) et les dosages dans les lichens en fonction des différents groupes de population – les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05

Sujets	Als	Ass	Cds	Cos	CrS	Hgs	Mns	Nis	Pbs	Vs	Zns	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMs	RIMu
Tous	0.12	-	-	0.08	-	-	0.07	-	-	0.33	0.05	-0.04	-0.07	-	-	0.00	0.08	0.09	-	-	0.00	-	-	-
Non-fumeurs	-	-	0.09	-	-	-	-	0.08	0.13	-	-	-	-	0.08	-	-	-	-	-	0.11	-	0.09	0.09	0.12
< 40 ans	-	0.02	-	-	0.06	0.04	-	0.00	0.26	-	-	-	-	0.24	0.15	-	-	-	0.21	0.19	-	-	0.10	-
≥ 40 ans	-	-0.08	-	-	-0.07	-0.17	-	0.11	0.02	-	-	-	-	0.05	0.02	-	-	-	0.02	0.06	-	-	0.04	-
Non-fumeurs < 40 ans	-	-	-	-	-	0.08	-	0.06	0.31	-	-	-	-	0.19	-	-	-	-	-	0.28	-	-	0.18	-
Non-fumeurs ≥ 40 ans	-	-	-	-	-	-0.18	-	0.08	0.02	-	-	-	-	-0.06	-	-	-	-	-	-0.03	-	-	0.01	-

Dans cette étude, nous avons utilisé des indices d'imprégnation multimétallique, dans le but d'étudier les relations en termes de contamination globale. A l'instar du RIM lichénique (voir Chapitre I, Partie 2), les RIM sanguin et urinaire sont proches des éléments qui les composent, ce qui témoigne de leur bonne capacité à synthétiser une imprégnation globale des populations par les ETM. Ces deux indices sont également corrélés entre eux. Lorsqu'on isole les individus non-fumeurs,

Le RIM urinaire présente un CS de 0.12 avec le RIM lichénique, mais celui-ci n'est pas significatif. Le même constat peut être fait pour le RIM urinaire des sujets de moins de 40 ans (CS : 0.10) et des sujets non-fumeurs de moins de 40 ans (CS : 0.18). Ainsi selon les critères définis dans le cadre de notre étude (influence du tabagisme et de l'âge pour RIMs et influence du tabagisme pour RIMu), le lien entre les charges globales en métaux dans l'environnement et les charges globales dans les liquides biologiques des populations n'est pas statistiquement identifiable à l'échelle du bassin de Dunkerque. Cependant, même si le RIM urinaire n'est pas influencé par l'âge, nous avons remarqué un CS significatif de 0.29 avec le RIM lichénique pour les sujets non-fumeurs de moins de 40 ans (tableau 19 Annexe I.e). Sur cette matrice de corrélation, nous pouvons également constater que l'indice multimétallique calculé pour les urines est significativement corrélé à 7 des 11 éléments mesurés chez les lichens : Al (CS : 0.27), Cd (CS : 0.25), Co (CS : 0.28), Mn (CS : 0.31), Pb (CS : 0.27), V (CS : 0.29) et Zn (CS : 0.25). Ainsi le lien entre les charges globales en métaux dans l'environnement et les charges globales dans les urines des populations semble identifiable à l'échelle du bassin de Dunkerque, notamment pour les personnes présentant un faible risque d'accumulation lié au tabagisme ou à l'âge.

Bien que les indices intégrés représentent une simplification de qualité environnementale, ils peuvent être considérés comme de puissants outils de traitement et d'analyse pour caractériser le statut général de qualité environnementale (Caeiro *et al.*, 2005 ; Golge *et al.*, 2013). Parmi les études employant la biosurveillance environnementale pour évaluer les liens entre la contamination des milieux et des effets sanitaires, plusieurs mettent en avant la difficulté d'associer une pathologie à un élément en particulier (Fuga *et al.*, 2008 ; Sarmiento *et al.*, 2008 ; Carreras *et al.*, 2009). L'utilisation de tels indices semble pouvoir pallier ces difficultés.

Pour 7 des 11 ETM mesurés au sein de la population, les concentrations urinaires sont liées aux concentrations sanguines. A l'échelle du bassin de Dunkerque, le plomb présente la plus forte similarité entre ces deux fluides biologiques (CS : 0.63). Lors de la mise en relation avec les concentrations mesurées dans les lichens, nous avons mis en évidence une corrélation significative pour cet élément. De plus, le niveau de cette relation croît lorsque nous prenons en considération les facteurs âge et tabac. Concernant la corrélation sang – lichens, le CS évolue de 0.13 (non significatif) pour les non-fumeurs à 0.26 pour les sujets de moins de 40 ans et à 0.31 pour les sujets de moins de 40 ans et non-fumeurs. La relation urines – lichens évolue de la même manière, avec un CS de 0.11 pour les non-fumeurs (non significatif), de 0.19 pour les sujets de moins de 40 ans et de 0.28 pour les sujets de moins de 40 ans et non-fumeurs. A l'instar du plomb, nous pouvons souligner un lien significatif entre les concentrations en nickel et cadmium dans l'environnement et celles retrouvées dans les urines des sujets jeunes (CS respectivement de 0.21 et 0.24). Finalement, le CS attribué entre le vanadium sanguin et le vanadium lichénique est le plus élevé (0.33). Ces résultats indiquent un rôle probable des charges environnementales dans l'imprégnation des populations par le plomb (sang et urines), le cadmium (urines) et le nickel (urines). Le plomb présente notamment les corrélations les plus importantes.

e. Conclusion

Cette étude avait pour objectif d'évaluer les liens entre les niveaux d'imprégnation de l'environnement par les ETM et les charges retrouvées dans les fluides biologiques de populations vivant à proximité. Grâce aux outils géostatistiques, nous avons pu identifier qu'une partie des concentrations biologiques en plomb (sang et urines), cadmium (urines) et nickel (urines) mesurées dans un échantillon de population du bassin de vie de Dunkerque peut être expliquée par une imprégnation de l'environnement par ces métaux. Le calcul d'indices intégrés d'imprégnation multimétallique montre également que les populations habitant dans un environnement globalement imprégné par les ETM présentent des niveaux de charge globale élevés dans les urines, notamment chez les populations jeunes et non-fumeurs. Ainsi, nous avons mis en évidence des degrés de corrélation entre les mesures populationnelles et environnementales plus importants pour

les concentrations urinaires. Cette matrice semble être un bon traceur des charges globales retrouvées dans l'environnement. Le RIM est ici utilisé pour décrire une imprégnation globale par les ETM. Il représente à ce titre une moyenne des 11 éléments. Cet indice est considéré comme un bon indicateur de suivi, mais ne peut se substituer à l'interprétation individuelle des éléments qui le composent.

Par ailleurs, l'imprégnation des populations peut aussi dépendre d'autres facteurs, tels que les habitudes alimentaires, l'habitat ou la profession exercée. Une étude multifactorielle semble nécessaire pour confirmer les liens identifiés dans notre étude, et caractériser la part de l'environnement dans l'imprégnation des populations par ces contaminants.

Concernant la biosurveillance environnementale, c'est la première fois à notre connaissance que des mesures d'imprégnation par les ETM au sein d'organismes sentinelles sont mises en relation à des mesures issues de la biosurveillance humaine. Plusieurs auteurs se sont consacrés à l'apparition de diverses pathologies (respiratoires, cardio-vasculaires, tumeurs, etc.) en lien avec les concentrations métalliques mesurées chez les mousses ou les lichens (Gailey & Lloyd, 1993 ; Wappelhorst *et al.*, 2000 ; Wolterbeek & Verburg, 2004 ; Fuga *et al.*, 2008 ; Sarmiento *et al.*, 2008 ; Carreras *et al.*, 2009). Bien que l'étude des corrélations avec les dosages de métaux dans les matrices humaines soit évoquée dans la plupart de ces articles, aucun d'entre eux ne traite ce type de données. En effet, malgré le fort intérêt de la biosurveillance environnementale, l'utilisation de ces méthodes dans le domaine de l'épidémiologie est assez récente (Rzepka & Cuny, 2008 ; Wolterbeek *et al.*, 2010).

Partie 5. Conclusion et perspectives du chapitre I

La biosurveillance environnementale de la qualité de l'air est une approche complémentaire de la surveillance physico-chimique. Elle mesure les impacts biologiques de la pollution atmosphérique sur des organismes sentinelles de l'environnement, tels que les lichens. L'utilisation de données de biosurveillance dans le domaine de l'épidémiologie est très récente. Beaucoup de travaux se basent sur les mesures physico-chimiques pour étudier les effets des polluants sur la santé, mais très peu utilisent ces mesures d'effets biologiques (Cuny, 2012). Du fait de l'extrapolation difficile du végétal à l'Homme, la biosurveillance ne peut être utilisée pour mesurer des impacts sanitaires, mais se rapproche plus du domaine de l'évaluation des risques, en tant qu'indicateur biologique de potentiel (éco)toxique du milieu (Van Haluwyn *et al.*, 2011). Les études menées dans le cadre de cette thèse présentent de nouveaux arguments concernant les apports de la biosurveillance et l'utilisation d'organismes sentinelles dans le domaine de la santé environnementale.

Dans ce chapitre, la capacité bioaccumulatrice des lichens a été mise en œuvre pour évaluer les disparités spatiales d'imprégnation de l'environnement par 18 ETM. Les mesures effectuées chez l'espèce *Xanthoria parietina* représentent la fraction biodisponible des éléments présents dans le milieu atmosphérique et bioaccumulée sur le long terme. Les charges métalliques retranscrivent ainsi une exposition chronique à la pollution de l'air. Les résultats des échantillons collectés au niveau de trois bassins de vie du Nord – Pas de Calais ont été répertoriés dans un SIG et cartographiés à l'aide des outils géostatistiques (méthode de krigeage empirique bayésien). Nos résultats révèlent des charges métalliques très élevées sur le bassin dunkerquois, notamment à proximité des zones industrielles, un peu moins élevées sur le bassin lillois et beaucoup plus faibles sur le bassin de Maubeuge.

Ce chapitre avait ensuite pour objectif de mettre en évidence si les populations les plus défavorisées sur un plan socioéconomique vivent dans un environnement fortement contaminé par les ETM et si ces populations sont également les plus imprégnées par ces mêmes éléments. Pour y répondre, deux études ont été menées sur le bassin de Dunkerque. La première a mis en relation un indicateur de défaveur social (l'Indice de Défaveur Localisé, développé par le laboratoire TVES ULCO) avec les données lichéniques, pour mettre en évidence des inégalités environnementales à l'échelle des Iris. La deuxième s'est intéressée aux inégalités environnementales de santé, en étudiant les corrélations entre les mesures effectuées chez les lichens et les dosages des mêmes éléments réalisés dans le sang et les urines d'un échantillon de population vivant sur ce bassin (données issues de l'étude IMEPOGE menée par le CERESTE, CHRU de Lille).

Lors de ces deux études, un indice intégré (le Ratio d'Imprégnation Moyen) fut développé pour décrire l'imprégnation multimétallique de l'environnement et des populations. Un lien statistique a été mis en évidence entre le RIM lichénique et l'IDL, témoignant la présence de populations défavorisées au niveau de territoires fortement imprégnés par l'ensemble des métaux. En parallèle, en isolant les individus présentant un faible risque de bioaccumulation des métaux (jeunes non-fumeurs), le RIM calculé pour les urines se trouve significativement corrélé au RIM lichénique, ce qui suppose que les populations les plus imprégnées par l'ensemble des éléments vivent dans un environnement également contaminé. Ainsi l'utilisation d'indices intégrés semble pouvoir pallier les difficultés d'établir une relation entre un effet de santé et une substance en particulier.

Suite à ces résultats, plusieurs perspectives s'offrent à nous. Le bassin de Dunkerque étant constitué de populations défavorisées vivant à proximité d'activités industrielles émettrices et donc susceptibles d'être plus soumises à des effets sanitaires, les conclusions présentées dans ce chapitre doivent être confrontées à celles observées sur d'autres typologies de bassin, tels que Lille (très urbanisé) et Maubeuge (essentiellement rural). De fait, les trois bases de données seront par la suite mises en relation sur chacun des bassins, afin d'évaluer la part de l'environnement et de la précarité sociale dans les inégalités d'imprégnation des populations par les métaux.

Lors de ces futures analyses, nous pourrons porter une attention particulière au plomb, pour lequel une relation significative est mise en évidence sur le dunkerquois entre les concentrations lichéniques, le niveau de défaveur sociale et les concentrations mesurées dans le sang et les urines des populations.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Albertini R, Bird M, Doerr N, Needham L, Robison S, Sheldon L, Zenick H. 2006. The use of biomonitoring data in exposure and human health risk assessments. *Environ Health Perspect*, 114(11): 1755-62.
- Angerer J, Ewers U, Wilhelm M. 2007. Human biomonitoring: state of the art. *Int J Hyg Environ Health*, 210(3-4): 201-28.
- Apostoli P, Baj A, Bavazzano P, Ganzi A, Neri G, Ronchi A, Soleo L, Di LL, Spinelli P, Valente T, Minoia C. 2002. Blood lead reference values: the results of an Italian polycentric study. *Sci. Total Environ*, 287(1-2): 1-11.
- Batariova A, Spevackova V, Benes B, Cejchanova M, Smid J, Cerna M. 2006. Blood and urine levels of Pb, Cd and Hg in the general population of the Czech Republic and proposed reference values. *Int. J. Hyg. Environ. Health*, 209: 359–366.
- Behera S, Xian H, Balasubramanian R. 2014. Human health risk associated with exposure to toxic elements in mainstream and sidestream cigarette smoke. *Science of the Total Environment*, 472: 947–956.
- Bernhard D, Rossmann A, Wick G. 2005. Metals in Cigarette Smoke – Critical review. *IUBMB Life*, 57(12): 805 – 809.
- Bharti N & Katyal D. 2011. Water quality indices used for surface water vulnerability assessment. *International Journal of Environmental Sciences*, 2 (1): 154-173.
- Boamponsem LK, Adam JI, Dampare SB, Nyarko BJB, Esumang DK. 2010. Assessment of atmospheric heavy metal deposition in the Tarkwa gold mining area of Ghana using epiphytic lichens. *Nuclear Instruments and Methods in Physics Research*, B 268: 1492–1501.
- Brulle RJ & Pellow DN. 2006. Environmental justice: Human health and environmental inequalities. *Annual Review of Public Health*, 27: 103-124.
- Caeiro S, Costa MH, Ramos TB, Fernandes F, Silveira N, Coimbra A, Medeiros G, Painho M. 2005. Assessing heavy metal contamination in Sado Estuary sediment: An index analysis approach. *Ecological Indicators*, 5: 151–169.
- Cairncross EK, John J, Zunckel M. 2007. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. *Atmos. Environ*, 41: 8442–8454.
- Carreras HA, Wannaz ED, Pignata ML. 2009. Assessment of human health risk related to metals by the use of biomonitors in the province of Cordoba Argentina. *Environmental Pollution*, 157 (1): 117-122.
- Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, Merlo J. 2006. Children’s exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *Journal of Epidemiology and Community Health*, 60 (3): 234-241.
- Chang MJ, Naworal JD, Walker K, Connell CT. 2003. Investigations on the direct introduction of cigarette smoke for trace elements analysis by inductively coupled plasma mass spectrometry. *Spectrochimica Acta Part B*, 58: 1979–1996.
- Chen TB, Zheng YM, Lei M, Huang ZC, Wu HT, Chen H, Fan KK, Yu K, Wu X, Tian QZ. 2005. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere*, 60: 542–551.

- Chiba M & Masironi R. 1992. Bull. World Health Organ, 70: 269 – 275.
- Cuny D. 2012. La biosurveillance végétale et fongique de la pollution atmosphérique : concepts et applications. Annales Pharmaceutiques Françaises,70: 182—187.
- Fuga A, Saiki M, Marcelli MP, Saldiva PHN. 2008. Atmospheric pollutants monitoring by analysis of epiphytic lichens. Environmental Pollution, 151: 334-340.
- Gailey FAY & Lloyd OLI. 1993. Spatial and temporal patterns of airborne metal pollution: The value of low technology sampling to an environmental epidemiology study. Sci. Total Environ, 113: 201–219.
- Glickman TS. 1994. Measuring environmental equity with geographic information systems. Renewable Resources Journal, 12 (3): 17-21.
- Golge M, Yenilmez F, Aksoy A. 2013. Development of pollution indices for the middle section of the Lower Seyhan Basin (Turkey). Ecological Indicators, 29: 6–17.
- Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. 2008. A small-area index of socioeconomic deprivation to capture health inequalities in France. Social Science and Medicine, 67(12): 2007-2016.
- Heitland P & Koster HD. 2006. Biomonitoring of 37 trace elements in blood samples from inhabitants of northern Germany by ICP-MS. J Trace Elem Med Biol, 20(4): 253-62.
- Jany-Catrice F & Zotti R. 2008. Les régions françaises face à leur santé sociale, contribution au débat. Institut pour le développement de l'information économique et sociale.
- McLeod H, Langford IH, Jones AP, Stedman JR, Day RJ, Lorenzoni I, Bateman IJ. 2000. The relationship between socio-economic indicators and air pollution in England and Wales: implications for environmental justice. Regional Environmental Change, 1 (2): 78-85.
- Nisse C. 2010. Imprégnation par des métaux de la population générale. Valeurs d'imprégnation par les métaux de la population générale du Nord - Pas de Calais : établissement de valeurs de base pour l'interprétation des indicateurs d'exposition en milieu professionnel. Rapport final, 109p.
- Ocelli F, Cuny M-A, Devred I, Deram A, Quarré S, Cuny D. 2013. Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques. Pollution atmosphérique [En ligne], 220 mis à jour le : 13/01/2014, URL : <http://lodel.irevues.inist.fr/pollution-atmospherique/index.php?id=2497>.
- Ocelli F, Bavdek R, Deram A, Hellequin A-P, Cuny M-A, Cuny D and Zwarterook I. Using lichen biomonitoring to assess environmental justice over neighborhoods in an industrial area of Northern France. Environmental Health Perspectives (A soumettre).
- Paustenbach D & Galbraith D. 2006a. Biomonitoring and biomarkers: exposure assessment will never be the same. Environ Health Perspect, 114(8): 1143-9.
- Paustenbach D & Galbraith D. 2006b. Biomonitoring: is body burden relevant to public health? Regul Toxicol Pharmacol, 44(3): 249-61.
- Rzepka MA & Cuny D. 2008. Biosurveillance végétale et fongique des éléments traces métalliques atmosphériques. Air Pur, 75: 66—77.
- Salo H, Bucko MS, Vahtovuori E, Limon J, Mäkinen J, Pesonen LJ. 2012. Biomonitoring of air pollution in SW Finland by magnetic and chemical measurements of moss bags and lichens. Journal of Geochemical Exploration, 115: 69–81.

Sargaonkar, A & Deshpande V. 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environmental Monitoring and Assessment*, 89: 43-67.

Sarmiento S, Wolterbeek HT, Verburg TG, Freitas MC. 2008. Correlating element atmospheric deposition and cancer mortality in Portugal: data handling and preliminary results. *Environmental Pollution*, 151: 314-351.

Singh RP, Nath S, Prasad SC, Nema AK. 2008. Selection of Suitable Aggregation Function for Estimation of Aggregate Pollution Index for River Ganges in India. *Journal of Environmental Engineering*, 134 (8): 689-701.

Tomlinson DL, Wilson JG, Harris CR, Jeffrey DW. 1980. Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgoländer Meeresunters*, 33: 566-575.

Van Haluwyn C, Cuny D, Garrec J-P. 2011. Introduction générale - Définitions, concepts et enjeux généraux de la biosurveillance de la qualité de l'air. *Pollution Atmosphérique*, Numéro spécial : 7-11.

Wappelhorst O, Kühn I, Oehlmann J, Markert B. 2000. Deposition and disease: a moss monitoring project as an approach to ascertaining potential connections. *The Science of the Total Environment*, 249: 243-256.

Wolterbeek HT & Verburg TG. 2004. Atmospheric metal deposition in a moss data correlation study with mortality and disease in the Netherlands. *The Science of the Total Environment*, 319: 53-64.

Wolterbeek B, Sarmiento S, Verburg T. 2010. Is there a future for biomonitoring of elemental air pollution? A review focused on a larger-scaled health-related (epidemiological) context. *J Radioanal Nucl Chem*, 286: 195-210.

Sites internet :

CITEPA : www.citepa.org/fr (consulté le 16/01/2014).

CHAPITRE II. DISPARITES SPATIALES D'INCIDENCE DE L'INSUFFISANCE RENALE CHRONIQUE TERMINALE EN NORD – PAS DE CALAIS

Partie 1. Introduction

Dans le chapitre précédent, nous avons associé des données de biosurveillance environnementale et humaine, c'est-à-dire des données de santé continues (concentration sanguine et urinaire) mesurées à l'échelle de l'individu. Ce second chapitre se consacre à l'interprétation de données mesurées à l'échelle de groupes de populations. Il s'agit de l'incidence d'une pathologie, calculée à partir d'une information binaire : malade / non malade.

Dans ce chapitre, nous avons étudié les disparités spatiales d'incidence de l'insuffisance rénale chronique terminale (IRCT), en intégrant les facteurs pouvant générer de telles disparités. La phase terminale de l'insuffisance rénale est synonyme d'un traitement de suppléance, par dialyse ou greffe. A l'échelle nationale, le Réseau Epidémiologie et Information en Néphrologie (REIN) est un registre de santé géré par l'Agence de Biomédecine, qui recense l'ensemble des patients débutant un traitement de suppléance rénale. Grâce à ce registre, des disparités d'incidence de l'IRCT entre certains départements français, ainsi que l'influence de plusieurs facteurs comme le statut socioéconomique, les pratiques médicales ou les maladies à risque comme le diabète ont été mises en évidence (Couchoud *et al.*, 2012). A travers cette étude, nous constatons que la région Nord – Pas de Calais figure parmi les zones de surincidence importante.

Le Registre Néphronor est la déclinaison régionale de REIN et recense l'ensemble des patients de la région Nord – Pas de Calais. Il existe depuis janvier 2005, sous la direction du Dr. F. Glowacki. La première partie de ce chapitre concerne la mise en évidence de variations spatiales d'incidence d'IRCT à plus fine résolution, au sein même de la région Nord – Pas de Calais. L'unité géographique choisie pour le calcul et la cartographie des indicateurs d'incidence est celle du canton-ville. L'étude de telles disparités entre des mailles plus fines que celles des départements nous a permis par la suite d'étudier de manière plus précise l'implication de certains facteurs étiologiques de la maladie, notamment la défaveur sociale, à travers une étude écologique. Cette partie fait l'objet d'un article, soumis en février 2014 à la revue *PLOS ONE* et intitulé « Mapping chronic kidney disease: spatial variations on small area level and relation to deprivation ».

Dans la deuxième partie, ces disparités d'incidence ont été caractérisées à l'aide de la statistique de scan (Kulldorff, 1997). Cette technique n'est que très peu utilisée dans le contexte des maladies rénales chroniques. Nous avons ainsi pu isoler des clusters atypiques d'incidence de l'IRCT, synonymes de points noirs sanitaires. Cette étude fait l'objet d'un article intitulé « End-Stage Renal Disease in Northern France: isotonic cluster detection », prochainement soumis à la revue *Nephrology Dialysis Transplantation*. Nous avons également utilisé une méthode isotonique de détection de clusters, assez novatrice dans ce type d'études (Kulldorff, 1999). Contrairement aux méthodes classiques, la version isotonique de SaTScan™ peut identifier un risque variable au sein d'un même cluster, ce qui apporte une information complémentaire et plus fine dans la perspective de recherche de facteurs étiologiques de la maladie. L'utilisation d'une telle méthode est originale, puisque peu d'études épidémiologiques s'intéressent actuellement à ce niveau de précision.

Ainsi dans ce chapitre, nous cherchons à mettre en évidence si les populations présentant un taux d'incidence élevé d'IRCT sont les plus défavorisées au sein de la région Nord – Pas de Calais. Dans chaque partie, les résultats de nos recherches sont présentés sous la forme d'un article, précédé d'une introduction détaillée.

Partie 2. Disparités spatiales d'incidence de l'IRCT en Nord – Pas de Calais en lien avec la défaveur sociale

a. Résumé de l'article

De fortes disparités d'incidence de l'IRCT ont récemment été observées entre les départements de France métropolitaine (Couchoud *et al.*, 2012). La région Nord – Pas de Calais présente les taux d'incidence les plus élevés. Compte tenu des caractéristiques présentées dans le chapitre de Matériels et méthodes (densité de population, hétérogénéité en termes d'environnement et de défaveur sociale), elle est une aire géographique d'intérêt pour une analyse spatiale fine, c'est-à-dire à l'échelle des cantons.

L'ensemble des cas incidents d'IRCT recensés dans le registre Néphronor entre 2005 et 2011 ont été récoltés, puis géoréférencés au canton de résidence au moment du diagnostic. Des analyses spatiales par cartographie de ratios d'incidence standardisés (SIR lissés, ajustés sur l'âge et le sexe) ont été menées pour mettre en évidence des disparités d'incidence entre les 170 cantons de la région. Une analyse de corrélations écologiques a ensuite été mise en œuvre pour évaluer les liens avec le niveau de défaveur sociale (modélisé par l'indice de Townsend) et d'autres facteurs pouvant générer de telles disparités spatiales.

Pour un total de 4597 cas incidents recensés entre 2005 et 2011, le taux d'incidence annuel régional est de 163 par million d'habitants (pmh). La région présente de fortes disparités entre les 170 cantons, avec une incidence variant de 38 à 432 pmh. La cartographie des SIR lissés met en exergue une hétérogénéité spatiale, avec des zones de surincidence significative et de sous-incidence significative.

Une corrélation significative des SIR lissés avec le niveau de défaveur sociale caractérisé par l'indice de Townsend est obtenue. En comparaison aux cantons les plus favorisés, le risque relatif est de 1,44 pour les cantons les plus défavorisés, avec un intervalle de confiance à 95% de 1,32 – 1,74. De telles corrélations ne sont pas obtenues avec les autres facteurs étudiés (pratiques médicales, répartition de néphropathies diabétiques ou vasculaires). Ainsi, d'importantes disparités spatiales d'incidence d'IRCT sont mises en évidence au sein même de la région Nord – Pas de Calais entre 2005 et 2011. Ces disparités peuvent en partie être expliquées par une hétérogénéité spatiale du niveau de défaveur sociale entre les cantons. Les points noirs d'inégalités sociales de santé sont observés au niveau des agglomérations de Maubeuge, Dunkerque et Lille-Roubaix-Tourcoing. A notre connaissance, un tel travail n'a pas d'équivalent en France.

Ce travail a fait l'objet d'un article soumis en février 2014 à la revue *PLOS ONE* et intégré à la suite de ce résumé. Ces résultats ont également été présentés lors d'une conférence anglophone (13^{ème} Journée André Verbert), organisée en septembre 2013 par l'Ecole Doctorale Biologie Santé de Lille, ainsi qu'à la conférence du Groupement des Néphrologues Francophones de Belgique, en mars 2013 à Tournai (Belgique). Ils seront enfin présentés lors de la 16^e Réunion Commune de la Société Francophone de Dialyse (SDF) et de la Société de Néphrologie (SN) organisée du 30 septembre au 3 octobre 2014 à Saint-Etienne.

b. [Article](#)

FULL TITLE

Mapping end-stage renal disease: spatial variations on small area level in France and relation to deprivation, intensity of renal replacement care and related diseases.

SHORT TITLE

Mapping chronic kidney disease: spatial variations on small area level and relation to deprivation.

AUTHORS AND AFFILIATIONS

Florent Occelli, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France.

Annabelle Deram, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France. Faculté Ingénierie et Management de la Santé (ILIS), Loos, France.

Michaël Génin, EA 2694, Université Lille Nord de France, Faculté de Médecine pôle Recherche, Lille, France.

Damien Cuny, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France.

François-Xavier Glowacki, EA 4483, Réseau Néphronor, Hôpital Huriez, Lille, France (www.nephronor.fr).

Néphronor Network, Réseau Néphronor, Hôpital Huriez, Lille, France (www.nephronor.fr).

KEYWORDS

Chronic Kidney Disease, End-Stage Renal Disease, Disease mapping, Deprivation index

ACKNOWLEDGEMENT

The authors would like to thank Vincent LEMAITRE, who developed the database of patients with chronic kidney disease for the Nord-Pas-de-Calais region of France. Furthermore, we gratefully acknowledge the outstanding technical support of Hasna Camara, Sebastien Gomis and the staff of the Nephronor network: AAZIB Larbi, ADDA Hassen, AL MORABITI Mustapha, AL MOUBARAK Imad, AZAR Raymond, BACRI Jean-Louis, BATAILLE Pierre, BEN HENDA Amaury, BENZIANE Abdelkader, BILLION Stephane, BINAUT Raynald, BLOCH Julie, BONNARD Guillaume, BONNE J-François, BOTTE-NOËL Alexandra, BOUBIA Toufik, BOUBIA Véronique, BOURDON Franck, BRASSEUR José, BURDA Guillaume, CARDON Gérard, CHLIH Bouchra, DAROUX Maïté, DECAMBRON Mélanie, DEHENNAULT Maud, DELVALLEZ Luc, DESPREZ Jonathan, DEVAUX Jean Philippe, DUFAY Alexandre, EL AMARI Aderrahim, FLEURY Dominique, FRIMAT Marie, GARSTKA Antoine, GHEERBRANDT J. Dominique, GLOWACKI Francois, GOSSET Brigitte, GOUBET Marlène, GUINCESTRE Thomas, GUYON – ROGER Tiphaine, HAMDINI Nasser, HAMMELIN Jean Philippe, HARDY-YVERNEAU Paule, JOMAA Zacharie, HAZZAN Marc, HOFFMANN Maxime, LABATUT Delphine, LAHOUCHE Annie, LAMOTTE Christian, LE MONIES Hervé, LEBAS Céline, LEBLEU Jacques, LEMOINE Corinne, LESSORE DE STE FOY Celia, LIONET Arnaud, LOUVET Daniel, MAC NAMARA Evelyne, MAISONNEUVE Nathalie, M'BARGA Roseline, MESBAH Rafik, MESSIER Gilles, MOULONGUET Florence, MOUSSALIEH Laura, NOEL Christian, NOVO Robert, PAGNIEZ Dominique, PAINCHART Bernard, PROVÔT Francois, RATSIMBAZAFY Anderson, READE Richard, REBEROLLE Matthieu, ROBITAILLE Géraldine, SEMJEN Elisabeth, SHAHAPUNI Irina, SHENOUDA Milad, SKALLI Amina, TAGHIPOUR TAMIJI Lili, TALASZKA Aline, VAIRON-CODACCIONI M. Xavière, VAN EGROO Anne, VANHILLE Philippe, VRIGNEAUD Laurence, WAJSBROT Lucie, WHEATLEY Pascal.

FIGURE LEGEND

Figure 1. Crude incidence rate of ESRD (pmi) by age and gender.

Figure 2. Smoothed SIRs of ESRD by cantons, 2005-2011.

Figure 3. Spatial distribution of Townsend deprivation index by cantons, 2009.

Figure 4. Relative risk (95% credible interval) of ESRD by Townsend quintile.

Figure 5. Smoothed SIRs of ESRD by cantons, 2005-2011, adjusted to the Townsend index covariable, 2009.

TABLES

Table 1. Summary statistics of canton characteristics.

ABSTRACT

Background

Strong geographic variations in the incidence of end-stage renal disease (ESRD) are observed in developed countries. The reasons for these variations are unknown. They may reflect regional inequalities in the population's sociodemographic characteristics, related diseases, or medical practice patterns. In France, at the district level, the highest incidence rates have been found in the Nord-Pas-de-Calais region. This area, with a high population density and homogeneous healthcare provision, represents a geographic situation which is quite suitable for the study, over small areas, of spatial disparities in the incidence of ESRD, together with their correlation with a deprivation index and other risk factors.

Methods

The Renal Epidemiology and Information Network is a national ESRD registry, which lists all patients who initiated renal replacement therapy in France. All cases included in the Nord-Pas-de-Calais registry between 2005 and 2011 were extracted. Adjusted crude incidence rate and smoothed standardized incidence ratio (SIR) were calculated for each of the 170 cantons. The correlation between ESRD incidence and deprivation was assessed using the Townsend index. Relative risk (RR) and credible intervals (CI) were estimated for each quintile.

Results

Significant spatial disparities in ESRD incidence were found within the Nord-Pas-de-Calais region. The sex- and age-adjusted, smoothed SIRs varied from 0.66 to 1.69. Although no correlation is found with diabetic or vascular nephropathy, the smoothed SIRs are correlated with the Townsend index (RR: 1.18, 95% CI [1.00–1.34] for Q2; 1.28, 95% CI [1.11–1.47] for Q3; 1.30, 95% CI [1.14–1.51] for Q4; 1.44, 95% CI [1.32–1.74] for Q5).

Conclusion

For the first time at this aggregation level in France, this study reveals significant geographic differences in ESRD incidence. Unlike the intensity of renal replacement care, deprivation is certainly a determinant in this phenomenon. This association is probably independent of the patients' financial ability to gain access to healthcare.

INTRODUCTION

In developed countries, the burdens of End-Stage Renal Disease (ESRD) and Renal Replacement Therapy (RRT) are continuously growing. Nevertheless, at the scale of a country, strong geographic variations in the incidence of treated ESRD have been observed (Rosansky *et al.*, 1990; Roderick *et al.*, 1999; Usami *et al.*, 2000; Wimmer *et al.*, 2003; Counil *et al.*, 2008; Hommel *et al.*, 2010; USRDS, 2013; Gilg *et al.*, 2012; Tanner *et al.*, 2013). In metropolitan France, the crude incidence rate of RRT also varies widely, from 80.4 to 238.6 per million inhabitants (pmi) in 71 districts, from 2006 to 2007 (Couchoud *et al.*, 2010), and from 85.8 to 225.5 pmi in 85 districts, from 2008 to 2009 (Couchoud *et al.*, 2012).

The reasons for these variations remain elusive. They may result from inter-regional variations in the population's sociodemographic characteristics (Roderick *et al.*, 1999; Ward, 2008; Hommel *et al.*, 2010; Couchoud *et al.*, 2012; Grace *et al.*, 2012), from other related diseases such as diabetes and cardiovascular diseases (Muntner *et al.*, 2003; Wimmer *et al.*, 2003; Couchoud *et al.*, 2012; Bell *et al.*, 2012; Huang *et al.*, 2013), or merely reflect differences in the timing of dialysis initiation (Couchoud *et al.*, 2010; Van de Luijtgarden *et al.*, 2012) and geographic distance from healthcare facilities (Boyle *et al.*, 1996).

On a smaller scale, the spatial variability of treated ESRD and its relationship to risk factors has also been assessed for US counties or census tracts (Foxman *et al.*, 1991 ; Moulton *et al.*, 1992 ; Young *et al.*, 1994 ; Joyce Fan *et al.*, 2007 ; Yan *et al.*, 2013; Volkova *et al.*, 2008), UK wards (Boyle *et al.*, 1996) and Australian postcodes (Grace *et al.*, 2012). There is no such equivalent in France. Such a spatial approach, made at the scale of homogeneous populations over a territory with similar health practices, may provide an improved knowledge of ESRD patterns, allowing an explanation to be found for these disparities. Such an outcome would then lead to a better understanding of environmental assumptions (Hellström *et al.*, 2001; Hodgson *et al.*, 2007; Muntner *et al.*, 2007).

The French Renal Epidemiology and Information Network (REIN) is a national Chronic Kidney Disease (CKD) registry, which lists all patients who initiated ESRD treatment since 2002, and is currently available in 22 regions of metropolitan France (Couchoud *et al.*, 2006). Among these, the Nord-Pas-de-Calais region has the highest ESRD incidence, with a standardized rate of 198 new cases pmi in 2011, as opposed to 149 new cases pmi for all of France (REIN, 2011). The large number of cases and the associated population density make the geographic situation of this region quite suitable for the study of disparities in ESRD incidence in small areas, and their correlation with sociodemographic status, as well as the quality of environmental media.

The aim of this study was to analyze the spatial variations of ESRD incidence over small areas, and to analyze the correlation between geographic variability and social discrepancies (assessed using the Townsend deprivation index). The study focuses on the Nord-Pas-de-Calais, a small region with 4 million inhabitants and homogeneous nephrological practices, for the period between 2005 and 2011.

MATERIAL & METHODS

Study area and sources of data

The Nord-Pas-de-Calais region has a surface area of 12 481 km² with approximately 4 033 000 inhabitants, including both rural, industrial and urban regions. The region's 170 cantons (a French small administrative unit) were used to represent distinct spatial units. These are referenced by the National Institute of Statistics and Economic Studies (INSEE), and in 2009 had an average population of 23 725 (extremes ranging from 4 991 to 226 827) and an average surface area of 73 km² (extremes ranging from 2 to 258 km²).

Cases were defined as all incident patients requiring RRT registered in the REIN (patients were registered on the first day of RRT) in the Nord-Pas-de-Calais region, from January 2005 to December 2011 (REIN, 2013). For the purpose of this study, all patients were grouped into cantons according to the postcode of their residence, determined at the time of their first RRT. They were ranked by sex and 5-year age group. The following characteristics of the patients were also collected: primary kidney disease, number of visits to a physician in the year preceding treatment, estimated Glomerular Filtration Rate (eGFR) by Modification of Diet in Renal Disease (MDRD) formula at the time of dialysis initiation (Froissart *et al.*, 2005).

Demographic and socioeconomic data were extracted from the 2009 national population census, provided by the INSEE. The population data was also ranked by sex and 5-year age groups. The Townsend deprivation index was computed from 4 variables: percentages of non owner-occupied households, unemployment, household overcrowding, and absence of access to a motor vehicle (Townsend, 1987).

Statistical analysis

Incidence rates and deprivation index

Firstly, the male and female crude incidence rates were estimated for the 170 cantons, by means of direct standardization. Crude incidence is the number of new patients divided by the total population at risk during the study period. The SIR, defined as the ratio of the number of observed cases to the expected number of cases, computed using indirect standardization, was then determined for each canton. Significant SIRs have a 95% credible interval, which does not contain the value 1. The method used to calculate the incidence rate denominators assumed the population in a given canton to remain constant over the study period.

On the basis of our registry, the proportion of cases with diabetic and/or vascular nephropathy was calculated for each canton. The Townsend index for each canton was classified into quintiles, with the first quintile (Q1) corresponding to the least deprived cantons, and the fifth quintile (Q5) corresponding to the most deprived cantons.

The Pearson correlation was used to assess the correlation between the logarithm of the SIR and the eGFR.

Spatial analysis

The centroid of each canton, defined by its geographical center (longitude and latitude), was used for the spatial analysis. In order to take the instability resulting from low frequencies and spatial autocorrelation effect into account, the SIR was smoothed using the hierarchical Bayesian model with three levels, proposed in Besag *et al.* (1991). At the first level, the observed number of cases in the i^{th} canton O_i is assumed to be Poisson distributed, with a mean $\theta_i E_i$, where θ_i is the relative risk (RR) associated with canton i , and E_i is the expected number of cases calculated by means of indirect standardization. At the second level, the a priori distribution of the logarithm of θ_i is defined as the sum of two random effects, the first corresponding to unstructured spatial heterogeneity and the second describing the correlation between the neighboring cantons (sharing a common boundary). At the third level, the variances associated with the previous random effects are assumed to be Gamma distributed.

In order to analyze the association between ESRD and deprivation index, the quintile of the latter was introduced at the second level of the hierarchical model.

The models were fitted using Markov Chain Monte Carlo methods with 25 000 iterations, following a burning step involving 5 000 iterations. All of the calculations were made using the WinBUGS Software (Spiegelhalter *et al.*, 2002), and the maps were produced using the ArcGIS® 10.1 software (<http://www.esri.com>). All statistical analyses were considered significant at the 0.05 type 1 error.

RESULTS

The study included 4 597 patients (57.2% men and 42.8% women), who began an RRT between 2005 and 2011 in the Nord-Pas-de-Calais. The overall crude annual incidence rate was 163 pmi for both sexes, 193 pmi for men and 135 pmi for women, with strong disparities over the 170 cantons, ranging from 38 to 432 pmi for both sexes combined, and from 0 to 424 pmi for men and 0 to 439 pmi for women (Table 1). As expected, the crude incidence rate increases sharply with age, and males have a relatively higher proportion of ESRD incidence (Fig. 1). The median age of incidents is 69 years for men, 72 years for women, and 71 years for both sexes combined.

Table 1. Summary statistics of canton characteristics.

	Mean	SD	Min	Q1	Median	Q3	Max
Number of patients	27	26	2	12	21	32	219
Total crude incidence (pmi)	158	53	38	124	157	186	432
Men crude incidence (pmi)	187	72	0	144	182	230	424
Women crude incidence (pmi)	130	61	0	89	123	169	439
Smoothed SIR	0.96	0.20	0.66	0.81	0.92	1.07	1.69
Townsend index	0.00	3.71	-6.01	-2.91	-0.97	2.82	10.80
Diabetic nephropathy (%)	27.5	11.7	0.0	20.0	27.6	33.3	75.0
Vascular nephropathy (%)	24.4	12.5	0.0	16.7	23.7	32.0	100.0
eGFR (ml/min/1.73 m ²)	9.05	1.46	5.59	8.05	8.92	9.84	14.40

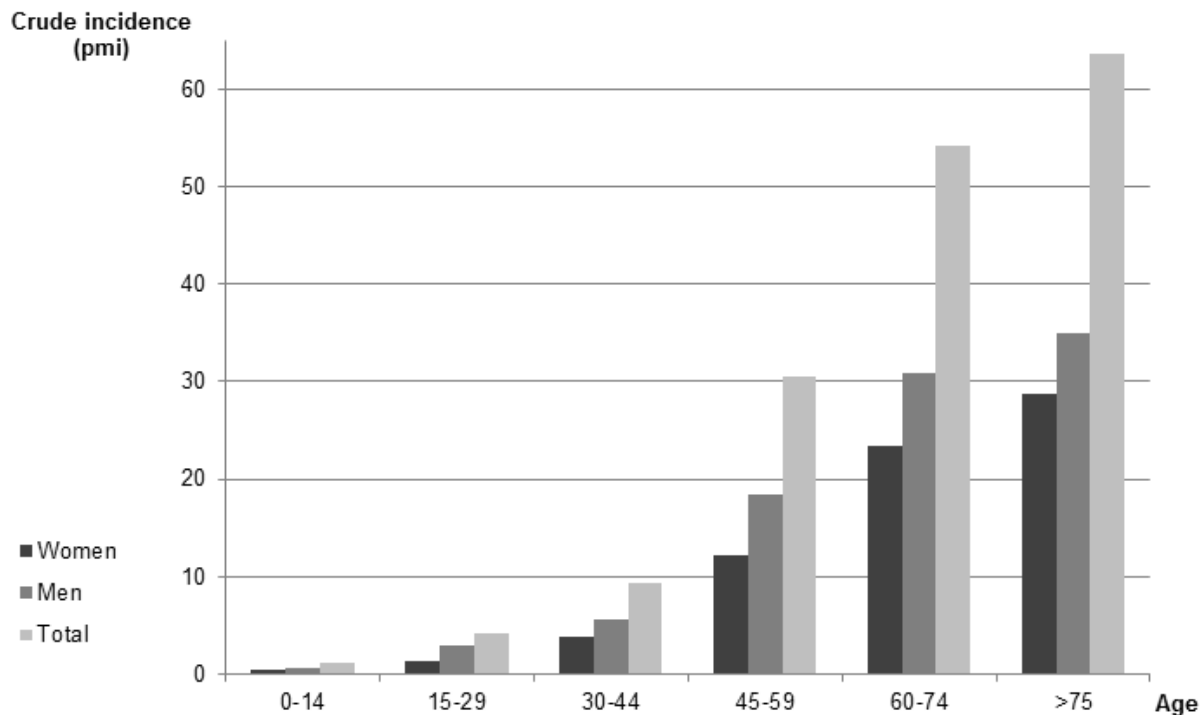


Figure 1. Crude incidence rate of ESRD (pmi) by age and gender.

The smoothed SIRs vary among cantons, from 0.66 to 1.64 (Table 1), and there is a significant spatial variability of SIR within the Nord-Pas-de-Calais region (Fig. 2). Cantons with significantly high incidence rates are observed in the South-East rural area of Maubeuge and Hautmont (where SIRs are respectively 1.64, 95% CI [1.32–2.01] and 1.50, 95% CI [1.18–1.88]), the industrial area of Grande-Synthe and Dunkerque (1.55, 95% CI [1.18–2.01] and 1.27, 95% CI [1.10–1.47] respectively), and the urban agglomerations of Roubaix, Tourcoing and Wattrelos (1.48, 95% CI [1.26–1.73], 1.32 95% CI [1.13–1.54] and 1.33, 95% CI [1.07–1.63] respectively). Cantons with significantly low incidence rates are observed in the South-West (SIR: 0.66, 95% CI [0.46–0.89]).

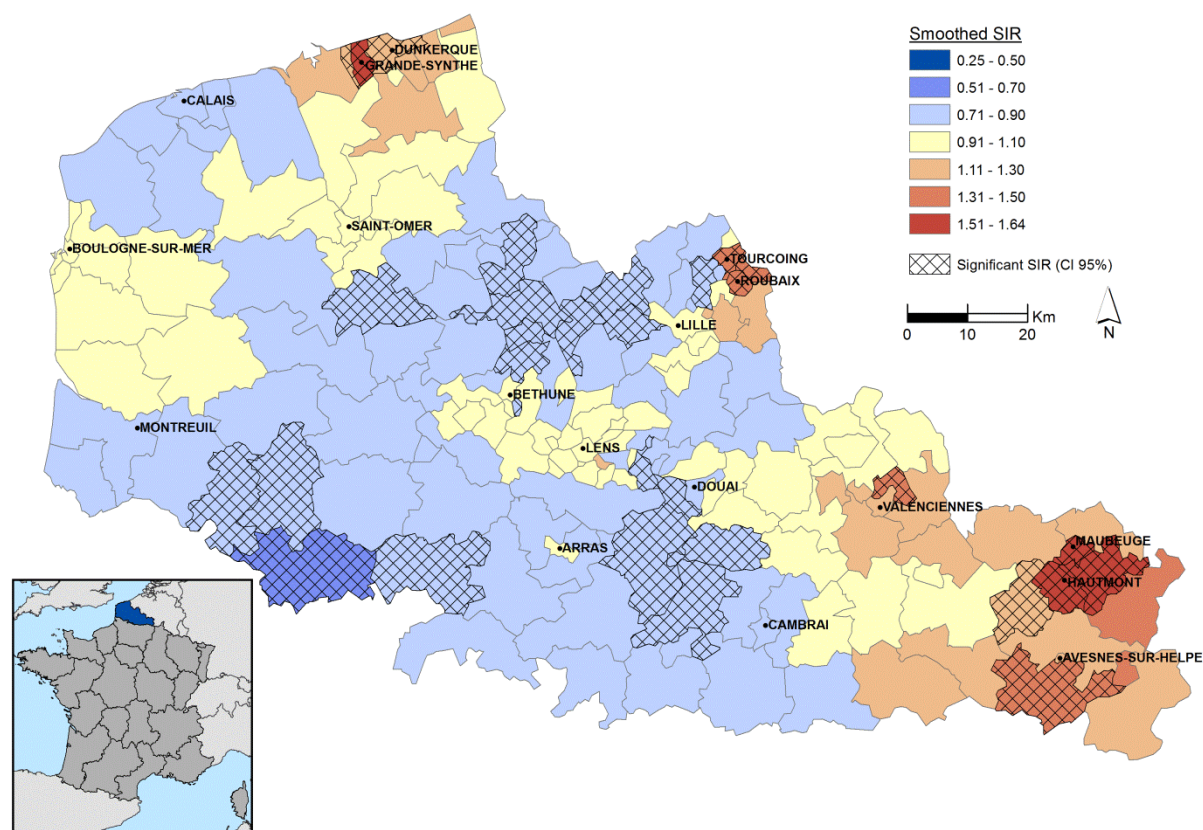


Figure 2. Smoothed SIRs of ESRD by cantons, 2005-2011.

Before assessing the relationship between deprivation and these spatial disparities, the influence of intensity of care was examined. Early dialysis initiation was determined by measuring the median eGFR for each canton (Table 1). Although it is significant ($p < 0.05$), the Pearson coefficient ($R = 0.23$) indicates a very weak positive correlation with the smoothed SIRs.

Disparities are observed in the proportions of diabetic or vascular nephropathy occurring in the different cantons within the region (Table 1). These are not correlated with the smoothed SIRs. The RR is 1.31, 95% CI [0.87-1.97] for diabetic nephropathy and 0.98, 95% CI [0.65-1.48] for vascular nephropathy.

The Townsend deprivation index varies strongly within the region (Table 1, Fig. 3), and there is a significant correlation between smoothed SIRs and the Townsend index quintiles. With Q1 taken as a reference, the relative risk (RR) of RRT was assessed for each level of deprivation (Fig. 4). Higher levels of deprivation are associated with an increase in RR: 1.18, 95% CI [1.00–1.34] for Q2, 1.28, 95% CI [1.11–1.47] for Q3, 1.30, 95% CI [1.14–1.51] for Q4 and 1.44, 95% CI [1.32–1.74] for Q5. In addition, the residential Townsend index for each collected case was compared with late referral, estimated by means of the number of visits to a physician in the year preceding treatment. No correlation is found ($p = 0.66$).

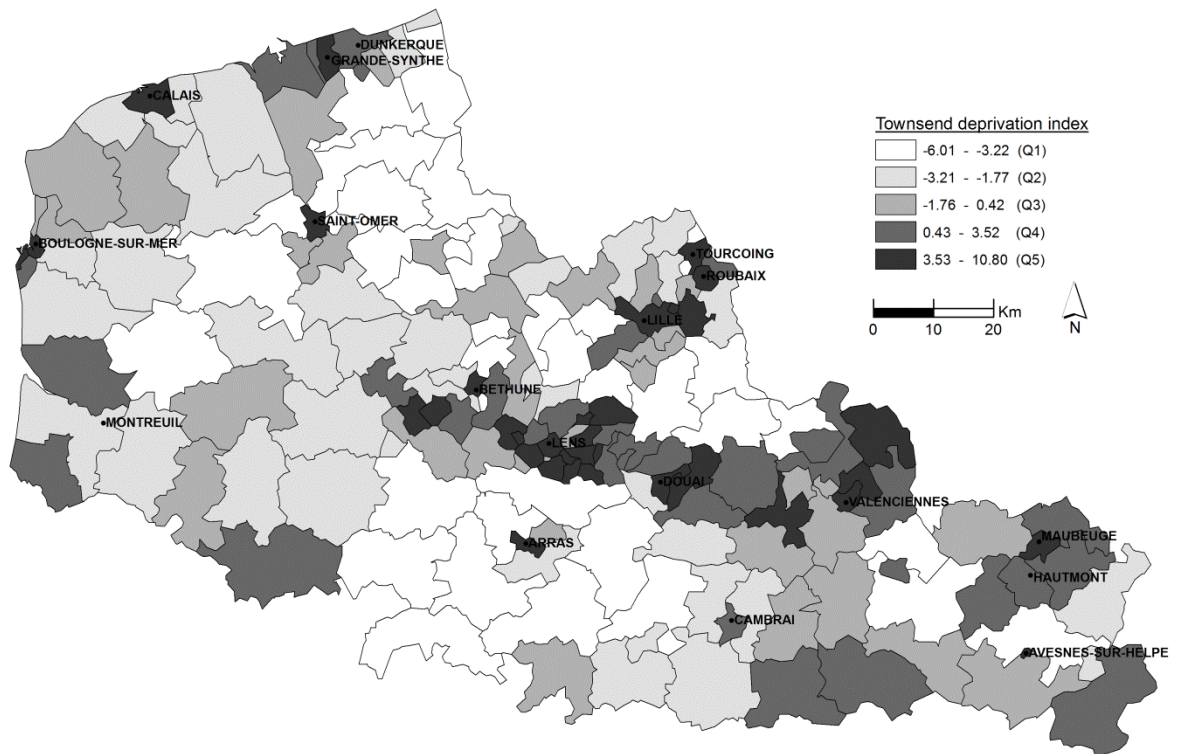


Figure 3. Spatial distribution of Townsend deprivation index by cantons, 2009.

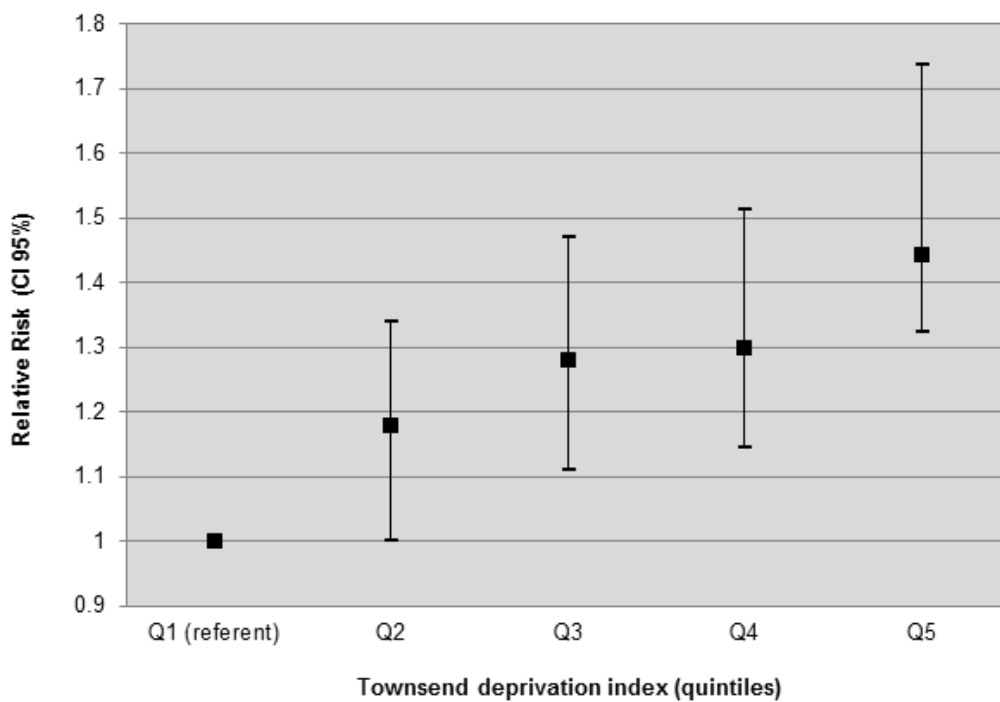


Figure 4. Relative risk (95% credible interval) of ESRD by Townsend quintile.

Analysis of the geographic patterns of the smoothed SIRs, adjusted to the Townsend covariable, still reveal three areas of significantly high risk in the north and south-east of the region, in the Lille metropolis and the agglomerations of Roubaix, Tourcoing and Wattrelos. Two significantly low-risk areas are also observed in the south-west of the region and the area between Saint-Omer and Lille (Fig. 5). The corresponding smoothed SIRs vary among cantons, from 0.66 to 1.69.

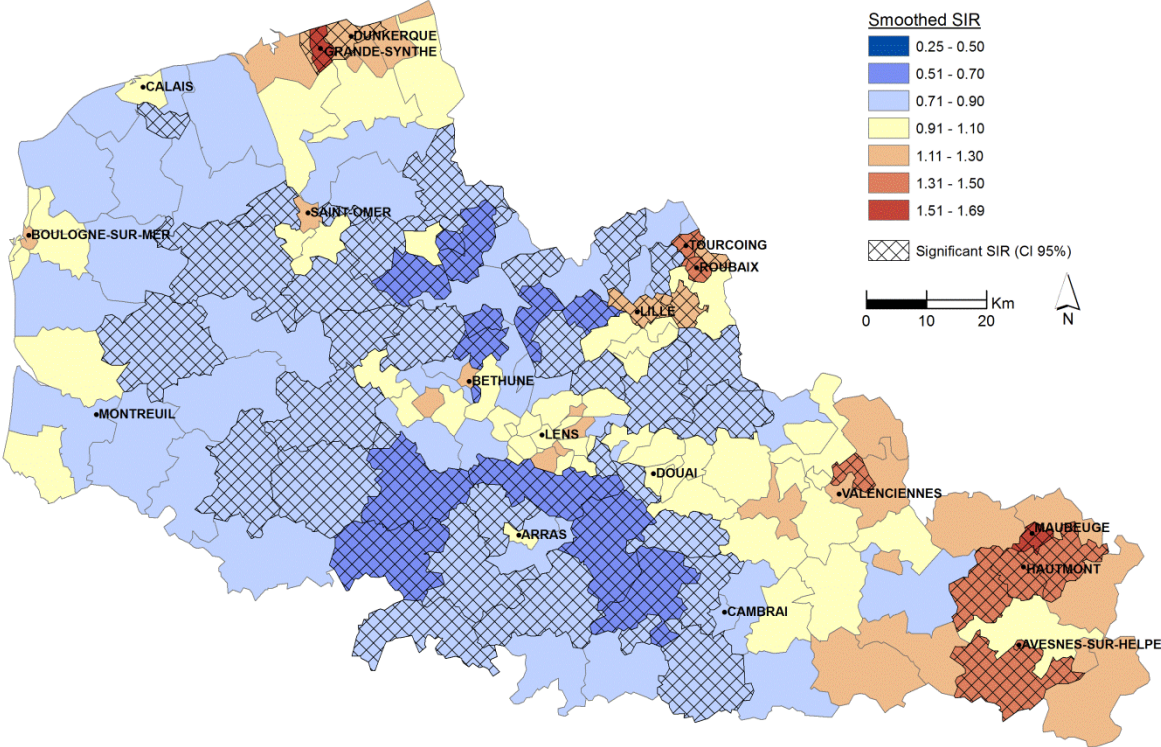


Figure 5. Smoothed SIRs of ESRD by cantons, 2005-2011, adjusted to the Townsend index covariable, 2009.

DISCUSSIONS

In this study, significant spatial disparities in ESRD incidence are revealed in the 170 cantons of the Nord-Pas-de-Calais region of France. Although such disparities have recently been observed among districts, they had never been seen at this spatial resolution in France. These results are consistent with variations observed at similar scales in other countries.

Several factors were analyzed, in an attempt to explain this phenomenon. It should be noted that any ecological correlations were made at the level of geographical areas, and not individuals. As a consequence, although causality cannot be assumed, etiological hypotheses can be proposed (Wakefield, 2008).

In France, Couchoud *et al.* (2010) found that the intensity of healthcare has a substantial impact on RRT incidence at the level of individual districts. Concerning the Nord-Pas-de-Calais region, a significant, but weak association was found between the median eGFR and the smoothed SIRs. A higher level of healthcare appears to make a negligible contribution to spatial disparities in the incidence of ESRD within this region. This result confirms the presence of uniform medical practice throughout this territory.

The spatial heterogeneity of ESRD is not related to the incidence of diabetic or vascular nephropathies. This result means that cantons with high incidence rates are not associated with a majority of diabetic or vascular nephropathies. The cases developing such CKD were not the cause of the observed phenomenon. However, several studies have identified ecological relationships between RRT incidence and the prevalence of diabetics or cardio-vascular diseases (Muntner *et al.*, 2003; Wimmer *et al.*, 2003; Couchoud *et al.*, 2012; Bell *et al.*, 2012; Huang *et al.*, 2013). As no data was available concerning the prevalence of these diseases at the scale of each canton, the present study data was used as a proxy, to determine the proportion of diabetic and vascular nephropathies in each canton.

On the other hand, deprivation is clearly associated with a higher ESRD incidence, and wealth with a lower ESRD incidence. These findings are consistent with results observed on a larger scale in France, since Couchoud *et al.* (2012) revealed a positive relationship between 82 districts, whatever the socio-economic factors used. The patterns shown here are similar to those found in other studies, which focused on deprivation indicators over larger heterogeneous geographic areas. Caskey *et al.* (2012) identified a correlation between RRT incidence and a country's macroeconomic factors, such as gross domestic product (GDP) per capita, percentage of GDP spent on health care, and dialysis facility reimbursement rate relative to GDP. Although Ward *et al.* (2008) highlighted a greater incidence of ESRD in patients living in Zoning Improvement Plan (ZIP) areas with a lower composite socioeconomic score, this trend was not uniform for all primary renal diseases. In the case of US counties, the incidence rate of treated ESRD has been shown to be inversely related to the level of income (Young *et al.* 1994). For similar, highly homogenous small area units, Grace *et al.* (2012) recently found a decreasing incidence of RRT with increasing area advantage in Australia. In the UK, deprivation is found to be a determinant of geographical variations in RRT, between wards or enumeration districts (Roderick *et al.*, 1999; Maheswaran *et al.*, 2003). Furthermore, Volkova *et al.* (2008) have revealed a strong correlation between incidence rates and neighborhood poverty, corresponding to populations living below the poverty level, for the case of the census tracts of Georgia, North Carolina, or South Carolina. Although the Townsend index is criticized for its urban view of deprivation, it is nevertheless widely used. Furthermore, some recently developed indexes have been shown to be strongly correlated with this one (Declercq *et al.*, 2004; Havard *et al.*, 2008).

In agreement with other studies, the assumption is made that deprivation is an obstacle to prevention, and that it supports the progression of CKD to ESRD (Ward *et al.*, 2008; Couchoud *et al.*, 2012). To substantiate this hypothesis, the influence of late referral from collected cases was assessed, by determining the number of visits to a physician during the year preceding treatment. There was no significant correlation between this number and the Townsend index. Patients who

were treated later do not live primarily in disadvantaged townships. Moreover, in France, access to healthcare and medical insurance coverage does not discriminate against poverty. The medical or hospital fees associated with CKD treatment are also completely covered. It is thus possible that the correlation between deprivation and ESRD incidence is not related to an individual's financial capacity to access healthcare. Other more relevant factors associated with deprivation, such as health literacy, acculturation or trust in healthcare providers, could explain these variations. This is supported by Lora *et al.* (2011), who report that lower levels of health literacy and acculturation are associated with differences in knowledge, attitude, and behavior, that may contribute to a poor outcome in patients with CKD.

In this paper, it is shown that SIR mapping can be used to highlight global spatial heterogeneities in ESRD incidence. Although this method is needed to reveal spatial patterns of interest, it cannot detect significant atypical spatial and space-time clusters in terms of ESRD incidence. Spatial and space-time scan statistics (Kulldorff 1997; Kulldorff *et al.* 1998) should thus be used to test for the presence of ESRD clusters, and to identify their location in space and time.

When the SIR map is adjusted to the Townsend covariable, areas of significantly high incidence still exist. These are not explained in the present study. Other etiological assumptions such as environmental contamination should be assessed in the future. As it includes industrial, urban and agricultural territories, the Nord-Pas-de-Calais region is suitable for such a study. This should be conducted at the level of small areas, to avoid a dilution of the spatial variations characterizing nearby, heterogeneous populations.

CONCLUSION

In developed countries, significant variations in ESRD incidence are observed over small areas (Roderick *et al.*, 1999; Volkova *et al.*, 2008; Grace *et al.*, 2012). The present study shows that this is also the case in France. Within the Nord-Pas-de-Calais territory, which has homogeneous healthcare provision, this phenomenon can be partially explained by deprivation. However, since access to healthcare is universal in France, and in the case of serious illnesses such as ESRD, this access is not affected by an individual's financial well-being, other deprivation-related factors may explain the observed correlation. Moreover, these disparities are not related to a specific type of medical practice, related to the initiation of extrarenal purification. The SIR can be used as a geographic tool, for decision-making in the management of dialysis units and the definition of prevention campaigns at local scales. Even when the deprivation factor is taken into account, spatial disparities in ESRD incidence remain, suggesting that environmental factors such as suspected heavy metal contamination (Hellström *et al.*, 2001; Hodgson *et al.*, 2007; Muntner *et al.*, 2007) still play a significant role. The authors plan to further investigate this aspect in future studies.

REFERENCES

- Bell EK, Gao L, Judd S, Glasser SP, McClellan W, *et al.* (2012) Blood pressure indexes and end-stage renal disease risk in adults with chronic kidney disease. *American Journal of Hypertension* 25(7): 789-96.
- Besag J, York J, Mollié A. (1991) Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43: 1–21.
- Boyle PJ, Kudlac H, Williams AJ. (1996) Geographical variation in the referral of patients with chronic end stage renal failure for renal replacement therapy. *QJM - Monthly Journal of the Association of Physicians* 89(2): 151-7.
- Caskey FJ, Kramer A, Elliott RF, Stel VS, Covic A, *et al.* (2011) Global variation in renal replacement therapy for end-stage renal disease. *Nephrology Dialysis Transplantation* 26(8): 2604-10.
- Couchoud C, Guihenneuc C, Bayer F, Lemaitre V, Brunet P, Stengel B. (2012) Medical practice patterns and socio-economic factors may explain geographical variation of end-stage renal disease incidence. *Nephrology Dialysis Transplantation* 27(6): 2312-22.
- Couchoud C, Guihenneuc C, Bayer F, Stengel B. (2010) The timing of dialysis initiation affects the incidence of renal replacement therapy. *Nephrology Dialysis Transplantation* 25(5): 1576-8.
- Couchoud C, Stengel B, Landais P, Aldigier J-C, De Cornelissen F, *et al.* (2006) The renal epidemiology and information network (REIN): a new registry for end-stage renal disease in France. *Nephrology Dialysis Transplantation* 21: 411–418.
- Counil É, Cherni N, Kharrat M, Achour A, Trimech H. (2008) Trends of incident dialysis patients in tunisia between 1992 and 2001. *American Journal of Kidney Diseases* 51(3): 463-70.
- Declercq C, Labbe E, Obein L, Poirier G, Lacoste O. (2004) Inégalités socio-spatiales de mortalité dans la région Nord - Pas-de-Calais. *Rapport Observatoire Régional de la Santé Nord – Pas de Calais*. 105 p.
- Fan ZJ, Lackland DT, Lipsitz SR, Nicholas JS, Egan BM, *et al.* (2007) Geographical patterns of end-stage renal disease incidence and risk factors in rural and urban areas of South Carolina. *Health and Place* 13(1): 179-87.
- Foxman B, Moulton LH, Wolfe RA, Guire KE, Port FK, Hawthorne VM. (1991) Geographic variation in the incidence of treated end-stage renal disease. *Journal of the American Society of Nephrology* 2(6): 1144-52.
- Froissart M, Rossert J, Jacquot C, Paillard M, Houillier P. (2005) Predictive performance of the modification of diet in renal disease and Cockcroft-Gault equations for estimating renal function. *Journal of the American Society of Nephrology* 16(3): 763-73.
- Gilg J, Rao A, Fogarty D (2013) UK Renal Replacement Therapy Incidence in 2012: National and Centre-specific Analyses. In: *UK Renal Registry 16th Annual Report* pp. 9-35.
- Grace BS, Clayton P, Cass A, McDonald SP. (2012) Socio-economic status and incidence of renal replacement therapy: A registry study of australian patients. *Nephrology Dialysis Transplantation* 27(11): 4173-80.
- Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. (2008) A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science and Medicine* 67(12): 2007-16.
- Hellström L, Elinder C, Dahlberg B, Lundberg M, Järup L, *et al.* (2001) Cadmium exposure and end-stage renal disease. *American Journal of Kidney Diseases* 38(5): 1001-8.
- Hodgson S, Nieuwenhuijsen MJ, Elliott P, Jarup L. (2007) Kidney disease mortality and environmental exposure to mercury. *American journal of epidemiology* 165(1): 72-7.

- Hommel K, Rasmussen S, Kamper A, Madsen M. (2010) Regional and social inequalities in chronic renal replacement therapy in Denmark. *Nephrology Dialysis Transplantation* 25(8): 2624-32.
- Huang Y, Cai X, Zhang J, Mai W, Wang S, *et al.* (2014) Prehypertension and incidence of ESRD: A systematic Review and Meta-analysis. *American Journal of Kidney Diseases* 63(1): 76-83.
- Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. (1998) Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *American journal of Public Health* 88(9): 1377-1380.
- Kulldorff M. (1997) A spatial scan statistic. *Communications in statistics: theory and methods* 26(6): 1481-1496.
- Lora CM, Gordon EJ, Sharp LK, Fischer MJ, Gerber BS, Lash JP. (2011) Progression of CKD in hispanics: Potential roles of health literacy, acculturation, and social support. *American Journal of Kidney Diseases* 58(2): 282-90.
- Maheswaran R, Payne N, Meechan D, Burden RP, Fryers PR, *et al.* (2000) Socioeconomic deprivation, travel distance, and renal replacement therapy in the Trent region, United Kingdom 2000: An ecological study. *Journal of epidemiology and community health* 57(7): 523-4.
- Moulton LH, Port FK, Wolfe RA, Foxman B, Guire KE. (1992) Patterns of low incidence of treated end-stage renal disease among the elderly. *American Journal of Kidney Diseases* 20(1): 55-62.
- Muntner P, Coresh J, Powe NR, Klag MJ. (2003) The contribution of increased diabetes prevalence and improved myocardial infarction and stroke survival to the increase in treated end-stage renal disease. *Journal of the American Society of Nephrology* 14(6): 1568-77.
- Muntner P, Menke A, Batuman V, Rabito FA, He J, Todd AC. (2007) Association of tibia lead and blood lead with end-stage renal disease: A pilot study of african-americans. *Environmental Research* 104(3): 396-401.
- Renal Epidemiology and Information Network (REIN). (2013) Guide de remplissage de DIADEM Informations sur la dialyse du registre REIN. 18 p.
- Réseau Epidémiologie et Information en Néphrologie, REIN. (2011) Rapport annuel 2011. 296 p.
- Roderick P, Clements S, Stone N, Martin D, Diamond I. (1999) What determines geographical variation in rates of acceptance onto renal replacement therapy in England? *Journal of Health Services Research and Policy* 4(3): 139-46.
- Rosansky SJ, Huntsberger TL, Jackson K, Eggers P. (1990) Comparative incidence rates of end-stage renal disease treatment by state. *American journal of nephrology* 10(3): 198-204.
- Spiegelhalter DJ, Thomas A, Best NG, Lunn D. (2003) WinBugs v. 1.4. User Manual. Cambridge: MRC Biostatistics. Unit. 60 p.
- Tanner RM, Gutiérrez OM, Judd S, McClellan W, Bowling CB, *et al.* (2013) Geographic variation in CKD prevalence and ESRD incidence in the United States: Results from the reasons for geographic and racial differences in stroke (REGARDS) study. *American Journal of Kidney Diseases* 61(3): 395-403.
- Townsend P. (1987) Deprivation. *Journal of Social Policy* 16(2): 125-146.
- U.S. Renal Data System, USRDS (2013) Incidence, Prevalence, Patient Characteristics, and Modalities. In: Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States Volume two. National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD, 2013. pp. 215-228.
- Usami T, Koyama K, Takeuchi O, Morozumi K, Kimura G. (2000) Regional variations in the incidence of end-stage renal failure in Japan. *Journal of the American Medical Association* 284(20): 2622-4.

- Van De Luijngaarden MWM, Noordzij M, Tomson C, Couchoud C, Cancarini G, *et al.* (2012) Factors influencing the decision to start renal replacement therapy: Results of a survey among european nephrologists. *American Journal of Kidney Diseases* 60(6): 940-8.
- Volkova N, McClellan W, Klein M, Flanders D, Kleinbaum D, *et al.* (2008) Neighborhood poverty and racial differences in ESRD incidence. *Journal of the American Society of Nephrology* 19(2): 356-64.
- Wakefield J. (2008) Ecologic studies revisited. *Annual review of public health* 29: 75–90.
- Ward MM. (2008) Socioeconomic status and the incidence of ESRD. *American Journal of Kidney Diseases* 51(4): 563-72.
- Wimmer F, Oberaigner W, Kramar R, Mayer G. (2003) Regional variability in the incidence of end-stage renal disease: An epidemiological approach. *Nephrology Dialysis Transplantation* 18(8): 1562-7.
- Yan G, Cheung AK, Ma JZ, Yu AJ, Greene T, *et al.* (2013) The associations between race and geographic area and quality-of-care indicators in patients approaching ESRD. *Clinical Journal of the American Society of Nephrology* 8(4): 610-8.
- Young EW, Mauger EA, Jiang K, Port FK, Wolfe RA. (1994) Socioeconomic status and end-stage renal disease in the united states. *Kidney International* 45(3): 907-11.

Partie 3. Détection de clusters atypiques d'IRCT en Nord – Pas de Calais

a. Introduction

Des disparités d'incidence de l'IRCT ont précédemment été mises en évidence entre les cantons de la région Nord – Pas de Calais (Ocelli *et al.*, soumise). Les techniques de cartographie des maladies sont employées pour évaluer l'hétérogénéité spatiale d'incidence des événements de santé. Bien qu'elles soient indispensables pour décrire la distribution géographique, elles ne permettent pas de détecter des zones d'incidence atypique, appelés clusters, et d'évaluer leur significativité. Pour identifier ces zones géographiques atypiques, les méthodes de détection d'agrégats peuvent être employées. La statistique de scan détecte des clusters d'événements spatiaux, temporels et spatio-temporels, sans biais de présélection (Kulldorff, 1997). Un cluster peut être défini comme une zone géographique au sein de laquelle le risque de développer la maladie est significativement plus faible ou plus élevé par rapport à l'ensemble de l'aire d'étude. La méthode classique de statistique de scan attribue un risque constant et homogène dans chaque cluster détecté. La version isotonique de SaTScan™ permet, de manière plus fine, d'estimer des risques non homogènes au sein de chaque cluster détecté (Kulldorff, 1999). Grâce à une fonction de régression isotonique, elle détermine plusieurs niveaux de risque à l'intérieur d'un même cluster, de manière décroissante en fonction de la distance au centre du cluster. Le cluster détecté est alors centré sur l'unité géographique présentant le risque le plus élevé. Les méthodes de détection de clusters atypiques d'événements sont une approche innovante, rarement employées dans le contexte des maladies rénales chroniques. Elles ne font l'objet d'aucune publication en France.

L'ensemble des cas incidents d'IRCT recensés dans le registre Néphronor entre 2005 et 2012 a été récolté, puis géoréférencé au canton de résidence au moment du premier traitement de suppléance. La cartographie des SIR lissés et ajustés sur l'âge et le sexe a été mise à jour en intégrant le recensement de 2012. La statistique de scan spatiale fut mise en œuvre selon la méthode isotonique, afin de détecter des clusters de surincidence et de sous-incidence.

Pour un total de 5362 cas incidents recensés entre 2005 et 2012, le taux d'incidence annuel régional est de 166 pmh. La région présente toujours de fortes disparités entre les 170 cantons, avec une incidence variant de 35 à 396 pmh. La cartographie des SIR lissés met en exergue une hétérogénéité spatiale, avec des zones de surincidence significative et de sous-incidence significative. La statistique de scan isotonique révèle 3 clusters de surincidence et aucun cluster de sous-incidence.

Chacun des clusters peut être divisé en plusieurs niveaux de risque. Le cluster le plus probable est centré sur le canton de Maubeuge, au sud-est de la région. Les clusters secondaires significatifs sont centrés sur les cantons de Roubaix-Wattrelos (métropole lilloise) pour le premier et Grande-Synthe (dunkerquois) pour le second. Contrairement aux méthodes classiques, la version isotonique de SaTScan™ peut identifier un risque variable au sein d'un même cluster, ce qui apporte une information complémentaire dans la perspective de recherche de facteurs étiologiques de la maladie.

Ce travail a fait l'objet d'un article prochainement soumis à la revue *Nephrology Dialysis Transplantation* et intégré à la suite de ce résumé. Ces résultats ont également été présentés lors de la conférence anglophone (13^{ème} Journée André Verbert), organisée en septembre 2013 par l'Ecole Doctorale Biologie Santé de Lille, ainsi qu'à la conférence du Groupement des Néphrologues Francophones de Belgique, en mars 2013 à Tournai (Belgique). Ils feront aussi l'objet d'une présentation orale lors de la 16^e Réunion Commune de la Société Francophone de Dialyse (SDF) et de la Société de Néphrologie (SN) organisée du 30 septembre au 3 octobre 2014 à Saint-Etienne.

b. [Article](#)

FULL TITLE

End-Stage Renal Disease in Northern France: isotonic cluster detection.

AUTHORS AND AFFILIATIONS

Florent Occelli, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France.

Michaël Génin, EA 2694, Université Lille Nord de France, Faculté de Médecine pôle Recherche, Lille, France.

Annabelle Deram, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France. Faculté Ingénierie et Management de la Santé (ILIS), Loos, France.

François-Xavier Glowacki, EA 4483, Réseau Néphronor, Hôpital Huriez, Lille, France (ww.nephronor.fr).

Damien Cuny, EA 4483, Université Lille Nord de France, Faculté de Pharmacie de Lille, Lille, France.

Néphronor Network, Réseau Néphronor, Hôpital Huriez, Lille, France (ww.nephronor.fr).

KEYWORDS

Cluster analysis, Disease mapping, End-Stage Renal Disease, Isotonic scan statistics

SUMMARY

Spatial variability in the incidence of end-stage renal disease is commonly assessed by standardized incidence rates maps. Spatial scan statistics, which are uncommon in the context of renal diseases, were used to detect atypical areas of incidence on small area level in the North of France. The isotonic version revealed 3 significant clusters including inhomogeneous high incidence rates.

ACKNOWLEDGEMENT

The authors would like to thank Vincent LEMAITRE, who developed the database of patients with chronic kidney disease for the Nord-Pas-de-Calais region of France. Furthermore, we gratefully acknowledge the outstanding technical support of Hasna Camara, Sebastien Gomis and the staff of the Nephronor network: AAZIB Larbi, ADDA Hassen, AL MORABITI Mustapha, AL MOUBARAK Imad, AZAR Raymond, BACRI Jean-Louis, BATAILLE Pierre, BEN HENDA Amaury, BENZIANE Abdelkader, BILLION Stephane, BINAUT Raynald, BLOCH Julie, BONNARD Guillaume, BONNE J-François, BOTTE-NOËL Alexandra, BOUBIA Toufik, BOUBIA Véronique, BOURDON Franck, BRASSEUR José, BURDA Guillaume, CARDON Gérard, CHLIH Bouchra, DAROUX Maité, DECAMBRON Mélanie, DEHENNAULT Maud, DELVALLEZ Luc, DESPREZ Jonathan, DEVAUX Jean Philippe, DUFAY Alexandre, EL AMARI Aderrahim, FLEURY Dominique, FRIMAT Marie, GARSTKA Antoine, GHEERBRANDT J. Dominique, GLOWACKI Francois, GOSSET Brigitte, GOUBET Marlène, GUINCESTRE Thomas, GUYON – ROGER Tiphaine, HAMDINI Nasser, HAMMELIN Jean Philippe, HARDY-YVERNEAU Paule, JOMAA Zacharie, HAZZAN Marc, HOFFMANN Maxime, LABATUT Delphine, LAHOUCHE Annie, LAMOTTE Christian, LE MONIES Hervé, LEBAS Céline, LEBLEU Jacques, LEMOINE Corinne, LESSORE DE STE FOY Celia, LIONET Arnaud, LOUVET Daniel, MAC NAMARA Evelyne, MAISONNEUVE Nathalie, M'BARGA Roseline, MESBAH Rafik, MESSIER Gilles, MOULONGUET Florence, MOUSSALIEH Laura, NOEL Christian, NOVO Robert, PAGNIEZ Dominique, PAINCHART Bernard, PROVÔT Francois, RATSIMBAZAFY Anderson, READE Richard, REBEROLLE Matthieu, ROBITAILLE Géraldine, SEMJEN Elisabeth, SHAHAPUNI Irina, SHENOUDA Milad, SKALLI Amina, TAGHIPOUR TAMIJI Lili, TALASZKA Aline, VAIRON-CODACCIONI M. Xavière, VAN EGROO Anne, VANHILLE Philippe, VRIGNEAUD Laurence, WAJSBROT Lucie, WHEATLEY Pascal.

FIGURE LEGEND

Figure 1. Principle of the isotonic scan statistic

Figure 2. End-Stage Renal Disease (ESRD) smoothed Standardized Incidence Rates (SIR), North of France, from 2005 to 2012.

Figure 3. End-Stage Renal Disease (ESRD) spatial clusters, North of France, from 2005 to 2012. Level 1 corresponds to the cluster's center, and following levels to the isotonic regression levels in the relative risk (RR).

TABLES

Table 1. Description of End-Stage Renal Disease (ESRD) clusters provided by isotonic analysis.

ABSTRACT

Background

While the burdens of End-Stage Renal Disease (ESRD) grow up steadily, strong geographic variations in the incidence of treated ESRD are observed on multiple scales in the world, including France. The spatial heterogeneity of diseases is commonly assessed by choropleth maps of standardized incidence rates. Although necessary to bring out spatial patterns, this approach did not enable detection of statistically significant geographical clusters, such as spatial scan statistics.

Methods

Cases were all incident patients registered in the Renal Epidemiology and Information Network (REIN) in the Nord – Pas de Calais (NPdC) region from 2005 to 2012. Patients were grouped by their resident canton (a small French administrative unit) at the time of the first treatment. Spatial scan statistics (SaTScan™ software) were employed to detect ESRD clusters and identify their location. The isotonic method was chosen to provide a multilevel risk cluster, adjusted for age and gender.

Results

Our study included 5362 patients with ESRD between 2005 and 2012 in the NPdC. The overall crude annual incidence rate was 166 per million inhabitants (pmi) and varied from 35 to 396 pmi among the 170 cantons. Isotonic spatial scan statistics revealed 3 significant clusters including inhomogeneous high incidence rates.

Conclusion

Spatial scan statistics help to distinguish contiguous areas of excess or deficit that are not likely to have arisen by chance. These are very uncommon in the context of chronic kidney diseases. The emergence of spatial clusters demonstrates significant disparities of ESRD incidence among the NPdC region, on small area aggregation level in France. This specificity may provide deeper knowledge of patterns that could explain such disparities and would afterwards supply better understanding of environmental assumptions.

INTRODUCTION

Despite a steady growth, strong geographic variations in the burdens of treated End-Stage Renal Disease (ESRD) have been found in developed countries (Usami *et al.*, 2000; Volkova *et al.*, 2008; Hommel *et al.*, 2010; Gilg *et al.*, 2012; Grace *et al.*, 2012; Yan *et al.*, 2013). In metropolitan France, the crude incidence rate varies widely from 85.8 to 225.5 per million inhabitants (pmi) in districts from 2008 to 2009 (Couchoud *et al.*, 2012). Substantial disparities were previously revealed within the North of France on small area aggregation level (Ocelli *et al.*, submitted).

The reasons for these geographic disparities remain elusive. While population's socioeconomic status and ethnicity are the most likely factors (Roderick *et al.*, 1999; Ward, 2008; Couchoud *et al.*, 2012; Grace *et al.*, 2012; Babayev *et al.*, 2013; Ocelli *et al.*, submitted), they also may result from variations in the geographic distribution of environmental contaminants (Hellström *et al.*, 2001; Hodgson *et al.*, 2007; Muntner *et al.*, 2007; Chowdhury *et al.*, 2014).

Geographic variation of ESRD provides clues for researchers to investigate possible environmental etiologic factors. The spatial heterogeneity of diseases is commonly assessed by choropleth (value-by-area) maps of standardized incidence rates (SIR). Using a Bayesian hierarchical model, we recently reported spatial heterogeneity in the incidence of ESRD with a predominance of the disease in most deprived areas (Ocelli *et al.*, submitted). Although the SIR is necessary to bring out spatial patterns, this approach does not enable the detection of statistically significant geographical clusters of incident cases. The Kulldorff's spatial scan statistics methods have been developed to ensure the detection of such clusters of events without pre-selection bias, allowing relevant statistical inference and adjustment for confounding factors (Kulldorff, 1997; Kulldorff *et al.* 1998). A cluster is defined as a geographical area in which the risk of disease is higher (or lower) than in the rest of the studied area. Following a first application for the detection of geographical clusters in cases of childhood leukemia and the identification of statistical excess of breast cancer mortality in the New-York-Philadelphia metropolitan area (Hjalmars *et al.* 1996; Kulldorff *et al.* 1997), this methodology has already demonstrated its ability to reveal contiguous areas of events' excess or deficit that are not likely to have arisen by chance (Cousens *et al.*, 2001; Sabel *et al.*, 2003; Viel *et al.*, 2005; Pedigo *et al.*, 2011; Génin *et al.*, 2013). Nevertheless, these are very uncommon in the context of Chronic Kidney Diseases (CKD).

The classical spatial scan statistics assume that the risk of disease is constant within the cluster. In case of cluster of large surface, this assumption is quite unrealistic. Thus, an extension of spatial scan statistics has been used, namely isotonic spatial scan statistics. The isotonic version provides a greater degree of information by calculating non-homogeneous relative risks within a detected cluster (Kulldorff, 1999). For a multilevel risk cluster, the isotonic spatial scan statistic could model an isotonic regression function with successively decreasing risk with increasing distance from the cluster center (Barlow *et al.*, 1972). Such a division in the detected cluster area will allow more spatial precision and provide useful information for etiology research (Boscoe *et al.*, 2003; Li *et al.*, 2011; Said-Ali *et al.*, 2013).

The Renal Epidemiology and Information Network (REIN) is a national CKD registry, which lists all patients who initiated ESRD treatment since 2002, and is currently available in 22 regions of metropolitan France (Couchoud *et al.*, 2006).

The aim of this study was to reveal significant atypical clusters in the incidence of ESRD, assessed using an isotonic spatial scan statistics, and focuses on the Nord-Pas-de-Calais region, which is characterized by a high incidence rate, for the period between 2005 and 2012.

MATERIAL AND METHODS

Study area and sources of data

This study was conducted in the Nord-Pas-de-Calais region in France. It has a surface area of 12 481 km² including both rural industrial and urban regions, and with approximately 4 033 000 inhabitants. The region's 170 cantons (a French small administrative unit) were used to represent distinct spatial units. These are referenced by the National Institute of Statistics and Economic Studies (INSEE), and in 2009 had an average population of 23 725 (extremes ranging from 4 991 to 226 827) and an average surface area of 73 km² (extremes ranging from 2 to 258 km²).

Cases were extracted from the REIN registry. For the purpose of this study, all incident patients registered on the first day of renal replacement therapy (RRT) in the Nord-Pas-de-Calais region from January 2005 to December 2012 were included. Patients were grouped into cantons according to the postcode of their residence, defined at the time of their first RRT. They were ranked by gender and 18 age groups (0-4 years; 5-9 years; 5-year-interval classes and ≥85 years).

Demographic data were extracted from national population census, provided by the National Institute of Statistics and Economic Studies (INSEE). The population data was also ranked by gender and 5-year age groups.

Statistical and spatial analysis

The centroid of each canton, defined by its geographical center (longitude and latitude), was used for the spatial analysis. The maps were produced using the ArcGIS® 10.1 software (<http://www.esri.com>). All statistical analyses were considered significant at the 0.05 type 1 error.

As the method defined in a previous study (Ocelli *et al.*, submitted), the crude incidence rates and SIRs were calculated, and smoothed SIRs were mapped (using Besag York and Mollié model (1991)). The maps of smoothed SIRs can be used to highlight global spatial heterogeneities in disease incidence. Nevertheless, this method cannot detect significant atypical geographic clusters in terms of ESRD incidence.

Spatial scan statistics were then used for the cluster detection, in an objective way. The isotonic version of scan statistics is divided into two steps: detection and inference (Kulldorff, 1999). During the detection step, the method uses a circular window of flexible size, which moves across the studied area, using as center the centroid of the cantons. At each position, the radius of the circular window varies from zero up to a maximum radius so that the window never contains more than 50% of the total number of cases. For each window, assuming that the number of cases in each canton is Poisson distributed, a likelihood function is calculated. This latter models the potential cluster using an isotonic regression function with successively decreasing risk with increasing distance from the cluster center. Actually, for a given window of radius d , the risk is modeled as a non-increasing function $r(d)$ of the distance of the centroid with multiple locations where the function takes a step down (Fig. 1).

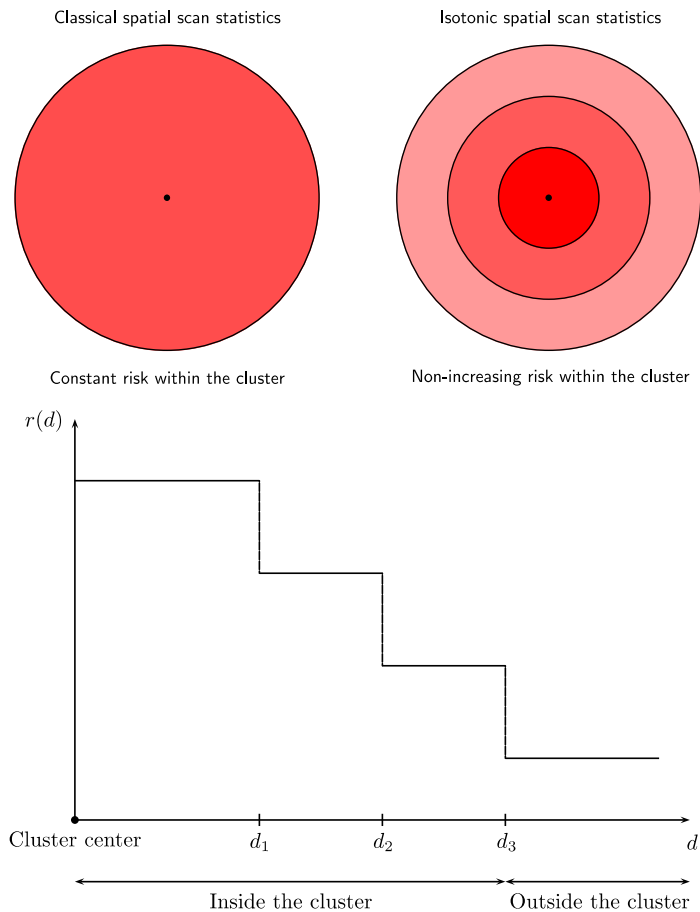


Figure 1. Principle of the isotonic scan statistic

The risk function is fitted by means of an isotonic regression and there is no *a priori* assumption about the number of steps. It is linked to the isotonic regression function chosen, among all possible non-increasing functions, thanks to maximum likelihood method. The window maximizing the likelihood function is defined as the Most Likely Cluster (MLC). During the inference step, the null hypothesis corresponds to the absence of clusters that is: “the risk of being affected by IRCT is constant throughout all cantons”. To test the null hypothesis, a likelihood ratio (LLR) test is used. The numerator of LLR is the likelihood function associated to the MLC and the denominator is the likelihood function under null hypothesis. The distribution of the LLR is unknown and is approximated using 9,999 Monte-Carlo random replications under null hypothesis to obtain P-values (Dwass, 1957). Secondary clusters (SC), which have high likelihood function and do not overlap the MLC, were taken into account.

RESULTS

The study included 5 362 patients (57.7% men and 42.3% women) who began an RRT between 2005 and 2012 in the Nord-Pas-de-Calais. From 2005 to 2012, the overall crude annual incidence rate was 166 pmi for both sexes, 199 pmi for men and 133 pmi for women. The crude incidence rate varied from 34.6 to 396.3 pmi over the 170 cantons (from 0 to 545 for men and from 0 to 276 for women). The choropleth map of age- and gender-adjusted smoothed SIRs suggested a strong spatial heterogeneity in the incidence of ESRD within the North of France (Fig. 2). The smoothed SIRs vary among cantons, from 0.71 to 1.56. Highest incidence rates are found in the North (1.55, 95% CI [1.17–2.00] for Grande-Synthe and 1.41, 95% CI [1.22–1.62] for Dunkerque), the South-East (1.56, 95% CI [1.25–1.92] for Maubeuge and 1.38, 95% CI [1.12–1.66] for Anzin) and in the North of Lille (1.54, 95% CI [1.32–1.79] for Roubaix and 1.34, 95% CI [1.15–1.56] for Tourcoing). Lowest incidence rates are instead in the South-West (0.72, 95% CI [0.50–0.98] for Auxi-le-Château) and the suburbs of Lille (0.71, 95% CI [0.49–0.97] for Loos, 0.73, 95% CI [0.61–0.99] for Mouveaux and 0.74, 95% CI [0.58–0.92] for Bondues).

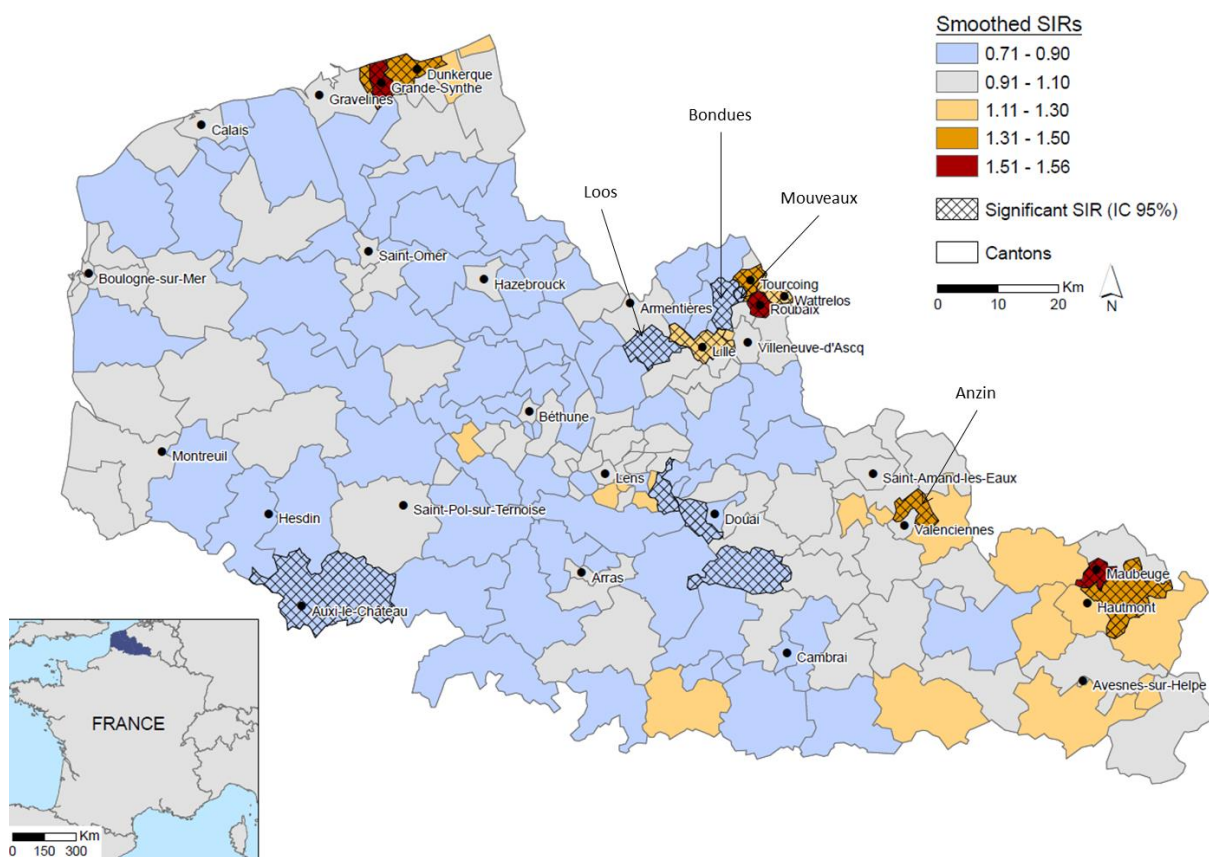


Figure 2. End-Stage Renal Disease (ESRD) smoothed Standardized Incidence Rates (SIR), North of France, from 2005 to 2012.

The isotonic spatial scan statistics revealed 3 significant clusters with inhomogeneous high incidence rates. No cluster with low incidence rates was significant. The spatial distribution of detected clusters is shown in Fig. 3. The color scale indicates the level of RR in each detected cluster. Red color corresponds to the highest RR (cluster center) and yellow to the lowest RR. The most likely cluster (MLC) was detected in the South-East rural area. Secondary clusters were then detected in the urban agglomeration of Lille (Secondary Cluster n°1: SC1) and the industrial area of Dunkerque (Secondary Cluster n°2: SC2).

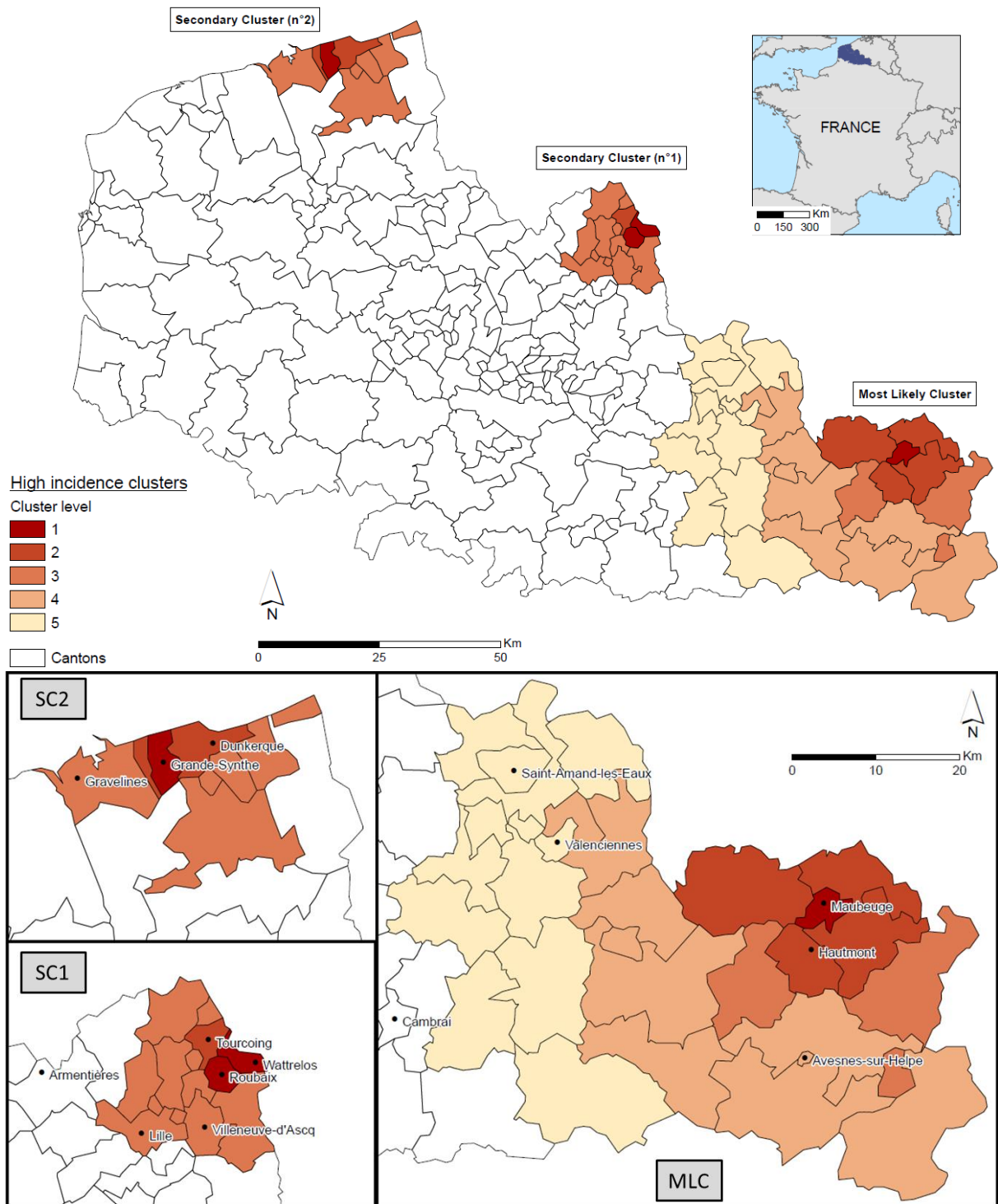


Figure 3. End-Stage Renal Disease (ESRD) spatial clusters, North of France, from 2005 to 2012. Level 1 corresponds to the cluster's center, and following levels to the isotonic regression levels in the relative risk (RR).

The description of significant cluster characteristics (observed and expected cases, P-value, isotonic level, incidence rate and RR) is provided in Table 1. All clusters are significant with a 0.0001 P-value. The MLC includes a total of 1060 cases for 877 expected. The RR varied from 1.83 for the cluster center (level 1 in Maubeuge) to 1.26 for the entire cluster (level 5) and the incidence decreases from 302 pmi to 201 pmi. This large area of 46.6 km radius is sparsely populated (about 635,000 inhabitants) and comprises 29 cantons. The SC1 is situated on a 14.7 km radius and densely populated area of 845,000 inhabitants. While 981 cases are expected, 1094 are observed during the period on 17 cantons. Three sub-clusters are observed within SC1. The cluster center comprises two units (Roubaix and Wattrelos), with a RR of 1.59 and an incidence rate of 261 pmi. When adding levels 2 and 3, the incidence rate is 186 pmi, with a RR of 1.14. For SC2, 361 cases are observed against 262 expected among a population of 195,000. The cluster of 7 units is also divided into three levels, from 2.04 (the cluster center in Grande-Synthe) to 1.41. Incidence rates decrease from 396 pmi to 230 pmi.

Table 1. Description of End-Stage Renal Disease (ESRD) clusters provided by isotonic analysis.

Cluster	P-value	Isotonic level	Radius (km)	Number of Cantons	Population	Cases	Expected	Incidence rate (pmi)	RR
MLC	0.0001	1	0	1	32,236	76	42	302	1.83
		2	10.3	5	117,513	231	156	244	1.50
		3	13.9	7	148,766	293	203	239	1.47
		4	29.8	17	316,168	576	438	218	1.35
		5	46.6	29	635,038	1060	877	201	1.26
SC1	0.0001	1	2.7	2	138,063	235	151	261	1.59
		2	4.9	3	230,475	384	256	252	1.54
		3	14.7	14	845,227	1094	981	186	1.14
SC2	0.0001	1	0	1	21,064	45	22	396	2.04
		2	3.3	2	103,711	218	138	236	1.60
		3	12.2	7	194,827	361	262	230	1.41

DISCUSSIONS

The study of spatial distribution of disease provides clues to suggest etiologic factors. We showed significant spatial heterogeneity in the incidence of ESRD on the Nord-Pas-de-Calais region in France, during the period from 2005 to 2012. We then revealed 3 atypical geographical clusters with high incidence rates. The MLC is situated on a rural area with a low density of population, which explains the large radius of 46.6 km. Secondary clusters are rather observed on highly urbanized areas. Another secondary cluster with high incidence rates and two clusters with low incidence rates were also detected, but not significant ($P > 0.05$).

The SaTScan methodology is regularly performed for disease surveillance and monitoring (Cousens *et al.*, 2001; Sabel *et al.*, 2003; Pedigo *et al.*, 2011; Génin *et al.*, 2013; Green *et al.*, 2013). Although the spatial variations in the incidence of ESRD have been widely documented (Foxman *et al.*, 1991; Wimmer *et al.*, 2003; Joyce Fan *et al.*, 2007; Volkova *et al.*, 2008; Caskey *et al.*, 2011; Couchoud *et al.*, 2012), very few focused on the detection of statistically significant geographical clusters. Such a spatial scan statistics approach on the clustering of ESRD incidence in France has never yet been published. In other countries, only two papers carried out a spatial scan analysis in the context of CKD. A Canadian study used scan statistics to identify clusters of underserved prevalent patients with CKD and select optimal locations for new healthcare facilities (Faruque *et al.*, 2012). In a study from US, scan statistics have been performed to detect geographical aggregations of treatment centers with low pre-dialysis care, estimated by late referral of patients with CKD into RRT (McClellan *et al.*, 2009). Spatial scan statistics should be employed in complement of incidence choropleth maps. These help to distinguish contiguous areas of excess or deficit that are not likely to have arisen by chance.

On the other hand, the isotonic version of scan statistics provided a greater degree of information by revealing non-homogeneous RR within each detected cluster (Kulldorff, 1999). The standard spatial scan statistic tends to reach the highest statistical power by identifying large areas with large populations, which may lead to low levels in risk. The isotonic scan statistic rather stratifies each detected cluster into sub-clusters. The modeled RR in each sub-region is graded using an isotonic regression function, depending on the distance to the high-risk kernel, which is in the central part of the whole cluster area.

Nevertheless, the results of the isotonic scan statistics should be interpreted with caution. Indeed this version defines a centered high-risk area and decreasing RR with increasing distance from the cluster center. It assumes that spatial variations in disease risks are isotropic, which means the same trends in all directions. However, owing to several risk factors in diseases, the real high-risk kernel is not mandatory situated on the central place in a whole cluster area (Li *et al.*, 2011). In these conditions, the choropleth map of incidence rates could help the interpretation of the cluster results. In our study, SC1 appears to be mainly due to three cantons (Wattrelos, Roubaix and Tourcoing), since other SIRs are close to 1. This is also the case for the SC2, with two cantons with high incidence rates (Grande-Synthe and Dunkerque). While these clusters seem to be centered, the MLC detected from our dataset includes two far areas of high incidence, one is centralized on Maubeuge (SIR: 1.56) and the other one on Anzin (SIR: 1.38). Lower rates are observed between these two hotspots.

We noted recent studies which tried to detect nonisotropic risks in clusters. In 2008, Chen *et al.* developed geovisual analytics approaches for distinguish between heterogeneous and homogeneous clusters and assess the stability of clusters across analytic scales. This method was recently performed using SaTScan by Said Ali *et al.* (2013), and consist in a successive detection of nested clusters inside each of the statistically significant larger clusters detected from the entire dataset. Then, Li *et al.* (2011) proposed a spatial scan statistic for a nonisotropic two-level risk cluster. This approach could bring out noncentralized high-risk kernel within a detected cluster. This extension of the standard spatial scan statistics is not available yet. Although such approaches help to reduce this bias, choropleth maps of incidence rates should be produced in parallel for more efficient interpretation of the SaTScan results.

CONCLUSION

In this study, we used the choropleth map of incidence rates to highlight spatial heterogeneity in the incidence of ESRD on the Nord-Pas-de-Calais region in France, during the period from 2005 to 2012, and the spatial scan statistics to reveal statistically significant geographical clusters. The isotonic version of SaTScan offers a centralized multi-level RR in each detected whole cluster. This approach should be interpreted with caution, since spatial variations in disease risks are nonisotropic. Choropleth maps of incidence rates and spatial clusters detection should be employed in parallel to ensure efficient disease surveillance and prevention. The existence of spatial disparities and atypical clusters suggests that risk factors of CKD are still at work in the Nord-Pas-de-Calais region.

REFERENCES

- Babayev R, Whaley-Connell A, Kshirsagar A, Klemmer P, Navaneethan S *et al.* 2013. Association of Race and Body Mass Index With ESRD and Mortality in CKD Stages 3-4: Results From the Kidney Early Evaluation Program (KEEP). *Am J Kidney Dis*, 61 (3): 404-412.
- Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD. 1972. *Statistical inference under restrictions: The theory and application of isotonic regression.* Wiley: England.
- Besag J, York J, Mollié A. 1991. Bayesian image restoration with two applications in spatial statistics. *Ann Inst Stat Math*, 43: 1–21.
- Boscoe FP, McLaughlin C, Schymura MJ, Kielb CL. 2003. Visualization of the spatial scan statistic using nested circles. *Health and Place*, 9 (3): 273-7.
- Caskey FJ, Kramer A, Elliott RF, Stel VS, Covic A, Cusumano A, *et al.* 2011. Global variation in renal replacement therapy for end-stage renal disease. *Nephrology Dialysis Transplantation*, 26 (8): 2604-10.
- Chen J, Roth RE, Naito AT, Lengerich EJ and MacEachren AM. 2008. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. *International Journal of Health Geographics*, 7: 57.
- Chowdhury R, Darrow L, McClellan W, Sarnat S, Steenland K. 2014. Incident ESRD Among Participants in a Lead Surveillance Program. *Am J Kidney Dis*, In Press.
- Couchoud C, Guihenneuc C, Bayer F, Lemaitre V, Brunet P, Stengel B. 2012. Medical practice patterns and socio-economic factors may explain geographical variation of end-stage renal disease incidence. *Nephrology Dialysis Transplantation*, 27 (6): 2312-22.
- Couchoud C, Stengel B, Landais P *et al.* 2006. The renal epidemiology and information network (REIN): a new registry for end-stage renal disease in France. *Nephrol Dial Transplant*, 21: 411–418.
- Cousens S, Smith PG, Ward H, Everington D, Knight RSG, Zeidler M, Stewart G, Smith-Bathgate EAB, Macleod MA, Mackenzie J, Will RG. 2001. Geographical distribution of variant Creutzfeldt-Jakob disease in Great Britain, 1994-2000. *The Lancet*, 357: 1002-1007.
- Dwass M. 1957. Modified randomization tests for nonparametric hypotheses. *Ann Math Stat*, 28 (1): 181–187.
- Faruque LI, Ayyalasomayajula B, Pelletier R, Klarenbach S, Hemmelgarn BR, Tonelli M. 2012. Spatial analysis to locate new clinics for diabetic kidney patients in the underserved communities in Alberta. *Nephrology Dialysis Transplantation*, 27 (11): 4102-9.
- Foxman B, Moulton LH, Wolfe RA, Guire KE, Port FK, Hawthorne VM. 1991. Geographic variation in the incidence of treated end-stage renal disease. *Journal of the American Society of Nephrology*, 2 (6): 1144-52.
- Génin M, Duhamel A, Preda C, Fumery M, Savoye G *et al.* 2013. Space-time cluster of Crohn's disease in northern France. *Journal of Public Health*, 21 (6): 497-504.
- Gilg J, Castledine C, Fogarty D. 2012. UK renal registry 14th annual report: Chapter 1 UK RRT incidence in 2010: National and centre-specific analyses. *Nephron - Clinical Practice*, 120 (SUPPL.1): c1-c27.
- Grace BS, Clayton P, Cass A, McDonald SP. 2012. Socio-economic status and incidence of renal replacement therapy: A registry study of Australian patients. *Nephrology Dialysis Transplantation*, 27 (11): 4173-80.
- Green C, Yu BN, Marrie RA. 2013. Exploring the implications of small-area variation in the incidence of multiple sclerosis. *Am J Epidemiol*, 178 (7): 1059-66.

- Hellström L, Elinder C-, Dahlberg B, Lundberg M, Järup L, Persson B, *et al.* 2001. Cadmium exposure and end-stage renal disease. *American Journal of Kidney Diseases*, 38 (5): 1001-8.
- Hjalmars ULF, Kulldorff M, Gustafsson G, Nagarwalla N. 1996. Childhood leukaemia in Sweden: using GIS and spatial scan statistic for cluster detection. *Statist Med*, 15 (7-9): 707-15.
- Hodgson S, Nieuwenhuijsen MJ, Elliott P, Jarup L. 2007. Kidney disease mortality and environmental exposure to mercury. *Am J Epidemiol*, 165 (1): 72-7.
- Hommel K, Rasmussen S, Kamper A, Madsen M. 2010. Regional and social inequalities in chronic renal replacement therapy in denmark. *Nephrology Dialysis Transplantation*, 25 (8): 2624-32.
- Joyce Fan Z, Lackland DT, Lipsitz SR, Nicholas JS, Egan BM, Tim Garvey W, *et al.* 2007. Geographical patterns of end-stage renal disease incidence and risk factors in rural and urban areas of south carolina. *Health and Place*, 13 (1): 179-87.
- Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. 1998. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health*, 88 (9): 1377-1380.
- Kulldorff M, Feuer EJ, Miller BA, Freedman LS. 1997. Breast cancer clusters in the northeast United States: a geographic analysis. *Am J Epidemiol*, 146 (2): 161-170.
- Kulldorff M. 1997. A spatial scan statistic. *Commun Statist Theory Meth*, 26 (6): 1481-1496.
- Kulldorff M. 1999. An isotonic spatial scan statistic for geographical disease surveillance. *Journal of the National Institute of Public Health*, 48: 94-101.
- Li X, Wang J, Yang W, Li Z, Lai S. 2011. A spatial scan statistic for multiple clusters. *Math Biosci.*, 233 (2): 135-42.
- McClellan W, Wasse H, McClellan A, Kipp A, Waller L, Rocco M. 2009. Treatment Center and Geographic Variability in Pre-ESRD Care Associate with Increased Mortality. *J Am Soc Nephrol*, 20: 1078-1085.
- Muntner P, Menke A, Batuman V, Rabito FA, He J, Todd AC. 2007. Association of tibia lead and blood lead with end-stage renal disease: A pilot study of african-americans. *Environ Res.*, 104 (3): 396-401.
- Occelli F, Deram A, Génin M, Cuny D, Glowacki F-X and Néphronor Network. Mapping end-stage renal disease (ESRD): spatial variations on small area level in France, and association with deprivation, earlier renal replacement care and related diseases. Submitted to Plos One.
- Pedigo A, Aldrich T, Odoi A. 2011. Neighborhood disparities in stroke and myocardial infarction mortality: a GIS and spatial scan statistics approach. *BMC Publ Health*, 11: 644.
- Roderick P, Clements S, Stone N, Martin D, Diamond I. 1999. What determines geographical variation in rates of acceptance onto renal replacement therapy in england? *Journal of Health Services Research and Policy*, 4 (3): 139-46.
- Sabel CE, Boyle PJ, Löytönen M, Gatrell AC, Jokelainen M, Flowerdew R, Maasilta P. 2003. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. *American Journal of Epidemiology*, 157: 898-905.
- Said-Ali Z, Grenouillet F, Knapp J, Bresson-Hadni S, Vuitton DA, Raoul F, *et al.* 2013. Detecting nested clusters of human alveolar echinococcosis. *Parasitology*, 140 (13): 1693-700.
- Usami T, Koyama K, Takeuchi O, Morozumi K, Kimura G. 2000. Regional variations in the incidence of end-stage renal failure in japan. *J Am Med Assoc*, 284 (20): 2622-4.
- Viel JF, Floret N, Mauny F. 2005. Spatial and space-time scan statistics to detect low clusters of sex ratio. *Environmental and Ecological Statistics*, 12: 289-299.

Volkova N, McClellan W, Klein M, Flanders D, Kleinbaum D, Soucie JM, *et al.* 2008. Neighborhood poverty and racial differences in ESRD incidence. *Journal of the American Society of Nephrology*, 19 (2): 356-64.

Ward MM. 2008. Socioeconomic status and the incidence of ESRD. *American Journal of Kidney Diseases*, 51 (4): 563-72.

Wimmer F, Oberaigner W, Kramar R, Mayer G. 2003. Regional variability in the incidence of end-stage renal disease: An epidemiological approach. *Nephrology Dialysis Transplantation*, 18 (8): 1562-7.

Yan G, Cheung AK, Ma JZ, Yu AJ, Greene T, Oliver MN, *et al.* 2013. The associations between race and geographic area and quality-of-care indicators in patients approaching ESRD. *Clinical Journal of the American Society of Nephrology*, 8 (4): 610-8.

Partie 4. Conclusion et perspectives du chapitre II

Les disparités spatiales d'incidence de l'IRCT sont régulièrement étudiées au Etats-Unis et en Grande-Bretagne et à différentes échelles géographiques (Young *et al.*, 1994 ; Roderick *et al.*, 1999 ; Maheswaran *et al.*, 2003 ; Ward *et al.*, 2008 ; Volkova *et al.*, 2008). En France, une seule étude s'est récemment intéressée à ce phénomène. Celle-ci montrait de fortes variations spatiales entre les départements français, ainsi qu'une corrélation avec le statut socioéconomique (Couchoud *et al.*, 2012). Il ressort également de cette étude que la région Nord – Pas de Calais présente des taux d'incidence très élevés. Dans ce chapitre, nous nous sommes intéressés aux variations spatiales d'incidence de l'IRCT au sein même de cette région. Cette analyse fut menée à résolution spatiale plus fine au niveau des cantons.

Les patients atteints d'IRCT et recensés par le réseau Néphronor (déclinaison régionale du registre REIN de l'Agence de Biomédecine) ont été sélectionnés pour les années 2005 à 2012. Les cartographies réalisées sur la base de ce recensement mettent en évidence de fortes disparités spatiales d'incidence d'IRCT au sein de la région. Cette variabilité peut être en partie expliquée par le niveau de défaveur sociale, renseigné ici par l'indice de Townsend. Les populations les plus touchées par cette maladie sont en effet les plus défavorisées. Ce résultat confirme donc les conclusions de l'étude nationale (Couchoud *et al.*, 2012), ainsi que d'autres menées à différentes échelles dans le monde (Hommel *et al.*, 2010 ; Grace *et al.*, 2012 ; Caskey *et al.*, 2012). Du fait de la prise en charge homogène de l'IRCT à l'échelle de la région Nord – Pas de Calais et du remboursement intégral des frais liés aux traitements médicaux et hospitaliers par l'Assurance Maladie en France, la relation entre défaveur sociale et incidence ne semble pas être basée sur un modèle financier d'accès aux soins. Le risque élevé d'atteindre le stade terminal de l'insuffisance rénale semble plutôt être lié à une faible prévention, notamment pour les populations ayant de faibles connaissances en santé (Lora *et al.*, 2011). Le niveau d'études tend à représenter un indicateur fiable de précarité dans le cadre de cette maladie rénale (Tupin, 2014). L'une des conclusions d'une récente journée thématique, organisée par la commission d'épidémiologie de la société de néphrologie, était d'ailleurs l'instauration d'une variable renseignant le niveau d'études des patients intégrés au registre REIN.

L'influence du diabète et des maladies cardiovasculaires n'a pas pu être clairement identifiée. A ce jour, les données d'incidence ou de prévalence de ces maladies à l'échelle des cantons ne sont pas disponibles. Nous avons donc utilisé la néphropathie initiale de la maladie rénale pour étudier ces facteurs. La proportion de néphropathie diabétique ou vasculaire au sein des cantons n'explique pas les disparités observées.

En parallèle des cartographies de SIR, la statistique de scan a été employée pour détecter des clusters atypiques d'incidence au sein de la région. Une première approche classique consistait à tester la présence de clusters spatiaux, temporels et spatio-temporels (Kulldorff, 1997; Kulldorff *et al.* 1998). Aucun regroupement atypique de cas n'est observé dans le temps ou l'espace-temps. En revanche, trois clusters spatiaux ont été identifiés, regroupant chacun plusieurs unités cantonales. Afin d'apporter une précision supplémentaire à cette détection, l'approche isotonique de la statistique de scan identifie une variabilité de risque au sein d'un même cluster (Kulldorff, 1999). Celle-ci estime le risque de manière décroissante, à partir de l'unité géographique centrale présentant l'incidence la plus élevée. L'approche isotonique de SaTScanTM est cependant peu utilisée du fait de cette centralisation du risque, non représentative de la distribution spatiale des maladies et remise en cause par plusieurs auteurs (Chen *et al.*, 2008 ; Li *et al.*, 2011). Une autre méthode semble alors porter ses fruits. Elle consiste à lancer une première phase de détection de clusters, sans limite de taille (la taille maximale d'un cluster est dans ce cas de 50% de la population présente sur la zone d'étude). Une deuxième phase de détection est ensuite lancée au sein même de chacun des clusters préalablement identifiés. Cette opération peut être renouvelée à plusieurs reprises et

permet de révéler différents niveaux de risque de manière non concentrique. Un autre avantage est la détection de plusieurs points noirs au sein d'un même cluster. Malgré tout, nous conseillons l'élaboration systématique d'une cartographie de SIR pour faciliter l'interprétation des clusters détectés.

Le dunkerquois, l'agglomération de Lille-Roubaix-Tourcoing, ainsi que le bassin de Maubeuge présentent les niveaux de risque les plus élevés. Ces derniers se traduisent par la présence de clusters significatifs, témoignant d'une surincidence atypique sur ces zones. A l'instar de l'étude nationale, le niveau de défaveur sociale des populations peut expliquer ces disparités, mais ne constitue pas le seul facteur influent. Nos résultats montrent en effet qu'il persiste des zones géographiques de sur- et de sous-incidence d'IRCT pour lesquelles les principaux déterminants restent inexpliqués. Dès lors, il est possible d'envisager d'autres déterminants comme l'exposition à des toxiques, notamment environnementaux, tels que des métaux. En effet, il a été démontré que l'exposition à certains métaux lourds, particulièrement le cadmium, est susceptible d'induire ou d'aggraver une pathologie rénale (Lauwerys *et al.*, 1993 ; Hellström *et al.*, 2001 ; Järup *et al.*, 2002). Ainsi, nous émettons l'hypothèse que l'exposition à un environnement dégradé et notamment imprégné par les métaux puisse être également responsable des disparités locales d'incidence d'IRCT que nous observons.

Pour étayer cette hypothèse, les cartographies environnementales du cadmium et d'autres éléments traces seront par la suite confrontées aux cartographies d'IRCT. Une étude épidémiologique sera également menée au niveau individuel. Celle-ci intégrera des mesures biologiques (concentrations sanguines en cadmium et dosages de biomarqueurs) réalisées chez des patients issus à la fois de zones de surincidence et de sous-incidence et la réalisation de nouvelles mesures environnementales.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Caskey FJ, Kramer A, Elliott RF, Stel VS, Covic A *et al.* 2011. Global variation in renal replacement therapy for end-stage renal disease. *Nephrology Dialysis Transplantation*, 26(8): 2604-10.
- Chen J, Roth RE, Naito AT, Lengerich EJ and MacEachren AM. 2008. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. *International Journal of Health Geographics*, 7: 57.
- Couchoud C, Guihenneuc C, Bayer F, Lemaitre V, Brunet P, Stengel B. 2012. Medical practice patterns and socio-economic factors may explain geographical variation of end-stage renal disease incidence. *Nephrology Dialysis Transplantation*, 27 (6): 2312-22.
- Grace BS, Clayton P, Cass A, McDonald SP. 2012. Socio-economic status and incidence of renal replacement therapy: A registry study of Australian patients. *Nephrology Dialysis Transplantation*, 27 (11): 4173-80.
- Hellström L, Elinder C-G, Dahlberg B, Lundberg M, Järup L, Persson B, Axelson O. 2001. Cadmium exposure and end-stage renal disease. *Am J Kidney Dis*, 38: 1001-8.
- Hommel K, Rasmussen S, Kamper A, Madsen M. 2010. Regional and social inequalities in chronic renal replacement therapy in Denmark. *Nephrology Dialysis Transplantation*, 25 (8): 2624-32.
- Järup L. 2002. Cadmium overload and toxicity. *Nephrol Dial Transplant*, 17: 35–9.
- Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. 1998. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health*, 88 (9): 1377–1380.
- Kulldorff M. 1997. A spatial scan statistic. *Communications in statistics: theory and methods*, 26 (6): 1481–1496.
- Kulldorff M. 1999. An isotonic spatial scan statistic for geographical disease surveillance. *Journal of the National Institute of Public Health*, 48: 94–101.
- Lauwerys RR, Bernard AM, Buchet J-P, Roels H. 1993. Assessment of the health impact of environmental exposure to cadmium: contribution of the epidemiologic studies carried out in Belgium. *Environ Res*, 62: 200-6.
- Li X, Wang J, Yang W, Li Z, Lai S. 2011. A spatial scan statistic for multiple clusters. *Math Biosci*, 233 (2): 135-42.
- Lora CM, Gordon EJ, Sharp LK, Fischer MJ, Gerber BS, Lash JP. 2011. Progression of CKD in hispanics: Potential roles of health literacy, acculturation, and social support. *American Journal of Kidney Diseases*, 58 (2): 282-90.
- Maheswaran R, Payne N, Meechan D, Burden RP, Fryers PR *et al.* 2000. Socioeconomic deprivation, travel distance, and renal replacement therapy in the Trent region, United Kingdom 2000: An ecological study. *Journal of epidemiology and community health*, 57 (7): 523-4.
- Occelli F, Deram A, Génin M, Cuny D, Glowacki F-X and Néphronor Network. Mapping end-stage renal disease (ESRD): spatial variations on small area level in France, and association with deprivation, earlier renal replacement care and related diseases. *Plos One* (soumis le 13.02.2014).
- Roderick P, Clements S, Stone N, Martin D, Diamond I. 1999. What determines geographical variation in rates of acceptance onto renal replacement therapy in England? *Journal of Health Services Research and Policy*, 4 (3): 139-46.
- Tupin P. 2014. Accès aux soins des patients sous CMU. Réunion thématique annuelle de la commission d'épidémiologie de la société de néphrologie. Paris, le 9 avril 2014 (Communication orale).

Volkova N, McClellan W, Klein M, Flanders D, Kleinbaum D, *et al.* 2008. Neighborhood poverty and racial differences in ESRD incidence. *Journal of the American Society of Nephrology*, 19 (2): 356-64.

Ward MM. 2008. Socioeconomic status and the incidence of ESRD. *American Journal of Kidney Diseases*, 51 (4): 563-72.

Young EW, Mauger EA, Jiang K, Port FK, Wolfe RA. 1994. Socioeconomic status and end-stage renal disease in the united states. *Kidney International*, 45 (3): 907-11.

CHAPITRE III. FACTEURS DE RISQUES ENVIRONNEMENTAUX DES MALADIES INFLAMMATOIRES CHRONIQUES DE L'INTESTIN (MICI)

Partie 1. Introduction

Dans le chapitre précédent, après l'étude des disparités d'incidence d'une maladie rénale chronique, nous avons confirmé l'influence du niveau de précarité. Les cartographies d'incidence de l'IRCT seront par la suite rapprochées à un facteur de risque environnemental connu, le cadmium. Dans ce troisième chapitre axé sur les MICI, les facteurs étiologiques environnementaux sont au contraire peu identifiés. De fait, nous allons nous intéresser à l'apport et à la puissance des outils d'analyse spatiale, non seulement pour vérifier l'influence de facteurs suspectés comme l'aluminium, mais aussi pour identifier d'autres pistes environnementales. Ce chapitre se consacre à l'étude d'une base de données environnementale caractérisant la qualité des eaux souterraines : la banque nationale d'accès aux eaux souterraines. Celle-ci permet une étude à grande échelle et sur le long terme. Enfin, cette recherche nous a permis de mener une réflexion sur la construction d'une base de données environnementale optimale, en vue d'une étude en santé – environnement.

La maladie de Crohn (MC) et la recto-colite hémorragique (RCH) sont des inflammations chroniques des voies gastro-intestinales, également connues sous l'appellation de maladies inflammatoires chroniques de l'intestin (MICI). Des disparités d'incidence des MICI sont observées à l'échelle mondiale, avec des taux d'incidence élevés au nord de l'Europe, en Nouvelle-Zélande et en Amérique du Nord et de faibles taux d'incidence en Europe de l'Est, en Amérique du Sud en Asie et en Afrique du Nord (Cosnes *et al.*, 2011). Quelques facteurs étiologiques des MICI sont identifiés, comme l'effet délétère du tabac pour la MC et protecteur pour la RCH, ou l'effet protecteur de l'appendicectomie réalisée avant l'âge de 20 ans pour la RCH (Birrenbach and Böcker, 2004). Des prédispositions génétiques (mutation des gènes NOD1 & 2) expliquent également une partie de l'occurrence de MICI (Hampe *et al.*, 2001). Cependant, la présence d'inégalités spatiales d'incidence à plusieurs échelles, accentuées par l'augmentation rapide de l'incidence de ces maladies dans les pays développés depuis la fin de la 2^e guerre mondiale, ou encore l'absence de concordance forte chez les jumeaux monozygotes plaident en faveur du rôle de facteurs d'environnement dans le développement de ces pathologies (Nerich *et al.*, 2006 ; Jantchou *et al.*, 2006 ; Spehlmann *et al.*, 2008 ; Declercq *et al.*, 2010). Bien qu'aucun facteur de risque environnemental ne soit clairement identifié, les micropolluants environnementaux, notamment l'aluminium, constituent l'une des pistes, puisqu'ils engendreraient un mauvais fonctionnement du circuit lymphatique intestinal (Beamish *et al.*, 2011 ; Lerner, 2007 ; 2012 ; Pineton de Chambrun *et al.*, 2014).

Le registre des maladies inflammatoires chroniques du tube digestif (EPIMAD) recense les cas de MICI depuis 1988 dans le Nord de la France, au sein de quatre départements : le Nord, le Pas de Calais, la Somme et la Seine-Maritime (Gower-Rousseau *et al.*, 1994). Les données issues du registre ont déjà fait l'objet de plusieurs analyses spatiales et temporelles, démontrant non seulement des variations d'incidence entre les cantons des quatre départements pour les patients recensés entre 1990 et 2006 (Declercq *et al.*, 2010 : figure 28), mais aussi la présence de clusters spatiaux et spatio-temporels de surincidence et de sous-incidence (Génin *et al.*, 2013 : figure 29).

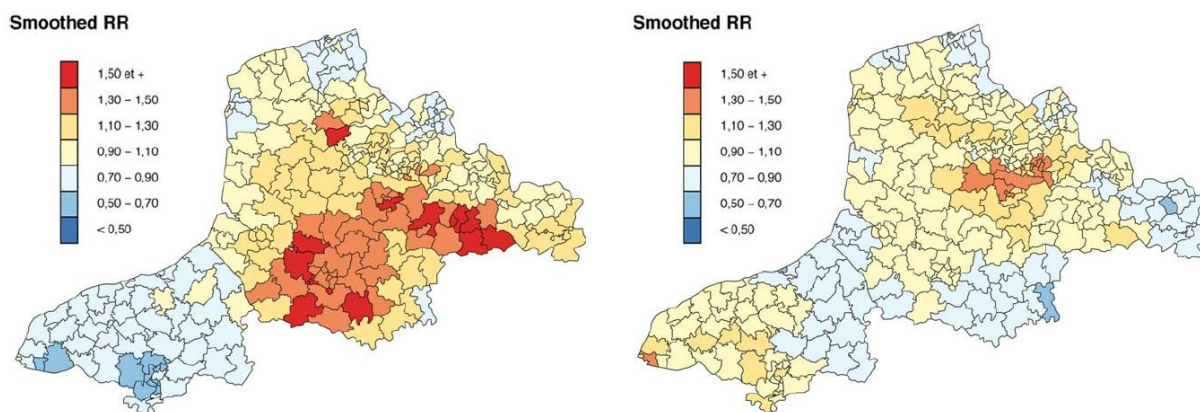


Figure 28. Risque relatif lissé de la MC (à gauche) et la RCH (à droite) par canton dans le nord de la France entre 1990 et 2003 (Declercq *et al.*, 2010)

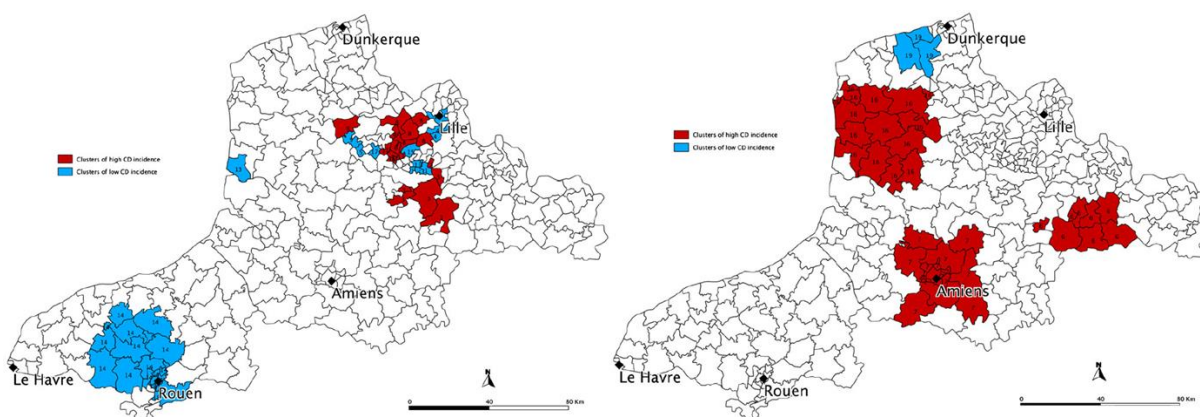


Figure 29. Risque relatif de la maladie de Crohn dans le nord de la France entre 1990 et 2006 – clusters spatiaux constants pendant la période (à gauche) et clusters spatio-temporels non constants pendant la période (à droite) (Génin *et al.*, 2013)

Dans le cadre de la recherche de facteurs environnementaux pouvant expliquer ces disparités d'incidence, une mise à jour de ces cartographies par l'équipe du Pr. Duhamel (EA2694 Université Lille 2) est en cours. Celle-ci intégrera les cas incidents de MC de 1990 à 2012 sur les 4 départements.

Suite à notre recensement des bases de données environnementales spatialisées, nous constatons que peu de ces données peuvent actuellement être mises en relation avec les cartographies des MICI. Pour l'air et le sol, aucune base de qualité des milieux n'est accessible à cette étendue géographique. Concernant le milieu aquatique, les mesures effectuées en sortie de robinet ne sont pas rendues publique sous un format de base de données. Dans ce contexte, la banque d'accès aux données des eaux souterraines (ADES Eau France, 2014) semble être la plus adéquate pour une telle étude. Elle a en effet l'avantage de proposer des résultats bruts d'analyses de contaminations chimiques des eaux sur l'ensemble du territoire du registre, et pour la totalité de la période d'étude.

Afin d'étayer ou suggérer des hypothèses environnementales des MICI, les cartographies d'incidence nécessitent d'être mises en relation avec des cartographies de qualité des milieux. Ce chapitre a pour objectifs d'évaluer le potentiel cartographique de la banque ADES, dans le cadre du suivi spatio-temporel de la qualité des eaux souterraines et d'une étude écologique en lien avec les MICI. Dans une première partie, nous nous sommes intéressés à l'optimisation d'une base de données environnementale, en vue de construire un outil cartographique d'aide à la décision. La deuxième

partie est axée sur l'analyse spatiale et spatio-temporelle des concentrations en aluminium dans les eaux souterraines, à l'échelle du registre EPIMAD. Dans une troisième partie, les cartographies de contamination des eaux souterraines sont confrontées à celles des MICI. Une discussion générale vient finalement conclure ce chapitre.

Partie 2. Optimisation d'une base de données environnementale en vue de construire un outil cartographique d'aide à la décision – cas de la contamination des eaux souterraines

a. Introduction

Les bases de données environnementales sont de puissants supports pour la construction d'outils d'aide à la décision, notamment lorsque les données sont géoréférencées. Bien qu'elles soient rarement conçues pour la mise en œuvre d'analyses spatialisées, certaines de ces bases incluent une masse non négligeable d'informations pour le développement d'outils cartographiques. A ce titre, la banque nationale ADES est un produit du SIE (Système d'Information sur l'Eau) disponible directement à partir du site internet public ADES Eau France (2014). Ce portail est animé par le Service géologique national (BRGM : Bureau de Recherches Géologiques et Minières) sous la coordination des ministères chargés de l'environnement et de la santé et de l'Office national de l'eau et des milieux aquatiques (Onema). Accessible depuis 2002, il fut mis en place afin de répondre aux enjeux de la Directive Cadre européenne sur l'eau concernant la surveillance des masses d'eau souterraines. De fait, il a pour objectif de constituer un outil de collecte et de conservation des données issues de ce milieu (Chery & Cattan, 2003).

La banque ADES recense l'ensemble des points d'eau de captage pour une alimentation en eau souterraine. Un point d'eau est un ouvrage souterrain dont le captage d'eau est destiné à alimenter des activités industrielles, agricoles, BTP, mais également le réseau de distribution des eaux de consommation humaine collectives et individuelles. Cette base permet un suivi quantitatif (niveau des nappes) et qualitatif (mesures de concentrations pour de nombreux paramètres chimiques dans l'eau) des eaux souterraines depuis 1900. Elle fournit également un accès aux fiches descriptives de l'ensemble des stations de mesures (opérateur de la station, mode d'utilisation, profondeur du forage, fréquence de mesures, localisation sur une carte, etc).

La complexité de cette base de données est liée à son mode collecte des informations. En effet, elle peut être alimentée par tout responsable d'un réseau de surveillance, d'un point d'eau ou d'un paramètre sur un point (Agences de l'Eau, Directions Régionales de l'Environnement, de l'Aménagement et du Logement (DREAL), Agences Régionales de Santé (ARS), collectivités territoriales, industriels dans le cadre du suivi des Installations Classées et Sites Pollués, etc.). Le fournisseur de données alimente régulièrement la banque, actualise et veille à la qualité des données communiquées. Par conséquent, les données peuvent varier en fonction de leur origine, leur nature, leur objectif, leur traçabilité ou leur niveau de validité.

L'objectif de cette étude est de réaliser une analyse critique de données issues de la banque ADES. L'enjeu est *in fine* de développer un outil cartographique d'aide à la décision, en termes de suivi de la qualité du milieu et de recherche en santé – environnement. Les cartographies générées à partir de ces données seront ainsi par la suite confrontées aux résultats cartographiques de MICI. Dans le cadre de nos travaux de recherche, les données utilisées concernent le suivi qualitatif des eaux souterraines. Celui-ci consiste à mesurer la concentration de paramètres chimiques directement à la sortie du captage, sans traitement préalable de l'eau. Ces données sont par la suite destinées à être confrontées aux cartographies d'incidence des MICI, grâce à une étude écologique (Declercq *et al.*, 2010; Génin *et al.*, 2013). C'est pourquoi nous nous sommes intéressés aux mesures de concentrations en aluminium, réalisées au niveau de l'étendue spatiale et temporelle du registre EPIMAD : le Nord – Pas de Calais, la Somme et la Seine – Maritime, entre 1990 et 2012.

b. Matériel et méthodes

La banque ADES met à disposition à ce jour les données reliées à 258 réseaux de surveillance contenant 74 941 points d'eau répartis sur tout le territoire français. Elles comportent 4 201 piézomètres et 71 977 qualitomètres (stations de mesure de la qualité des eaux). Cela représente plus de 10 millions de mesures de niveaux d'eau et 41 millions d'analyses d'eau consultables en ligne (ADES Eau France, 2014). Les informations recensées dans cette base sont disponibles par point d'eau, par réseau de mesure, par bassin hydrographique, région ou département et par aquifère. La recherche des analyses en aluminium s'est effectuée en fonction de localisation des points d'eau (sélection selon l'emprise géographique des 4 départements : figure 30). Le portail ADES propose ensuite d'exporter les données de ces points d'eau selon plusieurs critères. Les informations sélectionnées concernaient :

- les données issues de qualitomètres ;
- les analyses de micropolluants minéraux ;
- la période du 01/01/1990 au 31/12/2012 (période d'étude du registre EPIMAD) ;
- l'ensemble des réseaux de mesures et producteurs de données ;
- les mesures qualifiées selon le statut « indifférent ».



Figure 30. La zone d'étude, composée des départements du Nord, du Pas de Calais, de la Somme et de la Seine-Maritime (Chouraki *et al.*, 2011)

Un fichier d'export est finalement envoyé par mail au format texte. Ce fichier présente plus de 40 variables, comme le code national BSS (Banque du Sous-Sol) du point d'eau, la commune de localisation, le paramètre mesuré, le résultat de la mesure, le laboratoire accrédité ayant effectué l'analyse, la méthode d'analyse et le statut de fiabilité de l'analyse (ces variables seront détaillées par la suite). Pour des raisons de sécurité nationale, la localisation géographique exacte (coordonnées géographiques) des stations relatives à la qualité des eaux souterraines ne figure pas directement par défaut sur le portail de consultation, ni dans la base exportée. Ces coordonnées ont été retrouvées à partir du portail InfoTerreTM du BRGM (InfoTerre, 2014), *via* la base de données des ouvrages souterrains (BSS). Une jointure a permis d'agréger les deux bases téléchargées, selon un identifiant commun : le code BSS, attribué de manière unique à chaque point d'eau.

Plusieurs micropolluants minéraux sont recensés dans la base exportée. Seules les analyses en aluminium ont été conservées pour ce travail. Afin d'effectuer une analyse critique de la base téléchargée, des statistiques descriptives à l'aide du logiciel XLSTAT Addinsoft et des analyses cartographiques à l'aide du logiciel ArcGIS® v10.1 ont été réalisées.

c. Résultats

MODE D'UTILISATION DE L'OUVRAGE SOUTERRAIN

La base initiale exportée contient 8525 analyses en aluminium effectuées au niveau de 1439 points d'eau sur la période 1990 – 2012. Lors de la jointure de la base ADES avec la banque du sous-sol, l'une des variables de cette dernière nous informe sur l'utilisation de l'ouvrage, en d'autres termes, la destination des eaux captées. Compte tenu de nos objectifs (sanitaires), seuls les points de captage dont le mode d'utilisation constitue une possible voie de transfert directe (alimentation en eau potable) ou indirecte (transfert dans la chaîne alimentaire : eaux destinées à l'agriculture notamment) ont été conservés. Cela correspond à 70,8% des points d'eau et 69,8% des analyses de la base initiale. La base ainsi mise à jour sera nommée par la suite « B1 ». Le tableau 21 synthétise les différents modes d'utilisation des points d'eau et précise les modes conservés.

A la suite de cette étape, nous nous sommes intéressés aux variables permettant d'évaluer le niveau de fiabilité des informations. Nous avons dans ce cadre noté l'absence de métadonnées annexées à la base, ce qui rend difficile la compréhension de ces variables.

Tableau 21.

Sélection des points d'eau en fonction du mode d'utilisation

Mode d'utilisation	Conservation
Absorbant	Non
AEP	Oui
Agronomie	Non
Chauffage	Non
Chaussée	Non
Ciment	Non
Combustible	Non
Construction	Non
Dépollution	Non
Eau-agricole	Oui
Eau-alimentation	Oui
Eau-aspersion	Oui
Eau-cheptel	Oui
Eau-collective	Oui
Eau-domestique	Oui
Eau-individuelle	Oui
Eau-industrielle	Non
Eau-irrigation	Oui
Eau-minérale	Oui
Eau-pisciculture.	Oui
Eau-service-public	Oui
Embouteillage	Oui
Exhaure	Non
Fontaine	Oui
Granulat-béton	Non
Incendie	Non
Liant	Non
Piézomètre	Non
Pompe-a-chaleur	Non
Protection-cathodique	Non
Qualité-eau.	Non
Sonde-géothermique	Non
Thermalisme	Oui
Viabilité	Non
8525 Analyses	5954 Analyses
1439 Points d'eau	1019 Points d'eau

PRECISION INDUITE PAR LES TECHNIQUES D'ANALYSES

La précision de la donnée renseignée peut être évaluée à l'aide de la valeur du seuil de détection (LDD) ou de quantification (LDQ) des méthodes analytiques utilisées en laboratoire. Lorsqu'une mesure est inférieure à l'une de ces limites, le résultat attribué par défaut dans la base est soit la LDD en cas de non détection de la substance, soit la LDQ en cas de détection mais de non quantification de la substance. Cela concerne 66,5% des résultats d'analyses de la base B1.

Lorsqu'on s'intéresse à ces résultats, on observe une forte variabilité des LDD ou LDQ en fonction du laboratoire, allant de 1 à 10 000 µg/L pour l'aluminium. La figure 31 montre cette variabilité en fonction des laboratoires d'analyse (renommés de A à Q). Pour la plupart des laboratoires, les LDD ou LDQ sont situées entre 5 et 100 µg/L. Ainsi, même si les résultats de ces analyses sont peu précis, la valeur de concentration estimée reste inférieure à la norme sanitaire en termes de contamination des eaux souterraines, ce qui témoigne d'un milieu peu enrichi en aluminium. En effet, la limite de qualité des eaux brutes utilisées pour la production d'eau destinées à la consommation humaine s'élève à 200 µg/L pour l'aluminium (Arrêté du 11/01/07). En revanche, l'un des laboratoires présente des résultats d'analyses avec des limites de détection très élevées par rapport à cette norme (10000 µg/L), ce qui rend complexe l'interprétation du résultat. Il s'agit du laboratoire départemental de la Somme (193 analyses).

LDD / LDQ

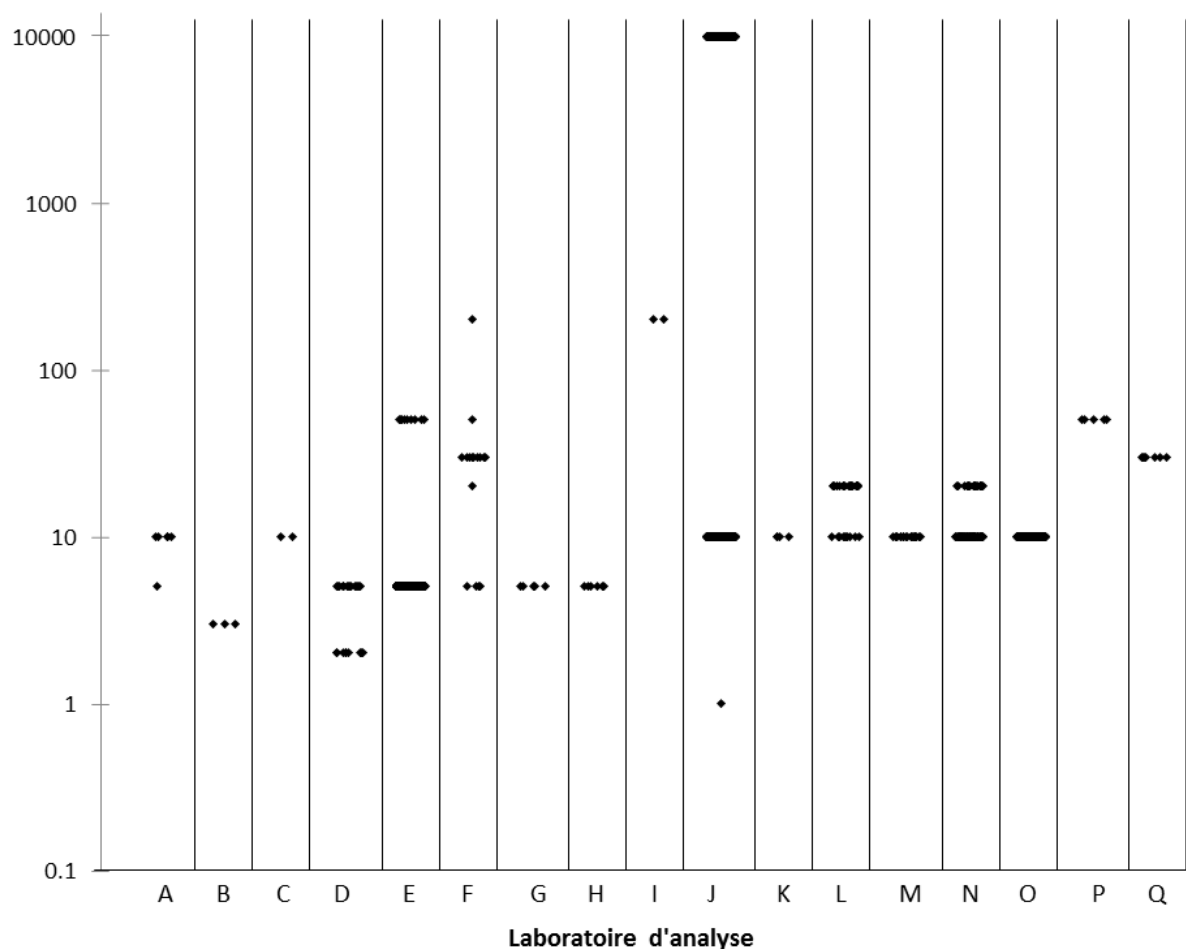


Figure 31. Variabilité des LDD ou LDQ (en µg/L d'aluminium) en fonction du laboratoire d'analyse

Si l'on s'intéresse à la variabilité temporelle de ces limites (figure 32), nous remarquons qu'en dehors des campagnes réalisées en 1994, 1995 et 1996 par le laboratoire départemental de la Somme, le seuil moyen (représenté par une croix rouge sur la figure) se situe entre 5 et 10 $\mu\text{g/L}$. Les performances analytiques restent donc stables dans le temps concernant l'aluminium.

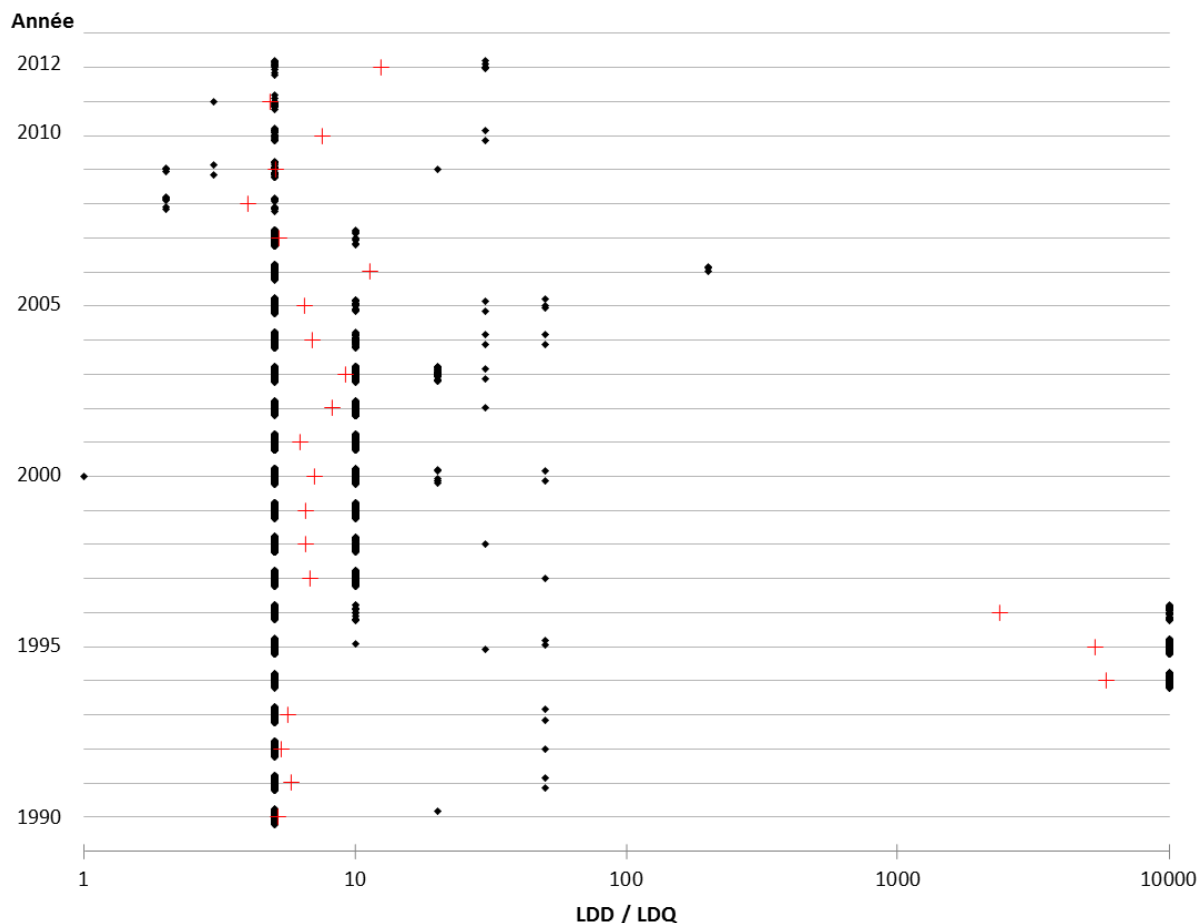


Figure 32. Variabilité des LDD et LDQ (en $\mu\text{g/L}$ d'aluminium) en fonction du temps

Les résultats concernés par des valeurs inférieures aux LDD ou LDQ peuvent induire une tendance à surestimer l'état de dégradation du milieu, car la limite haute maximise la valeur de concentration probable. Pour limiter cette surestimation, les résultats correspondants ont dans un premier temps été divisés par deux, ce qui permet d'induire le moins d'erreur possible. Par exemple, pour un résultat présentant une valeur inférieure à un LDD de 50 $\mu\text{g/L}$, la valeur considérée par la suite sera 25 $\mu\text{g/L}$. En plus de ce traitement, les résultats d'analyse pour lesquels la LDD ou LDQ est supérieure à la norme sanitaire de 200 $\mu\text{g/L}$ sont supprimés de la base. Ces données sont susceptibles d'engendrer un niveau d'incertitude trop important lors de l'analyse statistique (présence de valeurs extrêmes) et cartographique (regroupement de ces points au niveau du département de la Somme) et ne concernent que 1% des analyses de la base B1.

Par ailleurs, cette analyse des variables liées aux résultats analytiques a mis en évidence des concentrations de 0 $\mu\text{g/L}$ ou même parfois une absence de valeur de concentration. Ces résultats concernent moins de 1% de la base B1 et ont également été supprimés. La base mise à jour lors de cette étape sera nommée par la suite « B2 ».

STATUT ET QUALIFICATION DU RESULTAT

Les résultats fournis par l'organisme chargé de l'analyse chimique sont communiqués sur le portail sous la responsabilité du producteur de données. Ce dernier joint à ces résultats un niveau de fiabilité, évalué au regard de la connaissance et du contrôle du processus de production de la donnée. Il s'engage ainsi sur la vraisemblance et la représentativité du résultat d'analyse par rapport au milieu où a été réalisé le prélèvement. Les données peuvent avoir différents niveaux de validité, évalués à l'aide d'un couple statut/qualification (tableau 22). Le statut définit l'état de contrôle de la donnée, la qualification son niveau de fiabilité. Les définitions exactes des différents niveaux de statut et de qualification de la donnée sont disponibles en annexe II.

Tableau 22. Couples de valeurs statut/qualification autorisés dans le schéma Sandre et pourcentage de valeurs correspondantes dans la base aluminium

STATUT		QUALIFICATION				
		0	1	2	3	4
		Non définissable	Correcte	Incorrecte	Incertaine	Non qualifié
Brute	1	X (< 1%)				X (< 1%)
Contrôlée Niveau 1	2		X (96%)	X (0%)	X (0%)	
Contrôlée Niveau 2	3		X (1%)	X (0%)	X (0%)	
Interprétée	4		X (2%)	X (0%)	X (0%)	

Une donnée brute signifie que celle-ci n'a pas été contrôlée, ou bien que son niveau de fiabilité ne peut être défini (non définissable ou non qualifié). Au sein de la base B2, les données brutes représentent moins de 1% du nombre total d'analyses effectuées. Elles ont été supprimées par mesure de précaution. A l'inverse d'une donnée brute, le statut « contrôlée » ou « interprétée » signifie qu'un niveau de qualification du résultat est fourni par le producteur. Dans ce cas, la donnée est définie comme correcte, incorrecte ou incertaine. Aucune donnée classée incorrecte ou incertaine n'est observée dans notre base B2. Ainsi, toutes les valeurs contrôlées sont qualifiées de correctes, ce qui correspond à plus de 99% des analyses en aluminium. La base mise à jour sera nommée par la suite « B3 ».

TEST DE VALEURS EXTREMES

A la suite de ces étapes d'analyse des variables permettant d'évaluer le niveau de fiabilité des informations, nous nous sommes intéressés aux valeurs de concentrations en aluminium. Le scattergram montre la distribution des résultats analytiques (figure 33). On observe un grand nombre de valeurs de faibles concentrations et la présence de pics de valeurs extrêmes. Un test de valeurs extrêmes a donc été mis en œuvre pour identifier les valeurs statistiquement aberrantes. Le test de Grubbs unilatéral à droite (à un niveau de significativité de 1%), confirme que les valeurs supérieures à 2000 µg/L peuvent être considérées comme extérieures à la distribution et donc potentiellement aberrantes. Celles-ci correspondent à 13 mesures de la base B3 et ont été supprimées pour la suite de l'étude. La base mise à jour sera nommée par la suite « B4 ».

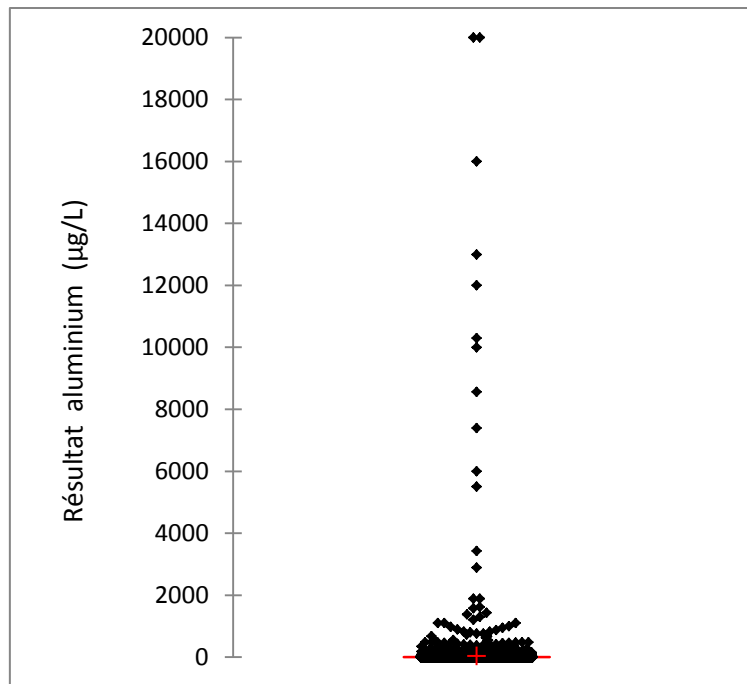


Figure 33. Distribution des résultats d'analyse en aluminium ($\mu\text{g/L}$) dans les eaux souterraines (base B3)

SYNTHESE DE LA VALIDATION DES DONNEES

Grâce aux étapes précédentes, nous avons pu mettre en forme, contrôler et valider une base de données de qualité des eaux souterraines extraite du portail ADES. Le tableau 23 synthétise les résultats issus de ces différentes étapes. Celles-ci peuvent représenter un arbre de décision quant à l'utilisation des qualitomètres de la banque ADES. La base finale contient 5695 analyses en aluminium effectuées au niveau de 995 points d'eau entre 1990 et 2012, ce qui représente 67% des analyses et 69% des points de la base initiale. Après sélection selon le mode d'utilisation du point d'eau (base B1), un peu plus de 4% des 5954 analyses ont été supprimées suite à l'interprétation du niveau de fiabilité des données. Ce chiffre témoigne de la prudence nécessaire lors de l'utilisation de ces données d'eau souterraine. La base B4 étant validée, nous pouvons à présent interpréter la distribution statistique des valeurs de concentrations.

Tableau 23. Synthèse des traitements de la base aluminium

Base	Etape de traitement	Nombre d'analyses après traitement (%)	Nombre de points d'eau après traitement (%)
Initiale	Aucun	8525 (100%)	1439 (100%)
B1	Mode d'utilisation point d'eau	5954 (69.8%)	1019 (70.8%)
B2	LDD et LDQ >200 $\mu\text{g/L}$ (norme)	5762 (67.6%)	996 (69.2%)
B3	Statut et qualification de la mesure	5708 (67%)	995 (69.1%)
B4 (validée)	Test valeurs extrêmes	5695 (66.8%)	995 (69.1%)

STATISTIQUES DESCRIPTIVES DE LA BASE VALIDEE

La figure 34 synthétise la distribution statistique des résultats de ces analyses pour l'aluminium. Le scattergram montre une distribution de forme pyramidale, c'est-à-dire que le nombre de valeurs décroît en fonction de la valeur de concentration. La majeure partie des valeurs se situe en dessous de 100 µg/L et quelques pics de concentrations sont mesurés.

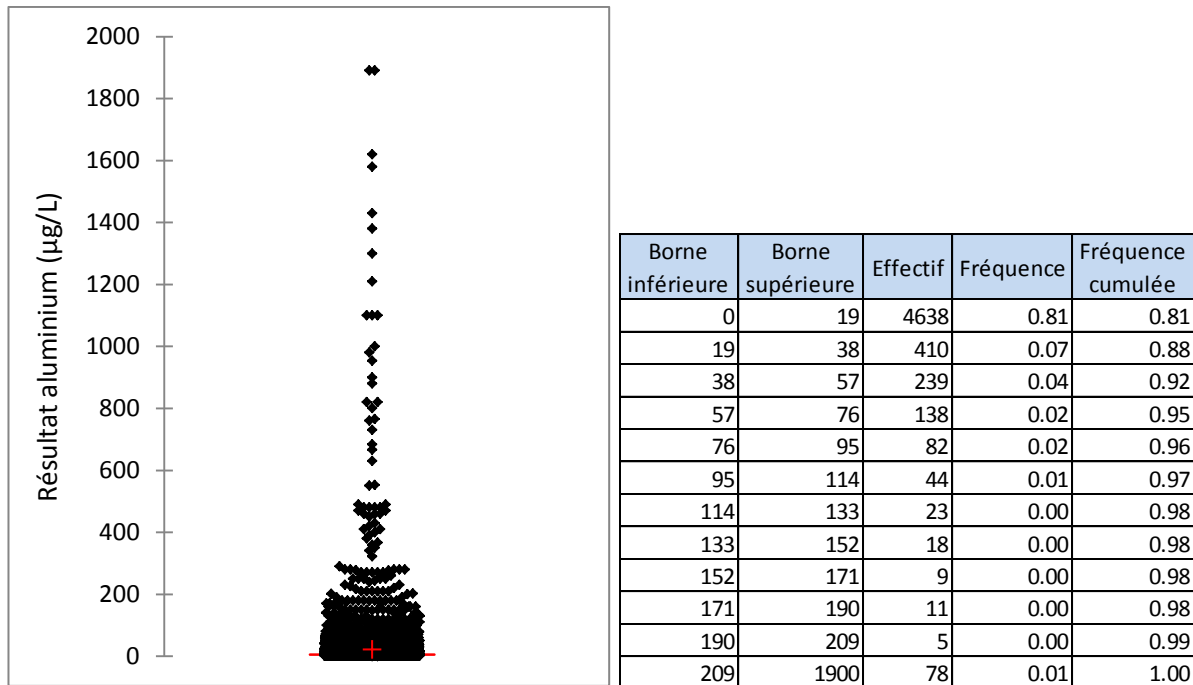


Figure 34. Distribution des résultats d'analyse en aluminium (µg/L) dans les eaux souterraines (base B4 validée)

Le tableau 24 indique un milieu globalement très peu enrichi en aluminium au regard de la limite de qualité des eaux brutes utilisées pour la production d'eau destinées à la consommation humaine (200 µg/L), puisque 90% des valeurs sont inférieures ou égales à 42 µg/L. Au moins 50% des valeurs sont mêmes inférieures à 5 µg/L, ce qui représente un niveau très faible de concentrations dans les eaux souterraines. La similarité des valeurs de premier quartile et de percentile 10 (2,5µg/L) témoigne de l'influence des limites de détection des appareils de mesures, situées pour la plupart à 5µg/L. On remarque ensuite une moyenne et une médiane inférieures à la norme, mais très éloignées (respectivement 21.07 µg/L et 5 µg/L). Malgré le nombre très élevé d'analyses, la distribution ne semble donc pas prendre une allure gaussienne, ce qui nécessitera l'utilisation de tests non-paramétriques. L'écart-type élevé (84.83 µg/L) confirme l'influence des valeurs extrêmes sur la moyenne. Finalement, le très faible nombre de valeurs supérieures à 200 µg/L laisse supposer la présence de pics de pollutions dans l'espace et le temps. Il nous faut donc maintenant évaluer la force d'échantillonnage selon ces deux dimensions.

Tableau 24. Statistiques descriptives de la distribution des résultats d'analyse en aluminium dans les eaux souterraines ($\mu\text{g/L}$)

Statistique	Résultat aluminium ($\mu\text{g/L}$)
Minimum	0.005
Maximum	1890
Percentile 10	2.5
1er Quartile	2.5
Médiane	5
3ème Quartile	10
Percentile 90	42
Moyenne	21.07
Variance (n)	7195.30
Ecart-type (n)	84.83

FORCE D'ÉCHANTILLONNAGE

Dans l'espace, nous constatons que la force d'échantillonnage, en termes de points d'eau, est variable en fonction du territoire (figure 35). Le département du Pas de Calais présente une forte densité de points d'eau ayant fait l'objet d'un suivi en aluminium (0.57/ha, 382 points). Cette densité est globalement plus faible sur les départements de la Somme et de la Seine-Maritime (respectivement 0.32/ha, 200 points et 0.37/ha, 234 points). Le département du Nord montre également une très faible densité de points (0.31/ha, 179 points), notamment sur sa moitié Nord. Sur le bassin lillois, la forte urbanisation limite les prélèvements d'eau. Plus au nord, la masse d'eau souterraine du bassin des Flandres est proche de la surface, ce qui lui vaut une extrême sensibilité aux infrastructures anthropiques. Par conséquent, les captages d'eau de cette zone ne sont pas destinés à alimenter l'eau potable.

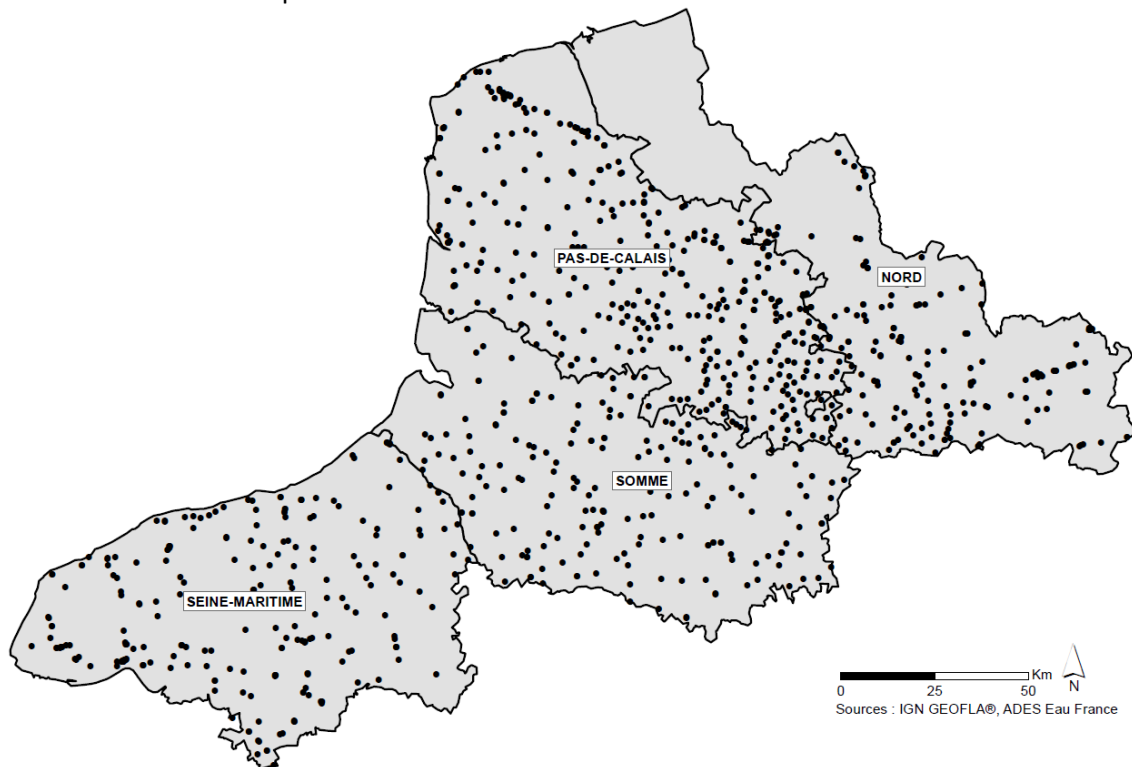


Figure 35. Répartition spatiale des points d'eau faisant l'objet d'un suivi en aluminium

Nous avons ensuite étudié la variabilité du nombre d'analyses en aluminium effectuées par année, en fonction du territoire (figure 36). Ce nombre n'est pas constant dans le temps. Sur l'ensemble du territoire d'étude (en noir), la courbe montre une densité croissante de mesures entre 1990 et 1999, avec un pic de surveillance entre 1997 et 2001 pour lesquelles plus de 500 échantillons sont prélevés par année. Ce nombre diminue ensuite jusqu'à atteindre moins de 100 analyses à partir de 2008. On remarque ensuite une forte hétérogénéité du nombre d'analyses en fonction des départements, ce qui confirme les observations précédentes en termes de densité de points d'eau. Le Pas de Calais (en bleu) est le territoire présentant une surveillance accrue de l'aluminium, notamment lors du pic entre 1997 et 2001. Le département de la Seine-Maritime (en vert) arrive en 2^e position, avec cependant quelques années dépourvues d'analyse en 1990, 1991, 2004 et 2006. A la différence de la situation précédente, c'est la Somme (en violet) qui fait l'objet du suivi le plus restreint, notamment avant 1997 et après 2004. Le nombre d'analyses est proche de 0 pour ces périodes. Enfin pour le département du Nord (en rouge), le nombre de mesures en aluminium est le plus élevé pour les 2 premières années de notre période d'étude. Après un vide d'échantillons entre 1992 et 1995, la surveillance se rapproche de celle de la Somme et de la Seine-Maritime, pour finalement être proche de 0 à partir des années 2000.

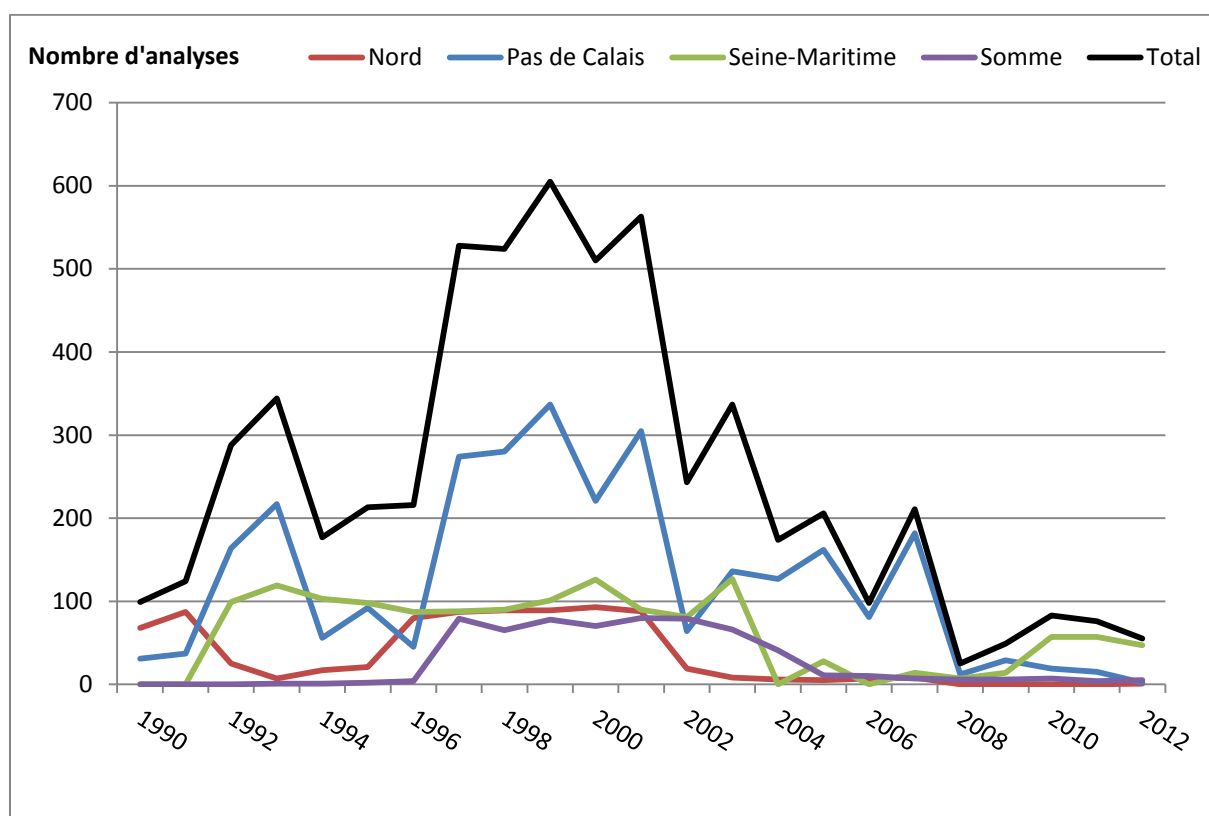


Figure 36. Nombre d'analyses effectuées par année en fonction du territoire étudié (source des données : ADES Eau France, 2014)

De cette analyse de la force d'échantillonnage en aluminium dans les eaux souterraines, il en ressort une hétérogénéité entre les 4 départements d'étude et une forte variabilité dans le temps. Un pic de surveillance entre 1997 et 2001 est observé pour l'ensemble des départements, ce qui pourrait correspondre à une période propice pour l'analyse spatiale de la contamination des eaux souterraines en aluminium.

d. Conclusion

Grâce à l'étude de plusieurs critères, nous avons pu mettre en forme, contrôler et valider une base de données de qualité des eaux souterraines extraite du portail ADES. L'analyse statistique de la base de données finale suggère la présence de pics de pollution en aluminium dans l'espace et le temps. Dans la partie suivante, nous nous focaliserons sur le potentiel cartographique de cette base pour le suivi de la contamination des eaux souterraines en aluminium.

Partie 3. Analyse spatiale et spatio-temporelle de la contamination en aluminium dans les eaux souterraines

a. Introduction

Nous avons vu précédemment que la banque ADES met à disposition à ce jour les données de 258 réseaux de surveillance contenant 74 941 points d'eau répartis sur tout le territoire français, comportant 4 201 piézomètres et 71 977 qualitomètres. Soit plus de 10 millions de mesures de niveaux d'eau et 41 millions d'analyses d'eau consultables en ligne (ADES Eau France, 2014).

Dans la partie précédente, la base de données de contamination en aluminium dans les eaux souterraines fut exportée concernant les points d'eau situés sur les départements du Nord, du Pas de Calais, de la Somme et de la Seine-Maritime. Les analyses concernent la période 1990 – 2012. Cette base a ensuite été optimisée pour être intégrée dans un outil cartographique d'aide à la décision en santé-environnement. Le présent travail a pour objectif de réaliser des analyses spatiales et spatio-temporelles des concentrations en aluminium issues de la banque ADES dans le but d'identifier des zones et des épisodes de contamination des eaux souterraines. L'enjeu sera par la suite de confronter ces résultats aux cartographies spatio-temporelles des MICI.

b. Matériel et méthodes

La base contrôlée et validée précédemment contient 5695 analyses en aluminium effectuées au niveau de 995 points d'eau entre 1990 et 2012. Ces analyses sont effectuées sur des échantillons d'eau prélevés directement à la sortie des points de captage, sans traitement préalable.

Afin d'étudier le comportement spatial de la contamination en aluminium dans les eaux souterraines, les résultats d'analyse ont été considérés en fonction des coordonnées spatiales des points d'eau. La concentration moyenne en aluminium a été calculée pour la période 1990 – 2012, à partir des différentes analyses effectuées pour chaque point d'eau. Une cartographie ponctuelle a été générée à l'aide du logiciel ArcGIS® v10.1. La détermination de classes de concentrations en aluminium sur les cartes fut réalisée à l'aide des seuils naturels de Jenks. Ils définissent des classes statistiques avec une variance maximale entre les classes et une variance minimale à l'intérieur de chaque classe. On obtient ainsi des groupes de valeurs homogènes. Certains points d'eau étant très proches dans l'espace (de l'ordre de quelques mètres), les entités ponctuelles représentées à grande échelle sur la carte peuvent se superposer. Les valeurs élevées ont été placées au-dessus.

La méthode de krigeage empirique bayésien de la boîte à outil Geostatistical Analyst (Krivoruchko, 2012) a ensuite été utilisée pour la production d'une cartographie géostatistique représentant la concentration moyenne en aluminium sur la période 1990-2012. Les concentrations ont également été classées selon des seuils de Jenks.

L'analyse spatio-temporelle a finalement été réalisée sous ArcGIS® à l'aide du module chronologique. Nous avons pour cela calculé la concentration moyenne en aluminium au niveau de chaque point d'eau, selon un pas de temps annuel. La date de prélèvement de l'échantillon d'eau était dans ce cas considérée. Une cartographie ponctuelle a ainsi été générée pour chaque année.

c. Résultats

ANALYSE SPATIALE SUR LA PERIODE 1990-2012

Cartographie ponctuelle

La cartographie ponctuelle de la concentration moyenne en aluminium par point d'eau, sur la période 1990 – 2012 est présentée en figure 37. Selon la classification de Jenks sept classes sur les dix concernant des concentrations inférieures à 100 µg/L. Seules les deux dernières classes intègrent les points d'eau de concentration supérieures à 200 µg/L.

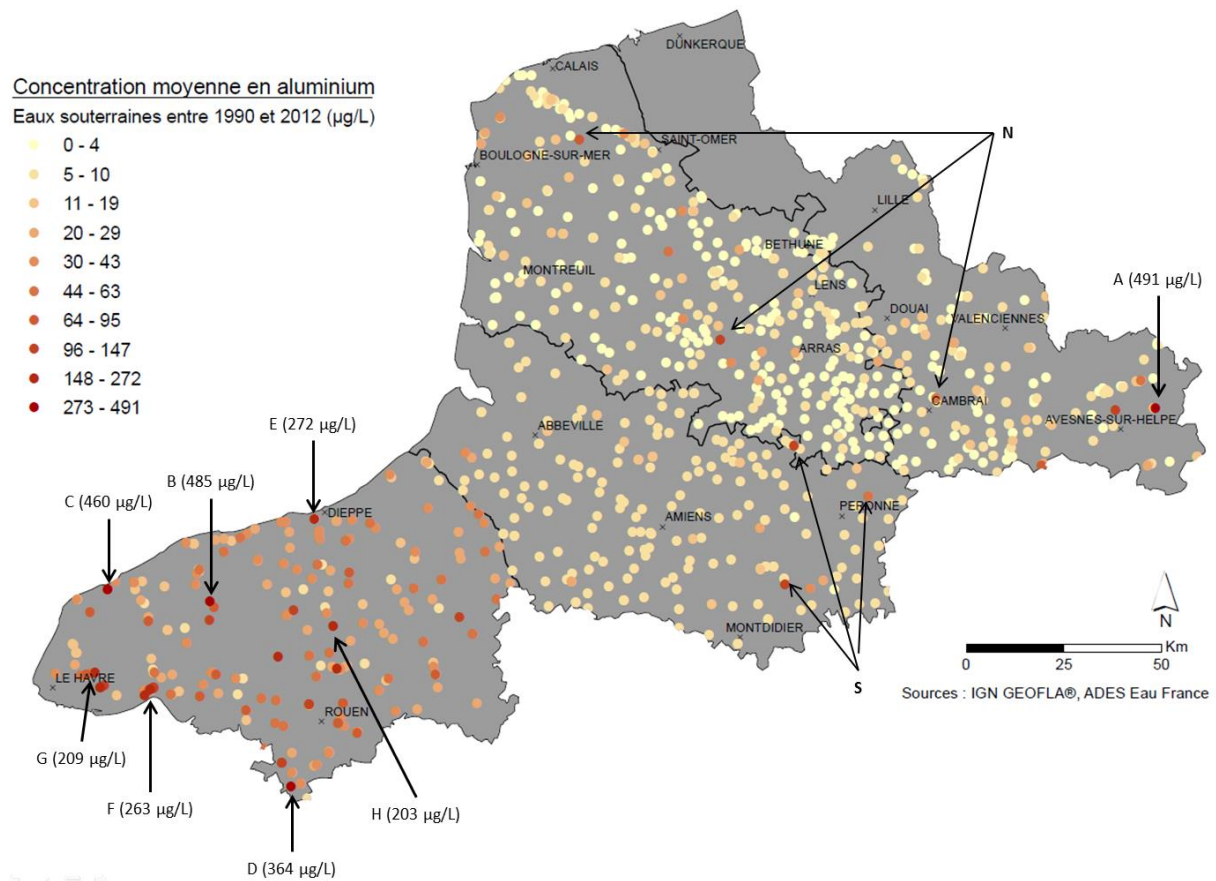


Figure 37. Concentration moyenne en aluminium par point d'eau entre 1990 et 2012 (µg/L) par cartographie ponctuelle – les classes sont déterminées à l'aide des seuils de Jenks – les points A à H sont supérieurs à 200 µg/L – les points N et S sont des pollutions ponctuelles sur les départements du Nord, du Pas de Calais et de la Somme

La figure 38 résume ensuite la distribution statistique des valeurs de ces points d'eau. Nous sommes en présence d'un grand nombre de valeurs faibles (< 50 µg/L). La médiane (5µg/L) et la moyenne (16.58 µg/L) de ces points indiquent que le niveau global de concentration en aluminium dans les eaux souterraines est très faible sur la zone d'étude. Huit points sur 995 sont situés au-dessus de 200 µg/L, ce qui suppose des zones ponctuelles de contamination. Ces huit points, nommés de A à H, sont détaillés sur la carte à l'aide de flèches. Parmi eux, sept sont situés sur le département de la Seine-Maritime. Le point d'eau à proximité d'Avesnes-sur-Helpe (A : à l'extrême Est de la carte) présente la concentration moyenne la plus élevée (491 µg/L). Une description détaillée de ces points est présentée par la suite.

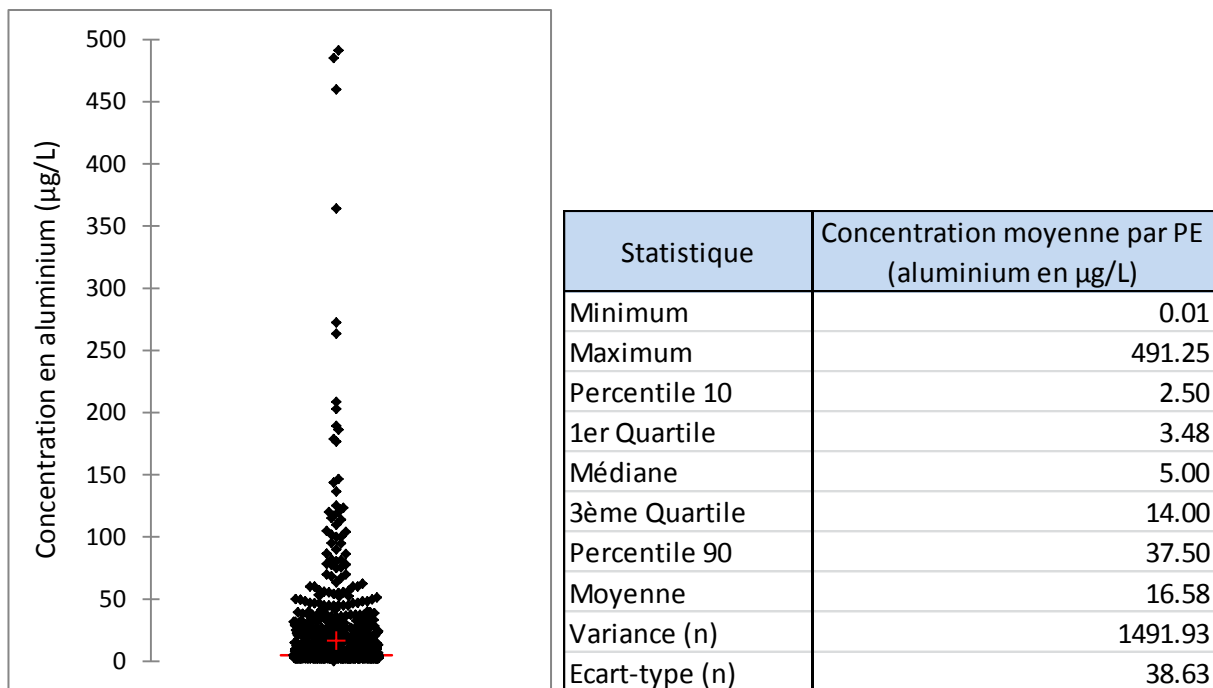


Figure 38. Distribution des concentrations moyennes en aluminium par point d'eau entre 1990 et 2012 (µg/L)

Sur la carte, 84% des valeurs supérieures à 64 µg/L (à partir du 7^e niveau de Jenks) sont situées sur le territoire de la Seine-Maritime. Les concentrations retrouvées sur le département de la Somme sont situées entre la médiane (5 µg/L) et la moyenne (17 µg/L) et sont très peu variables. Seuls trois points ressortent (marqués par un « S » sur la carte), à proximité de la frontière du Pas de Calais (100 µg/L), au niveau de Péronne (56 µg/L) et au Nord de Montdidier (120 µg/L). Les départements du Nord et du Pas de Calais présentent en revanche une plus grande variabilité, avec un grand nombre de concentrations très faibles (<10 µg/L) et quelques points d'eau de concentration plus élevée (marqués par un « N » sur la carte), notamment à proximité de Cambrai (57 µg/L), à l'Ouest d'Arras (100 µg/L) et entre Boulogne-sur-Mer et St-Omer (101 µg/L).

Les différences entre les départements ont été étudiées à l'aide du test multiple de Dunn. Il existe une différence significative ($p < 0.0083$) des concentrations moyennes en aluminium entre chaque département (tableau 25). Le département de la Seine-Maritime montre des concentrations en moyenne beaucoup plus élevées par rapport aux autres territoires (47.12 ± 61.6 µg/L), alors que le Pas de Calais présente la moyenne la plus faible (5.71 ± 8.57 µg/L).

Tableau 25. Résultats du test multiple de Dunn pour la comparaison des concentrations en aluminium entre les départements (niveau de signification : $p < 0.0083$)

Département	Moyenne (µg/L) ± écart-type	Test multiple de Dunn (p value)			
		Nord	Pas de Calais	Seine-Maritime	Somme
Nord	9.9 ± 38.01	0.003	< 0.0001	< 0.0001	< 0.0001
Pas de Calais	5.71 ± 8.57				
Seine-Maritime	47.12 ± 61.6	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Somme	7.59 ± 11.82	< 0.0001	< 0.0001	< 0.0001	< 0.0001

Cartographie géostatistique

Notre objectif est la recherche de facteurs environnementaux pouvant expliquer les disparités d'incidence de MICI. Afin de pouvoir confronter ces données environnementales aux résultats cartographiques d'incidence, les zones dépourvues d'échantillons nécessitent d'être renseignées. A ce titre, nous avons montré dans le chapitre I que les outils géostatistiques d'interpolation spatiale peuvent être utilisés pour pallier ce manque d'information. En complément de la cartographie ponctuelle, une cartographie géostatistique a donc été réalisée (figure 39). Sur celle-ci, seules les deux dernières classes représentent des concentrations élevées en aluminium, la dernière classe concernant les valeurs supérieures à la norme sanitaire. La carte montre des niveaux de concentrations globalement faibles sur les départements du Nord, du Pas de Calais et de la Somme. Les niveaux sont inférieurs à 100 µg/L d'aluminium. Nous pouvons cependant remarquer la présence des pics de pollutions au niveau d'Avesnes-sur-Helpe, ainsi qu'au centre du département du Pas de Calais (valeurs situées entre 85 et 200 µg/L). Sur le département de la Seine-Maritime, les niveaux de concentration sont plus élevés, puisque l'ensemble de ce territoire est concerné par des valeurs situées dans les 3 dernières classes. La moitié Ouest présente d'ailleurs des valeurs supérieures à 85 µg/L. Nous pouvons finalement isoler deux zones de très fortes concentrations, l'une à l'Est du Havre et l'autre plus au Nord.

Pour accompagner cette cartographie géostatistique, l'outil de krigeage propose de représenter l'erreur standard d'estimation (figure 40). Sur la carte, nous pouvons constater d'une manière attendue que l'erreur d'estimation est plus élevée au niveau des zones dépourvues d'échantillons, comme sur le bassin dunkerquois, où le modèle du krigeage estime l'erreur potentielle à hauteur de 57 µg/L au maximum. De même, cette erreur est plus importante à proximité des points présentant des valeurs élevées de concentrations et plus faible à proximité des zones de faibles concentrations. Au niveau du département de la Seine-Maritime qui présente les concentrations les plus élevées, l'erreur est également estimée 57 µg/L au maximum, alors qu'elle est de moins de 3 µg/L entre Abbeville et Amiens, où les concentrations ne dépassent pas 10 µg/L d'aluminium.

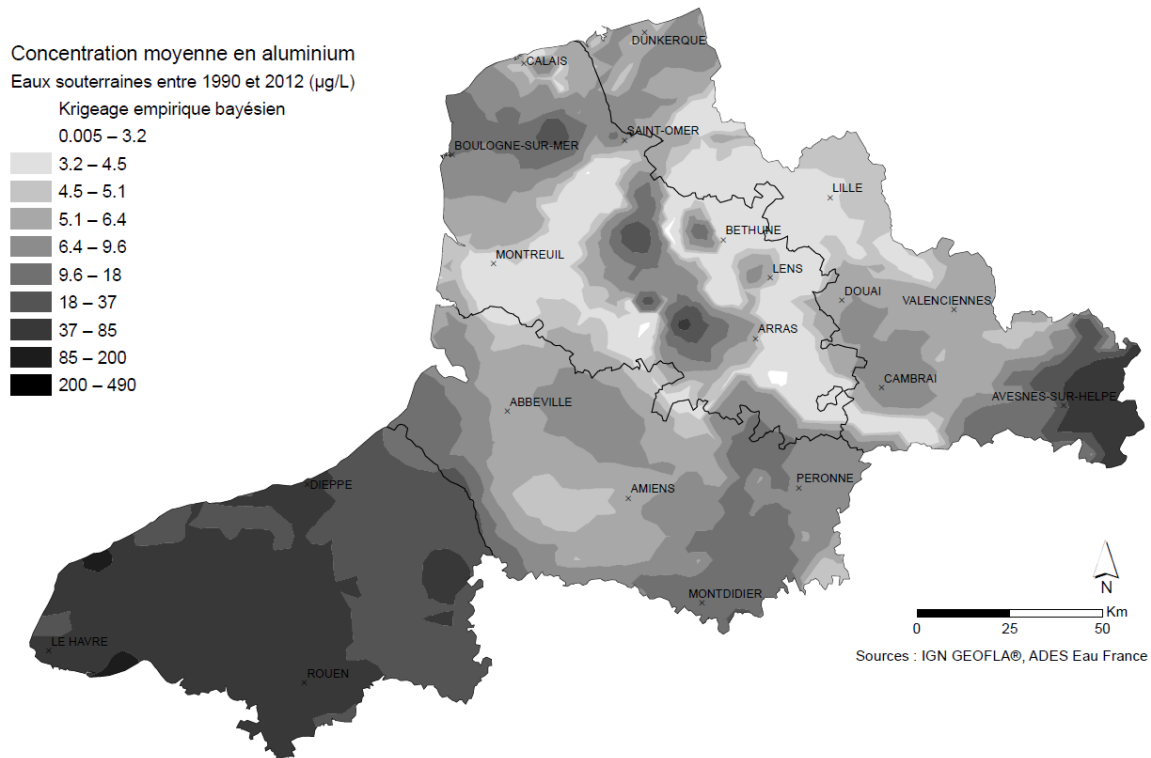


Figure 39. Concentration moyenne en aluminium par point d'eau entre 1990 et 2012 ($\mu\text{g/L}$) par cartographie géostatistique (logiciel ArcGIS®)

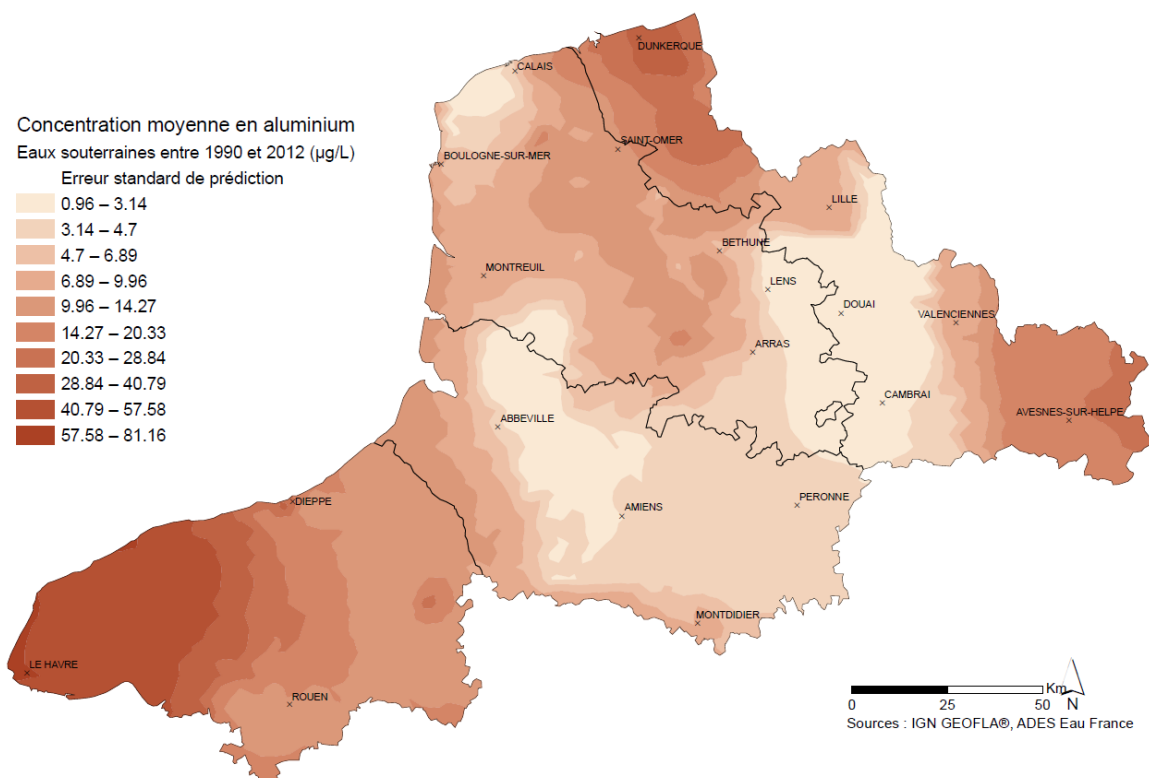


Figure 40. Erreur standard de prédiction de la concentration moyenne en aluminium entre 1990 et 2012 ($\mu\text{g/L}$) (logiciel ArcGIS®)

ANALYSE TEMPORELLE

Afin d'évaluer la représentativité de ces résultats en termes de contamination moyenne sur l'ensemble de la période d'étude, nous nous sommes ensuite intéressés au nombre d'échantillons prélevés par point d'eau. Le tableau 26 montre qu'en moyenne 5,8 analyses par point sont réalisées sur la période 1990 – 2012. Plus de 10% des points ne disposent que d'une seule valeur sur l'ensemble de la période et la moitié ont été échantillonnés 4 fois au maximum. Ces résultats montrent un échantillonnage irrégulier dans le temps en fonction des points.

Tableau 26. Statistiques descriptives de la distribution du nombre d'analyses en aluminium effectuées au niveau des points d'eau

Statistique	Nombre d'analyses par PE
Minimum	1
Maximum	52
Eff. du minimum	113
Eff. du maximum	1
Percentile 10	1
1er Quartile	3
Médiane	4
3ème Quartile	8
Percentile 90	11
Moyenne	5.8
Variance (n)	26.3
Ecart-type (n)	5.1

Cette faible force d'échantillonnage pour un même point d'eau peut supposer une forte influence de valeurs élevées et donc la présence de pics de concentration dans le temps. Pour le vérifier, les 8 points d'eau présentant les concentrations moyennes supérieures à 200 µg/L ont été isolés, afin de réaliser une description temporelle. Sous le portail ADES, il est possible de représenter l'évolution des concentrations dans le temps sous forme de graphique (figure 41). Ceux-ci étant générés de manière individuelle, les échelles temporelles (en abscisse) et de concentrations (en ordonnées) ne sont pas les mêmes.

Nous pouvons distinguer plusieurs situations. Dans un premier cas (graphiques A et E), le point d'eau est peu suivi (deux ou trois échantillons) et une valeur très élevée est observée (980 µg/L pour A et 800 µg/L pour E). La faible surveillance de ces points ne nous permet pas de définir si cette pollution est ponctuelle ou chronique. Au contraire, dans un second cas (graphiques B, D et H) le point fait l'objet d'un suivi plus régulier dans le temps (quatre ou cinq échantillons). La valeur élevée de concentration mesurée sur l'un de ces échantillons témoigne d'une pollution ponctuelle (1890 µg/L pour B, 1380 µg/L pour D et 900 µg/L pour H). Dans un troisième cas (graphiques F et G), le point ne fait pas l'objet d'un suivi régulier. Parmi les trois échantillons mesurés sur chaque point, deux présentent des résultats supérieurs à 200 µg/L. Nous sommes alors dans une situation de pollution chronique, mais peu surveillée. Finalement, le quatrième cas concerne le graphique C. Celui-ci fait l'objet d'une surveillance accrue de la concentration en aluminium (15 échantillons répartis sur trois années). Parmi les 15 valeurs mesurées, neuf sont supérieures à 200 µg/L. Nous sommes ici dans une situation de pollution chronique et surveillée. Une telle variabilité de surveillance et de contamination des points d'eau nous incite donc à réaliser une analyse spatio-temporelle.

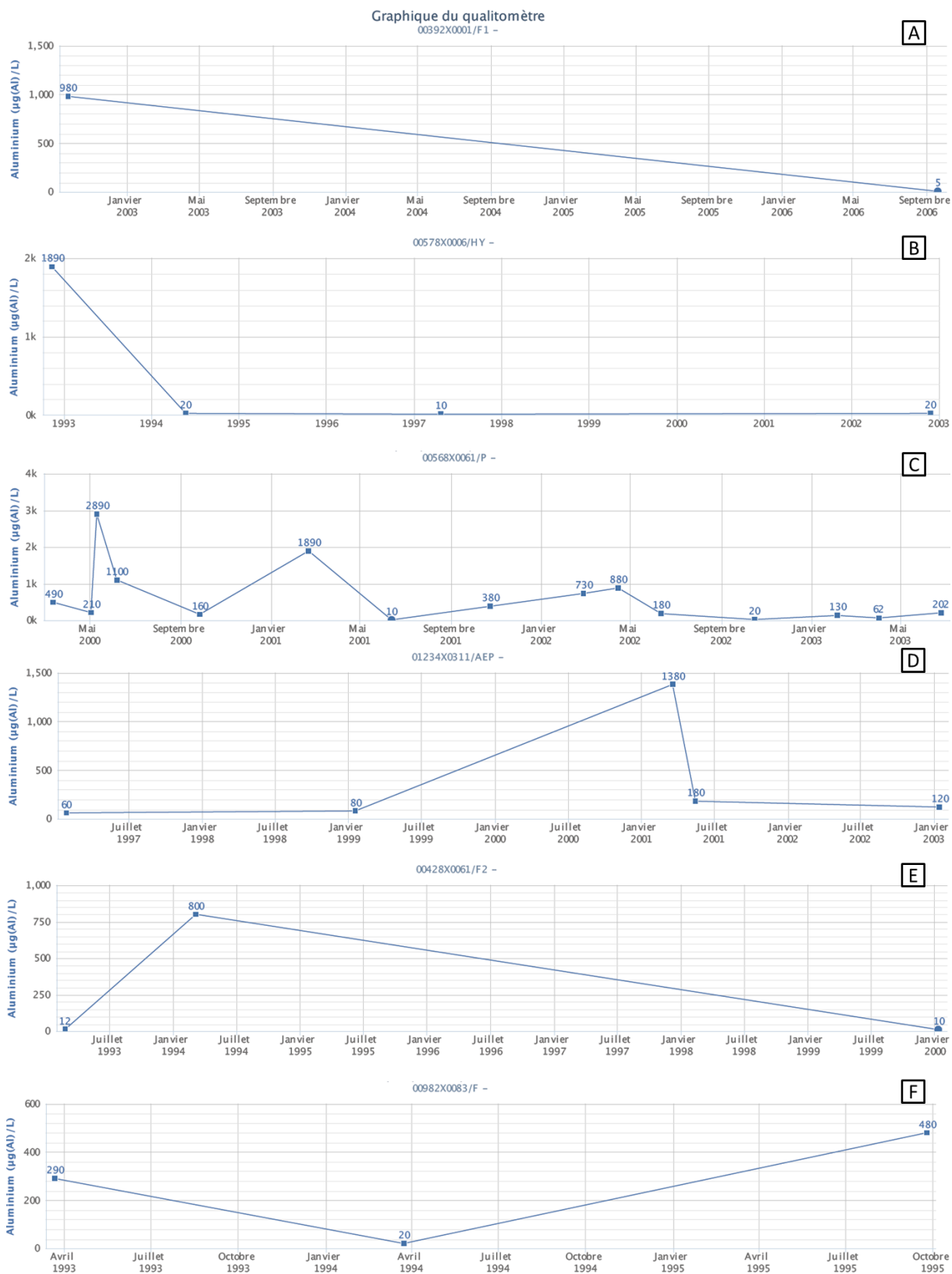


Figure 41. Suivi temporel des concentrations en aluminium dans les eaux souterraines (ADES Eau France, 2014)

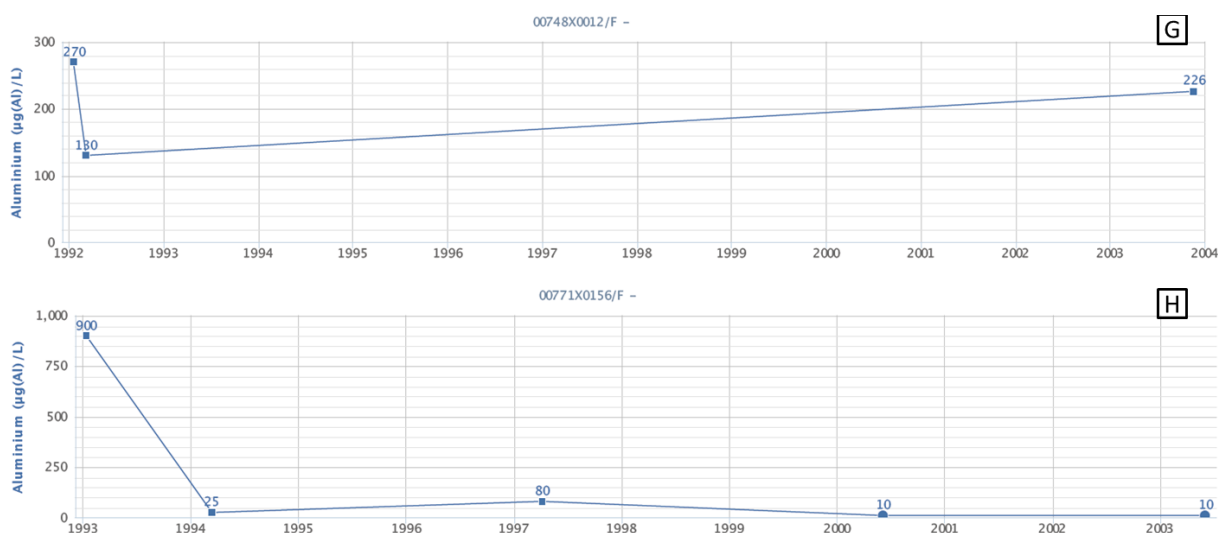


Figure 41 (suite). Suivi temporel des concentrations en aluminium dans les eaux souterraines (ADES Eau France, 2014)

ANALYSE SPATIO-TEMPORELLE

La variabilité spatiale et temporelle de l'échantillonnage et la présence de pollutions ponctuelles observées précédemment nous ont mené à réaliser des cartographies spatio-temporelles de la contamination en aluminium. La suite de notre propos porte sur l'atlas de 22 cartographies de concentration moyenne annuelle en aluminium (annexe III), dont un extrait est présenté ici (figure 42). Sur ces cartes, les concentrations ont été discrétisées selon 3 classes définies de façon manuelle (0-100 µg/L ; 101-200 µg/L ; >200 µg/L). Ces cartographies nous permettent d'isoler les épisodes ponctuels de pollution, représentés par les points de couleur rouge.

En 1990 et 1991, les points d'eau faisant l'objet d'un suivi sont localisés sur les départements du Nord et du Pas de Calais, tandis qu'aucune campagne n'est menée sur les départements de la Somme et la Seine-Maritime. Les points concernés par un échantillonnage sont majoritairement situés dans la moitié Sud du département du Nord. Les concentrations mesurées sont toutes inférieures à 100 µg/L.

En 1992 et 1993, ces deux départements font également l'objet d'un suivi, un plus grand nombre de points est cette fois-ci observé sur le Pas de Calais. Les concentrations mesurées sont toujours inférieures à 100 µg/L. En plus de ces deux territoires, la Seine-Maritime fait aussi l'objet d'un suivi, contrairement à la Somme. Plusieurs points présentent des concentrations élevées, 7 entre 100 et 200 µg/L et 5 >200 µg/L. Parmi ces 5 points, 4 concernent les graphiques C, F, G et H de la figure 41 précédente.

Entre 1994 et 1996, des points d'eau apparaissent sur les 4 territoires, mais très peu concernent la Somme. A l'instar de la période précédente, nous retrouvons des concentrations élevées sur la Seine-Maritime, avec notamment 9 mesures supérieures à 200 µg/L en 1994, 4 en 1995 et 3 en 1996. A ce titre, nous constatons un épisode de contamination chronique en aluminium dans les eaux souterraines situées à l'Est du Havre entre 1992 et 1996. Concernant les autres départements, le seul point échantillonné en 1994 sur la Somme est situé entre 100 et 200 µg/L. Pour le Nord et le Pas de Calais, on observe une mesure >200 µg/L en 1994 à l'Ouest d'Arras et 2 mesures moyennes très éloignées en 1996.

A partir de 1997, l'échantillonnage devient plus dense sur l'ensemble des départements. Ce constat est valable jusque 2003, avec une légère diminution de la concentration de points pour les deux dernières années, notamment sur le département du Nord. Entre 1997 et 1999, les concentrations

mesurées sur la Seine-Maritime sont globalement plus faibles que pour la période précédente. Des concentrations élevées sont retrouvées pour les années 2000 (4 points), 2001 (3 points), 2002 (2 points) et 2003 (2 points). Pour l'un de ces points, situé au centre du territoire, la contamination persiste entre 2000 et 2003. Des concentrations élevées sont toujours retrouvées à l'Est du Havre. Sur les autres départements, les niveaux de concentration en aluminium restent très faibles. Seuls de rares épisodes ponctuels sont observés, au centre du Pas de Calais en 1997, à proximité de Cambrai en 2001 et à proximité d'Avesnes-sur-Helpe en 2002. Ce dernier correspond au graphique A de la figure 41.

De 2004 à 2007, en dehors du Pas de Calais, nous constatons un manque d'échantillons sur les 3 autres départements. Pour le premier, les concentrations sont toujours largement en dessous de la norme. Sur cette période, seuls 2 points dépassent les 200 µg/L en 2005 sur la Seine-Maritime, dont l'un est encore situé à l'Est du Havre.

A partir de 2008 et jusque la fin de la période d'étude, le nombre de points échantillonnés est très faible sur la zone. Seul un point montre une concentration élevée en 2011 au centre de la Seine-Maritime.

d. Conclusion

Les résultats de cette étude du comportement spatio-temporel de la contamination des eaux souterraines par l'aluminium confirment dans un premier temps une surveillance très irrégulière de la qualité des eaux souterraines en France. Le faible nombre de mesures dépassant la valeur de la norme environnementale, ainsi que leur distribution spatiale et temporelle témoignent ensuite du comportement ponctuel de cette pollution. Le calcul d'une concentration moyenne en aluminium sur une longue période ne semble pas représentatif d'une pollution de fond.

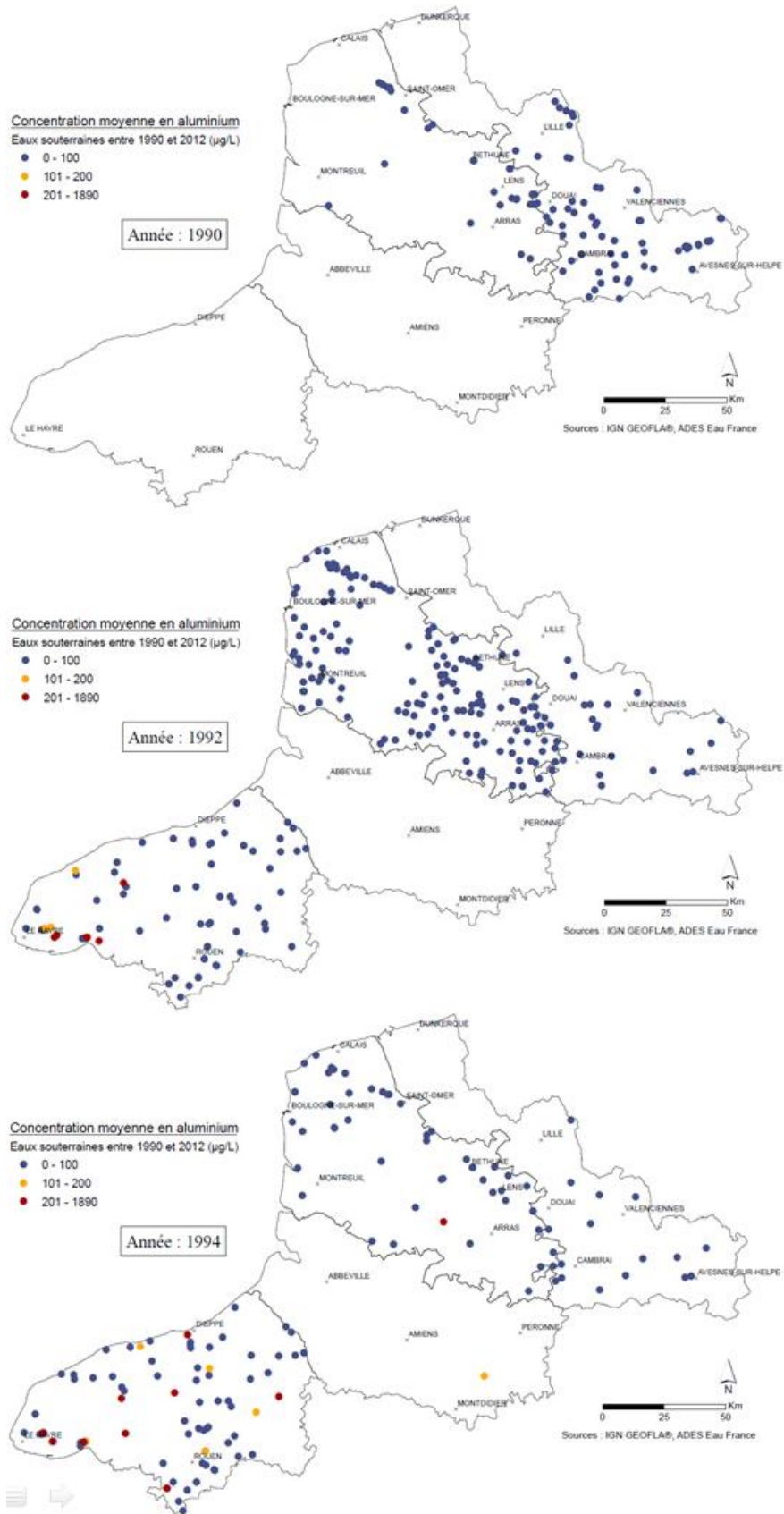


Figure 42. Variabilité spatiale et temporelle de l'échantillonnage et des concentrations en aluminium dans les eaux souterraines (suite)

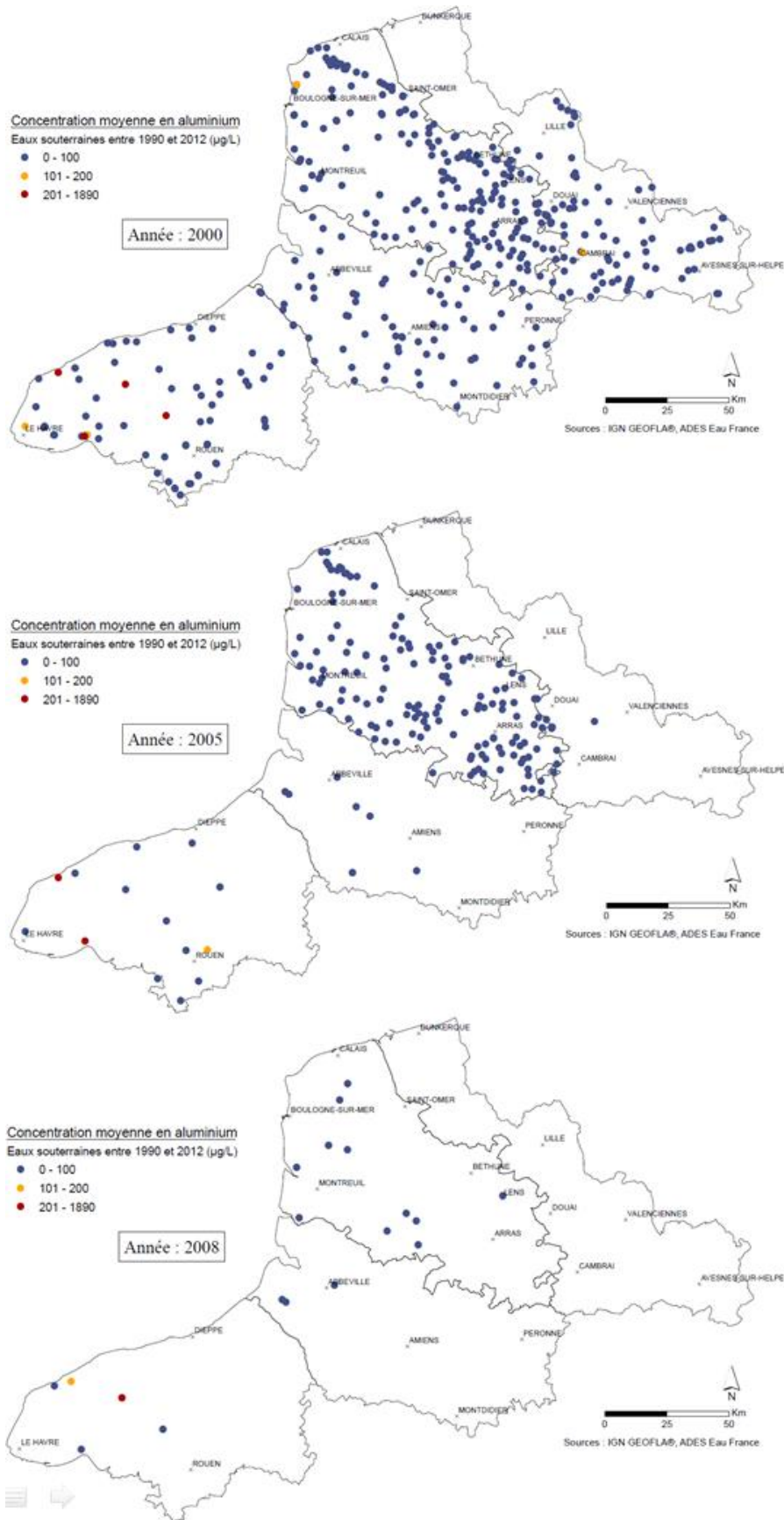


Figure 42 (suite). Variabilité spatiale et temporelle de l'échantillonnage et des concentrations en aluminium dans les eaux souterraines (suite)

Partie 4. ADES comme outil d'aide à la décision en santé – environnement dans le contexte des MICI

a. Introduction

Dans les parties précédentes, nous avons effectué l'analyse critique d'une base de données de qualité des eaux souterraines téléchargée à partir du portail ADES, en vue de réaliser une analyse spatio-temporelle de la contamination en aluminium. Sur la base des résultats obtenus, cette dernière partie se consacre à la mise en relation du résultat cartographique aux données d'incidence des MICI.

Pour rappel, le registre EPIMAD recense les cas de MICI depuis 1988 dans le Nord de la France, au sein de quatre départements : le Nord, le Pas de Calais, la Somme et la Seine-Maritime (Gower-Rousseau *et al.*, 1994). Les données issues du registre ont permis d'étudier les variations spatiales d'incidence de la Maladie de Crohn (MC) et ont déjà fait l'objet de plusieurs analyses spatiales et spatio-temporelles, démontrant non seulement des variations d'incidence entre les cantons des quatre départements pour les patients recensés entre 1990 et 2006 (Declercq *et al.*, 2010), mais aussi la présence de clusters spatiaux et spatio-temporels de surincidence et de sous-incidence (Génin *et al.*, 2013).

Ainsi notre recherche s'intéresse à l'utilisation des données de la banque ADES en tant qu'outil d'aide à la décision en santé – environnement, notamment pour la recherche de facteurs étiologiques dans le contexte de la MC. Les facteurs de risques environnementaux de la MC étant peu identifiés, la puissance des outils SIG est ici mise en œuvre pour vérifier l'influence d'un facteur suspecté (l'aluminium) mais aussi pour identifier d'autres pistes (la contamination multimétallique).

Une analyse de clusters spatiaux (version isotonique de la statistique de scan) et spatio-temporels de la MC par l'équipe du Pr. Duhamel (EA2694 Université Lille 2) est actuellement en cours. Celle-ci intégrera les cas incidents de 1990 à 2012 sur les 4 départements. En attendant les résultats de ces analyses, nous avons utilisé les données publiées.

b. Matériel et méthodes

Les données concernant la MC sont les résultats cartographiques des analyses spatio-temporelles de clusters (Génin *et al.*, 2013). Celles-ci ont révélé la présence de 14 clusters spatiaux significatifs sur l'ensemble de la période d'étude (six concernent des zones de sous-incidence et huit des zones de surincidence) et quatre clusters spatiaux significatifs sur une partie de la période d'étude (un de sous-incidence et trois de surincidence). Les données environnementales sont les résultats d'analyses en aluminium dans les eaux souterraines. Celles-ci sont issues de la banque ADES et ont précédemment fait l'objet d'une optimisation pour la création d'une cartographie géostatistique, représentant la concentration moyenne observée entre 1990 et 2012. Lors de ce travail, ces deux jeux de données ont été superposés dans le SIG (ArcGIS® v10.1) pour une interprétation visuelle concomitante. La mise en relation statistique fera l'objet d'une étude écologique future, intégrant les clusters de MC détectés sur la période 1990 – 2012.

A l'instar du Chapitre 1, un indice intégré a ensuite été utilisé pour décrire la contamination multimétallique des eaux souterraines. Le Ratio d'Imprégnation Moyen (RIM) a été calculé à partir de l'aluminium et des huit métaux les plus renseignés dans la banque ADES et faisant partie des substances prioritaires de la qualité des eaux (arsenic, cadmium, chrome, cuivre, mercure, nickel, plomb et zinc). Les données concernant ces ETM ont été exportées, puis contrôlées et optimisées de la même manière que pour l'aluminium. Les normes environnementales utilisées comme critère de décision sont la limite de qualité des eaux brutes utilisées pour la production d'eau destinées à la consommation humaine ou la valeur seuil des eaux souterraines lorsque cette dernière n'existe pas (tableau 27).

Tableau 27. Normes environnementales utilisées pour le calcul du RIM eaux souterraines (Arrêté du 11/01/07)

Élément	Limite de qualité des eaux brutes utilisées pour la production d'eau destinées à la consommation humaine	Valeur seuil des eaux souterraines
Aluminium		200 µg/L
Arsenic	100 µg/L	10 µg/L
Cadmium	5 µg/L	5 µg/L
Chrome	50 µg/L	50 µg/L
Cuivre		2 mg/L
Mercure	1 µg/L	1 µg/L
Nickel		20 µg/L
Plomb	50 µg/L	10 µg/L
Zinc	5 mg/L	5 mg/L

La concentration moyenne de chaque élément sur la période 1990-2012 a enfin été calculée au niveau des points d'eau, afin de réaliser une analyse spatiale. L'analyse spatio-temporelle de la contamination multimétallique nécessite une phase de traitement beaucoup plus longue, qui sera mise en œuvre sur un plus long terme. Pour rappel, le RIM est un indice d'enrichissement du milieu, représentant le niveau de contamination par rapport aux normes environnementales. Les normes présentées ci-dessus ont également été utilisées comme valeur de comparaison pour la construction de l'indice. Pour chaque élément dosé en un point d'eau, le ratio de la concentration mesurée (C_i) sur la valeur de référence (B_i) est dans un premier temps calculé. La moyenne de ces 9 ratios représente le RIM. Comme pour l'aluminium, la cartographie géostatistique de cet indicateur a été générée et comparée aux clusters de MC.

$$RIM = \frac{\sum C_i/B_i}{9}$$

c. Résultats

La figure 43 montre les cartographies de clusters de MC, de concentrations moyennes en aluminium, ainsi que le résultat de leur superposition. Nous pouvons constater que les zones présentant des pics de concentrations ne correspondent pas aux clusters de surincidence de MC. La Seine-Maritime présente par exemple des niveaux de concentrations en aluminium régulièrement élevés au niveau d'un cluster de sous-incidence de la MC, significatif sur l'ensemble de la période d'étude. Lorsque l'on s'intéresse aux clusters spatio-temporels de MC, la variabilité des forces d'échantillonnage en eau souterraine complique la mise en relation des données. Pour le cluster de sous-incidence au nord de la zone (significatif sur la période 1990-1995), nous avons vu qu'aucun point d'eau n'est échantillonné sur cette zone en raison de la nappe affleurante. Les points situés à proximité montrent de faibles concentrations pendant cette période (annexe III). Pour le cluster de surincidence situé sur la zone de Montreuil (significatif pour la période 1996-2006), la plupart des analyses indiquent des concentrations inférieures à 100 µg/L. Quelques points situés au nord de Boulogne-sur-Mer et à l'Est de ce cluster sont sujets à des pics de concentrations pendant cette période. Pour le cluster de surincidence de la région d'Amiens (significatif pour la période 1994-2005), aucun pic de concentration n'est observé. Finalement, le cluster de surincidence situé au sud de la région Nord – Pas de Calais (significatif pendant la période 1992-2004) est concerné par deux pics de concentrations en aluminium. Le premier est observé en 1996 tout au sud de la région (150 µg/L) et le second en 2000 et 2001 à proximité de Cambrai (128 et 230 µg/L).

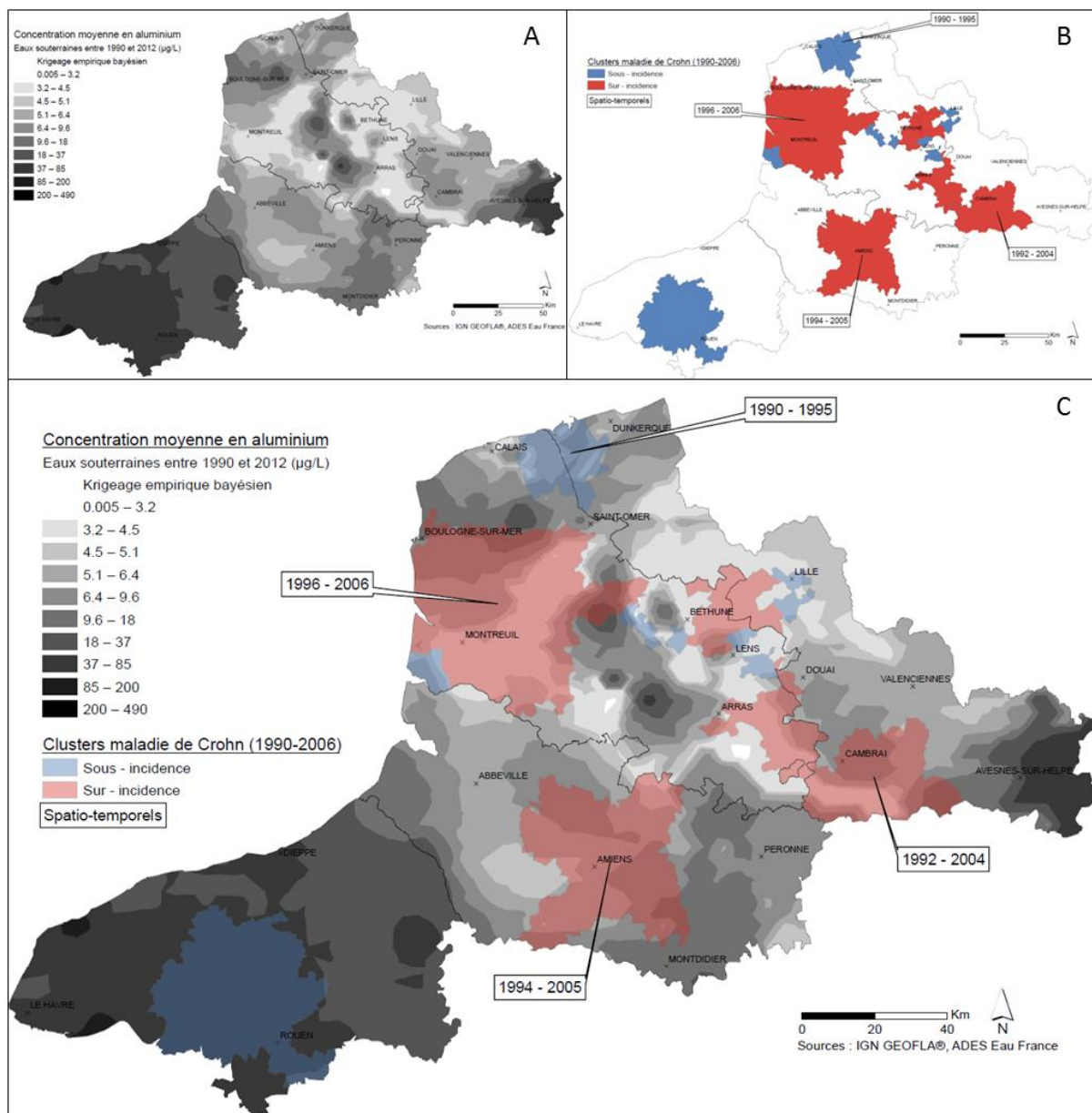


Figure 43. Concentration moyenne en aluminium par point d'eau ($\mu\text{g/L}$) entre 1990 et 2012 (A), clusters de MC entre 1990 et 2006 (B : repris à partir de Génin *et al.*, 2013) et superposition des deux cartes (C)

La superposition de ces données environnementales et sanitaires ne montre pour le moment aucun lien entre l'incidence de la MC et la contamination des eaux souterraines par l'aluminium. Dans le Chapitre I, nous avons souligné l'apport des indices intégrés pour la mise en évidence de liens entre les effets sur la santé et la contamination globale de l'environnement. Afin d'étudier l'influence sur les MICI de l'état des eaux souterraines en termes de pollution globale, nous avons utilisé le RIM, calculé à partir de neuf éléments traces. La figure 44 montre la superposition des cartographies de RIM (1990-2012) et de clusters spatiaux et spatio-temporels de MC (1990-2006), repris à partir de Génin *et al.* (2013). A l'instar de l'aluminium, l'imprégnation multimétallique des eaux souterraines est plus importante sur le département de la Seine-Maritime et à l'est de la région Nord – Pas de Calais. Des concentrations plus faibles sont retrouvées sur le bassin lillois, ainsi qu'au sud d'Arras. Les zones de surincidence de MC ne correspondent pas aux zones de contamination multimétallique.

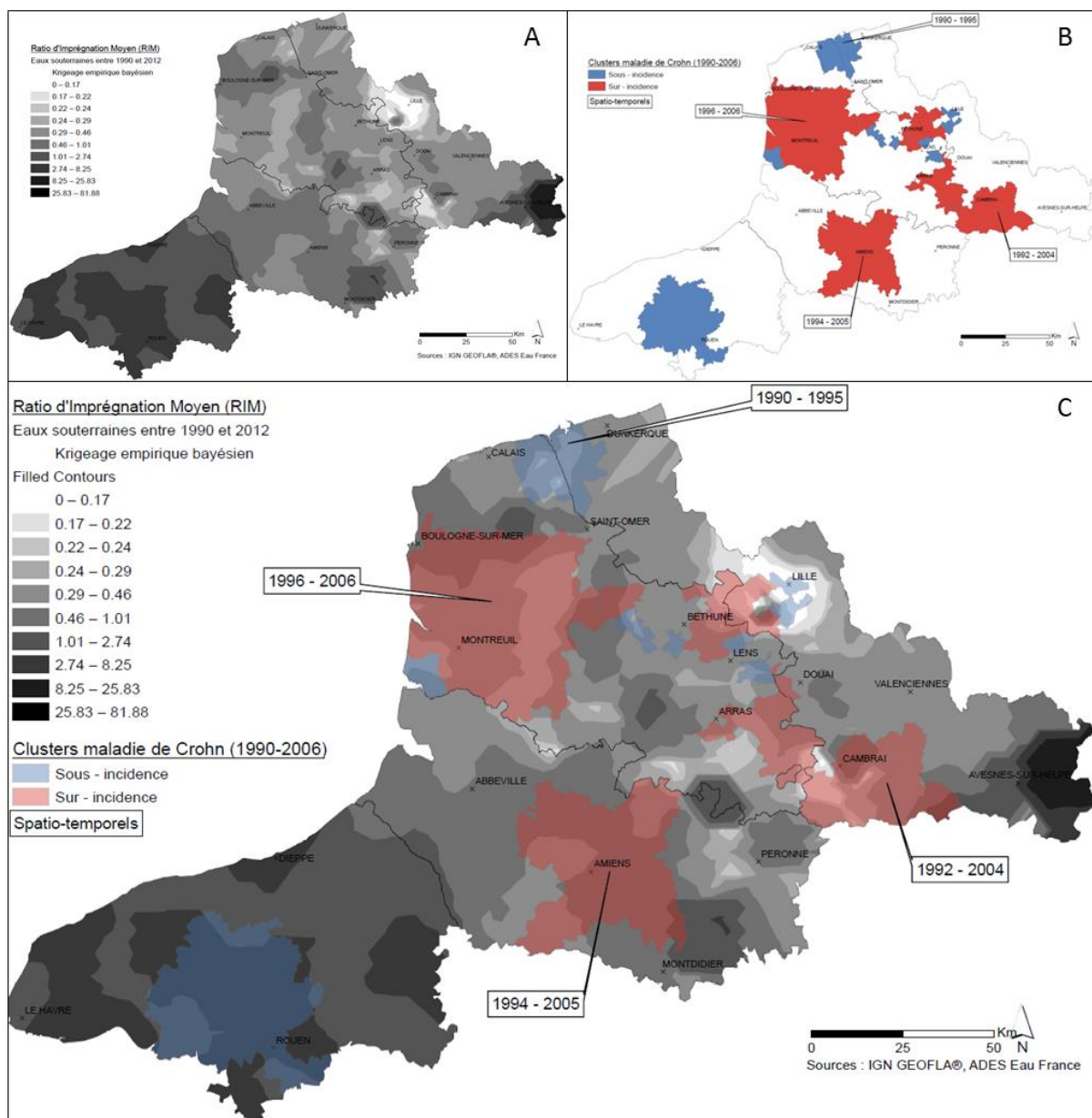


Figure 44. RIM par point d'eau entre 1990 et 2012 (A), clusters de MC entre 1990 et 2006 (B : repris à partir de Génin *et al.*, 2013) et superposition des deux cartes (C)

d. Conclusion

Bien que les cartographies des clusters spatiaux et spatio-temporels de la MC soient en cours de réactualisation, ces résultats ne permettent pas à ce jour d'étayer l'hypothèse de l'aluminium dans les eaux souterraines comme facteur explicatif d'incidence de cette maladie. La contamination multimétallique des eaux souterraines ne semble pas non plus être un facteur explicatif.

Partie 5. Discussion générale et perspectives du chapitre III

a. La banque ADES comme outil d'aide à la décision en environnement

La banque ADES est une source considérable d'informations concernant l'état quantitatif et qualitatif des masses d'eau souterraines (ADES Eau France, 2014). Parmi les bases de données environnementales publiques, elle constitue certainement la plus complète pour caractériser la qualité du milieu, de par sa transparence et son accès libre. Les données communiquées sont directement exportables suivant une arborescence de critères de sélection, comme la localisation géographique, l'entité hydrogéologique, la masse d'eau, le code BSS, le réseau de surveillance, l'organisme producteur ou le paramètre mesuré. Cet export est réalisé sous format informatique, ce qui facilite la mise en forme de la base, ainsi que son traitement statistique et cartographique. Bien que les coordonnées géographiques précises des points de captage ne soient pas communiquées au sein du portail pour des raisons de sécurité nationale, elles peuvent être retrouvées par ailleurs, dans la banque du sous-sol (InfoTerre, 2014). Concernant les mesures environnementales, la diffusion publique des coordonnées spatiales de points d'échantillonnage est peu commune, en raison du secret statistique (INSEE, 2010).

Le véritable atout de cette base de données est la transmission des données à l'état brut, c'est-à-dire l'ensemble des substances mesurées en un lieu et à une date précise. Contrairement à la plupart des bases de données environnementales publiques, aucun prétraitement statistique (communication d'une gamme de concentrations, plutôt que la valeur brute mesurée) ou d'agrégation (moyenne sur un pas de temps ou sur une zone géographique) n'est effectué pour la communication des données (Algros & Charissou, 2010). Pour chaque analyse, la banque recense également les informations nécessaires à l'identification du point d'eau (exploitant, mode d'utilisation, etc.) et identifie l'appréciation qualitative du résultat (réseau de surveillance, organisme responsable de l'analyse, méthode d'analyse, statut de la mesure, limites de détection et quantification, etc.). Finalement, elle procure des informations sur le long terme, à partir de 1900, ce qui concerne peu de bases de données environnementales.

Cette transparence de l'état qualitatif brut des eaux souterraines est donc un point fort, mais peut très vite devenir une source d'erreurs lors du traitement et de l'interprétation de ces données. En effet, l'analyse critique menée dans cette étude en vue de développer un outil cartographique, nous a permis d'identifier plusieurs obstacles. Nous nous sommes ici intéressés aux données brutes issues des analyses de concentrations en aluminium, effectuées entre 1990 et 2012 au niveau des points de captage de quatre départements : le Nord, le Pas de Calais, la Somme et la Seine – Maritime.

Le premier obstacle à l'interprétation de cette base est l'absence de métadonnées. Cela complique la compréhension des différentes variables et des valeurs renseignées. Une recherche approfondie sur le portail peut néanmoins apporter certains compléments d'information. Bien que la qualité des eaux souterraines ne soit pas concernée par les thèmes de la Directive INSPIRE (élaborée par la Direction générale de l'environnement de la Commission européenne en 2007), qui impose la communication de métadonnées en annexe des données géographiques numériques, une telle base de données devrait proposer un document d'aide.

La multitude des acteurs et professionnels de l'eau autorisés à alimenter la banque ADES peut ensuite occasionner des incertitudes quant à la fiabilité des données. Ces données peuvent avoir différents niveaux de validité (couple statut/qualification), depuis une donnée brute jusqu'à une donnée élaborée et contrôlée. Nous avons en effet remarqué que le niveau de contrôle des résultats d'analyses peut être hétérogène, puisque certains résultats sont communiqués sans être vérifiés par le producteur. Cela concerne néanmoins une faible part des analyses (moins de 1%). Par ailleurs, aucune valeur considérée comme incorrecte n'apparaît dans la base exportée. L'intégralité des analyses chimiques est de plus réalisée par des organismes accrédités, ce qui est un gage de qualité (Arrêté du 19/09/11 fixant la liste des laboratoires agréés par le ministère chargé de la santé pour la réalisation des prélèvements et des analyses du contrôle sanitaire des eaux). Nous pouvons

néanmoins nous interroger sur cette hétérogénéité et le manque de gestion coordonnée des entrées dans la base. Il paraît nécessaire d'établir des critères de validité d'entrée des données.

L'indicateur le plus contraignant en termes d'analyse et d'interprétation est la variabilité des limites de détection et de quantification des appareils de mesure. En effet, plus des ⅔ des analyses sont concernées par des valeurs inférieures aux LDD ou LDQ. Nous avons à ce titre remarqué que les performances analytiques varient en fonction des organismes missionnés pour la réalisation des analyses. Cette diversité des méthodes analytiques utilisées entraîne une forte hétérogénéité de précision, de l'ordre de 10^5 (celle-ci passe à 10^2 lorsque les analyses issues du laboratoire de la Somme ne sont pas prises en compte). Des limites de détections extrêmes ont par exemple été retrouvées au sein du département de la Somme entre 1994 et 1996 (10 000 $\mu\text{g/L}$). Les valeurs concernées sont en intégralité issues de campagnes menées dans le cadre du réseau SISE-Eaux (Note 02.07.10 SISE Eaux). Ce réseau dépend du ministère chargé de la Santé, il représente le contrôle sanitaire des eaux souterraines captées pour la production d'eau potable. Parmi les résultats inférieurs aux LDD ou LDQ, moins de 1% sont susceptibles de dépasser la norme sanitaire de concentration en aluminium de 200 $\mu\text{g/L}$ (limite de qualité des eaux brutes utilisées pour la production d'eau destinées à la consommation humaine : Arrêté du 11/01/07). Ce résultat témoigne d'un milieu peu contaminé en aluminium. Afin d'homogénéiser les techniques analytiques, un arrêté ministériel fixe depuis 2003 les performances des méthodes d'analyses appliquées dans le cadre de la surveillance de la qualité des eaux (Arrêté du 17/09/03). Pour l'aluminium, les LDD et LDQ sont respectivement de 10 et 30 $\mu\text{g/L}$.

Des valeurs très élevées sont aussi présentes dans la base. Même si ces résultats sont issus de méthodes analytiques validées, des incertitudes telles que les erreurs de saisies peuvent être envisagées. Un test de valeurs extrêmes a donc été employé afin d'identifier et retirer les valeurs considérées comme potentiellement aberrantes.

Suite à l'ensemble de ces constats, nous proposons un arbre de décision pour l'utilisation et l'interprétation des données de qualimètres de la banque ADES dans le cadre d'une analyse spatiale et spatio-temporelle (figure 45). Cet arbre est basé sur cinq critères :

- la portée de l'étude ;
- les limites de détection et de quantifications ;
- le couple de statut et qualification de la mesure ;
- le résultat de la mesure ;
- la variabilité temporelle du résultat ;
- le suivi des points d'eau.

Selon ces critères, 33.2% des 8525 analyses initiales n'ont pas été retenues pour la suite de notre recherche. Le premier critère constitue le plus sélectif, avec 30.2% des analyses non retenues. Les autres rejets concernent essentiellement des valeurs absentes ou statistiquement aberrantes.

Suite à l'optimisation de ces données, nous avons réalisé une analyse spatiale et spatio-temporelle de la contamination en aluminium. Notre recherche montre dans un premier temps une forte hétérogénéité du suivi de la qualité des eaux souterraines dans l'espace et le temps. En France, le suivi de la qualité des eaux souterraines est programmé depuis 2010 par les schémas directeurs d'aménagement et de gestion des eaux (SDAGE) à l'échelle de six bassins hydrographiques. Ces plans de gestion français de la Directive Cadre Eau européenne (DCE, 2000) définissent pour chaque bassin la densité minimale de points d'eau à suivre en fonction du type de masse d'eau et le nombre d'échantillons à analyser par année en fonction de l'importance des substances recherchées (Arrêté du 25/01/10). Nous disposerons ainsi d'un échantillonnage homogène dans l'espace et le temps. La figure 46 montre par exemple les points d'eau concernés pour le bassin Artois – Picardie. Au total, 186 points de mesure sont suivis, avec au minimum deux prélèvements annuels⁹.

⁹ L'Agence de l'Eau responsable du bassin hydrographique peut éventuellement compléter ces suivis en finançant des réseaux complémentaires départementaux, dont les programmes sont définis annuellement avec les Agences Régionales de Santé.

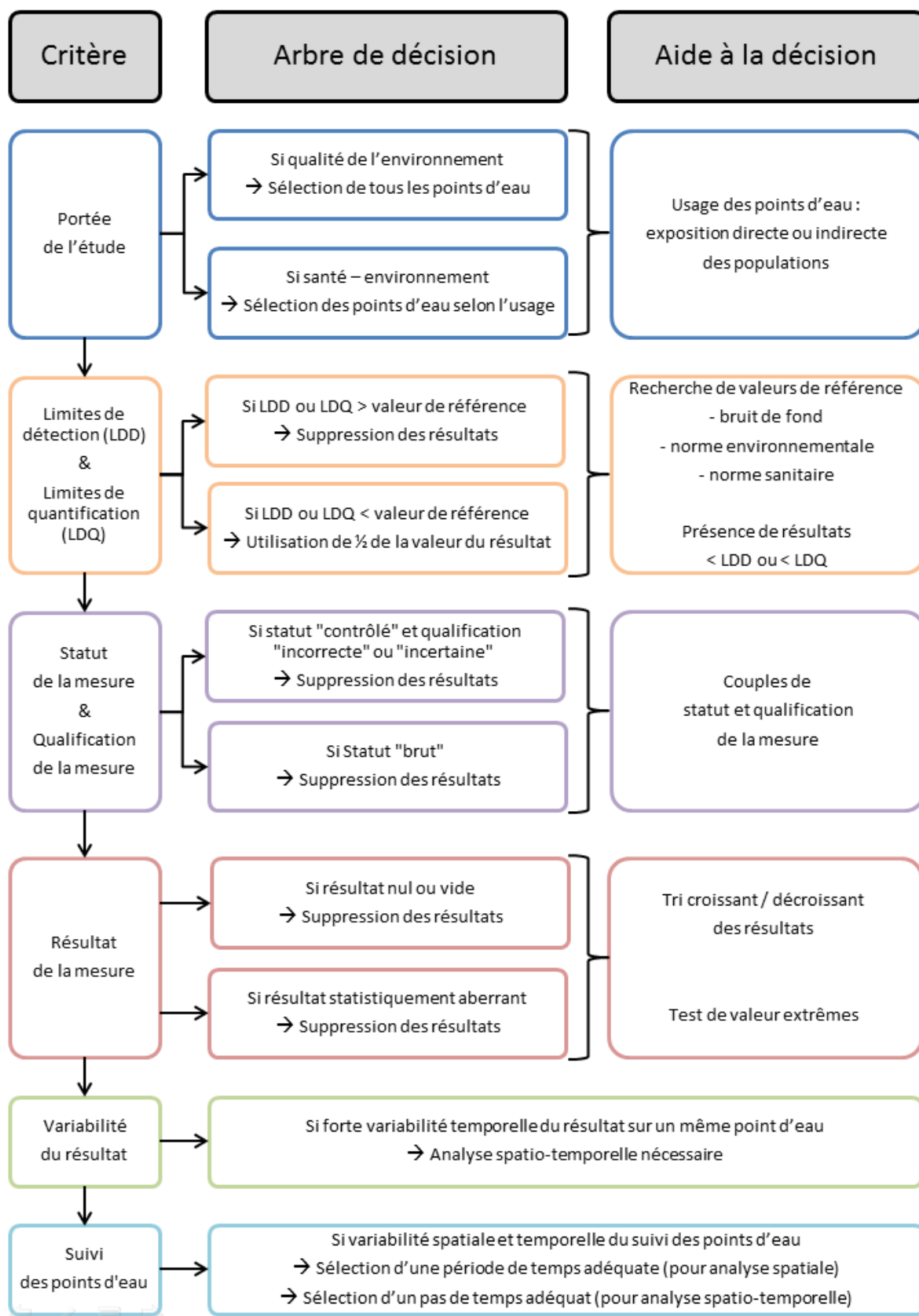


Figure 45. Arbre de décision pour l'utilisation et l'interprétation des données de qualimètres de la banque ADES dans le cadre d'une analyse spatiale et spatio-temporelle

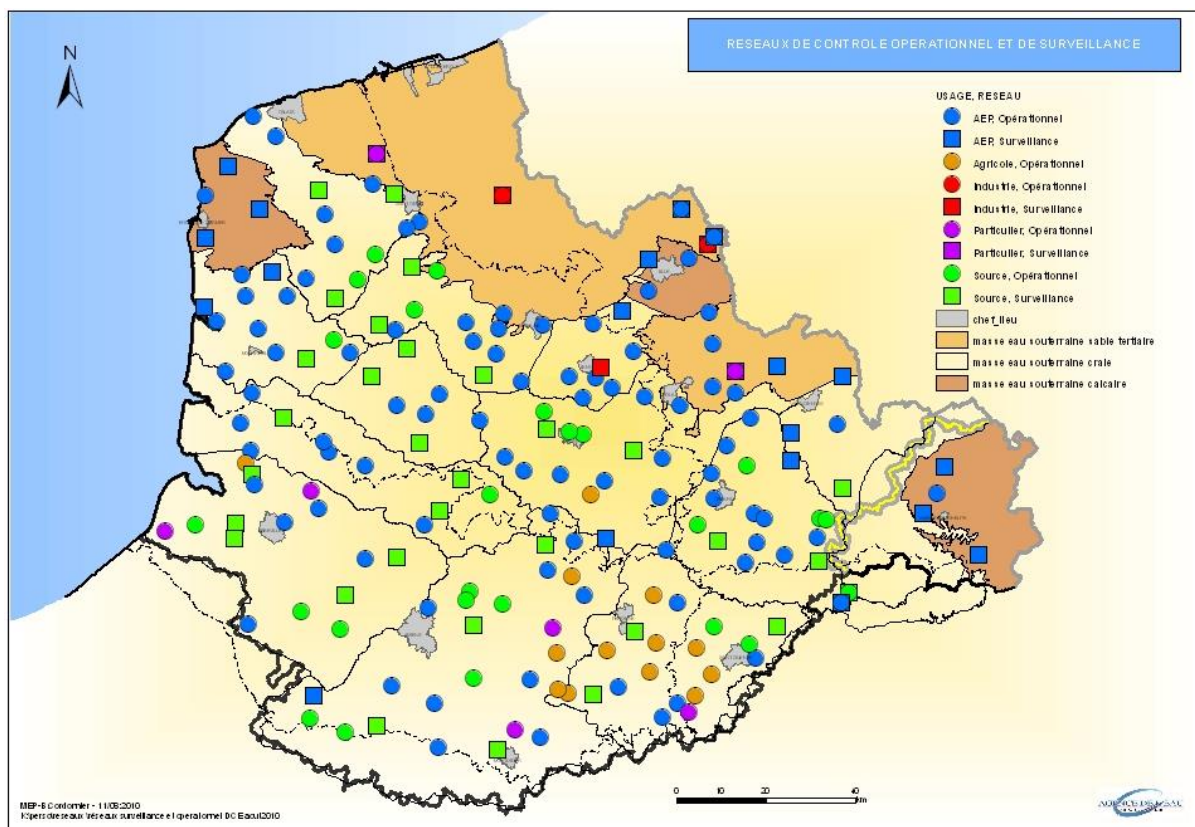


Figure 46. Réseaux de contrôle opérationnel et de surveillance des masses d'eau souterraines (www.artois-picardie.eaufrance.fr)

Les analyses spatiales et temporelles de la concentration en aluminium indiquent ensuite que le calcul d'une concentration moyenne basée sur plusieurs années est peu représentatif d'une pollution de fond, en raison de pics de contamination observés sur de courtes périodes. Une analyse spatio-temporelle est donc à privilégier. Celle-ci met finalement en évidence un niveau de contamination des eaux souterraines en aluminium globalement faible sur 3 départements (Nord, Pas de Calais, Somme), avec de rares épisodes ponctuels de pollution. Le territoire de la Seine-Maritime fait en revanche l'objet d'une contamination plus importante en aluminium. Un total de 42 points dépasse la norme sanitaire sur l'ensemble de la période et 48 points montrent des niveaux de concentrations entre 100 et 200 µg/L. Parmi ces points, 11 ayant une concentration >200 µg/L et 7 dont les concentrations sont entre 100 et 200 µg/L sont situés entre le port du Havre et la zone industrielle de Port-Jérôme (Tancarville). Cette zone géographique borde la rive nord de la Seine juste en amont de son embouchure. La masse d'eau souterraine est donc en affleurement avec la surface du sol, ce qui la rend très sensible aux contaminations. La zone industrielle héberge deux raffineries, ainsi que d'autres activités telles que la pétrochimie ou l'incinération d'ordures ménagères. Elle est une source d'émissions de nombreux contaminants environnementaux, notamment de métaux (ARS Haute-Normandie, 2010). Les points d'eau de cette zone géographique sont destinés à alimenter le réseau d'eau potable de la ville du Havre.

A l'image des indices intégrés de qualité des eaux présentés dans le chapitre d'Introduction générale, la Directive européenne du 12 décembre 2006 sur la protection des eaux souterraines contre la pollution et la détérioration propose une méthode d'évaluation de l'état des masses d'eau souterraine. Une masse d'eau est alors considérée comme étant en bon état chimique lorsque les paramètres suivis par le programme de surveillance ne dépassent en aucun point de cette masse d'eau les normes de qualité et les valeurs seuils pertinentes (contrôle de surveillance et contrôle opérationnel).

b. La banque ADES comme outil d'aide à la décision en santé – environnement

Dans le domaine de la santé environnementale, la banque ADES semble pouvoir constituer un outil d'aide à la décision, notamment dans le contexte de pathologies multifactorielles et chroniques, comme les MICI. Elle renseigne les résultats de dosages de plus de 450 substances chimiques sur l'ensemble du territoire métropolitain, suivant une dimension temporelle précise. Il est également possible d'effectuer une sélection des points d'eau étudiés en fonction de leur mode d'utilisation. Ainsi, nous avons pu isoler les points dont l'eau est destinée à la consommation humaine (eau potable) ou à l'agriculture (irrigation), ce qui peut constituer un vecteur d'exposition des populations.

Les premiers résultats d'analyse de la contamination des eaux souterraines et leur superposition aux clusters de MC confirment la difficulté d'établir un lien environnement – santé. En effet, aucune conclusion de notre recherche ne vient valider ou étayer l'hypothèse de l'aluminium comme facteur de risque environnemental de la MC. Il en est de même pour les suspicions retrouvées dans la littérature concernant le rôle des micropolluants minéraux (Beamish *et al.*, 2011 ; Lerner, 2007 ; 2012 ; Pineton de Chambrun *et al.*, 2014).

Néanmoins, les données renseignées dans la banque ADES retracent la qualité de l'eau brute et non celle de l'eau consommée par la population, alimentée par les réseaux de distribution d'eau potable. Il n'existe donc pas de lien direct entre une eau souterraine dégradée au niveau d'un point de captage et l'exposition des populations voisines à cette eau. En effet, les prélèvements réalisés dans le cadre de la contamination des eaux souterraines sont effectués directement à la sortie du captage. Les résultats ne tiennent pas compte des possibles traitements de ces eaux réalisés pour atteindre un bon état sanitaire. Des sels d'aluminium sont par exemple utilisés dans les procédés physico-chimiques de traitement pour éliminer les particules organiques en suspension dans l'eau. La dégradation de certaines masses d'eau impose également aux distributeurs le recours à des ressources plus lointaines et donc à des regroupements de production (mélanges, dilutions). Ainsi, le réseau d'eau potable d'une aire urbaine peut être alimenté par un mélange d'eaux issues de plusieurs points de captage, parfois distants de plusieurs dizaines de kilomètres. C'est le cas notamment de la métropole lilloise (figure 47). Autre exemple, le secteur des Flandres étant dépourvu de ressource en eau souterraine potable, celle-ci provient de la nappe de la craie grâce à des adductions de près de 50 km. Ces méthodes d'adduction et de distribution des eaux à partir des points de captage ne sont pas connues à grande échelle et évoluent régulièrement en fonction de l'état qualitatif des masses d'eau. Finalement, l'étude du profil temporel de cette contamination montre que la concentration moyenne calculée sur la base de 22 années n'est pas synonyme d'une contamination de fond. L'observation concomitante des données provenant de cette base avec des indicateurs de santé nécessite donc certaines précautions.

La difficulté d'établir un lien entre l'incidence de la MC et la contamination de l'environnement par l'aluminium peut également être liée à la multiplicité des sources d'exposition de l'Homme aux contaminants. Exley (2013) cible par exemple six autres sources d'exposition à l'aluminium (vaccins, produits cosmétiques, médicaments anti-acides, emballages alimentaires et ustensiles de cuisines, eau de distribution). L'environnement ne semble donc pas occuper une place prépondérante dans l'exposition globale des populations. Concernant les eaux de distribution, un contrôle sanitaire de la qualité de l'eau potable au robinet du consommateur est mis en place par le Ministère chargé de la santé et réalisé par les agences régionales de santé. L'accès aux résultats de ce contrôle est possible via un portail internet. Cet accès est cependant très limité, puisque les données sont disponibles à la commune, mais ne sont pas exportables à grande échelle sous forme de base de données. Une saisie manuelle est donc nécessaire. De plus, les micropolluants (organiques et minéraux) sont trop rarement renseignés.

A l'opposé des eaux souterraines, les populations sont plus directement exposées au compartiment atmosphérique, celui-ci étant notamment responsable de multiples impacts cardio-respiratoires (OMS, 2014). Concernant la MC, la pollution de l'air est depuis peu considérée comme source possible d'inflammations intestinales. Beamish *et al.* (2011) montrent par exemple les différentes

voies de transfert des polluants atmosphériques vers le système digestif. La clairance mucociliaire peut ainsi entraîner une ingestion des polluants suite à leur inhalation, tandis que les retombées atmosphériques particulières peuvent se transférer dans la chaîne alimentaire. Pour la suite de nos recherches concernant les facteurs de risques environnementaux de la MC, nous nous focaliserons donc sur la qualité de l'air.

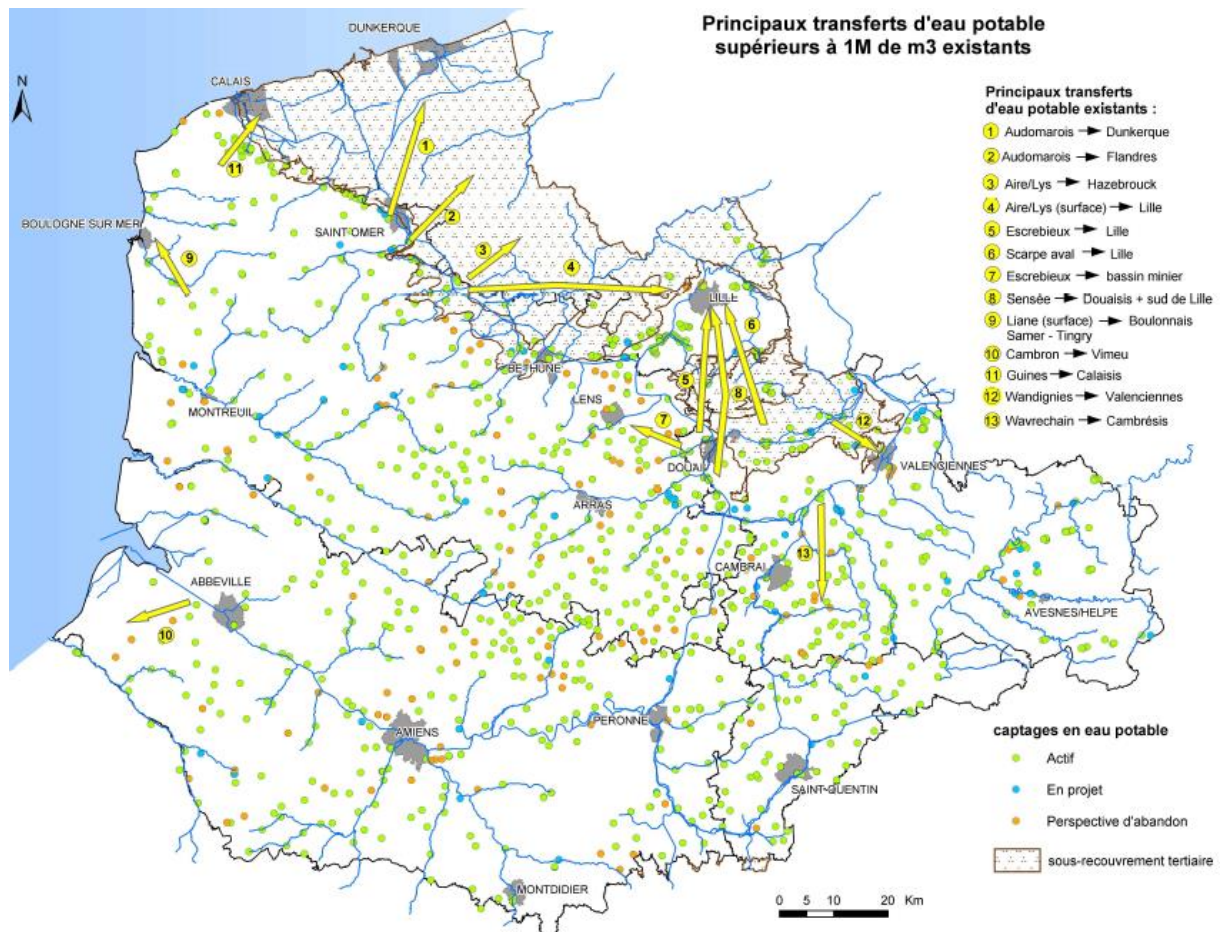


Figure 47. Principaux transferts d'eau potable sur le bassin Artois – Picardie (ADES Eau France, 2014)

c. Propositions pour la construction d'une base de données environnementale optimale

Pour conclure ce chapitre, si nous devons définir les caractéristiques d'une base de données idéale pour la réalisation d'une étude spatialisée en santé – environnement, cette base nécessiterait :

- de communiquer les informations brutes issues de l'analyse des milieux ;
- de préciser le niveau de fiabilité de chaque mesure ;
- de fournir les coordonnées spatiales des points d'échantillonnage ;
- de réaliser les analyses selon des méthodes normalisées et homogènes sur l'ensemble de la zone d'étude ;
- une force d'échantillonnage homogène dans l'espace et suivant un pas de temps régulier ;
- d'être exportable sous forme de table ;
- de présenter un fichier de métadonnées aidant l'utilisateur à comprendre les variables renseignées.

REFERENCES BIBLIOGRAPHIQUES

Algros E & Charissou AM. 2010. Réalisation d'un avis d'experts : « bruit de fond » des opluants dans l'environnement, analyse des données existantes et identification des lacunes. Etude RECORD N° 09-0812/1A, 163p.

Arrêté du 17/09/03 relatif aux méthodes d'analyse des échantillons d'eau et à leurs caractéristiques de performance.

Arrêté du 11/01/07 relatif aux limites et références de qualité des eaux brutes et des eaux destinées à la consommation humaine mentionnées aux articles R. 1321-2, R. 1321-3, R. 1321-7 et R. 1321-38 du code de la santé publique.

Arrêté du 25/01/10 établissant le programme de surveillance de l'état des eaux en application de l'article R. 212-22 du code de l'environnement.

Arrêté du 19/09/11 fixant la liste des laboratoires agréés par le ministère chargé de la santé pour la réalisation des prélèvements et des analyses du contrôle sanitaire des eaux.

ARS Haute-Normandie. 2010. Environnement & santé sur Port-Jérôme. Evaluation des risques sanitaires liés aux activités industrielles – résultats de l'étude réalisée de 2007 à 2010. 8p.

Beamish L, Osornio-Vargas A, Wine E. 2011. Air pollution: an environmental factor contributing to intestinal disease. *Journal of Crohn's and Colitis*, 5: 279–286.

Birrenbach T & Bocker U. 2004. Inflammatory bowel disease and smoking: a review of epidemiology, pathophysiology, and therapeutic implications. *Inflamm Bowel Dis*, 10: 848–59.

Chery L & Cattan A. 2003. ADES : la banque de données nationale sur les eaux souterraines. *La houille blanche*, 2: 5p.

Chouraki V, Savoye G, Dauchet L *et al.* 2011. The changing pattern of Crohn's disease incidence in northern France: a continuing increase in the 10- to 19-year-old age bracket (1988–2007). *Aliment Pharmacol Ther*, 33 (10): 1133–1142.

Cosnes J. 2004. Tobacco and IBD: relevance in the understanding of disease mechanisms and clinical practice. *Best Pract Res Clin Gastroenterol*, 18: 481–496.

Declercq C, Gower-Rousseau C, Vernier-Massouille G *et al.* 2010. Mapping of Inflammatory Bowel Disease in Northern France: Spatial Variations and Relation to Affluence. *Inflamm Bowel Dis* Volume, 16 (5): 807-812.

Directive Cadre sur l'Eau. Directive 2000/60/CE du Parlement européen et du Conseil établissant un cadre pour une politique communautaire dans le domaine de l'eau.

Directive 2006/118/CE du 12 décembre 2006 sur la protection des eaux souterraines contre la pollution et la détérioration.

Exley C. 2013. Human exposure to aluminium. *Environmental Science: Processes & Impacts*, 15: 1807-1816.

Génin M, Duhamel A, Preda C, Fumery M, Savoye G *et al.* 2013. Space-time clusters of Crohn's disease in northern France. *J Public Health*, DOI 10.1007/s10389-013-0580-9.

Gower-Rousseau C, Salomez JL, Dupas JL *et al.* 1994. Incidence of inflammatory bowel disease in northern France (1988-1990). *Gut*, 35 (10): 1433-8.

Hampe J, Cuthbert A, Croucher P, Mirza M, Mascheretti S *et al.* 2001. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *The Lancet*, 357 (9272): 1925 – 1928.

INSEE. 2010. Guide du secret statistique. 26p.

Jantchou P, Monnet E, Carbonnel F. 2006. Les facteurs d'environnement dans la maladie de Crohn et la rectocolite hémorragique (tabac et appendicectomie exclus). *Gastroenterol Clin Biol*, 30: 859-867.

Krivoruchko K. 2012. Empirical Bayesian Kriging implemented in ArcGIS Geostatistical Analyst. *ArcUser*, 15 (4): 6-10.

Lerner, A. 2007. Aluminum is a potential environmental factor for Crohn's disease induction: extended hypothesis. *Ann NY Acad Sci*, 1107: 329–345.

Lerner A. 2012. Aluminum as an adjuvant in Crohn's disease induction. *Lupus Volume*, 21 (2): 231-238.

Nerich V, Monnet E, Etienne A *et al.* 2006. Geographical Variations of Inflammatory Bowel Disease in France: A Study Based on National Health Insurance Data. *Inflamm Bowel Dis*, 12 (3): 218-226.

Note de service DGS/EA4 n° 2010-241 du 02/07/10 relative aux données du contrôle sanitaire des eaux destinées à la consommation humaine du système d'information en santé-environnement sur les eaux (SISE-Eaux).

OMS. 2014. Burden of disease from the joint effects of Household and Ambient Air Pollution for 2012. Geneva.

Pineton de Chambrun G, Body-Malapel M, Frey-Wagner I, Djouina M, Deknuydt F *et al.* 2014. Aluminum enhances inflammation and decreases mucosal healing in experimental colitis in mice. *Mucosal Immunology*, 7 (3): 589-601.

Spehlmann ME, Begun AZ, Burghardt J, Lepage P, Raedler A, Schreiber S. 2008. Epidemiology of inflammatory bowel disease in a German twin cohort: results of a nationwide study. *Inflamm Bowel Dis*, 14 (7): 968-76.

Sites internet :

ADES : www.ades.eaufrance.fr (consulté le 22.04.2014).

Artois-Picardie : www.eau-artois-picardie.fr (consulté le 22.04.2014).

InfoTerre: www.infoterre.brgm.fr (consulté le 27.01.2014).

DISCUSSION GENERALE

Ces travaux de recherche portent sur les inégalités environnementales et sociales de santé. Ils ont pour objectifs de caractériser la qualité des milieux de vie des populations grâce aux bases de données environnementales disponibles et aux outils SIG, puis de rapprocher les cartographies générées avec des indicateurs sanitaires et socio-économiques par des études écologiques. Ainsi, en intégrant les paramètres sociodémographiques, nous avons cherché à évaluer si les populations ayant un état de santé dégradé vivent dans un environnement dégradé. Ce questionnement a été étudié à travers trois axes de recherche. Ce dernier chapitre propose de discuter les travaux effectués dans la thèse. Après avoir retracé l'apport des outils SIG pour notre recherche en santé environnementale, une discussion générale se consacre aux liens santé – environnement identifiés dans nos travaux. Cette discussion est suivie d'un bilan général des avancées scientifiques obtenues et se conclut par l'identification de perspectives de recherche.

Dans le cadre de cette thèse, les connaissances de l'état qualitatif de l'environnement, du statut socio-économique et de l'état de santé des populations, ainsi que de leurs interactions ont été enrichies à l'aide des outils SIG. Les SIG sont très utiles pour la récolte, le traitement et la spatialisation de bases de données issues de sources très diverses (Vine *et al.*, 1997 ; Pornon, 2007). Ils sont également le support de base pour la mise en œuvre de puissants outils d'analyse statistique spatio-temporels, adaptés en fonction de la nature des données.

Dans le domaine de la santé, les outils d'épidémiologie spatiale sont basés sur des indicateurs standardisés d'incidence, calculés de manière agrégée au sein d'unités géographiques à partir de registres exhaustifs de maladies (Waller & Gotway, 2004 ; Gorja *et al.*, 2011). Les méthodes employées pour cartographier ces indicateurs utilisent des modèles mathématiques de lissage, qui tiennent compte de l'information apportée par les unités voisines pour estimer la distribution géographique des événements de santé (Clayton & Kaldor, 1987 ; Elliott & Wartenberg, 2004 ; Auchincloss *et al.*, 2012). De manière plus précise, les outils de détection d'agrégats, telle que la statistique de scan, identifient les zones géographiques et/ou temporelles atypiques de risque, en intégrant les facteurs confondants (Kulldorff 1997 ; Fritz *et al.*, 2013). Ces outils sont de plus en plus utilisés pour la description des disparités d'incidence de pathologies et l'identification de points noirs sanitaires à différentes résolutions spatiales (Cousens *et al.*, 2001; Sabel *et al.*, 2003; Viel *et al.*, 2005; Pedigo *et al.*, 2011; Génin *et al.*, 2013). Les résultats cartographiques sont considérés comme des images fiables de la variabilité spatiale des risques ou de la fréquence des maladies et permettent la prise de décision pour la réduction des inégalités territoriales de santé et la recherche de facteurs étiologiques, telles que les contaminations environnementales. Dans nos travaux de recherche, ces outils ont été utilisés pour évaluer les disparités d'incidence de l'IRCT, à l'aide de la cartographie des SIR et de la détection de clusters spatiaux.

Dans le domaine de l'environnement, les bases de données issues de la surveillance physico-chimique et biologique des milieux reposent sur des mesures réalisées au niveau d'un échantillon de points géolocalisés. Dans ce contexte, les méthodes d'interpolations spatiales et notamment les géostatistiques sont de puissants outils pour pallier la répartition hétérogène des points dans l'espace (Krige, 1951 ; Ripley, 1981 ; Cressie, 1993 ; Li et Heap, 2014). Ces outils modélisent la valeur d'une variable environnementale en tous points de l'espace, ce qui offre une vision globale du phénomène (Hengl, 2007). Leur utilisation facilite ainsi la caractérisation des milieux environnementaux (Cuny *et al.*, 2004 ; Bourennane *et al.*, 2010 ; Finzgar *et al.*, 2014 ; Roberts *et al.*, 2014), mais permet également la construction de proxys d'exposition des populations dans le cadre d'études épidémiologiques (Nuckols *et al.*, 2004 ; Vienneau *et al.*, 2009 ; Caudeville *et al.*, 2012). Les outils géostatistiques ont été utilisés dans nos travaux de biosurveillance pour évaluer les variations

spatiales de l'imprégnation des lichens par les ETM et de la contamination des eaux souterraines par l'aluminium. Dans le premier chapitre, ces outils ont également permis de modéliser les valeurs de concentrations métalliques dans les lichens au niveau du lieu de vie d'un échantillon de population.

En outre, les SIG permettent l'étude concomitante de ces différentes informations. Dans le troisième chapitre, nous avons ainsi superposé la cartographie des clusters de maladie de Crohn à celle de la contamination des eaux souterraines en aluminium, ce qui nous a permis d'obtenir une première interprétation visuelle. Cependant, des fonctions plus puissantes permettent d'aller plus loin que la simple superposition de couches de données, telle que l'agrégation de données environnementales au sein d'unités spatiales pour estimer l'exposition de groupes de population aux contaminations (Auchincloss *et al.*, 2012).

Nous avons en effet souligné les difficultés que peuvent représenter des différences de nature des informations dans le traitement des données en santé – environnement. D'un côté, les données caractérisant l'état de santé ou le statut socio-économique des populations sont de nature agrégée, c'est-à-dire qu'elles caractérisent un groupe de population, représenté par une unité spatiale administrative. D'un autre côté, les données environnementales sont de nature géostatistique et sont représentées par une matrice de points, au sein de laquelle chaque point peut prendre une valeur différente de la variable. Afin de pouvoir étudier les covariations spatiales de ces différentes informations, les SIG peuvent être utilisés pour adapter les données, en l'occurrence celles caractérisant l'environnement. Ainsi, une méthode de transformation des données de biosurveillance lichénique a été mise en œuvre dans le premier chapitre pour passer d'un format géostatistique à un format agrégé à l'Iris. Elle consiste dans un premier temps à modéliser la valeur de la variable pour chaque point d'un maillage fin et régulier, puis à calculer la valeur moyenne de cette variable pour chaque unité spatiale à partir des mailles situées à l'intérieur de celle-ci. Nous avons au final obtenu la concentration métallique moyenne pour chacun des Iris.

Par conséquent, les SIG sont très utiles pour la réalisation d'études écologiques. Celles-ci visent à évaluer les corrélations entre les indicateurs sanitaires, socio-économiques et environnementaux, afin de suggérer ou étayer des hypothèses concernant les facteurs explicatifs ou aggravants des pathologies (Best *et al.*, 2001 ; Wakefield & Salway, 2001). Dans nos travaux, ces études nous ont permis d'associer le niveau de défaveur sociale des populations à l'imprégnation multimétallique de l'environnement à l'échelle des Iris du bassin de Dunkerque et de révéler une influence de la précarité sociale sur les niveaux d'incidence de l'IRCT à l'échelle des cantons du Nord – Pas de Calais. De par l'agrégation des données populationnelles, le rapport de causalité exprimé par les études écologiques est moins robuste que celui que peut traduire une étude menée à l'échelle individuelle (cohorte, cas-témoins). En effet, les études écologiques cherchent d'abord à mettre en évidence un effet sanitaire au niveau d'un groupe au moyen d'une analyse statistique, afin d'élaborer des hypothèses sur les potentiels facteurs de risque d'apparition de cet effet (Richardson, 2000). La nature causale de cet effet ne peut être clairement définie et extrapolée à l'individu en raison de biais écologiques. Ces biais ont été largement décrits dans la littérature (Piantadosi *et al.*, 1988 ; Greenland & Morgenstern, 1989 ; Morgenstern, 1995 ; King, 1997 ; Wakefield, 2007 ; 2008). Le biais de spécification est lié à la variabilité des niveaux d'exposition aux facteurs de risque ou des facteurs de confusion entre les individus d'un même groupe (Greenland, 1992). Le biais d'agrégation implique que les individus ayant présenté un effet ne sont pas forcément ceux qui ont été exposés au facteur de risque étudié. Enfin le biais de confusion évoque la possibilité qu'il existe un autre facteur de risque, non mesuré, pouvant influencer sur l'effet observé (Wakefield, 2008). Ce facteur de risque peut par exemple être lié aux migrations journalières (déplacements domicile – travail) ou aux habitudes de vie (alimentation). Pour une interprétation plus fiable des résultats, les groupes de population comparés nécessitent donc d'être les plus homogènes possibles en termes de facteurs de confusion. En revanche, les études à l'échelle de groupes d'individus sont, en comparaison aux études individuelles, bien plus faciles et rapides à mettre en œuvre puisqu'elles utilisent des informations généralisées, comme des indicateurs d'incidence de maladies et la présence de sources d'exposition

aux facteurs de risques environnementaux. C'est la raison pour laquelle elles sont développées en amont d'études plus fines, lorsque les facteurs étiologiques sont peu identifiés. De plus, pour renforcer le lien écologique dans ces études et réduire les risques de biais, certains auteurs proposent d'intégrer, lorsqu'elles sont disponibles, des données individuelles (Haneuse & Wakefield, 2008). A ce titre, les mesures de biosurveillance des populations semblent prendre une place privilégiée.

L'impact que peuvent avoir les différentes composantes de l'environnement sur la santé est reconnu depuis de nombreuses années (OMS, 2014). Pourtant, ce lien reste difficile à caractériser car l'état de santé d'une population peut être affecté par un ensemble complexe de déterminants individuels et de facteurs sociaux et environnementaux (OMS, 2012 ; Besse *et al.*, 2014). En effet, l'OMS (1946) définit la santé comme « un état de complet bien-être physique, mental et social », qui « ne consiste pas seulement en une absence de maladie ou d'infirmité ». Cette définition n'a pas été modifiée depuis.

Lors de ce travail de thèse, le lien santé – environnement a été considéré comme l'association spatiale de populations ayant un état de santé dégradé et d'un environnement de vie dégradé par les pollutions. Dans ce contexte, l'objectif était de caractériser la qualité des milieux de vie des populations grâce aux bases de données environnementales disponibles et aux outils SIG, puis de rapprocher les cartographies générées aux indicateurs sanitaires et socio-économiques par des études écologiques. La difficulté de mettre en évidence le lien santé – environnement s'est confirmée dans les trois axes de recherche développés.

Dans le premier chapitre consacré à l'étude des charges métalliques dans le sang et l'urine d'un échantillon de population du bassin de Dunkerque, nous avons obtenu des corrélations significatives pour cinq variables (Cdu, Pbs, Pbu, Niu et Vs) parmi vingt-deux avec les charges environnementales mesurées dans les lichens. Ces résultats supposent une influence de l'environnement sur l'imprégnation des populations pour ces éléments. Néanmoins, les faibles coefficients de corrélation témoignent de la multiplicité des facteurs d'exposition des populations aux métaux. Les résultats de cette partie montrent *in fine* que la biosurveillance environnementale reflète peu les imprégnations en métaux. La part de l'environnement dans l'exposition globale des populations semble ainsi être masquée. Cette exposition peut par exemple être d'origine professionnelle (Kakkar & Jaffery, 2005) ou être liée aux habitudes de vie, notamment le tabagisme (Bernhard *et al.*, 2005). De même, l'ingestion d'eau de boisson et l'alimentation peuvent constituer une source d'exposition aux ETM. En effet, du fait de leur persistance, de très nombreux travaux ont démontré l'accumulation des métaux dans l'environnement en général et dans les organismes en particulier, tout au long des réseaux trophiques (Lim *et al.*, 2008 ; Vidovic *et al.*, 2005 ; Wilhelm *et al.*, 2007). De plus, d'autres facteurs confondants viennent impacter les niveaux d'imprégnation des populations par les métaux. Ces derniers étant des éléments non dégradés par l'organisme et cumulatifs, les personnes plus âgées sont susceptibles de présenter des niveaux de concentrations plus élevés que les personnes jeunes. Ainsi, pour établir ces corrélations, les facteurs âge et tabac nécessitent d'être pris en considération. Par ailleurs, la recherche menée en parallèle sur le bassin de Dunkerque signale également que les populations vivant sur les territoires les plus contaminés par ces métaux sont les plus défavorisées. Nous pouvons ainsi nous poser la question du rôle de la défaveur sociale dans l'imprégnation des populations par les ETM.

Les résultats du second chapitre indiquent justement un rôle important du statut socio-économique dans la survenue de l'IRCT. Cette conclusion se confirme d'ailleurs à l'échelle nationale et internationale (Couchoud *et al.*, 2012 ; Usami *et al.*, 2000 ; Counil *et al.*, 2008 ; Hommel *et al.*, 2010 ; Stel *et al.*, 2011 ; Caskey *et al.*, 2011). De plus, même si dans nos travaux le diabète et les maladies cardiovasculaires ne semblent pas influencer l'incidence de cette pathologie, plusieurs études les

considèrent comme des facteurs de risque importants (Muntner *et al.*, 2003; Wimmer *et al.*, 2003; Couchoud *et al.*, 2012; Bell *et al.*, 2012; Huang *et al.*, 2014). Malgré tout, suite à la prise en compte du facteur social dans l'analyse cartographique, il persiste des zones géographiques de sur- et de sous-incidence d'IRCT pour lesquelles les principaux déterminants restent inexplicables. Ainsi, nous émettons l'hypothèse que l'exposition à un environnement dégradé et notamment imprégné par les métaux puisse être également responsable des disparités locales d'incidence. En effet, il a été démontré que l'exposition au cadmium est susceptible d'induire ou d'aggraver une pathologie rénale (Lauwerys *et al.*, 1993 ; Hellström *et al.*, 2001 ; Järup *et al.*, 2002). Cependant, plutôt qu'un facteur de causalité, nous supposons que l'environnement doit être considéré comme un facteur aggravant, qui accélère le processus d'initialisation de la phase terminale de l'insuffisance rénale chronique. A ce jour, la difficulté de l'étude de facteurs environnementaux est surtout liée à l'absence de données environnementales accessibles à l'échelle des cantons du Nord – Pas de Calais.

Dans le troisième chapitre, notre recherche s'est portée sur l'aluminium, un facteur environnemental suspecté des MICI. Actuellement, très peu de littérature concerne ce sujet et l'aluminium est considéré comme un facteur potentiel d'induction de la maladie de Crohn (Beamish *et al.*, 2011 ; Lerner, 2012 ; Pineton de Chambrun *et al.*, 2014). Ce niveau de connaissances nécessite donc la mise en œuvre de recherches exploratoires, telle qu'une étude écologique géographique. Notre recherche s'est dans un premier temps portée sur les bases de données environnementales susceptibles d'être rapprochées avec les cartographies d'incidence et de clusters de la MC. A ce jour, très peu de bases de données permettent de caractériser l'état qualitatif des milieux à ces étendues géographiques et temporelles, notamment concernant l'aluminium. Suite à cette recherche, la banque ADES semble adéquate pour une étude écologique prospective, du fait de l'étendue spatiale et temporelle des données et du nombre de paramètres mesurés dont font partie les métaux. Les résultats de l'analyse critique et cartographique de celle-ci mettent en évidence des difficultés d'interprétation de l'état de contamination des eaux souterraines, en raison du mode d'alimentation de la banque. De plus, l'exposition des populations aux eaux souterraines ne peut être identifiée, ce qui nécessite un niveau d'extrapolation important à partir des mesures effectuées au niveau des points de captage. Finalement, la superposition des cartographies de l'aluminium avec les clusters de MC ne permet pas d'étayer cette hypothèse. L'ensemble de ces résultats confirme donc la difficulté d'établir un lien causal entre les polluants environnementaux et les effets sur la santé dans le cadre d'études spatialisées. Cette difficulté est principalement liée à la rareté des bases de données disponibles et permettant de caractériser la qualité des milieux environnementaux. La construction, le suivi et la publication de telles informations sont donc nécessaires.

Ce lien santé – environnement est d'autant plus complexe lorsque l'on s'intéresse à une substance particulière. En dehors de quelques exceptions, comme le saturnisme principalement lié à l'exposition au plomb ou le mésothéliome de la plèvre suite à une exposition à l'amiante, peu de maladies peuvent être indubitablement associées à un polluant en particulier. Pour pallier cette difficulté, nous proposons l'utilisation d'indices intégrés de qualité des milieux dans les études en santé – environnement. Ces outils peuvent être considérés comme de puissants outils de traitement et d'analyse à partir de données environnementales brutes. En plus des nombreuses mesures de concentrations, un indice intégré caractérise à lui seul le statut général de pollution (Caeiro *et al.*, 2005 ; Pykh *et al.*, 2000 ; Golge *et al.*, 2013). Il constitue donc une opportunité pour faciliter l'identification du rôle de l'environnement en général dans les effets de santé.

Ces indices sont de plus complémentaires aux outils SIG. En effet, ces outils sont conçus pour aider à leur construction et à leur renforcement. Dans un contexte de contamination multiple, les géostatistiques peuvent être employées pour estimer les valeurs de concentrations en polluants au niveau de points d'échantillonnage pour lesquels tous les paramètres n'ont pas été mesurés. Cette procédure peut donc réduire certains facteurs limitants, comme l'hétérogénéité d'un échantillonnage. Dans l'idée d'une analyse multi-milieux, ce procédé a récemment été mis en œuvre pour estimer les valeurs d'indices de qualité de l'air aux coordonnées géographiques de stations de

mesure de la qualité des eaux et alimenter un indice de contamination multi-milieu (Golge *et al.*, 2013).

L'utilisation concomitante des indices intégrés et des outils SIG constitue une solution optimale et incontournable pour caractériser la qualité des milieux environnementaux. Néanmoins, leur utilisation nécessite certaines précautions. En effet, ils représentent une simplification de la qualité environnementale et engendrent nécessairement une perte d'information. Quelle que soit la méthode de calcul utilisée, de tels indices lissent les valeurs des variables intégrées, par exemple lors de l'utilisation d'une fonction moyenne, ce qui peut masquer la variabilité intrinsèque des paramètres. De ce fait, leur utilisation ne doit surtout pas se substituer à une interprétation mono-paramètre.

Dans le cadre de ces travaux de recherche, le Ratio d'Imprégnation Moyen (RIM) a été utilisé pour décrire la contamination multimétallique. Le RIM répond à plusieurs critères définis dans la littérature pour évaluer la pertinence des indices intégrés de qualité des milieux (Caeiro *et al.*, 2005 ; Cairncross *et al.*, 2007 ; Singh *et al.*, 2008 ; Bharti & Katyal, 2011). Il est simple, facile à interpréter, robuste, sensible et représentatif d'une contamination globale par les métaux. Les résultats de son intégration dans nos recherches en santé – environnementale se révèlent également concluants. En effet, à l'instar des éléments qui le composent, le RIM calculé sur la base des mesures lichéniques en 18 ETM est significativement corrélé à l'indice de défaveur sociale sur le bassin de Dunkerque. Sur ce même bassin, nous avons souligné la corrélation entre le RIM lichénique et le RIM urinaire, tous deux calculés sur la base de 11 ETM mesurés dans les lichens et les urines.

Nous pouvons ensuite établir un bilan des avancées scientifiques obtenues dans le cadre de cette thèse. Un état des lieux des sources de données environnementales sur le territoire de la région Nord – Pas de Calais a été réalisé. Il est important de souligner une forte hétérogénéité d'accessibilité et de contenu de ces bases en fonction du milieu. Ainsi, le compartiment atmosphérique fait l'objet de la surveillance la plus régulière et la plus complète par les AASQA, grâce notamment aux stations de mesures de la qualité de l'air et aux différents outils de modélisation. Cependant, les données issues de la surveillance physico-chimique ne sont pas adaptées pour une analyse cartographique fine. Dans ce contexte, la biosurveillance représente une source importante et complémentaire d'information pour le suivi spatial et temporel de la qualité de l'air.

Le milieu aquatique est également concerné par une surveillance accrue de l'ensemble des eaux (souterraines, de surface, de baignade, de distribution), notamment depuis l'instauration des SDAGE. Mais peu de bases de données de qualité des eaux sont actuellement accessibles. La banque ADES représente le point fort du dispositif de suivi de ce milieu, puisqu'elle communique les résultats bruts des mesures de qualité effectuées en tout point de captage au niveau national. Nous avons cependant mis en évidence qu'il convient d'adopter une certaine prudence lors de leur traitement. Contrairement aux autres matrices, le sol ne fait pas l'objet d'un suivi régulier. Malgré tout, ce milieu est concerné par le plus grand nombre de bases de données (Algros & Charissou, 2010). Mais peu d'entre elles sont basées sur la mesure de polluants. Bien que le RMQS fasse office de référence pour la qualité des sols en France, les données sont disponibles au public uniquement sous forme de cartographies. De plus, le maillage de 16 km de côté limite les possibilités d'études à fine résolution spatiale.

Ainsi, nous avons constaté que très peu de bases de données environnementales sont actuellement accessibles et encore moins utilisables en l'état pour une étude spatiale ou spatio-temporelle de l'état qualitatif du milieu. Ces résultats corroborent le bilan effectué par le réseau RECORD (Algros & Charissou, 2010). La mise en relation future de données environnementales et sanitaires à l'aide d'études spatialisées devra donc passer par la construction, le suivi régulier et la communication de bases de données environnementales.

Deux bases de données environnementales ont été utilisées dans nos travaux pour caractériser la qualité des milieux. La première concerne les mesures d'imprégnation par les ETM, réalisées dans les lichens au sein de trois territoires du Nord – Pas de Calais. Nous avons pu générer un total de 18 cartographies monométalliques et une cartographie multimétallique, basée sur un indice intégré. Une méthode innovante d'interpolation spatiale a pour cela été employée. Le Krigeage Empirique Bayésien, développé par Krivoruchko (2012), a en effet rarement été utilisé pour cartographier une variable environnementale. Finzgar *et al.* (2014) ont employé cette méthode pour cartographier les concentrations en plomb et zinc dans les sols avant et après remédiation, Deng *et al.* (2013) pour le carbone organique des sols et Roberts *et al.* (2014) pour les concentrations atmosphériques en particules fines et ozone. Par ailleurs, l'indice intégré (le Ratio d'Imprégnation Moyen) a permis de décrire le niveau d'imprégnation global du milieu par les métaux. L'ensemble de ces cartographies révèle des variations spatiales d'imprégnation des lichens entre les bassins et au sein d'un même bassin. Des charges métalliques très élevées sont observées sur le bassin dunkerquois, notamment à proximité des deux zones industrielles. Sur le bassin urbain de Lille, les charges sont globalement moins importantes et mettent en évidence les axes routiers majeurs. Enfin, sur le bassin rural de Maubeuge, les niveaux de concentration sont beaucoup plus faibles. En comparaison avec les valeurs de fond, les niveaux d'imprégnation de l'environnement sont différents en fonction des ETM. Ces résultats confirment de précédentes observations du laboratoire (Cuny *et al.*, 2004).

La deuxième base utilisée est la banque de qualité des eaux souterraines ADES. Une analyse critique de celle-ci conclut à sa nécessaire consolidation lors de son interprétation spatiale. Cela permet d'avoir conscience et potentiellement d'éliminer les sources de variabilité dues à un manque de gestion coordonnée des entrées dans la base et à la forte irrégularité du suivi dans l'espace et le temps. Sur la base de ce travail, nous proposons un arbre de décision facilitant le traitement spatial et spatio-temporel de ce type de données.

Une fois consolidée, des cartographies géostatistiques ont été réalisées pour l'aluminium et le RIM. Elles représentent les concentrations moyennes en métaux, mesurées entre 1990 et 2012 au niveau de quatre départements. Une analyse spatio-temporelle employant des cartographies annuelles des concentrations en aluminium a également été réalisée pour cette période. Grâce aux outils SIG, nous pouvons souligner la différence de répartition des mesures dans l'espace et le temps. Certains points d'eau font l'objet d'une surveillance régulière dans le temps, alors que d'autres sont suivis de manière ponctuelle. En ce qui concerne les concentrations en aluminium, les résultats indiquent une faible contamination des eaux souterraines sur le Nord, le Pas de Calais et la Somme, avec des points remarquables de contamination ponctuelle. Sur le département de la Seine-Maritime, la pollution est plus chronique, notamment à proximité du port du Havre. Les SIG constituent donc des outils de caractérisation des pollutions, mais aussi de gestion pour le suivi de la qualité des milieux. Ils facilitent l'exploitation de bases de données conséquentes, telle que la banque ADES.

Etudiées en parallèle, les données qui caractérisent la santé des populations sont quant à elles bancarisées et structurées dans des registres exhaustifs de patients, ou bien issues de cohortes en population générale. Les registres de santé, établis sur une étendue géographique et temporelle définie, constituent des bases de données homogènes puisqu'elles sont basées sur le suivi régulier des patients et renseignent des informations sur le lieu d'habitation et l'âge du patient au moment de l'évènement de santé, des indicateurs relatifs à la pathologie et parfois certains facteurs confondants.

Lors de ces travaux de recherche, nous avons pu travailler à partir du registre de l'insuffisance rénale chronique terminale (Réseau REIN Néphronor : Couchoud *et al.*, 2006). Des cartographies d'incidence ont été produites à l'échelle des cantons du Nord – Pas de Calais, à partir des cas recensés entre 2005 et 2012. Celles-ci montrent d'importantes disparités, avec des risques nettement plus élevés au niveau de Dunkerque, Maubeuge et de l'agglomération Lille-Roubaix-Tourcoing. A cette maille fine, ce travail n'a aujourd'hui pas d'équivalent en France. En parallèle, nous avons utilisé une méthode de

détection de clusters atypiques d'incidence, peu commune dans le contexte de cette pathologie. La statistique de scan (Kulldorff, 1997) identifie la présence de trois clusters significatifs de surincidence au niveau des territoires précédemment évoqués. Aucun cluster temporel ou spatio-temporel n'a en revanche été détecté sur la base de nos données. De manière plus fine que la méthode classique, la version isotonique de la statistique de scan (Kulldorff, 1999) a été mise en œuvre pour révéler une variabilité du risque entre les unités spatiales d'un même cluster. Cette méthode apporte une information complémentaire dans la perspective de recherche de facteurs étiologiques de la maladie, car elle permet de hiérarchiser les risques au sein d'un même cluster. Peu d'études épidémiologiques spatialisées s'intéressent actuellement à ce niveau de précision.

La finalité de nos travaux était la caractérisation des liens environnement – santé, à partir des résultats cartographiques précédemment générés. L'état de santé d'une population pouvant être affecté par un ensemble complexe de déterminants individuels et de facteurs sociaux et environnementaux (Besse *et al.*, 2014 ; OMS, 2012), des indicateurs de défaveur sociale et des facteurs confondants ont été intégrés à nos analyses.

Dans le premier chapitre, des inégalités environnementales ont dans un premier temps été identifiées à l'échelle des Iris du bassin industriel de Dunkerque. Ces résultats sont issus de la mise en relation des données d'imprégnation lichéniques par les ETM avec un indice de défaveur sociale (l'Indice de Défavorisation Localisé). Pour ce faire, la mise au point d'une méthode de transformation des données environnementales nous a permis de les représenter de manière agrégée, c'est-à-dire d'obtenir une valeur moyenne par unité géographique. Cela nous a permis de conclure que sur ce territoire, les populations les plus défavorisées vivent à proximité de zones fortement contaminées par les éléments traces. Des différences significatives de niveaux d'imprégnation multimétallique (RIM) sont d'ailleurs observées entre les groupes extrêmes de défaveur sociale. L'étude des inégalités environnementales à l'aide de données de biosurveillance a peu d'équivalent à notre connaissance. Dans un second temps, ces données de biosurveillance ont été associées aux mesures d'imprégnation par les ETM, réalisées dans le sang et les urines d'un échantillon de population du bassin de Dunkerque. Des corrélations partielles, intégrant l'âge et le statut tabagique des individus, ont été mises en évidence entre les niveaux d'imprégnation dans le sang (plomb) ou les urines (cadmium, nickel, plomb) et les niveaux de concentrations dans les lichens. A ce titre, le RIM a été calculé pour chacune des matrices à partir de 11 éléments traces communs. Chez les sujets les moins à risques (jeunes et non-fumeurs), une corrélation significative est observée entre l'imprégnation multimétallique dans les urines et chez les lichens. Ainsi, les charges environnementales peuvent expliquer une partie de l'imprégnation des populations.

Pour résumer le travail réalisé à ce jour, la biosurveillance lichénique a été employée pour évaluer le niveau d'imprégnation du milieu par les ETM. Après avoir cartographié et transformé ces données, nous avons révélé la présence d'inégalités environnementales par l'étude du lien entre la contamination multimétallique de l'environnement et la présence de population défavorisées, notamment à l'aide d'un indice intégré (le RIM) et un indice de défaveur sociale (l'IDL). Nous avons ensuite mis en évidence une possible influence de cette contamination sur l'imprégnation des populations par ces métaux. Ces résultats démontrent le rôle important de la biosurveillance environnementale et humaine pour l'évaluation des liens santé – environnement. La suite de nos travaux se consacrera à la recherche de situations à risque à partir de ces trois sources d'information.

Dans le chapitre consacré à l'IRCT, des disparités spatiales d'incidence ont été révélées à fine résolution spatiale, entre les cantons de la région Nord – Pas de Calais. Les variations observées peuvent être en partie expliquées par le niveau de précarité sociale des populations, estimé à l'aide de l'indice de défaveur de Townsend. Nos résultats sont novateurs à cette échelle en France et confirment les conclusions de l'étude nationale menée à l'échelle des départements (Couchoud *et al.*, 2012), ainsi que d'autres menées à différentes échelles dans le monde (Hommel *et al.*, 2010 ;

Grace *et al.*, 2012 ; Caskey *et al.*, 2012). A ce jour, l'hypothèse environnementale de cette pathologie est en cours d'étude.

Concernant les facteurs de risques environnementaux des maladies inflammatoires chroniques de l'intestin (MICI) et notamment la maladie de Crohn (MC), le troisième chapitre se consacre aux contaminations métalliques de eaux souterraines. Les cartographies géostatistiques de l'aluminium et du RIM ont été superposées aux clusters de MC détectés sur la période 1990 – 2006. Bien que l'aluminium et de manière globale les micropolluants métalliques soient suspectés d'avoir un rôle dans le développement de la MC (Beamish *et al.*, 2011 ; Lerner, 2007 ; 2012 ; Pineton de Chambrun *et al.*, 2014), nos résultats ne nous permettent pas de valider ou d'étayer ces hypothèses.

En termes de perspectives, ces travaux de recherche nous ont permis d'identifier un certain nombre de pistes à court, moyen et long terme. Dans la suite des travaux entamés dans le premier chapitre, les inégalités environnementales et les inégalités environnementales de santé seront évaluées à court terme pour les bassins de Lille et Maubeuge, de la même manière que pour le bassin de Dunkerque. L'objet de cette recherche sera de révéler si les conclusions établies sur le bassin industriel du dunkerquois sont les mêmes lorsque l'on s'intéresse à des bassins de typologie différente. Une analyse intégrant l'ensemble des bassins sera également réalisée. A plus long terme et en fonction des résultats précédents, nos recherches pourront se focaliser sur les situations à risques, combinant imprégnation de l'environnement, imprégnation des populations et précarité sociale. L'exposition des personnes est en effet très complexe et la biosurveillance n'est que peu prédictive des imprégnations. Néanmoins, elle se révèle être un outil pratique de mise en évidence de telles situations à risque, potentiellement à impacts sanitaires importants. Nous démontrons et formalisons ainsi un concept seulement pressenti en 2004 (Cuny *et al.*, 2004). Ces conclusions replacent la biosurveillance comme sentinelle d'effets biologique, très utile en amont de l'épidémiologie.

Dans le contexte des inégalités spatiales d'incidence de l'IRCT sur la région Nord – Pas de Calais, après l'étude du facteur social, nous nous focaliserons sur les facteurs environnementaux et particulièrement le cadmium. Nous utiliserons pour cela les données de biosurveillance lichénique. Par ailleurs, dans ce chapitre nous avons utilisé la version isotonique de détection de clusters d'incidence (Kulldorff, 1999). Nous souhaitons y associer une méthode alternative qui consiste à renouveler la procédure de détection de clusters au sein même d'un cluster préalablement mis en évidence. Cette méthode nécessite encore une mise au point qui sera réalisée en collaboration avec l'équipe du Pr. Duhamel (EA2694, Université Lille 2).

Concernant les facteurs environnementaux des MICI, la difficulté d'établir un lien d'exposition direct entre la qualité des eaux souterraines et la population oriente actuellement nos travaux vers la qualité de l'air. Nous avons à ce titre pris contact avec le réseau de surveillance de la qualité de l'air ATMO Nord – Pas de Calais pour l'acquisition de données d'émissions en polluants atmosphériques sur le territoire et la période d'étude du registre EPIMAD. A plus long terme, des campagnes de biosurveillance végétale et fongique seront menées au niveau de zones contrastées d'incidence de MC, pour le suivi de la qualité globale de l'air (indices de biodiversité) et de divers polluants inorganiques et organiques (bioaccumulation). Des indices intégrés seront également mis en œuvre pour évaluer le niveau d'imprégnation global du milieu par l'ensemble de ces polluants.

D'un point de vue fondamental, nous nous intéresserons à l'application des méthodes de statistique de scan aux données environnementales. L'enjeu est ici de détecter des clusters spatio-temporels statistiquement atypiques de contamination des milieux, signalant la présence de points noirs environnementaux.

En parallèle des résultats présentés dans ce document de thèse, d'autres programmes de recherche ont été débutés. Les méthodes et outils développés dans nos travaux seront par exemple employés pour évaluer les disparités d'incidence des évènements cardiovasculaires (Registre MONICA de Lille) à fine échelle sur la Communauté Urbaine de Lille (LMCU) et l'influence de contaminations environnementales. Les résultats montrent également un manque de renseignement des bases de données environnementales par rapport aux registres de santé. Un travail est donc nécessaire afin de définir le contenu de ces bases, de caractériser la qualité des milieux et ainsi faciliter l'étude de l'interaction de l'Homme avec son environnement.

Finalement, nous avons souligné la difficulté d'accès aux bases de données environnementales existantes et notamment aux données brutes concernant la qualité des milieux. Pour pallier ce manque de données et le retard des bases environnementales par rapport aux bases de santé, nous avons entamé un travail visant à définir le contenu de ces bases, nécessaire pour caractériser la qualité des milieux et faciliter l'étude de l'interaction de l'Homme avec son environnement. A cette occasion, nous avons débuté la récolte des données environnementales non concernées par les bases existantes. Par exemple, les résultats des diagnostics de pollution effectués lors de la réhabilitation de sites potentiellement pollués font l'objet de récents partenariats avec la Ville de Lille, la Communauté Urbaine de Lille et le bureau d'étude Tauw France. Toujours dans le cadre de la qualité des sols, les données de biosurveillance générées par le laboratoire n'ont pas à ce jour fait l'objet d'analyse spatialisée.

REFERENCES BIBLIOGRAPHIQUES

- Algos E & Charissou AM. 2010. Réalisation d'un avis d'experts : « bruit de fond » des opluants dans l'environnement, analyse des données existantes et identification des lacunes. Etude RECORD N° 09-0812/1A. 163p.
- Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. 2012. A review of spatial methods in epidemiology, 2000-2010. *Annu Rev Public Health*, 33: 107-122.
- Beamish L Osornio-Vargas A Wine E. 2011. Air pollution: an environmental factor contributing to intestinal disease. *Journal of Crohn's and Colitis*, 5: 279–286.
- Bell EK, Gao L, Judd S, Glasser SP, McClellan W *et al.* 2012. Blood pressure indexes and end-stage renal disease risk in adults with chronic kidney disease. *American Journal of Hypertension*, 25 (7): 789-96.
- Bernhard D, Rossmann A, Wick G. 2005. Metals in Cigarette Smoke – Critical review. *IUBMB Life*, 57 (12): 805 – 809.
- Besse G, Caudeville J, Michelot F, Prisse N, Trugeon A. 2014. Inégalités territoriales, environnementales et sociales de santé. Regards croisés en régions : de l'observation à l'action. SGMAS & CGDD, 68p.
- Best NG, Cockings S, Bennett JE, Wakefield JC, Elliott P. 2001. Ecological regression analysis of environmental benzene exposure and childhood leukaemia: sensitivity to data inaccuracies, geographical scale and ecological bias. *Journal of the Royal Statistical Society, Series A* (164): 155-74.
- Bharti N, Katyal D. 2011. Water quality indices used for surface water vulnerability assessment. *International Journal of Environmental Sciences*, 2 (1): 154-173.
- Bourenane H, Douay F, Sterckeman T, Villanneau E, Ciesielski F, King D, Baize D. 2010. Mapping of anthropogenic trace elements inputs in agricultural topsoil from Northern France using enrichment factors. *Geoderma*, 157: 165–174.
- Caeiro S, Costa MH, Ramos TB, Fernandes F, Silveira N, Coimbra A, Medeiros G, Painho M. 2005. Assessing heavy metal contamination in Sado Estuary sediment: An index analysis approach. *Ecological Indicators*, 5: 151–169.
- Cairncross EK, John J, Zunckel M. 2007. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. *Atmos. Environ*, 41: 8442–8454.
- Caskey FJ, Kramer A, Elliott RF, Stel VS, Covic A *et al.* 2011. Global variation in renal replacement therapy for end-stage renal disease. *Nephrology Dialysis Transplantation*, 26 (8): 2604-10.
- Caudeville J, Bonnard R, Boudet C, Denys S, Govaert G, Cicilella A. 2012. Development of a spatial stochastic multimedia exposure model to assess population exposure at a regional scale. *Science of the Total Environment*, 432: 297–308.
- Clayton D & Kaldor J. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43: 671-681.
- Couchoud C, Stengel B, Landais P, Aldigier J-C, De Cornelissen F, *et al.* 2006. The renal epidemiology and information network (REIN): a new registry for end-stage renal disease in France. *Nephrology Dialysis Transplantation*, 21: 411–418.
- Couchoud C, Guihenneuc C, Bayer F, Lemaitre V, Brunet P, Stengel B. 2012. Medical practice patterns and socio-economic factors may explain geographical variation of end-stage renal disease incidence. *Nephrology Dialysis Transplantation*, 27 (6): 2312-22.

- Council É, Cherni N, Kharrat M, Achour A, Trimech H. 2008. Trends of incident dialysis patients in tunisia between 1992 and 2001. *American Journal of Kidney Diseases*, 51 (3): 463-70.
- Cousens S, Smith PG, Ward H, Everington D, Knight RSG, Zeidler M, Stewart G, Smith-Bathgate EAB, Macleod MA, Mackenzie J, Will RG. 2001. Geographical distribution of variant Creutzfeldt-Jakob disease in Great Britain, 1994-2000. *The Lancet*, 357: 1002-1007.
- Cressie NA. 1993. *Statistics for Spatial Data* (revised edition). John Wiley & Sons, Inc., New York.
- Cuny D, Davranche L, Thomas P, Kempa M, Van Haluwyn C. 2004. Spatial and temporal variations of trace element contents in *Xanthoria parietina* thalli collected in a highly industrialized area in Northern France as an element for a future epidemiological study. *Journal of Atmospheric Chemistry*, 49: 391-401.
- Deng F, Minasny B, Knadel M, McBratney A, Heckrath G, Greve MH. 2013. Using vis-NIR spectroscopy for monitoring temporal changes in soil organic carbon. *Soil Sci*, 178 (8): 389-99.
- Elliott P & Wartenberg D. 2004. Spatial Epidemiology: Current Approaches and Future Challenges. *Environmental Health Perspectives*, 112: 998-1006.
- Finzgar N, Jez E, Volgar D, Lestan D. 2014. Spatial distribution of metal contamination before and after remediation in the Meza Valley, Slovenia. *Geoderma*, 217-218: 135-143.
- Fritz CE, Schuurman N, Robertson C, Lear S. 2013. A scoping review of spatial cluster analysis techniques for point-event data. *Geospatial Health*, 7 (2): 183-198.
- Génin M, Duhamel A, Preda C, Fumery M, Savoye G *et al.* 2013. Space-time clusters of Crohn's disease in northern France. *J Public Health*, DOI 10.1007/s10389-013-0580-9.
- Golge M, Yenilmez F, Aksoy A. 2013. Development of pollution indices for the middle section of the Lower Seyhan Basin (Turkey). *Ecological Indicators*, 29: 6-17.
- Goria S, Stempfelet M, de Crouy-Chanel P. 2011. *Introduction aux méthodes statistiques et aux systèmes d'information géographique en santé environnement- Application aux études écologiques*. Saint-Maurice: Institut de veille sanitaire : 65 p.
- Grace BS, Clayton P, Cass A, McDonald SP. 2012. Socio-economic status and incidence of renal replacement therapy: A registry study of australian patients. *Nephrology Dialysis Transplantation*, 27 (11): 4173-80.
- Greenland S & Morgenstern H. 1989. Ecological bias, confounding and effect modification. *Int. J. Epidemiol*, 18: 269-74.
- Greenland S. 1992. Divergent biases in ecologic and individual level studies. *Stat. Med*, 11: 1209-23.
- Haneuse S & Wakefield J. 2008. Geographic-based ecological correlation studies using supplemental case-control data. *Stat Med*, 27 (6): 864-87.
- Hellström L, Elinder C-G, Dahlberg B, Lundberg M, Järup L, Persson B, Axelson O. 2001. Cadmium exposure and end-stage renal disease. *Am J Kidney Dis*, 38: 1001-8.
- Hengl T. 2007. *A Practical Guide to Geostatistical Mapping of Environmental Variables*. Office for Official Publication of the European Communities, Luxembourg, 143p.
- Hommel K, Rasmussen S, Kamper A, Madsen M. 2010. Regional and social inequalities in chronic renal replacement therapy in Denmark. *Nephrology Dialysis Transplantation*, 25 (8): 2624-32.
- Huang Y, Cai X, Zhang J, Mai W, Wang S, *et al.* 2014. Prehypertension and incidence of ESRD: A systematic Review and Meta-analysis. *American Journal of Kidney Diseases*, 63 (1): 76-83.
- Järup L. 2002. Cadmium overload and toxicity. *Nephrol Dial Transplant*, 17: 35-9.

- Kakkar P & Jaffery FN. 2005. Biological markers for metal toxicity. *Environ Toxicol Pharmacol*, 19 (2): 335-49.
- King G. 1997. *A Solution to the Ecological Inference Problem*. Princeton, NJ: Princeton Univ. Press
- Krige DG. 1951. A statistical approach to some mine valuations problems at the Witwatersrand. *J. Chem. Metall. Min. Soc. S. Afr*, 52: 119-139.
- Krivoruchko K. 2012. Empirical Bayesian Kriging implemented in ArcGIS Geostatistical Analyst. *ArcUser*, 15 (4): 6-10.
- Kulldorff M. 1997. A spatial scan statistic. *Communication in Statistics - Theory and Methods*, 26: 1481–1496.
- Lauwerys RR, Bernard AM, Buchet J-P, Roels H. 1993. Assessment of the health impact of environmental exposure to cadmium: contribution of the epidemiologic studies carried out in Belgium. *Environ Res*, 62: 200-6.
- Lerner A. 2012. Aluminum as an adjuvant in Crohn's disease induction. *Lupus*, 21 (2): 231-238.
- Li J & Heap AD. 2014. Spatial interpolation methods applied in the environmental sciences: a review. *Environmental Modelling & Software*, 53: 173-189.
- Morgenstern H. 1995. Ecologic studies in epidemiology: concepts, principles, and methods. *Annu. Rev. Public Health*, 16: 61–81.
- Muntner P, Coresh J, Powe NR, Klag MJ. 2003. The contribution of increased diabetes prevalence and improved myocardial infarction and stroke survival to the increase in treated end-stage renal disease. *Journal of the American Society of Nephrology*, 14 (6): 1568-77.
- Nuckols JR, Ward MH, Jarup L. 2004. Using geographic information systems for exposure assessment in environmental epidemiology studies. *Environ Health Perspect*, 112 (9): 1007-15.
- OMS. 1946 Préambule à la Constitution de l'Organisation mondiale de la Santé, tel qu'adopté par la Conférence internationale sur la Santé, New York, 19-22 juin 1946; signé le 22 juillet 1946 par les représentants de 61 Etats. Actes officiels de l'Organisation mondiale de la Santé, n°. 2, p. 100 et entré en vigueur le 7 avril 1948.
- OMS. 2012. *Environmental health inequalities in Europe*. Copenhagen, OMS Europe. 190p.
- OMS. 2014. *Burden of disease from the joint effects of Household and Ambient Air Pollution for 2012*. Geneva.
- Pedigo A, Aldrich T, Odoi A. 2011. Neighborhood disparities in stroke and myocardial infarction mortality: a GIS and spatial scan statistics approach. *BMC Publ Health*, 11: 644.
- Piantadosi S, Byar DP, Green SB. 1988. The ecological fallacy. *Am. J. Epidemiol*, 127: 893–904.
- Pineton de Chambrun G, Body-Malapel M, Frey-Wagner I, Djouina M, Deknuydt F *et al.* 2014. Aluminum enhances inflammation and decreases mucosal healing in experimental colitis in mice. *Mucosal Immunology*, 7 (3): 589-601.
- Pornon H. 2007. Bilan et perspectives de 20 années de Géomatique. *Géomatique Expert*, 57: 36-46.
- Pykh YA, Kennedy ET, Grant WE. 2000. An overview of systems analysis methods in delineating environmental quality indices. *Ecol. Model.*, 130: 25–38.
- Richardson S. 2000. Problèmes méthodologiques dans les études écologiques santé–environnement. *Life Sciences*, 323: 611–616.
- Ripley BD. 1981. *Spatial Statistics*. New York: Wiley.

- Roberts JD, Voss JD, Knight B. 2014. The association of ambient air pollution and physical inactivity in the United States. *Plos One*, 9 (3): e90143.
- Sabel CE, Boyle PJ, Löytönen M, Gatrell AC, Jokelainen M, Flowerdew R, Maasilta P. 2003. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. *American Journal of Epidemiology*, 157: 898-905.
- Singh RP, Nath S, Prasad SC, Nema AK. 2008. Selection of Suitable Aggregation Function for Estimation of Aggregate Pollution Index for River Ganges in India. *Journal of Environmental Engineering*, 134 (8): 689-701.
- Stel VS, van de Luijngaarden MWM, Wanner C, Jager KJ. 2011. The 2008 ERA–EDTA Registry Annual Report—a précis. *NDT Plus*, 4: 1–13.
- Usami T, Koyama K, Takeuchi O, Morozumi K, Kimura G. 2000. Regional variations in the incidence of end-stage renal failure in Japan. *Journal of the American Medical Association*, 284 (20): 2622-4.
- Viel JF, Floret N, Mauny F. 2005. Spatial and space-time scan statistics to detect low clusters of sex ratio. *Environmental and Ecological Statistics*, 12: 289-299.
- Vienneau D, de Hoogh K, Briggs D. 2009. A GIS-based method for modelling air pollution exposures across Europe. *Science of the Total Environment*, 408: 255–266.
- Vine M, Degnan D, Hanchette C. 1997. Geographic Information Systems: Their Use in Environmental Epidemiologic Research. *Environmental Health Perspectives*, 105 (6): 598-605.
- Wakefield JC & Salway R. 2001. A statistical framework for ecological and aggregate studies. *Journal of the Royal Statistical Society, series A* (164): 119-37.
- Wakefield JC. 2007. Disease mapping and spatial regression with count data. *Biostatistics*, 8: 158–83.
- Wakefield JC. 2008. Ecologic studies revisited. *Annual Review of Public Health*, 29: 75–90.
- Waller LA & Gotway CA. 2004. *Applied spatial statistics for public health data*. Wiley, Hoboken. 494p.
- Wimmer F, Oberaigner W, Kramar R, Mayer G. 2003. Regional variability in the incidence of end-stage renal disease: An epidemiological approach. *Nephrology Dialysis Transplantation*, 18 (8): 1562-7.

REFERENCES BIBLIOGRAPHIQUES

AFNOR. 2003. NF ISO 10381-1. Qualité du sol - Échantillonnage - Partie 1 : lignes directrices pour l'établissement des programmes d'échantillonnage.

AFNOR. 2008. NF X43-903 : Biosurveillance de l'air - Détermination d'un indice biologique de lichens épiphytes (IBLE). 19p.

AFNOR. 2013. NF X 43-904 : Biosurveillance passive de la qualité de l'air à l'aide des lichens autochtones : de la récolte à la préparation des échantillons.

Agnan Y. 2013. Bioaccumulation et bioindication par les lichens de la pollution atmosphérique actuelle et passée en métaux et en azote en France : sources, mécanismes et facteurs d'influence. Thèse de Doctorat Université de Toulouse. 306p.

Albertini R, Bird M, Doerrer N, Needham L, Robison S, Sheldon L, Zenick H. 2006. The use of biomonitoring data in exposure and human health risk assessments. *Environ Health Perspect*, 114(11): 1755-62.

Algros E & Charissou AM. 2010. Réalisation d'un avis d'experts : « bruit de fond » des opluants dans l'environnement, analyse des données existantes et identification des lacunes. Etude RECORD N° 09-0812/1A. 163p.

Andersen A, Hovmand M, Johnsen IB. 1978. Atmospheric heavy metal deposition in the Copenhagen area. *Environ. Pollut*, 17 : 133-151.

Angerer J, Ewers U, Wilhelm M. 2007. Human biomonitoring: state of the art. *Int J Hyg Environ Health*, 210(3-4): 201-28.

Apostoli P, Baj A, Bavazzano P, Ganzi A, Neri G, Ronchi A, Soleo L, Di LL, Spinelli P, Valente T, Minoia C. 2002. Blood lead reference values: the results of an Italian polycentric study. *Sci. Total Environ*, 287(1-2): 1-11.

Arrêté du 17/09/03 relatif aux méthodes d'analyse des échantillons d'eau et à leurs caractéristiques de performance.

Arrêté du 11/01/07 relatif aux limites et références de qualité des eaux brutes et des eaux destinées à la consommation humaine mentionnées aux articles R. 1321-2, R. 1321-3, R. 1321-7 et R. 1321-38 du code de la santé publique.

Arrêté du 19/09/11 fixant la liste des laboratoires agréés par le ministère chargé de la santé pour la réalisation des prélèvements et des analyses du contrôle sanitaire des eaux.

Arrêté du 25/01/10 établissant le programme de surveillance de l'état des eaux en application de l'article R. 212-22 du code de l'environnement.

ARS Haute-Normandie. 2010. Environnement & santé sur Port-Jérôme. Evaluation des risques sanitaires liés aux activités industrielles – résultats de l'étude réalisée de 2007 à 2010. 8p.

Auchincloss AH, Gebreab SY, Mair C, Diez Roux AV. 2012. A review of spatial methods in epidemiology, 2000-2010. *Annu Rev Public Health*, 33: 107-122.

Aznar JC, Richer-Lafleche M, Cluis D. 2008. Metal contamination in the lichen *Alectoria sarmentosa* near the copper smelter of Murdochville, Quebec. *Environ. Pollut*, 156-76.

Baillargeon S. 2005. Le krigeage : revue de la théorie et application à l'interpolation spatiale de données de précipitations. Mémoire Faculté des études supérieures de l'Université Laval. 128p.

- Balabanova B, Stafilov T, Sajin R, Baeeva K. 2012. Characterisation of Heavy Metals in Lichen Species *Hypogymnia physodes* and *Evernia prunastri* due to Biomonitoring of Air Pollution in the Vicinity of Copper Mine. *Int. J. Environ. Res.*, 6(3): 779-794.
- Bargagli, R. 1998. Trace Elements in Terrestrial Plants, an Ecophysiological Approach to Biomonitoring and Biorecovery. Springer, Berlin, 324p.
- Bari A, Minciardi M, Troiani F, Bonotto F, Paonessa F. 1998. Lichens and mosses in air quality monitoring: a biological model proposal. *Govt. Reports Announcements & Index*, Issue 16.
- Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD. 1972. Statistical inference under restrictions: The theory and application of isotonic regression. Wiley: England.
- Batariova A, Spevackova V, Benes B, Cejchanova M, Smid J, Cerna M. 2006. Blood and urine levels of Pb, Cd and Hg in the general population of the Czech Republic and proposed reference values. *Int. J. Hyg. Environ. Health*, 209: 359–366.
- Beale L, Abellan JJ, Hodgson S, Jarup L. 2008. Methodologic Issues and Approaches to Spatial Epidemiology. *Environmental Health Perspectives*, Vol 116 (8): 1105-1110.
- Beale L, Hodgson S, Abellan JJ, LeFevre S, Jarup L. 2010. Evaluation of spatial relationships between health and the environment: the Rapid Inquiry Facility. *Environ Health Perspect*, 118(9): 1306–1312.
- Beamish L, Osornio-Vargas A, Wine E. 2011. Air pollution: an environmental factor contributing to intestinal disease. *Journal of Crohn's and Colitis*, 5: 279–286.
- Behera S, Xian H, Balasubramanian R. 2014. Human health risk associated with exposure to toxic elements in mainstream and sidestream cigarette smoke. *Science of the Total Environment*, 472: 947–956.
- Bell EK, Gao L, Judd S, Glasser SP, McClellan W *et al.* 2012. Blood pressure indexes and end-stage renal disease risk in adults with chronic kidney disease. *American Journal of Hypertension*, 25 (7): 789-96.
- Bergamaschi L, Rizzio E, Giaveri G, Loppi S, Gallorini M. 2007. Comparison between the accumulation capacity of four lichen species transplanted to a urban site. *Environmental Pollution*, 148: 468-476.
- Bergeron M. 1992. *Vocabulaire de la géomatique*. Office de la langue française, Québec. 41 p.
- Bernhard D, Rossmann A, Wick G. 2005. Metals in Cigarette Smoke – Critical review. *IUBMB Life*, 57 (12): 805 – 809.
- Béron P, Valiquette L, Patty G, Briere F. 1982. Water quality indices. *Trib. Cebedeau*, 35: 385-391.
- Besag J & Newell J. 1991. The detection of clusters in rare diseases. *Journal of the Royal Statistical Society, Series A*, 154: 143–155.
- Besag J, York J, Mollié A. 1991. Bayesian image restoration, with two applications in spatial statistics (with Discussion). *Annals of the Institute of Statistical Mathematics*, 43(1): 1-59.
- Besse G, Caudeville J, Michelot F, Prisse N, Trugeon A. 2014. Inégalités territoriales, environnementales et sociales de santé. Regards croisés en régions : de l'observation à l'action. *SGMAS & CGDD*, 68p.
- Best N, Richardson S, Thomson A. 2005. A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res*, 14(1): 35-59.
- Best NG, Cockings S, Bennett JE, Wakefield JC, Elliott P. 2001. Ecological regression analysis of environmental benzene exposure and childhood leukaemia: sensitivity to data inaccuracies, geographical scale and ecological bias. *Journal of the Royal Statistical Society, Series A* (164): 155-74.

- Bezuglaya EY, Shchutskaya AB, Smirnova IV. 1993. Air pollution index and interpretation of measurements of toxic pollutant concentrations. *Atmospheric Environment*, Vol. 27A, No. 5: 773-779.
- Bhargava DS. 1983. Use of a water quality index for river classification and zoning of Ganga river. *J. Environ. pollut., Series B*, 6: 51–67.
- Bharti N, Katyal D. 2011. Water quality indices used for surface water vulnerability assessment. *International Journal of Environmental Sciences*, 2 (1): 154-173.
- Birrenbach T & Bocker U. 2004. Inflammatory bowel disease and smoking: a review of epidemiology, pathophysiology, and therapeutic implications. *Inflamm Bowel Dis*, 10: 848–59.
- Bithell JF. 2000. A classification of disease mapping methods. *Statist. Med*, 19: 2203-2215.
- Boamponsem LK, Adam JI, Dampare SB, Nyarko BJB, Essumang DK. 2010. Assessment of atmospheric heavy metal deposition in the Tarkwa gold mining area of Ghana using epiphytic lichens. *Nuclear Instruments and Methods in Physics Research*, B 268: 1492–1501.
- Bosser P. 2012. Interpolation spatiale. *Ouvrage Ecole Nationale des Sciences Géographiques*. 53p.
- Bourennane H, Douay F, Sterckeman T, Villanneau E, Ciesielski F, King D, Baize D. 2010. Mapping of anthropogenic trace elements inputs in agricultural topsoil from Northern France using enrichment factors. *Geoderma*, 157: 165–174.
- Bouyer I, Hémon D, Cordier S, Derriennic F, Stüker I, Stengel B, Clavel J. 1993. *Epidémiologie : principes et méthodes quantitatives*. Les Éditions INSERM. 498p.
- Brown RM, McClelland NI, Deininger RA, Tozer RG. 1970. A water quality index—Do we dare? *Water Sewage Works*, 117(10): 339–343.
- Brulle RJ & Pellow DN. 2006. Environmental justice: Human health and environmental inequalities. *Annual Review of Public Health*, 27: 103-124.
- Burrough PA & McDonnell RA. 1998. *Principles of geographical information systems*. Oxford University Press, Oxford.
- Caeiro S, Costa MH, Ramos TB, Fernandes F, Silveira N, Coimbra A, Medeiros G, Painho M. 2005. Assessing heavy metal contamination in Sado Estuary sediment: An index analysis approach. *Ecological Indicators*, 5: 151–169.
- Cairncross EK, John J, Zunckel M. 2007. A novel air pollution index based on the relative risk of daily mortality associated with short-term exposure to common air pollutants. *Atmos. Environ*, 41: 8442–8454.
- Carreras HA, Wannaz ED, Pignata ML. 2009. Assessment of human health risk related to metals by the use of biomonitors in the province of Cordoba Argentina. *Environmental Pollution*, 157 (1): 117-122.
- Carstairs V, Morris R. 1990. Deprivation and health in Scotland. *Health Bull*, 48 (4): 162-175
- Caskey FJ, Kramer A, Elliott RF, Stel VS, Covic A *et al.* 2011. Global variation in renal replacement therapy for end-stage renal disease. *Nephrology Dialysis Transplantation*, 26 (8): 2604-10.
- Catalyud A & Barreno E. 2004. Response to ozone in two lettuce varieties on chlorophyll a fluorescence, photosynthetic pigments and lipids peroxidation. *Plant Physiology and Biochemistry*, 42: 549-555.
- Caudeville J, Bonnard R, Boudet C, Denys S, Govaert G, Cicilella A. 2012. Development of a spatial stochastic multimedia exposure model to assess population exposure at a regional scale. *Science of the Total Environment*, 432: 297–308.

- Chaix B, Gustafsson S, Jerrett M, Kristersson H, Lithman T, Boalt A, Merlo J. 2006. Children's exposure to nitrogen dioxide in Sweden: investigating environmental injustice in an egalitarian country. *Journal of Epidemiology and Community Health*, 60 (3): 234-241.
- Chang MJ, Naworal JD, Walker K, Connell CT. 2003. Investigations on the direct introduction of cigarette smoke for trace elements analysis by inductively coupled plasma mass spectrometry. *Spectrochimica Acta Part B*, 58: 1979–1996.
- Chapman P. 1990. The sediment quality Triad approach to determining pollution-induced degradation. *Sci. Total Environ*, 97/98: 815–825.
- Chen J, Roth RE, Naito AT, Lengerich EJ and MacEachren AM. 2008. Geovisual analytics to enhance spatial scan statistic interpretation: an analysis of U.S. cervical cancer mortality. *International Journal of Health Geographics*, 7: 57.
- Chen TB, Zheng YM, Lei M, Huang ZC, Wu HT, Chen H, Fan KK, Yu K, Wu X, Tian QZ. 2005. Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China. *Chemosphere*, 60: 542–551.
- Cheng JL, Shi Z, Zhu YW. 2007. Assessment and mapping of environmental quality in agricultural soils of Zhejiang Province, China. *Journal of Environmental Sciences*, 19: 50–54.
- Chery L & Cattani A. 2003. ADES : la banque de données nationale sur les eaux souterraines. *La houille blanche*, 2: 5p.
- Chiba M & Masironi R. 1992. *Bull. World Health Organ*, 70: 269 – 275.
- Chouraki V, Savoye G, Dauchet L *et al.* 2011. The changing pattern of Crohn's disease incidence in northern France: a continuing increase in the 10- to 19-year-old age bracket (1988–2007). *Aliment Pharmacol Ther*, 33 (10): 1133–1142.
- Cislaghi C & Nimis PL. 1997. Air pollution and lung cancer. *Nature*, 387: 463-464.
- Clarke KC. 2001. *Getting Started with Geographic Information Systems* (3rd ed.). Upper Saddle River, NJ: Prentice Hall. 352p.
- Clayton D & Kaldor J. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, 43: 671-681.
- Cloquet C, De Muyneck D, Signoret J, Vanhaecke F. 2009. Urban/peri-urban aerosol survey by determination of the concentration and isotopic composition of Pb collected by transplanted lichen *Hypogymnia physodes*. *Environmental Science and Technology*, 43 (3): 623-9.
- Conti ME & Cecchetti G. 2001. Biological monitoring: lichens as bioindicators of air pollution assessment: a review. *Environ Pollut*, 114: 471-92.
- Cosnes J. 2004. Tobacco and IBD: relevance in the understanding of disease mechanisms and clinical practice. *Best Pract Res Clin Gastroenterol*, 18: 481–496.
- Couchoud C, Guihenneuc C, Bayer F, Lemaitre V, Brunet P, Stengel B. 2012. Medical practice patterns and socio-economic factors may explain geographical variation of end-stage renal disease incidence. *Nephrology Dialysis Transplantation*, 27 (6): 2312-22.
- Couchoud C, Stengel B, Landais P, Aldigier J-C, De Cornelissen F, *et al.* 2006. The renal epidemiology and information network (REIN): a new registry for end-stage renal disease in France. *Nephrology Dialysis Transplantation*, 21: 411–418.
- Counil É, Cherni N, Kharrat M, Achour A, Trimech H. 2008. Trends of incident dialysis patients in tunisia between 1992 and 2001. *American Journal of Kidney Diseases*, 51 (3): 463-70.
- Cousens S, Smith PG, Ward H, Everington D, Knight RSG, Zeidler M, Stewart G, Smith-Bathgate EAB, Macleod MA, Mackenzie J, Will RG. 2001. Geographical distribution of variant Creutzfeldt-Jakob disease in Great Britain, 1994-2000. *The Lancet*, 357: 1002-1007.

- Cressie NA. 1993. *Statistics for Spatial Data* (revised edition). John Wiley & Sons, Inc., New York.
- Crous KI, Vandermeiren K, Ceulemans R. 2006. Physiological responses to cumulative ozone uptake in two white clover (*Trifolium repens* L. cv. Regal) clones with different ozone sensitivity. *Environmental and Experimental Botany*, 58(1-3): 169-179.
- Cuny D, Van Haluwyn C, Pesch R. 2001. Biomonitoring of trace elements in air and soil compartments along the major motorway in France. *Water, Air, and Soil Pollution*, 125: 273-289.
- Cuny D, Pignata ML, Kranner I, Beckett R. 2002. Biomarkers of pollution-induced oxidative stress and membrane damage in lichens. In : Nimis PL, Scheidegger C, Wolseley PA. *Monitoring with lichens – Monitoring lichens*. Kluwer, The Netherlands.
- Cuny D, Davranche L, Kempa M, Vandamme L, Pailleux N, Caous AS, Dossin M, Silvie JP, Van Haluwyn C. 2003. Comparison of lichen flora (between 1995 and 2002) in two cities in the north of France : contribution of nitrogen pollution ? In *Nitrogen in the Environment*, Nettlecombe Court, Taunton, Angleterre, 24-27/02/03.
- Cuny D, Davranche L, Thomas P, Kempa M, Van Haluwyn C. 2004. Spatial and temporal variations of trace element contents in *Xanthoria parietina* thalli collected in a highly industrialized area in Northern France as an element for a future epidemiological study. *Journal of Atmospheric Chemistry*, 49: 391-401.
- Cuny D, Rzepka MA, Van Haluwyn C. 2008. Introduction à la biosurveillance végétale et fongique de la qualité de l'air, concept et intérêt. *Spectra Analyse*, 264: 31-35.
- Cuny D, Quarre S, Devred I, Jeanne B, Reynaert A, Rzepka-Cuny MA. 2010. Évolution spatiale et temporelle des concentrations en Éléments Traces Métalliques de thalles de *Xanthoria parietina* récoltés sur la zone industrialo-portuaire de Dunkerque. *Air Pur*, 77 : 43-49.
- Cuny D. 2012. La biosurveillance végétale et fongique de la pollution atmosphérique : concepts et applications. *Annales Pharmaceutiques Françaises*, 70: 182—187.
- De Blomac F, Gal R, Hubert M, Richard D, Tourret C. 1994. *Arc/Info, concepts et applications en géomatique*. Paris, Hermès, 256 p.
- De Temmerman L, Bell JNB, Garrec JP, Klumpp A, Krause GHM, Tonneijck AEG. 2004. Biomonitoring of air pollutants with plants. In Klumpp A, Ansel W, Klumpp G, editors. *Urban Air Pollution, bioindication and environmental awareness*. Göttingen: Cuvillier Verlag: 337—74.
- Declercq C, Gower-Rousseau C, Vernier-Massouille G, Salleron J, Baldé M *et al.* 2010. Mapping of inflammatory bowel disease in northern France: spatial variations and relation to affluence. *Inflamm. Bowel Dis.*, 16 (5): 807-12.
- Declercq C. 2007. Géographie de la santé : le point de vue de l'épidémiologiste. Séminaire ORS Île de France « De la mesure des expositions à l'évaluation des conséquences pour la santé : le traitement spatialisé des données ». Paris, 7 septembre 2007. Communication orale.
- Decocq C. 2010. *Friches industrielles et pollutions historiques*. Lille Métropole Communauté Urbaine, 98p.
- Delzenne-Van Haluwyn C. 1973. Contribution à l'étude de la distribution des lichens épiphytes dans le nord de la France : application au problème de la pollution atmosphérique. Thèse de Doctorat Faculté de Pharmacie Lille2. 162p.
- Deng F, Minasny B, Knadel M, McBratney A, Heckrath G, Greve MH. 2013. Using vis-NIR spectroscopy for monitoring temporal changes in soil organic carbon. *Soil Sci*, 178 (8): 389-99.
- Directive 2006/118/CE du 12 décembre 2006 sur la protection des eaux souterraines contre la pollution et la détérioration.

- Directive Cadre sur l'Eau. 2000. Directive 2000/60/CE du Parlement européen et du Conseil établissant un cadre pour une politique communautaire dans le domaine de l'eau.
- DREAL. 2014a. L'environnement en Nord – Pas de Calais – Tome 1 Etat des lieux. 78p
- DREAL. 2014b. Plan de prévention de l'atmosphère Nord – Pas de Calais. 252p.
- Duigou N & Baize D. 2010. Nouvelle collecte nationale d'analyses d'éléments en traces dans les sols (horizons de surface) - (Cd, Cr, Cu, Hg, Ni, Pb, Se, Zn). Contrat ADEME. 284 p.
- Echaubard M. 1995. Les animaux comme bioindicateurs biologiques de pollutions. Marqueurs Biologiques de Pollution. Chinon, 335-338.
- Elliott P & Wartenberg D. 2004. Spatial Epidemiology: Current Approaches and Future Challenges. *Environmental Health Perspectives*, Vol 112: 998-1006.
- Elliott P, Wakefield JC, Best NG, Briggs DJ. 2000. Spatial epidemiology: methods and applications. In: Elliott P, Wakefield JC, Best NG, Briggs DJ, (dir.). *Spatial epidemiology: methods and applications*. Oxford: Oxford University Press, 3-14.
- Exley C. 2013. Human exposure to aluminium. *Environmental Science: Processes & Impacts*, 15: 1807-1816.
- Fairey R, Long ER, Roberts CA, Anderson BS, Phillips BM, Hunt JW, Puckett HR, Wilson CJ. 2001. An evaluation of methods for calculating mean sediment quality guideline quotients as indicators of contamination and acute toxicity to amphipods by chemical mixture. *Environ. Toxicol. Chem.*, 20: 2276–2286.
- Fernandez-Falcon E, Strittholt JR, Alobaida AI, Schmidley RW, Bossler JD, Ramirez JR. 1993. A Review of Digital Geographic Information Standards for the State/Local User. *URISA Journal*, Vol 5 (2): 21-27.
- Ferry BW, Baddeley MS, Hawksworth DL. 1973. *Air Pollution and Lichens*. Athlone Press, London: 389 pp.
- Finzgar N, Jez E, Volgar D, Lestan D. 2014. Spatial distribution of metal contamination before and after remediation in the Meza Valley, Slovenia. *Geoderma*, 217–218: 135–143.
- Freitas MC, Reis MA, Alves LC, Wolterbeek HT. 1999. Distribution in Portugal of some pollutants in the lichen *Parmelia sulcata*. *Environmental Pollution*, 106 (2): 229-235.
- Fritz CE, Schuurman N, Robertson C, Lear S. 2013. A scoping review of spatial cluster analysis techniques for point-event data. *Geospatial Health*, 7 (2): 183-198.
- Fuga A, Saiki M, Marcelli MP, Saldiva PHN. 2008. Atmospheric pollutants monitoring by analysis of epiphytic lichens. *Environmental Pollution*, 151: 334-340.
- Gadal S. 2008. *Les systèmes d'informations géographiques libres*. Agence universitaire de la francophonie. Ouvrage numérique.
- Gadsdon S, Dagley JR, Wolseley PA, Power SA. 2010. Relationships between lichen community composition and concentrations of NO₂ and NH₃. *Environmental Pollution*, 158: 2553-2560.
- Gailey FAY & Lloyd OLI. 1993. Spatial and temporal patterns of airborne metal pollution: The value of low technology sampling to an environmental epidemiology study. *Sci. Total Environ*, 113: 201–219.
- Garrec J-P, Van Haluwyn C. 2002. *Biosurveillance végétale de la qualité de l'air*. Tec & Doc, Paris.
- Garty J. 1993. Lichens as biomonitors for heavy metal pollution. in B. Market (ed.), *Plants as Biomonitors, Indicators for Heavy Metals in the Terrestrial Environment*, VCH, Weinheim: 193–263.
- Garty J. 2001. Biomonitoring atmospheric heavy metals with lichens: Theory and application. *Crit. Rev. Plant Sci*, 20(4): 309–371.

- Gaveriaux J-P. 1995. Lichens et la bioindication de la qualité de l'air (les). Avion : Lycée Pablo Picasso. 52 p.
- Génin M, Duhamel A, Preda C, Fumery M, Savoye G *et al.* 2013. Space-time clusters of Crohn's disease in northern France. *J Public Health*, DOI 10.1007/s10389-013-0580-9.
- Gerdol R, Marchesini R, Iacuminb P, Brancaleoni L. 2014. Monitoring temporal trends of air pollution in an urban area using mosses and lichens as biomonitors. *Chemosphere*, 108: 388-395.
- Giljanovic NS. 1999. Water quality evaluation by index in Dalmatia. *Water Res.*, 33(16): 3423–3440.
- Glaz J, Pozdnyakov V, Wallenstein S. 2009. *Scan Statistics Methods and Applications*. Statistics for Industry and Technology, Birkhauser, 394p.
- Glickman TS. 1994. Measuring environmental equity with geographic information systems. *Renewable Resources Journal*, 12 (3): 17-21.
- Golge M, Yenilmez F, Aksoy A. 2013. Development of pollution indices for the middle section of the Lower Seyhan Basin (Turkey). *Ecological Indicators*, 29: 6–17.
- Gombert S, Asta J, Seaward MRD. 2004. Assessment of lichen diversity by index of atmospheric purity (IAP), index of human impact (IHI) and other environmental factors in an urban area (Grenoble, southeast France). *Sci Total Environ*, 324: 183-99.
- Gomez-Rubio G, Ferrandiz-Ferragud J, Lopez-Quilez A. 2005. Detecting clusters of disease with R. *J Geograph Syst*, 7: 189–206.
- Goovaerts P. 1997. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Goria S, Stempfelet M, de Crouy-Chanel P. 2011. Introduction aux méthodes statistiques et aux systèmes d'information géographique en santé environnement – Application aux études écologiques. Saint-Maurice: Institut de veille sanitaire. 65 p. Disponible à partir de l'URL : <http://www.invs.sante.fr>.
- Gower-Rousseau C, Salomez JL, Dupas JL, Marti R, Nuttens M-C *et al.* 1994. Incidence of inflammatory bowel disease in northern France (1988-1990). *Gut.*, 35 (10): 1433-8.
- Grace BS, Clayton P, Cass A, McDonald SP. 2012. Socio-economic status and incidence of renal replacement therapy: A registry study of Australian patients. *Nephrology Dialysis Transplantation*, 27 (11): 4173-80.
- Gratton Y. 2002. Le krigeage : la méthode optimale d'interpolation spatiale. *Les articles de l'Institut d'Analyse Géographique*, 4p.
- Greenland S & Morgenstern H. 1989. Ecological bias, confounding and effect modification. *Int. J. Epidemiol*, 18: 269–74.
- Greenland S. 1992. Divergent biases in ecologic and individual level studies. *Stat. Med*, 11: 1209–23.
- Guéguen F, Stille P, Geagea ML, Boutin R. 2012. Atmospheric pollution in an urban environment by tree bark biomonitors – Part I: Trace element analysis. *Chemosphere*, 86: 1013–1019.
- Hakanson L. 1980. An ecological risk index for aquatic pollution control. A sedimentological approach. *Water Res*, 14: 975–1001.
- Hampe J, Cuthbert A, Croucher P, Mirza M, Mascheretti S *et al.* 2001. Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations. *The Lancet*, 357 (9272): 1925 – 1928.
- Haneuse S & Wakefield J. 2008. Geographic-based ecological correlation studies using supplemental case-control data. *Stat Med*, 27 (6): 864-87.

- Hardouin M. 1983. Les lichens épiphytes de la région Sud-Est de Lille: généralités, distribution et estimation de la pollution atmosphérique acide. Thèse de Doctorat en Pharmacie. 137p.
- Havard S, Deguen S, Bodin J, Louis K, Laurent O, Bard D. 2008. A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science and Medicine*, 67(12): 2007-2016.
- Hawksworth DL & Rose F. 1970. Qualitative scale for estimating sulphur dioxide air pollution in England and Wales using epiphytic lichens. *Nature*, 227: 145-8.
- Heitland P & Koster HD. 2006. Biomonitoring of 37 trace elements in blood samples from inhabitants of northern Germany by ICP-MS. *J Trace Elem Med Biol*, 20(4): 253-62.
- Hellström L, Elinder C-G, Dahlberg B, Lundberg M, Järup L, Persson B, Axelson O. 2001. Cadmium exposure and end-stage renal disease. *Am J Kidney Dis*, 38: 1001-8.
- Hengl T. 2007. A Practical Guide to Geostatistical Mapping of Environmental Variables. Office for Official Publication of the European Communities, Luxembourg, 143p.
- Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. 1996. Childhood leukemia in Sweden: Using GIS and a spatial scan statistic for cluster detection. *Statistics in Medicine*, 15: 707-715.
- Hommel K, Rasmussen S, Kamper A, Madsen M. 2010. Regional and social inequalities in chronic renal replacement therapy in Denmark. *Nephrology Dialysis Transplantation*, 25 (8): 2624-32.
- Horton RK. 1965. An index number system for rating water quality. *J. Water Pollut. Control Fed.*, 37(3): 300-306.
- Huang Y, Cai X, Zhang J, Mai W, Wang S, *et al.* 2014. Prehypertension and incidence of ESRD: A systematic Review and Meta-analysis. *American Journal of Kidney Diseases*, 63 (1): 76-83.
- IARC. 2013. Outdoor air pollution a leading environmental cause of cancer deaths. Press release N°221. 4p.
- Ikem A, Egiebor NO, Nyavor K. 2003. Trace elements in water, fish and sediment from Tuskegee Lake, Southeastern USA. *Water Air Soil Pollut*, 149: 51-75.
- INSEE. 2010. Guide du secret statistique. 26p.
- INSEE. 2013. Le développement durable en Nord – Pas de Calais. 180p.
- Jantchou P, Monnet E, Carbonnel F. 2006. Les facteurs d'environnement dans la maladie de Crohn et la rectocolite hémorragique (tabac et appendicectomie exclus). *Gastroenterol Clin Biol*, 30: 859-867.
- Jany-Catrice F & Zotti R. 2008. Les régions françaises face à leur santé sociale, contribution au débat. Institut pour le développement de l'information économique et sociale.
- Järup L. 2002. Cadmium overload and toxicity. *Nephrol Dial Transplant*, 17: 35-9.
- Joseph WD. 2002. Comparative assessment and harmonization of the US EPA air quality index (AQI) with related air quality and pollutant standard indices in other nations—phase 2. Work assignment 5-15 of Contract no. 68-D-98-030.
- Kakkar P & Jaffery FN. 2005. Biological markers for metal toxicity. *Environ Toxicol Pharmacol*, 19 (2): 335-49.
- King G. 1997. A Solution to the Ecological Inference Problem. Princeton, NJ: Princeton Univ. Press
- Krige DG. 1951. A statistical approach to some mine valuations problems at the Witwatersrand. *J. Chem. Metall. Min. Soc. S. Afr*, 52: 119-139.
- Krivoruchko K. 2012. Empirical Bayesian Kriging implemented in ArcGIS Geostatistical Analyst. *ArcUser*, 15 (4): 6-10.

- Kulldorff M & Nagarwalla N. 1995. Spatial disease clusters: Detection and inference. *Statistics in Medicine*, 14: 799–810.
- Kulldorff M. 1997. A spatial scan statistic. *Communications in statistics: theory and methods*, 26 (6): 1481–1496.
- Kulldorff M, Feuer EJ, Miller BA, Freedman LS. 1997. Breast cancer in northeastern United States: A geographical analysis. *American Journal of Epidemiology*, 146: 161-170.
- Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. 1998. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *Am J Public Health*, 88 (9): 1377–1380.
- Kulldorff M. 1999. An isotonic spatial scan statistic for geographical disease surveillance. *Journal of the National Institute of Public Health*, 48: 94–101.
- Kulldorff M. 2010. SaTScan User Guide for version 9.0. 109p.
- Kumar D & Alappat BJ. 2004. Selection of the appropriate aggregation function for calculating leachate pollution index. *Pract. Period. Hazard. Toxic Radioact. Waste Manage.*, 8(4): 253–264.
- Lagadic L, Caquet T, Amiard JC. 1997. Biomarqueurs en écotoxicologie : principes et définitions. Dans Lagadic L, Caquet, Amiard JC, Ramade F. *Biomarqueurs en écotoxicologie – Aspects fondamentaux*. Masson, Paris: 1-9.
- Landwehr JM & Deininger RA. 1976. A comparison of several water quality indexes. *J. Water Pollut. Control Fed.*, 48(5): 954–958.
- Larsen RS, Bell JNB, James PW, Chimonides PJ, Rumsey FJ, Tremper A, Purvis OW. 2007. Lichen and bryophyte distribution on oak in London in relation to air pollution and bark acidity. *Environmental Pollution*, 146: 332-340.
- Lauwerys RR, Bernard AM, Buchet J-P, Roels H. 2006. Assessment of the health impact of environmental exposure to cadmium: contribution of the epidemiologic studies carried out in Belgium. *Environ Res* 1993, 62: 200-6.
- Lemarchand O, Jeannée N. 2009. Méthodes de cartographie et approche géostatistique – La cartographie de la pollution au dioxyde d’azote en Alsace. *Cahier des thèmes transversaux ArScAn*, 9: 203-214.
- Lerner A. 2012. Aluminum as an adjuvant in Crohn's disease induction. *Lupus*, 21 (2): 231-238.
- Lerner, A. 2007. Aluminum is a potential environmental factor for Crohn’s disease induction: extended hypothesis. *Ann NY Acad Sci*, 1107: 329–345.
- Lerond M. 1981. Les lichens épiphytes en Normandie orientale, distribution, sociologie et application à la cartographie de la pollution atmosphérique, *Actes Muséum de Rouen*: 1-295.
- Li J & Heap AD. 2014. Spatial interpolation methods applied in the environmental sciences: a review. *Environmental Modelling & Software*, 53: 173-189.
- Li X, Wang J, Yang W, Li Z, Lai S. 2011. A spatial scan statistic for multiple clusters. *Math Biosci*, 233 (2): 135-42.
- Llop E, Pinho P, Matos P, Pereira MJ, Branquinho C. 2012. The use of lichen functional groups as indicators of air quality in a Mediterranean urban environment. *Ecol Indicator*, 13: 215-21.
- Loi sur l'air et l'utilisation rationnelle de l'énergie. 1996. Loi N° 96-1236 du 30 décembre 1996 sur l'air et l'utilisation rationnelle de l'énergie.
- Long ER, MacDonald DD. 1998. Recommended uses of empirically derived, sediment quality guidelines for marine and estuarine ecosystems. *Hum. Ecol. Risk Assess.*, 4: 1019–1039.

- Loppi S & Nascimbene J. 2010. Monitoring H₂S air pollution caused by the industrial exploitation of geothermal energy: The pitfall of using lichens as bioindicators. *Environmental Pollution*, 158: 2635-2639.
- Loppi S, Frati L, Paoli L, Bibabli V, Rossetti C, Bruscoli C, Corsini A. 2004. Biodiversity of epiphytic lichens and heavy metal contents of *Flavoparmelia caperata* thalli as indicators of temporal variations of air pollution in the town of Montecatini Terme (central Italy). *Science of the Total Environment*, 326: 113–122.
- Lora CM, Gordon EJ, Sharp LK, Fischer MJ, Gerber BS, Lash JP. 2011. Progression of CKD in hispanics: Potential roles of health literacy, acculturation, and social support. *American Journal of Kidney Diseases*, 58 (2): 282-90.
- Macoveanu M. 2005. Methods and techniques of ecological impact assessment (in Romanian). Iasi, Romania: Ecozone Publishing House, 2nd Edition.
- Madejon P, Maranon T, Murillo JM. 2006. Biomonitoring of trace elements in the leaves and fruits of wild olive and holm oak trees. *Science of the Total Environment*, 355: 187– 203.
- Maheswaran R, Payne N, Meechan D, Burden RP, Fryers PR *et al.* 2000. Socioeconomic deprivation, travel distance, and renal replacement therapy in the Trent region, United Kingdom 2000: An ecological study. *Journal of epidemiology and community health*, 57 (7): 523-4.
- Manier N, Deram A, Lecurieux F, Marzin D. 2009. Comparison between new wild plant *Trifolium repens* and *Vicia faba* on their sensitivity in detecting the genotoxic potential of heavy metal solutions and heavy metal-contaminated soils. *Water, Air, and Soil Pollution*, 202 (1-4): 343-352.
- Matheron G. 1962. *Traité de géostatistique appliquée*, Tome I. Mémoires du Bureau de Recherches Géologiques et Minières, No.14. Editions Technip, Paris.
- Matheron G. 1963a. Principles of geostatistics. *Economic Geology*, 58: 1246–1266.
- Matheron G. 1963b. *Traité de géostatistique appliquée*, II : Le Krigeage. Mémoires du Bureau de Recherches Géologiques et Minières, No.24. Editions B. R. G. M., Paris.
- Maynard RL, Coster SM. 1999. Informing the public about air pollution. In: Holgate ST *et al.* (Eds.) *Air Pollution and Community Health*. Academic Press, Sydney: 1019–1033. (ISBN: 0-12-352335-4).
- McLeod H, Langford IH, Jones AP, Stedman JR, Day RJ, Lorenzoni I, Batement IJ. 2000. The relationship between socio-economic indicators and air pollution in England and Wales: implications for environmental justice. *Regional Environmental Change*, 1 (2): 78-85.
- Merwade V. 2009. Effect of spatial trends on interpolation of river bathymetry. *Journal of Hydrology*, 371: 169–181.
- Misik M, Ma TH, Nersesyanyan A, Monarca S, Kim JK, Knasmueller S. 2011. Micronucleus assays with *Tradescantia pollen* tetrads: an update. *Mutagenesis*, 26(1): 215-21.
- Moran P. 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2): 17–23.
- Morgenstern H. 1995. Ecologic studies in epidemiology: concepts, principles, and methods. *Annu. Rev. Public Health*, 16: 61–81.
- Moschino V, Delaney E, Da Ros L. 2012. Assessing the significance of *Ruditapes philippinarum* as a sentinel for sediment pollution: Bioaccumulation and biomarker responses. *Environmental Pollution*, 171: 52-60.
- Muntner P, Coresh J, Powe NR, Klag MJ. 2003. The contribution of increased diabetes prevalence and improved myocardial infarction and stroke survival to the increase in treated end-stage renal disease. *Journal of the American Society of Nephrology*, 14 (6): 1568-77.

- Murena F. 2004. Measuring air quality over large urban areas: development and application of an air pollution index at the urban area of Naples. *Atmospheric Environment*, 38: 6195–6202.
- National Research Council, Committee on the institutional means for assessment of risks to public health. 1983. Risk assessment in the Federal government: managing the process. National Academy Press Washington, D.C. 191p.
- Nerich V, Monnet E, Etienne A *et al.* 2006. Geographical Variations of Inflammatory Bowel Disease in France: A Study Based on National Health Insurance Data. *Inflamm Bowel Dis*, 12 (3): 218-226.
- Niober EA, Richardson DHS, Tomassini FD. 1978. Mineral Uptake and Release by Lichens: An Overviewed. *Bryologist*, 81: 226-246.
- Nisse C. 2010. Imprégnation par des métaux de la population générale. Valeurs d'imprégnation par les métaux de la population générale du Nord - Pas de Calais : établissement de valeurs de base pour l'interprétation des indicateurs d'exposition en milieu professionnel. Rapport final, 109p.
- Note de service DGS/EA4 n° 2010-241 du 02/07/10 relative aux données du contrôle sanitaire des eaux destinées à la consommation humaine du système d'information en santé-environnement sur les eaux (SISE-Eaux).
- Nuckols JR, Ward MH, Jarup L. 2004. Using geographic information systems for exposure assessment in environmental epidemiology studies. *Environ Health Perspect*, 112 (9): 1007-15.
- Nylander W. 1866. Les lichens du Jardin du Luxembourg. *Bulletin de la Société botanique de France*, 13: 364-372.
- Ocelli F, Cuny M-A, Devred I, Deram A, Quarré S, Cuny D. 2013. Étude de l'imprégnation de l'environnement de trois bassins de vie de la région Nord-Pas-de-Calais par les éléments traces métalliques. *Pollution atmosphérique* [En ligne], 220 mis à jour le : 13/01/2014, URL : <http://lodel.irevues.inist.fr/pollution-atmospherique/index.php?id=2497>.
- Ocelli F, Deram A, Génin M, Cuny D, Glowacki F-X and Néphronor Network. Mapping end-stage renal disease (ESRD): spatial variations on small area level in France, and association with deprivation, earlier renal replacement care and related diseases. *Plos One* (soumis le 13.02.2014).
- Ocelli F, Bavdek R, Deram A, Hellequin A-P, Cuny M-A, Cuny D and Zwarterook I. Using lichen biomonitoring to assess environmental justice over neighborhoods in an industrial area of Northern France. *Environmental Health Perspectives* (A soumettre).
- OMS. 1946 Préambule à la Constitution de l'Organisation mondiale de la Santé, tel qu'adopté par la Conférence internationale sur la Santé, New York, 19-22 juin 1946; signé le 22 juillet 1946 par les représentants de 61 Etats. Actes officiels de l'Organisation mondiale de la Santé, n°. 2, p. 100 et entré en vigueur le 7 avril 1948.
- OMS. 2011. Burden of disease attributable to outdoor air pollution. Geneva
- OMS. 2012. Environmental health inequalities in Europe. Copenhagen, OMS Europe. 190p.
- OMS. 2014. Burden of disease from the joint effects of Household and Ambient Air Pollution for 2012. Geneva.
- Openshaw S, Charlton M, Wymer C, Craft AW. 1987. A mark I geographical analysis machine for the automated analysis of point data sets. *International Journal of Geographical Information Systems*, 1: 335–358.
- ORS Nord – Pas de Calais. 2011. Indice de défaveur sociale de Townsend par commune en 2006. Disponible à partir de www.orsnpdc.org.
- Ott W. 1978. Environmental Indices—Theory and Practice. Ann Arbor Science, Michigan, USA, 371p.

- Otvös E, Pazmandi T, Tuba Z. 2003. First national survey of atmospheric heavy metal deposition in Hungary by the analysis of mosses. *The Science of the Total Environment*, 309: 151-160.
- Pascal M, Corso M, Chanel O, Declercq C, Badaloni C, Cesaroni G, Henschel S, Maister K, Haluza D, Martin-Olmedo P, Medina S. 2013. Assessing the public health impact of urban air pollution in 25 European cities: results of the Aphekom project. *Science of the Total Environment*, 449: 390-400.
- Paustenbach D & Galbraith D. 2006a. Biomonitoring and biomarkers: exposure assessment will never be the same. *Environ Health Perspect*, 114(8): 1143-9.
- Paustenbach D & Galbraith D. 2006b. Biomonitoring: is body burden relevant to public health? *Regul Toxicol Pharmacol*, 44(3): 249-61.
- Pedigo A, Aldrich T, Odoi A. 2011. Neighborhood disparities in stroke and myocardial infarction mortality: a GIS and spatial scan statistics approach. *BMC Publ Health*, 11: 644.
- Piantadosi S, Byar DP, Green SB. 1988. The ecological fallacy. *Am. J. Epidemiol*, 127: 893–904.
- Pignata ML, Gudino GL, Wannaz ED, Pla RR, Gonzalez CM, Carreras HA, Orellana L. 2002. Atmospheric quality and distribution of heavy metals in Argentina employing *Tillandsia capillaris* as a biomonitor. *Environmental Pollution*, 120: 59-68.
- Pineton de Chambrun G, Body-Malapel M, Frey-Wagner I, Djouina M, Deknuydt F *et al.* 2014. Aluminum enhances inflammation and decreases mucosal healing in experimental colitis in mice. *Mucosal Immunology*, 7 (3): 589-601.
- PNSE1. 2004. Plan National Santé Environnement 2004-2008. 88p.
- PNSE2. 2009. Deuxième Plan National Santé-Environnement 2009-2013. 67p.
- Pornon H. 2007. Bilan et perspectives de 20 années de Géomatique. *Géomatique Expert*, 57: 36-46.
- PRSE2. 2011. Plan Régional Santé Environnement Nord – Pas de Calais 2011-2014. 63p.
- Prüss-Üstün A & Corvalán C. 2006. Preventing disease through healthy environments: Towards an estimate of the environmental burden of disease. WHO Library Cataloguing-in-Publication Data. 104p.
- Pykh YA, Kennedy ET, Grant WE. 2000. An overview of systems analysis methods in delineating environmental quality indices. *Ecol. Model.*, 130: 25–38.
- Qiao Q, Huang B, Zhang, C, Piper JDA, Pan Y, Sun Y. 2013. Assessment of heavy metal contamination of dustfall in northern China from integrated chemical and magnetic investigation. *Atmospheric Environment*, 74: 182-193.
- Qingjie G, Jun D, Yunchuan X, Qingfei W, Liqiang Y. 2008. Calculating Pollution Indices by Heavy Metals in Ecological Geochemistry Assessment and a Case Study in Parks of Beijing. *Journal of China University of Geosciences*, 19 (3): 230–241.
- Rai R & Agrawal M. 2008. Evaluation of physiological and biochemical responses of two rice (*Oryza sativa* L.) cultivars to ambient air pollution using open top chambers at a rural site in India. *Science of the Total Environment*, 407 (1): 679-691.
- Rajagopal B, Shrihari S, Dwarakish GS. 2010. Global Water Quality Indices for River Gurpur, Karnataka State, India. *International Journal of Earth Sciences and Engineering*, 3 (6): 833-840.
- Remy C. 2011. Utilisation de la modification de la biodiversité lichénique sur résineux pour détecter la pollution azotée d'origine routière. *Pollut Atmos, Numéro spécial « Biosurveillance de la pollution atmosphérique »*: 31-4.

- Rican S. 2007. Représentation cartographique des données sanitaires. Séminaire ORS Île de France « De la mesure des expositions à l'évaluation des conséquences pour la santé : le traitement spatialisé des données ». Paris, 7 septembre 2007. Communication orale.
- Richardson S. 1992. Modélisation statistique des variations géographiques en épidémiologie. *Rev Epidemiol Sante Publique*, 40: 33-45.
- Richardson S. 2000. Problèmes méthodologiques dans les études écologiques santé–environnement. *Life Sciences*, 323: 611–616.
- Ripley BD. 1981. *Spatial Statistics*. New York: Wiley.
- Roberts JD, Voss JD, Knight B. 2014. The association of ambient air pollution and physical inactivity in the United States. *Plos One*, 9 (3): e90143.
- Robertson C, Nelson TA, MacNab YC, Lawson AB. 2010. Review of methods for space–time disease surveillance. *Spatial and Spatio-temporal Epidemiology*, 1: 105–116.
- Roderick P, Clements S, Stone N, Martin D, Diamond I. 1999. What determines geographical variation in rates of acceptance onto renal replacement therapy in England? *Journal of Health Services Research and Policy*, 4 (3): 139-46.
- Rojanschi V. 1991. Global estimation methodology for ecosystem state. *The Environment (in Romanian: Mediul înconjurător)*, 2: 45-52.
- Rzepka MA & Cuny D. 2008. Biosurveillance végétale et fongique des éléments traces métalliques atmosphériques. *Air Pur*, 75: 66–77.
- Rzepka MA & Cuny D. 2011. Recherches préliminaires sur la mise en place d'une biosurveillance végétale de la pollution atmosphérique en zone industrielle complexe : applications du ray-grass dans la zone industrialoportuaire de Dunkerque. *Acta Botanica Gallica*, 158(3): 387-400.
- Sabel CE, Boyle PJ, Löytönen M, Gatrell AC, Jokelainen M, Flowerdew R, Maasilta P. 2003. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. *American Journal of Epidemiology*, 157: 898-905.
- Sackett DL, Haynes RB, Guyatt GH, Tugwell P. 1985. *Clinical Epidemiology, a Basic Science for Clinical Medicine*. Little, Brown and Co, Boston, 1985.
- Salo H, Bucko MS, Vaahtovuori E, Limo J, Mäkinen J, Pesonen LJ. 2012. Biomonitoring of air pollution in SW Finland by magnetic and chemical measurements of moss bags and lichens. *Journal of Geochemical Exploration*, 115: 69–81.
- Sargaonkar, A & Deshpande V. 2003. Development of an overall index of pollution for surface water based on a general classification scheme in Indian context. *Environmental Monitoring and Assessment*, 89: 43-67.
- Sarmiento S, Wolterbeek HT, Verburg TG, Freitas MC. 2008. Correlating element atmospheric deposition and cancer mortality in Portugal: data handling and preliminary results. *Environmental Pollution*, 151: 314-351.
- Scerbo R, Possenti L, Lampugnani L, Ristori T, Barale R, Barghigiani C. 1999. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Livorno Province (Tuscany, Italy). *The Science of the Total Environment*, 241: 91-106.
- Scerbo R, Ristori T, Possenti L, Lampugnani L, Barale R, Barghigiani C. 2002. Lichen (*Xanthoria parietina*) biomonitoring of trace element contamination and air quality assessment in Pisa Province (Tuscany, Italy). *The Science of the Total Environment*, 286: 27-40.
- Shin PKS & Lam WKC. 2001. Development of a marine sediment pollution index. *Environ. Pollut.*, 113: 281–291.

- Silva DT, Meirelles ST, Moraes RM. 2012. Relationship between ozone, meteorological conditions, gas exchange and leaf injury in *Nicotiana tabacum* Bel-W3 in a sub-tropical region. *Atmospheric Environment*, 60: 211-216.
- Singh M, Müller G, Singh IB. 2002. Heavy metals in freshly deposited stream sediments of rivers associated with urbanization of the Ganga plain, India. *Water, Air, and Soil Pollution*, 141: 35–54.
- Singh RP, Nath S, Prasad SC, Nema AK. 2008. Selection of Suitable Aggregation Function for Estimation of Aggregate Pollution Index for River Ganges in India. *Journal of Environmental Engineering*, 134 (8): 689-701.
- Sloof JE & Wolterbeek HT. 1991. National trace-element air pollution monitoring survey using epiphytic lichens. *Lichenologist*, 23(2): 139-165.
- Smith DG. 1990. A better water quality indexing system for rivers and streams. *Water Res.*, 24(10): 1237–1244.
- Snow J. 1855. *On the Mode of Communication of Cholera*. 2nd ed. London: Churchill.
- Spehlmann ME, Begun AZ, Burghardt J, Lepage P, Raedler A, Schreiber S. 2008. Epidemiology of inflammatory bowel disease in a German twin cohort: results of a nationwide study. *Inflamm Bowel Dis*, 14 (7): 968-76.
- Stel VS, van de Luijngaarden MWM, Wanner C, Jager KJ. 2011. The 2008 ERA–EDTA Registry Annual Report—a précis. *NDT Plus*, 4: 1–13.
- Swamee PK & Tyagi A. 2000. Describing water quality with aggregate index. *J. Environ. Eng.*, 126(5): 451–455.
- Takahashi T & Miyajima Y. 2010. Effects of roads on alpine and subalpine plant species distribution along an altitudinal gradient on Mount Norikura, central, Japan. *Journal of Plant Research*, 123: 741-749.
- Tango T. 1995. A class of tests for detecting 'general' and 'focused' clustering of rare diseases. *Stat Med*, 14(21-22): 2323-34.
- Tango T. 2000. A test for spatial disease clustering adjusted for multiple testing. *Stat Med*, 19(2): 191-204.
- Thimonier A, Dupouey JL, Bost F, Becker M. 1994. Simultaneous eutrophication and acidification of a forest ecosystem in North-East France. *New Phytologist*, 126: 533-9.
- Tobler W. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46: 234–40.
- Tomlinson DL, Wilson JG, Harris CR, Jeffrey DW. 1980. Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgoländer Meeresunters*, 33: 566-575.
- Townsend P. 1987. Deprivation. *Journal of Social Policy*, 16 (2) : 125-146.
- Tupin P. 2014. Accès aux soins des patients sous CMU. Réunion thématique annuelle de la commission d'épidémiologie de la société de néphrologie. Paris, le 9 avril 2014 (Communication orale).
- US EPA. 1998. Air quality index reporting; proposed rules. *Federal Register*, 63 (236).
- US EPA. 1999. Air quality index reporting; final rule. *Federal Register*, 64 (149).
- Usami T, Koyama K, Takeuchi O, Morozumi K, Kimura G. 2000. Regional variations in the incidence of end-stage renal failure in Japan. *Journal of the American Medical Association*, 284 (20): 2622-4.
- Usero J, Gonzalez-Regalado E, Gracia I. 1996. Trace metals in the bivalve mollusc *Chamelea gallina* from the Atlantic Coast of Southern Spain. *Mar. Pollut. Bull.*, 32 (3): 305–310.

- Van Haluwyn C. 1993. Guide des lichens. Paris: Masson.
- Van Haluwyn C. 1994. Niveaux de la qualité de l'air sur le littoral Calais-Dunkerque ; évaluation en fonction de la diversité lichénique. Rapport SPPPI du Littoral : 57-74.
- Van Haluwyn C & Cottard S. 1995. Cartographie de la qualité de l'air de la communauté urbaine de Lille à partir de l'observation des lichens. Rapport AREMA LM. 14p.
- Van Haluwyn C & Cuny D. 1997. Cartographie de la pollution de l'air par certains métaux lourds sur le littoral Calais-Dunkerque. Rapport SPPPI du Littoral: 25-40.
- Van Haluwyn C. 1998. La bioindication : applications dans notre région. Air Pur, 55: 23-25.
- Van Haluwyn C, Cuny D, Garrec J-P. 2011. Introduction générale - Définitions, concepts et enjeux généraux de la biosurveillance de la qualité de l'air. Pollution Atmosphérique, Numéro spécial : 7-13.
- Viel JF, Floret N, Mauny F. 2005. Spatial and space-time scan statistics to detect low clusters of sex ratio. Environmental and Ecological Statistics, 12: 289-299.
- Vienneau D, de Hoogh K, Briggs D. 2009. A GIS-based method for modelling air pollution exposures across Europe. Science of the Total Environment, 408: 255–266.
- Vine M, Degnan D, Hanchette C. 1997. Geographic Information Systems: Their Use in Environmental Epidemiologic Research. Environmental Health Perspectives, 105 (6): 598-605.
- Volkova N, McClellan W, Klein M, Flanders D, Kleinbaum D, *et al.* 2008. Neighborhood poverty and racial differences in ESRD incidence. Journal of the American Society of Nephrology, 19 (2): 356-64.
- Wakefield JC & Salway R. 2001. A statistical framework for ecological and aggregate studies. Journal of the Royal Statistical Society, series A (164): 119-37.
- Wakefield JC. 2007. Disease mapping and spatial regression with count data. Biostatistics, 8: 158–83.
- Wakefield JC. 2008. Ecologic studies revisited. Annual Review of Public Health, 29: 75–90.
- Waller LA & Gotway CA. 2004. Applied spatial statistics for public health data. Wiley, Hoboken. 494p.
- Walter SD & Birnie SE. 1991. Mapping mortality and morbidity patterns: an international comparison. Int J Epidemiol, 20: 678–689.
- Wappelhorst O, Kühn I, Oehlmann J, Markert B. 2000. Deposition and disease: a moss monitoring project as an approach to ascertaining potential connections. The Science of the Total Environment, 249: 243-256.
- Ward MM. 2008. Socioeconomic status and the incidence of ESRD. American Journal of Kidney Diseases, 51 (4): 563-72.
- Wimmer F, Oberaigner W, Kramar R, Mayer G. 2003. Regional variability in the incidence of end-stage renal disease: An epidemiological approach. Nephrology Dialysis Transplantation, 18 (8): 1562-7.
- Wolterbeek B, Sarmiento S, Verburg T. 2010. Is there a future for biomonitoring of elemental air pollution? A review focused on a larger-scaled health-related (epidemiological) context. J Radioanal Nucl Chem, 286: 195–210.
- Wolterbeek HT & Verburg TG. 2004. Atmospheric metal deposition in a moss data correlation study with mortality and disease in the Netherlands. The Science of the Total Environment, 319: 53-64.
- Young EW, Mauger EA, Jiang K, Port FK, Wolfe RA. 1994. Socioeconomic status and end-stage renal disease in the united states. Kidney International, 45 (3): 907-11.

Zaharia C & Surpateanu M. 2006. Environmental impact assessment using the method of global pollution index applied for heat and power co-generation plant. Environ. Eng. Manag. J., 5: 1141–1152.

Zaharia C & Murarasu I. 2009. Environmental impact assessment induced by an industrial unit of basic chemical organic compounds synthesis using the alternative method of global pollution index. Environ. Eng. Manag. J., 8: 107–112.

Zhao S, Feng C, Yang Y, Niu J, Shen Z. 2012. Risk assessment of sedimentary metals in the Yangtze Estuary: New evidence of the relationships between two typical index methods. Journal of Hazardous Materials, 241– 242: 164– 172.

Sites internet :

ADES : www.ades.eaufrance.fr (consulté le 22.04.2014).

Agreste : www.agreste.agriculture.gouv.fr (consulté le 22.04.2014).

Artois-Picardie : www.eau-artois-picardie.fr (consulté le 22.04.2014).

ATMO Nord – Pas de Calais : www.atmo-npdc.fr (consulté le 22.04.2014).

BASIAS : www.basias.brgm.fr (consulté le 22.04.2014).

BASOL : www.basol.developpement-durable.gouv.fr (consulté le 22.04.2014).

CITEPA : www.citepa.org/fr (consulté le 16/01/2014).

Eau France : www.eaufrance.fr (consulté le 11/03/2014).

ESRI : www.esrifrance.fr (consulté le 08/04/2014).

Fédération ATMO France : www.atmo-france.org (consulté le 11/03/2014).

InfoTerre: www.infoterre.brgm.fr (consulté le 27.01.2014).

INDIQUASOL : www.indiquasol.gissol.fr (consulté le 23.04.2014).

INSEE : www.insee.fr (consulté le 22/04/2014).

MEDDE : www.installationsclassees.developpement-durable.gouv.fr (consulté le 22.04.2014).

ORS Nord – Pas de Calais. www.orsnpdc.org (consulté le 22.04.2014).

UK Air Quality Archive : www.airquality.co.uk (consulté le 11/03/2014).

SaTScan : www.satscan.org (consulté le 03/04/2014).

ANNEXES

TABLE DES ANNEXES :

Annexe I. Matrices de corrélation de Spearman des indicateurs sanguins, urinaires et lichéniques	II
Annexe II. Statut de contrôle et niveau de qualification des données dans la banque ADES	V
Annexe III. Cartographies spatio-temporelles de la contamination en aluminium dans les eaux souterraines sur la période 1990-2012 et pour les départements du Nord, du Pas de Calais, de la Somme et de la Seine-Maritime (ADES eau France).	VII

ANNEXE I. MATRICES DE CORRELATION DE SPEARMAN DES INDICATEURS SANGUINS, URINAIRES ET LICHENIQUES

Annexe I.a. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) et urinaires (u) pour l'ensemble des individus. Les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05.

	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu
Als	1.00	0.14	0.01	0.19	0.14	0.01	0.02	0.10	0.00	0.33	0.09	0.33	0.08	-0.05	0.01	0.00	0.20	-0.02	0.01	0.01	-0.04	0.04	-0.06	0.04
Ass	0.14	1.00	-0.05	0.16	0.33	0.45	0.00	-0.06	0.18	0.09	0.02	0.55	-0.05	0.59	0.07	0.04	0.00	-0.02	0.02	0.04	0.02	0.05	-0.09	0.14
Cds	0.01	-0.05	1.00	0.14	0.06	-0.13	0.00	0.08	0.23	0.19	0.03	0.35	0.08	-0.03	0.36	0.21	0.02	-0.04	-0.02	0.23	0.26	0.11	0.16	0.24
Cos	0.19	0.16	0.14	1.00	0.23	-0.03	0.29	0.19	-0.05	0.04	-0.12	0.38	0.11	0.07	0.09	0.42	0.00	0.08	0.07	0.30	-0.13	-0.04	-0.14	0.11
Crs	0.14	0.33	0.06	0.23	1.00	0.16	0.04	-0.08	0.12	0.23	-0.05	0.52	0.15	0.15	0.11	0.00	-0.13	-0.01	-0.08	0.11	0.02	0.15	-0.02	0.09
Hgs	0.01	0.45	-0.13	-0.03	0.16	1.00	0.01	-0.26	0.15	0.00	0.03	0.36	-0.09	0.35	0.03	-0.06	-0.14	0.18	0.00	-0.17	0.04	-0.01	-0.03	0.07
Mns	0.02	0.00	0.00	0.29	0.04	0.01	1.00	-0.04	-0.08	0.23	0.08	0.22	0.07	-0.03	0.03	0.22	0.03	0.04	0.04	0.07	-0.08	0.03	-0.05	0.03
Nis	0.10	-0.06	0.08	0.19	-0.08	-0.26	-0.04	1.00	-0.03	0.13	0.12	0.10	0.02	-0.09	0.00	0.06	0.03	-0.14	0.13	0.28	-0.01	-0.14	0.02	-0.01
Pbs	0.00	0.18	0.23	-0.05	0.12	0.15	-0.08	-0.03	1.00	-0.08	0.23	0.40	-0.01	0.09	0.20	-0.11	0.01	-0.07	0.00	-0.01	0.63	0.10	0.14	0.23
Vs	0.33	0.09	0.19	0.04	0.23	0.00	0.23	0.13	-0.08	1.00	0.23	0.61	0.21	0.12	0.37	0.06	0.49	0.17	-0.07	0.12	0.01	0.12	0.07	0.21
Zns	0.09	0.02	0.03	-0.12	-0.05	0.03	0.08	0.12	0.23	0.23	1.00	0.16	0.04	-0.02	0.06	-0.20	0.18	-0.06	0.26	0.02	0.06	-0.06	0.19	0.10
RIMs	0.33	0.55	0.35	0.38	0.52	0.36	0.22	0.10	0.40	0.61	0.16	1.00	0.10	0.29	0.23	0.20	0.02	0.05	0.02	0.17	0.26	0.15	0.04	0.33
Alu	0.08	-0.05	0.08	0.11	0.15	-0.09	0.07	0.02	-0.01	0.21	0.04	0.10	1.00	0.03	0.30	0.15	0.16	0.22	0.31	0.15	0.17	0.26	0.14	0.47
Asu	-0.05	0.59	-0.03	0.07	0.15	0.35	-0.03	-0.09	0.09	0.12	-0.02	0.29	0.03	1.00	0.10	0.18	0.12	0.04	0.07	0.14	0.17	0.19	0.17	0.33
Cdu	0.01	0.07	0.36	0.09	0.11	0.03	0.03	0.00	0.20	0.37	0.06	0.23	0.30	0.10	1.00	0.23	0.23	0.15	0.07	0.32	0.42	0.39	0.32	0.64
Cou	0.00	0.04	0.21	0.42	0.00	-0.06	0.22	0.06	-0.11	0.06	-0.20	0.20	0.15	0.18	0.23	1.00	0.20	0.24	0.04	0.59	0.23	0.23	0.17	0.47
Cru	0.20	0.00	0.02	0.00	-0.13	-0.14	0.03	0.03	0.01	0.49	0.18	0.02	0.16	0.12	0.23	0.20	1.00	0.05	0.11	0.27	0.20	0.26	0.12	0.42
Hgu	-0.02	-0.02	-0.04	0.08	-0.01	0.18	0.04	-0.14	-0.07	0.17	-0.06	0.05	0.22	0.04	0.15	0.24	0.05	1.00	0.18	0.11	0.24	0.13	0.16	0.37
Mnu	0.01	0.02	-0.02	0.07	-0.08	0.00	0.04	0.13	0.00	-0.07	0.26	0.02	0.31	0.07	0.07	0.04	0.11	0.18	1.00	0.18	0.12	-0.08	0.17	0.35
Niu	0.01	0.04	0.23	0.30	0.11	-0.17	0.07	0.28	-0.01	0.12	0.02	0.17	0.15	0.14	0.32	0.59	0.27	0.11	0.18	1.00	0.22	0.19	0.26	0.50
Pbu	-0.04	0.02	0.26	-0.13	0.02	0.04	-0.08	-0.01	0.63	0.01	0.06	0.26	0.17	0.17	0.42	0.23	0.20	0.12	0.22	1.00	0.38	0.39	0.62	0.62
Vu	0.04	0.05	0.11	-0.04	0.15	-0.01	0.03	-0.14	0.10	0.12	-0.06	0.15	0.26	0.19	0.39	0.23	0.26	0.13	-0.08	0.19	0.38	1.00	0.38	0.60
Znu	-0.06	-0.09	0.16	-0.14	-0.02	-0.03	-0.05	0.02	0.14	0.07	0.19	0.04	0.14	0.17	0.32	0.17	0.12	0.16	0.17	0.26	0.39	0.38	1.00	0.56
RIMu	0.04	0.14	0.24	0.11	0.09	0.07	0.03	-0.01	0.23	0.21	0.10	0.33	0.47	0.33	0.64	0.47	0.42	0.37	0.35	0.50	0.62	0.60	0.56	1.00

Annexe I.b. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour l'ensemble des individus. Les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05.

	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu
AlL	0.12	0.03	0.05	0.04	0.03	0.04	0.07	0.04	0.13	0.15	0.11	0.09	-0.04	-0.03	0.10	-0.01	-0.01	0.09	0.07	0.08	0.08	-0.02	0.11	0.12
AsL	0.09	-0.02	0.14	0.01	-0.08	-0.08	-0.04	0.02	0.13	0.27	0.08	0.06	-0.01	-0.07	0.12	0.06	0.08	0.11	0.03	0.11	0.16	0.02	0.11	0.17
CdL	0.17	-0.06	0.07	0.01	0.02	-0.03	0.00	0.04	0.11	0.25	0.06	0.07	0.10	-0.08	0.16	-0.03	0.06	0.10	0.09	0.10	0.13	-0.06	0.11	0.15
CoL	0.16	-0.04	0.09	0.08	0.00	-0.04	0.05	0.03	0.08	0.32	0.09	0.10	0.03	-0.07	0.13	0.09	0.04	0.10	0.05	0.12	0.13	0.02	0.12	0.18
CrL	0.13	0.02	-0.02	-0.01	0.02	-0.08	-0.06	0.10	0.10	-0.12	-0.09	-0.04	-0.06	-0.11	0.01	-0.16	0.00	0.04	0.03	0.04	0.02	-0.05	0.01	0.00
HgL	0.17	-0.09	0.10	0.08	-0.04	-0.06	0.05	0.03	0.08	0.31	0.08	0.07	0.01	-0.13	0.13	0.07	0.06	0.08	0.08	0.10	0.12	-0.02	0.08	0.14
MnL	0.20	-0.04	0.07	0.09	0.01	-0.02	0.07	0.03	0.09	0.32	0.14	0.09	0.00	-0.07	0.13	0.06	0.09	0.12	0.09	0.13	0.11	0.00	0.10	0.18
NiL	0.12	-0.18	0.09	0.04	-0.02	-0.14	0.05	0.06	0.03	0.21	0.04	0.00	0.02	-0.19	0.11	-0.01	0.02	0.03	0.04	0.12	0.06	-0.07	0.09	0.07
PbL	0.13	-0.03	0.12	0.08	0.00	-0.04	0.03	0.04	0.13	0.20	0.13	0.09	-0.01	-0.04	0.12	0.05	0.03	0.08	0.08	0.12	0.12	0.01	0.11	0.17
VL	0.18	-0.03	0.09	0.07	-0.01	-0.02	0.03	0.02	0.12	0.33	0.13	0.10	0.01	-0.05	0.13	0.06	0.08	0.11	0.06	0.12	0.13	0.00	0.11	0.19
ZnL	0.16	-0.03	0.11	0.10	0.02	-0.08	0.01	0.09	0.15	0.06	0.05	0.07	-0.07	-0.09	0.10	-0.02	0.00	0.04	0.09	0.10	0.10	-0.02	0.09	0.11
RIML	0.20	-0.03	0.08	0.07	0.00	-0.05	0.03	0.06	0.11	0.22	0.10	0.08	-0.01	-0.09	0.12	0.00	0.05	0.10	0.10	0.11	0.09	-0.03	0.09	0.14

Annexe I.c. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus fumeurs (A) et non-fumeurs (B). Les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05.

A	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu
AlL	0.28	0.12	-0.07	0.18	0.01	0.10	0.19	-0.02	0.07	0.16	0.16	0.09	-0.01	0.06	0.15	0.02	0.04	0.11	0.15	0.09	0.05	0.04	0.07	0.16
AsL	0.23	0.08	-0.01	0.11	0.09	0.03	0.07	-0.06	0.20	0.69	0.14	0.12	0.06	0.02	0.27	0.09	0.19	0.22	0.05	0.20	0.24	0.13	0.20	0.31
CdL	0.31	-0.07	0.04	0.01	0.00	-0.03	0.05	0.11	0.00	0.42	-0.01	-0.03	0.12	-0.09	0.28	0.03	0.08	0.15	0.23	0.23	0.16	0.03	0.16	0.22
CoL	0.31	0.02	0.01	0.12	0.09	0.08	0.14	-0.06	0.06	0.18	0.08	0.08	0.01	-0.03	0.23	0.10	0.12	0.13	0.13	0.18	0.14	0.05	0.11	0.22
CrL	0.34	0.12	-0.03	0.03	0.09	-0.08	-0.05	0.02	0.00	0.23	-0.11	0.00	0.02	-0.06	0.20	-0.14	0.00	0.23	0.13	0.11	0.03	0.14	0.01	0.10
HgL	0.33	-0.04	-0.02	0.10	0.07	-0.01	0.15	-0.02	0.05	0.37	0.12	0.05	0.03	-0.13	0.24	0.10	0.15	0.15	0.18	0.20	0.16	0.03	0.12	0.22
MnL	0.34	0.04	-0.06	0.14	0.07	0.10	0.14	-0.05	0.11	0.40	0.19	0.08	0.03	-0.02	0.21	0.02	0.18	0.16	0.17	0.13	0.12	0.03	0.10	0.21
NiL	0.28	-0.13	0.09	-0.02	0.00	-0.09	0.13	-0.02	-0.06	0.02	-0.04	-0.04	0.03	-0.17	0.24	0.03	0.10	0.12	0.20	0.23	0.11	0.05	0.11	0.17
PbL	0.25	0.01	-0.01	0.13	0.09	0.06	0.11	-0.08	0.11	0.41	0.17	0.08	0.02	-0.03	0.17	0.05	0.12	0.14	0.12	0.17	0.13	0.05	0.11	0.18
VL	0.33	0.00	-0.03	0.12	0.08	0.07	0.12	-0.03	0.14	0.43	0.16	0.08	0.05	-0.03	0.21	0.05	0.17	0.16	0.13	0.16	0.17	0.04	0.12	0.23
ZnL	0.28	-0.08	0.02	0.03	0.02	-0.04	0.06	0.03	0.03	0.01	0.07	-0.02	-0.03	-0.15	0.16	-0.04	0.05	0.09	0.19	0.17	0.10	0.03	0.06	0.11
RIML	0.38	0.02	-0.04	0.10	0.05	0.03	0.10	-0.01	0.07	0.44	0.15	0.04	0.03	-0.07	0.22	0.00	0.15	0.17	0.20	0.16	0.11	0.05	0.11	0.20
B	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu
AlL	0.01	-0.02	0.14	-0.03	0.04	0.01	0.01	0.08	0.16	0.15	0.08	0.08	-0.05	-0.07	0.07	-0.04	-0.05	0.08	0.04	0.07	0.10	-0.06	0.13	0.11
AsL	-0.01	-0.06	0.18	-0.02	-0.16	-0.11	-0.10	0.06	0.09	0.04	0.04	0.02	-0.05	-0.11	0.04	0.04	0.03	0.05	0.00	0.07	0.11	-0.03	0.06	0.10
CdL	0.07	-0.05	0.09	0.01	0.03	-0.03	-0.03	0.00	0.16	0.18	0.08	0.11	0.07	-0.08	0.08	-0.06	0.03	0.08	0.02	0.04	0.13	-0.10	0.07	0.12
CoL	0.07	-0.08	0.12	0.06	-0.04	-0.09	-0.01	0.07	0.10	0.35	0.10	0.09	0.04	-0.10	0.07	0.07	-0.01	0.08	0.01	0.09	0.12	-0.01	0.10	0.15
CrL	-0.01	-0.03	-0.01	-0.02	-0.01	-0.08	-0.05	0.16	0.14	-0.28	-0.08	-0.05	-0.10	-0.13	-0.08	-0.17	0.00	-0.05	-0.02	0.01	0.01	-0.14	0.01	-0.05
HgL	0.06	-0.11	0.13	0.07	-0.09	-0.08	-0.01	0.06	0.09	0.27	0.06	0.06	0.00	-0.12	0.06	0.05	0.00	0.04	0.01	0.05	0.09	-0.05	0.04	0.10
MnL	0.12	-0.08	0.17	0.06	-0.02	-0.07	0.01	0.08	0.08	0.32	0.11	0.09	-0.01	-0.10	0.08	0.07	0.03	0.10	0.04	0.13	0.12	-0.02	0.09	0.16
NiL	0.03	-0.19	0.08	0.07	-0.02	-0.15	0.00	0.08	0.06	0.26	0.07	0.00	0.00	-0.19	0.02	-0.04	-0.02	-0.01	-0.03	0.06	0.02	-0.14	0.06	0.01
PbL	0.07	-0.05	0.15	0.06	-0.05	-0.08	-0.03	0.10	0.13	0.08	0.11	0.10	-0.02	-0.04	0.08	0.05	-0.01	0.05	0.04	0.07	0.11	0.00	0.10	0.17
VL	0.09	-0.05	0.14	0.04	-0.05	-0.06	-0.03	0.04	0.11	0.28	0.11	0.11	-0.02	-0.07	0.07	0.05	0.02	0.08	0.02	0.10	0.13	-0.01	0.10	0.17
ZnL	0.09	-0.02	0.12	0.13	0.02	-0.10	-0.01	0.12	0.19	0.03	0.03	0.10	-0.09	-0.06	0.05	-0.01	-0.04	0.02	0.03	0.07	0.08	-0.05	0.09	0.11
RIML	0.09	-0.05	0.15	0.06	-0.02	-0.07	-0.01	0.09	0.14	0.16	0.08	0.09	-0.03	-0.10	0.05	0.00	0.00	0.06	0.04	0.08	0.09	-0.07	0.08	0.12

Annexe I.d. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus de moins de 40 ans (A) et de 40 ans et plus (B). Les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05.

A	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu
AlL	0.13	0.07	0.08	0.16	0.05	0.08	0.08	0.06	0.17	-0.08	0.05	0.14	0.03	0.02	0.17	0.07	-0.01	0.14	0.15	0.19	0.12	-0.02	0.14	0.24
AsL	0.17	0.02	0.09	0.07	-0.01	0.04	-0.04	-0.06	0.23	0.26	0.08	0.10	0.06	-0.05	0.19	0.06	0.03	0.12	0.06	0.15	0.18	0.07	0.11	0.23
CdL	0.27	-0.12	0.09	0.04	0.02	-0.02	0.03	0.04	0.11	0.39	0.01	0.05	0.13	-0.08	0.24	0.04	0.17	0.15	0.06	0.22	0.17	0.03	0.14	0.24
CoL	0.27	0.00	0.09	0.15	0.03	0.08	0.03	-0.03	0.18	0.20	0.04	0.14	0.09	-0.05	0.23	0.15	0.07	0.17	0.08	0.24	0.20	0.06	0.15	0.27
CrL	0.13	0.01	-0.03	0.05	0.06	0.01	-0.06	0.06	0.07	0.12	-0.06	0.01	-0.08	-0.06	0.01	-0.14	-0.02	0.17	0.11	0.08	0.07	0.07	0.06	0.05
HgL	0.24	-0.07	0.08	0.13	0.01	0.04	0.05	-0.03	0.17	0.24	0.03	0.07	0.08	-0.11	0.20	0.11	0.06	0.15	0.09	0.21	0.17	0.01	0.10	0.21
MnL	0.29	-0.02	0.07	0.17	0.04	0.07	0.09	-0.01	0.17	0.24	0.10	0.11	0.10	-0.04	0.23	0.13	0.14	0.16	0.11	0.25	0.17	0.04	0.13	0.29
NiL	0.20	-0.18	0.09	0.07	-0.01	-0.02	0.07	0.00	0.05	0.13	-0.02	0.00	0.01	-0.17	0.15	0.02	0.08	0.11	0.03	0.21	0.13	-0.05	0.15	0.13
PbL	0.24	-0.02	0.11	0.10	0.01	0.07	-0.04	0.01	0.26	0.14	0.14	0.12	0.04	-0.01	0.19	0.05	0.01	0.12	0.14	0.18	0.19	0.01	0.16	0.24
VL	0.30	-0.02	0.06	0.12	0.02	0.06	0.01	-0.02	0.21	0.31	0.10	0.11	0.09	-0.02	0.20	0.09	0.13	0.16	0.09	0.20	0.18	0.06	0.14	0.27
ZnL	0.27	-0.07	0.11	0.08	-0.01	0.02	-0.05	0.05	0.20	0.06	0.05	0.08	-0.05	-0.05	0.12	-0.03	0.02	0.14	0.12	0.14	0.21	0.01	0.16	0.18
RIML	0.29	-0.03	0.08	0.12	0.02	0.03	0.03	0.01	0.17	0.26	0.09	0.10	0.06	-0.06	0.18	0.05	0.11	0.17	0.12	0.20	0.16	0.02	0.13	0.25
B	Als	Ass	Cds	Cos	Crs	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu
AlL	0.10	-0.05	-0.01	-0.13	-0.03	-0.05	0.07	0.04	0.02	0.49	0.16	0.01	-0.14	-0.07	-0.01	-0.08	0.01	0.06	0.00	-0.05	0.02	-0.03	0.09	-0.02
AsL	0.03	-0.08	0.20	-0.05	-0.18	-0.20	-0.04	0.12	0.03	0.30	0.08	0.00	-0.09	-0.12	0.05	0.07	0.15	0.10	0.01	0.08	0.14	-0.03	0.11	0.10
CdL	0.04	-0.02	0.03	-0.06	0.00	-0.09	-0.04	0.05	0.11	0.27	0.11	0.07	0.05	-0.11	0.05	-0.11	-0.07	0.05	0.11	-0.04	0.09	-0.14	0.05	0.05
CoL	0.06	-0.08	0.06	0.00	-0.03	-0.16	0.06	0.08	0.02	0.49	0.16	0.07	-0.07	-0.11	0.04	0.02	0.04	0.02	0.01	0.00	0.07	-0.01	0.08	0.08
CrL	0.12	-0.04	0.00	-0.07	-0.07	-0.23	-0.02	0.17	0.05	-0.26	-0.16	-0.14	-0.01	-0.14	-0.03	-0.13	0.02	-0.02	-0.01	0.01	-0.07	-0.04	-0.05	-0.09
HgL	0.09	-0.10	0.09	0.01	-0.09	-0.17	0.06	0.10	0.04	0.42	0.14	0.08	-0.10	-0.15	0.06	0.02	0.04	0.00	0.07	-0.01	0.08	-0.04	0.08	0.08
MnL	0.12	-0.07	0.05	-0.01	-0.04	-0.11	0.05	0.09	0.00	0.46	0.18	0.07	-0.14	-0.10	0.02	0.00	0.05	0.07	0.07	0.01	0.06	-0.05	0.08	0.05
NiL	0.05	-0.16	0.06	0.00	-0.01	-0.26	0.03	0.11	0.05	0.37	0.12	0.01	-0.01	-0.22	0.07	-0.06	-0.04	-0.08	0.04	0.02	0.00	-0.08	0.01	0.01
PbL	0.05	-0.03	0.08	0.06	-0.01	-0.13	0.12	0.08	0.02	0.28	0.12	0.09	-0.10	-0.08	0.05	0.04	0.06	0.04	0.02	0.04	0.06	0.03	0.04	0.10
VL	0.08	-0.03	0.09	0.01	-0.03	-0.10	0.06	0.08	0.03	0.41	0.15	0.10	-0.11	-0.08	0.04	0.03	0.04	0.06	0.05	0.04	0.09	-0.06	0.07	0.09
Z																								

Annexe I.e. Matrice de corrélation (coefficient de Spearman) des indicateurs sanguins (s) / urinaires (u) et lichéniques (L) pour les individus en fonction des quatre groupes de risque (A : non-fumeur et moins de 40 ans ; B : non-fumeur et plus de 40 ans ; C : fumeur et moins de 40 ans ; D : fumeur et plus de 40 ans). Les valeurs en gras sont différentes de 0 à un niveau de signification alpha de 0.05.

A	Als	Ass	Cds	Cos	CrS	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu	
AIL	0.00	0.02	0.21	0.08	0.00	0.08	0.01	0.14	0.20	-0.20	-0.07	0.13	-0.04	0.02	0.14	0.09	-0.01	0.15	0.15	0.18	0.25	-0.03	0.17	0.27	
AsL	0.15	0.06	0.20	0.05	-0.10	0.04	-0.13	-0.01	0.28	-0.04	0.01	0.15	-0.05	-0.02	0.10	0.04	-0.01	0.09	0.03	0.02	0.22	0.02	0.04	0.17	
CdL	0.31	0.00	0.12	0.07	0.04	-0.03	-0.07	0.04	0.22	0.31	0.01	0.14	0.10	-0.01	0.19	0.01	0.20	0.14	-0.01	0.17	0.29	0.05	0.10	0.25	
CoL	0.25	0.03	0.24	0.13	-0.06	0.06	-0.09	0.03	0.29	0.13	0.00	0.19	0.09	-0.02	0.19	0.13	0.05	0.16	0.05	0.20	0.33	0.10	0.13	0.28	
CrL	0.07	0.04	-0.09	0.03	0.08	0.02	-0.15	0.16	0.15	-0.12	-0.05	0.06	-0.13	-0.01	-0.08	-0.15	0.08	0.05	0.13	0.07	0.20	-0.23	0.05	0.09	
HgL	0.16	-0.03	0.23	0.13	-0.08	0.08	-0.08	0.02	0.28	0.13	-0.03	0.12	0.04	-0.05	0.16	0.09	0.00	0.12	0.05	0.13	0.29	0.01	0.04	0.19	
MnL	0.25	-0.02	0.31	0.17	-0.02	0.03	0.01	0.06	0.24	0.07	0.01	0.17	0.04	-0.03	0.22	0.16	0.08	0.15	0.09	0.26	0.32	0.06	0.11	0.31	
NiL	0.18	-0.15	0.13	0.08	0.00	-0.03	-0.07	0.06	0.15	0.12	-0.01	0.04	0.02	-0.15	0.10	-0.02	0.08	0.03	-0.05	0.18	0.20	-0.11	0.12	0.09	
PbL	0.13	0.01	0.23	0.09	-0.05	0.06	-0.12	0.09	0.31	-0.09	0.07	0.20	0.02	0.04	0.14	0.07	-0.03	0.11	0.12	0.16	0.28	0.02	0.13	0.27	
VL	0.26	0.03	0.20	0.10	-0.06	0.05	-0.10	0.01	0.30	0.12	0.06	0.18	0.03	0.04	0.17	0.09	0.08	0.14	0.08	0.11	0.31	0.10	0.14	0.29	
ZnL	0.22	0.04	0.12	0.12	0.02	0.03	-0.16	0.09	0.27	-0.03	0.01	0.19	-0.04	0.02	0.08	0.01	0.01	0.13	0.07	0.14	0.33	0.01	0.17	0.24	
RIML	0.24	0.04	0.20	0.13	-0.01	0.04	-0.07	0.06	0.25	0.05	0.00	0.18	0.03	0.01	0.15	0.08	0.09	0.17	0.08	0.18	0.32	0.02	0.11	0.29	
B	Als	Ass	Cds	Cos	CrS	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu	
AIL	0.02	-0.09	-0.01	-0.14	0.04	-0.08	0.04	0.06	0.04	0.63	0.19	0.01	-0.08	-0.15	-0.05	-0.11	-0.08	0.06	-0.03	-0.03	-0.05	-0.07	0.10	-0.04	
AsL	-0.10	-0.17	0.16	-0.09	-0.23	-0.23	-0.08	0.11	-0.04	0.06	0.08	-0.08	-0.07	-0.20	-0.03	0.04	0.08	0.02	-0.02	0.10	0.03	-0.09	0.06	0.03	
CdL	-0.19	-0.11	0.01	-0.05	0.01	-0.05	0.00	-0.04	0.14	0.24	0.15	0.04	0.06	-0.17	-0.06	-0.12	-0.12	0.04	0.04	-0.09	-0.03	-0.20	0.03	-0.01	
CoL	-0.08	-0.13	0.04	0.00	-0.01	-0.17	0.04	0.06	0.01	0.54	0.19	0.03	-0.04	-0.16	-0.04	-0.02	-0.04	0.00	-0.04	-0.02	-0.03	-0.08	0.05	0.02	
CrL	-0.08	-0.14	0.00	-0.06	-0.14	-0.21	-0.08	0.21	0.06	0.21	0.06	-0.28	-0.14	-0.20	-0.07	-0.22	-0.11	-0.14	-0.08	-0.12	-0.10	-0.02	-0.18	-0.01	-0.18
HgL	-0.02	-0.14	0.06	0.01	-0.09	-0.18	0.02	0.06	0.01	0.44	0.15	0.04	-0.08	-0.17	-0.03	-0.01	0.00	-0.03	0.00	-0.05	-0.06	-0.11	0.04	0.02	
MnL	0.02	-0.12	0.05	-0.03	-0.02	-0.15	0.01	0.08	-0.02	0.57	0.20	0.04	-0.09	-0.16	-0.04	-0.01	-0.01	0.03	0.01	0.02	-0.04	-0.09	0.06	0.01	
NiL	-0.08	-0.19	0.05	0.05	-0.01	-0.21	0.05	0.08	0.06	0.44	0.14	-0.02	-0.05	-0.22	-0.05	-0.10	-0.11	-0.06	-0.04	-0.06	-0.13	-0.17	-0.03	-0.09	
PbL	0.06	-0.07	0.07	0.05	-0.03	-0.17	0.07	0.10	0.02	0.18	0.13	0.06	-0.09	-0.10	0.04	0.00	0.00	0.00	0.00	-0.02	-0.03	-0.02	0.04	0.08	
VL	-0.04	-0.09	0.07	-0.02	-0.03	-0.13	0.02	0.07	0.00	0.50	0.16	0.06	-0.08	-0.14	-0.03	0.01	-0.01	0.02	-0.02	0.04	-0.02	-0.11	0.05	0.04	
ZnL	0.00	-0.09	0.09	0.13	0.00	-0.20	0.11	0.16	0.13	0.07	0.04	0.01	-0.13	-0.15	0.01	-0.04	-0.08	-0.09	-0.01	0.02	-0.12	-0.09	0.01	-0.02	
RIML	-0.06	-0.13	0.06	-0.01	-0.05	-0.18	0.05	0.12	0.04	0.42	0.13	0.01	-0.09	-0.20	-0.08	-0.06	-0.07	-0.01	0.00	-0.01	-0.11	-0.12	0.04	-0.06	
C	Als	Ass	Cds	Cos	CrS	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu	
AIL	0.25	0.14	-0.07	0.26	0.11	0.08	0.13	-0.01	0.10	0.09	0.21	0.10	0.12	-0.01	0.16	0.03	-0.02	0.13	0.18	0.20	-0.10	0.02	0.08	0.17	
AsL	0.23	-0.05	0.01	0.08	0.14	0.02	0.07	-0.09	0.24	0.71	0.18	0.08	0.20	-0.08	0.36	0.09	0.11	0.18	0.07	0.32	0.13	0.12	0.24	0.34	
CdL	0.18	-0.28	0.03	0.03	-0.04	0.04	0.16	0.05	-0.02	0.60	0.00	-0.17	0.14	-0.19	0.27	0.10	0.11	0.18	0.19	0.31	0.01	0.05	0.23	0.22	
CoL	0.27	-0.02	0.01	0.16	-0.03	0.05	0.10	-0.10	0.07	0.08	0.10	0.02	0.10	-0.10	0.26	0.12	0.11	0.15	0.12	0.38	0.01	0.01	0.17	0.23	
CrL	0.24	-0.02	0.00	0.08	0.07	0.00	0.04	-0.02	-0.08	0.56	-0.09	-0.05	-0.03	-0.13	0.17	-0.12	-0.14	0.21	0.02	0.09	-0.12	0.15	0.06	0.01	
HgL	0.31	-0.10	-0.03	0.13	0.12	-0.05	0.13	-0.08	0.05	0.40	0.13	-0.04	0.15	-0.18	0.26	0.12	0.14	0.20	0.17	0.29	0.01	0.01	0.19	0.24	
MnL	0.32	0.00	-0.07	0.16	0.11	0.10	0.11	-0.08	0.14	0.63	0.24	0.01	0.17	-0.07	0.26	0.04	0.19	0.15	0.13	0.23	-0.03	0.02	0.14	0.23	
NiL	0.19	-0.20	0.08	0.04	-0.06	-0.03	0.21	-0.10	-0.09	0.06	-0.05	-0.14	0.02	-0.22	0.20	0.06	0.08	0.22	0.15	0.24	0.01	0.03	0.21	0.16	
PbL	0.35	-0.05	-0.04	0.12	0.08	0.03	0.03	-0.09	0.20	0.47	0.23	0.00	0.06	-0.07	0.23	0.01	0.06	0.11	0.11	0.25	0.05	0.01	0.17	0.17	
VL	0.32	-0.08	-0.02	0.11	0.12	0.04	0.08	-0.05	0.19	0.64	0.19	0.01	0.18	-0.08	0.23	0.06	0.19	0.17	0.11	0.25	0.04	0.03	0.17	0.24	
ZnL	0.30	-0.20	-0.02	0.01	-0.09	-0.04	0.05	-0.01	0.05	0.04	0.11	-0.19	-0.08	-0.17	0.13	-0.09	0.03	0.14	0.18	0.14	0.00	0.01	0.14	0.05	
RIML	0.37	-0.08	-0.05	0.11	0.08	-0.01	0.11	-0.06	0.07	0.69	0.20	-0.03	0.11	-0.16	0.22	0.00	0.14	0.17	0.16	0.23	-0.06	0.06	0.14	0.18	
D	Als	Ass	Cds	Cos	CrS	Hgs	Mns	Nis	Pbs	Vs	Zns	RIMs	Alu	Asu	Cdu	Cou	Cru	Hgu	Mnu	Niu	Pbu	Vu	Znu	RIMu	
AIL	0.35	0.03	-0.06	-0.08	-0.14	0.16	0.28	-0.07	-0.01	0.03	0.06	-0.01	-0.25	0.21	0.09	-0.01	0.21	0.08	0.07	-0.14	0.34	0.08	0.00	0.11	
AsL	0.25	0.19	0.04	0.08	-0.01	-0.01	0.05	0.05	0.11	0.58	0.07	0.07	-0.19	0.25	0.08	0.04	0.41	0.26	0.01	-0.07	0.41	0.06	0.06	0.22	
CdL	0.50	0.24	0.08	-0.07	0.08	-0.09	-0.13	0.23	0.07	0.32	-0.02	0.13	0.03	0.09	0.31	-0.14	0.06	0.06	0.28	0.03	0.48	-0.01	0.01	0.26	
CoL	0.38	0.08	0.01	-0.01	0.02	0.14	0.23	0.06	0.05	0.23	0.06	0.14	-0.20	0.13	0.16	0.01	0.23	0.06	0.10	-0.03	0.39	0.14	0.01	0.22	
CrL	0.47	0.28	-0.04	-0.08	0.17	-0.32	-0.30	0.10	0.02	-0.29	-0.25	0.01	0.14	0.16	0.17	-0.18	0.29	0.35	0.28	0.06	0.20	-0.06	-0.15	0.15	
HgL	0.34	-0.02	0.02	0.02	0.01	0.08	0.20	0.13	0.05	0.32	0.11	0.12	-0.19	-0.01	0.11	0.02	0.24	0.04	0.14	0.03	0.39	0.08	-0.03	0.19	
MnL	0.32	0.06	-0.05	0.06	0.00	0.16	0.20	0.03	0.08	0.06	0.14	0.09	-0.26	0.12	0.07	-0.05	0.22	0.18	0.18	-0.06	0.38	0.03	0.00	0.18	
NiL	0.34	-0.08	0.11	-0.16	0.09	-0.20	0.02	0.14	0.00	0.17	-0.02	0.07	0.07	-0.07	0.27	-0.07	0.24	-0.11	0.28	0.17	0.24	0.11	0.00	0.22	
PbL	0.14	0.10	-0.01	0.12	0.12	0.18	0.30	-0.01	0.06	0.32	0.06	0.19	-0.14	0.06	0.08	0.11	0.26	0.26	0.06	0.06	0.35	0.19	-0.02	0.30	
VL	0.30	0.11	-0.03	0.08	0.01	0.14	0.19	0.03	0.06	0.06	0.15	0.11	-0.17	0.12	0.11	0.00	0.14	0.16	0.14	-0.01	0.40	0.06	0.03	0.25	
ZnL	0.31	0.13	0.07	0.08	0.21	-0.04	0.07	0.13	0.05	0.03	0.00	0.22	0.02	-0.12	0.21	-0.02	0.23	-0.03	0.21	0.23	0.33	0.12	-0.02	0.32	
RIML	0.42	0.12	0.01	0.05	0.01	0.06	0.06	0.10	0.05	-0.03	0.08	0.08	-0.14	0.09											

ANNEXE II. STATUT DE CONTROLE ET NIVEAU DE QUALIFICATION DES DONNEES DANS LA BANQUE ADES

Statut de la donnée

Code	mnémonique	Libellé	Définition
1	Brute	Donnée brute	Données issues du processus d'acquisition n'ayant subi aucun examen. Ex. : donnée directement issue de l'appareil de mesure
2	NV1	Donnée contrôlée niveau 1 (données contrôlées)	Données ayant subi un ou plusieurs contrôles (au bureau, par un système expert ou par une personne physique) en fonction du contexte de la mesure. Le producteur examine les résultats par rapport à la connaissance qu'il a sur la station et le point de prélèvement. Exemple : comparaison par rapport à des seuils min- max classiquement rencontrés sur le point, utilise un système expert qui compare les résultats entre eux.
3	NV2	Donnée contrôlée niveau 2 (données validées)	Données ayant subi un contrôle par une comparaison avec une mesure manuelle sur le terrain (contrôle de la chaîne d'acquisition avec correction des dérives) (contrôle spatio-temporel, contrôle des capteurs et de leur dérive, contrôle de la chaîne d'acquisition,...). Le producteur regarde et vérifie l'ensemble de la chaîne d'acquisition et la cohérence des données (par exemple : depuis le prélèvement, conditionnement, flaconnage, transport, mesure en laboratoire).
4	MP	Donnée interprétée	La valeur a été utilisée dans un rapport ou valorisée. Cette mise en perspective de l'information permet de consolider son niveau de validité et détecter les dernières erreurs. Par exemple : diagrammes binaires, comparaison faciès, etc....

Qualification de la donnée

Code	mnémonique	Libellé	Définition
0	Non définissable	Qualification non définissable	Une valeur sera non définissable lorsque le producteur est dans l'impossibilité d'obtenir les informations nécessaires évaluer la conformité de la donnée. Il s'agit par exemple de données historiques récupérées des archives dont on a perdu toute information sur la façon dont elles ont été produites.
1	correcte	Correcte	Une valeur est déclarée « Correcte » lorsque elle est estimée valide au stade de validation indiquée dans l'information « statut de la donnée » et vis-à-vis de la finalité recherchée.
2	incorrecte	Incorrecte	Une valeur est déclarée « Incorrecte » lorsque elle est estimée erronée au stade de validation indiquée dans l'information « statut de la donnée » et vis-à-vis de la finalité recherchée.
3	incertaine	Incertaine	Une valeur sera déclarée « Incertaine » si la validité de la donnée reste « douteuse » au stade de validation indiquée dans l'information « statut de la donnée ». Dans la mesure du possible, la qualification « Douteuse » doit être une étape transitoire de la validation de la donnée et doit être réservé à des avancements intermédiaires de la validation.
4	Non qualifié	Non qualifié	Etat initial de la mesure qui n'a encore subi aucun audit ou interprétation du producteur de données en vue de sa validation.

**ANNEXE III. CARTOGRAPHIES SPATIO-TEMPORELLES DE LA CONTAMINATION EN ALUMINIUM
DANS LES EAUX SOUTERRAINES SUR LA PERIODE 1990-2012 ET POUR LES DEPARTEMENTS DU
NORD, DU PAS DE CALAIS, DE LA SOMME ET DE LA SEINE-MARITIME (ADES EAU FRANCE)**

