



HAL
open science

CARTOGRAPHIE DE LA POLLUTION PARTICULAIRE EN VILLE

Anne-Lise Beaulant

► **To cite this version:**

Anne-Lise Beaulant. CARTOGRAPHIE DE LA POLLUTION PARTICULAIRE EN VILLE. Sciences de l'environnement. Université Paris-Diderot - Paris VII, 2006. Français. NNT: . pastel-00949078

HAL Id: pastel-00949078

<https://pastel.hal.science/pastel-00949078v1>

Submitted on 19 Feb 2014

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.



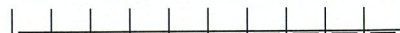
MINES PARIS

Collège doctoral

UNIVERSITÉ
PARIS 7 - DENIS DIDEROT



N° attribué par la bibliothèque



THESE

pour obtenir le grade de
Docteur en Sciences de l'Université Paris 7 – Denis Diderot
Ecole doctorale : Sciences de l'Environnement d'Ile de France
Spécialité : Méthodes physiques en télédétection

sera présentée et soutenue publiquement par
Anne-Lise BEAULANT

le 27 septembre 2006

CARTOGRAPHIE DE LA POLLUTION PARTICULAIRE EN VILLE

Directeur de thèse : Lucien WALD

Jury :

M. Stéphane JACQUEMOUDPrésident
Mme Christiane WEBER Rapporteur
M. Nicolaos SIFAKIS Rapporteur
M. Denis SARIGIANNIS Examineur
M. Lucien WALD Examineur
M. Gilles PERRON Examineur

Avant-propos

Cette thèse a été réalisée au Centre Energétique et Procédés (CEP) à l'Ecole des Mines de Paris (ENSMP) à Sophia Antipolis, au sein de l'équipe Observation, Modélisation et Décision (OMD).

Je tiens tout d'abord à remercier Didier Mayer, directeur adjoint, pour m'avoir accueillie au sein du Centre Energétique et Procédés de l'Ecole des Mines de Paris à Sophia-Antipolis.

Je remercie vivement Lucien Wald, professeur à l'ENSMP et directeur de ma thèse, pour l'exemplarité de son encadrement durant ces trois années, pour son attention et sa disponibilité bienveillantes, son soutien et les nombreux conseils qu'il m'a prodigué tout au long de mes travaux de recherche.

J'exprime toute ma gratitude aux membres de mon jury : Nicolaos Sifakis, Associate Professor à l'Institute for Space and Applications d'Athènes, et Christiane Weber, directeur de recherche CNRS, pour avoir accepté d'être les rapporteurs de mes travaux et pour s'être déplacés pour ma soutenance ; Stéphane Jacquemoud, professeur à l'Université Paris VII, pour avoir accepté de présider le jury ; Denis Sarrigiannis, chercheur à l'Institute for Environment à Ispra, pour être venu à ma soutenance et pour avoir accepté d'être examinateur ; et enfin Gilles Perron, ingénieur ASPA, pour son aide et ses conseils et pour avoir accepté de participer au jury.

Je remercie les membres de l'ex-groupe T&M devenu OMD pour avoir contribué au bon déroulement de cette thèse : Michel Albuisson, Thierry Ranchin, Roseline Wald, Lionel Menard, François Cauneau, Mireille Lefèvre, Ludovic Basly et Thierry Pogglio.

J'exprime ma sincère sympathie à mes deux co-bureau pendant cette thèse : Sylvain Cros et Mohammed-Bassam.Ben_ticha, ainsi qu'aux membres du trio du midi : Nicolas Fichaux, Pierre Pinson et Olivier Rhétoré avec qui j'ai passés des moments inoubliables.

Je pense chaleureusement à tous mes collègues, thésards pour la plupart, qui sont devenus docteurs ou qui le seront un jour : Nils Siebert, Alexis Bocquet, Juliette Korhel, Charles-Emile Hubert, Yasmine Masmoudi, Jean-Damien Rollier, Silvère Viriot, Luis Costa, Florent Ficher, Maryline Moreno, Jérémie Juban, Julien Marie, Guillaume Petitpas, Edouardo Izquierdo et tous les autres.

Je pense également aux personnes que j'ai rencontrées en venant ici faire une thèse et qui sont devenus mes amis : LN, Gaetane, Vincent, Lematt, Simon, Maca et Fabien et bien sûr Géraldine et Ludo.

Une pensée affectueuse est pour mes parents dont le soutien est inconditionnel.

Enfin, ma dernière pensée va à Manu, pour son tendre soutien, ses encouragements et sa présence à mes côtés dans tout ce que j'entreprends.

Table des matières

Table des matières	5
Chapitre 1	11
Introduction	11
1 Cadre du sujet	13
2 Problématique	15
3 Objectifs de la thèse	16
4 Démarche entreprise	17
Bibliographie	20
Chapitre 2	18
Méthodes actuelles de cartographie – Apports connus de l'imagerie satellitaire	18
1 Méthodes classiques de cartographie	25
1.1. Interpolation des mesures ponctuelles	25
1.2. Limite principale des méthodes d'interpolation	29
1.3. Modèles numériques de dispersion/transport des polluants statistiques et déterministes	31
1.4. Conclusion	36
2 Méthodes originales proposées récemment	36
2.1. Utilisation de l'imagerie satellitaire	36
2.2. Méthode des « champs typiques » originelle (Ung 2003)	39
2.2.1. Principe	39
2.2.2. Définition et construction d'un champs typique	41
2.2.3. Discussion	42
2.3. Méthode de densification du réseau de mesures	44
2.3.1. Les premiers travaux (Basly 2000)	44
2.3.2. Les apports de Ung (2003)	45
2.3.3. Discussion	47
Bibliographie	48
Chapitre 3	49
Méthode des « champs typiques »	49
1 Contexte et objectifs	57
2 Hypothèse fondatrice	58
3 Outils mathématiques	61
3.1. Notions de résolution et d'échelle	61
3.2. Transformée en ondelettes	61
3.3. Analyse multirésolution et algorithme « à trous »	62
4 Présentation de la méthode	64
4.1. Principe général de la méthode	67
5 Mise en œuvre et résultats	69

5.1.	Inventaire des données disponibles et création d'un champ typique	69
5.2.	Création d'un champ typique pour une situation donnée	71
5.3.	Construction d'une carte de pollution à l'échelle locale	72
6	Conclusion et discussion	73
	Bibliographie	76
Chapitre 4		74
Méthode de densification virtuelle du réseau de mesures		74
1	Introduction	83
2	Cartes et résultats	84
3	Conclusion et perspectives	89
4	Démarche de validation des stations virtuelles	90
5	Article soumis à <i>Environment International</i>	91
Chapitre 5		98
Détection des aérosols urbains (PM) à l'aide d'images satellitaires		98
1	Démarche	109
2	Approche simplifiée du problème et hypothèses	111
2.1.	Effets atmosphériques	111
2.2.	Mesures par satellite	113
2.3.	Outils expérimentaux	114
2.3.1.	Plans d'expériences numériques	114
2.3.2.	Modèle de transfert radiatif 6S	116
3	Effets moyens et contributions des paramètres atmosphériques	117
3.1.	Effets moyens	117
3.2.	Contributions	122
4	Etude de la sensibilité des réflectances aux PM par simulations numériques	124
4.1.	Détermination de la sensibilité spectrale du capteur	124
4.2.	Quantification de la sensibilité	126
5	Relation entre concentrations en PM et réflectances observées	128
6	Validation des stations virtuelles	130
7	Conclusion	135
	Bibliographie	137
Chapitre 6		132
Conclusion générale et perspectives		132
Annexe A		ii
Publications		ii
Annexe B		vi

<i>Analyse multirésolution et transformée en ondelettes</i>		<i>vi</i>
1	Nomenclature	<i>vi</i>
2	Transformée en ondelettes	<i>vi</i>
3	Analyse multirésolution	<i>viii</i>

Chapitre 1

Introduction

1 Cadre du sujet

Les polluants atmosphériques émis par l'homme, au travers de ses activités, sont variés et présents à l'état de gaz ou de particules dans l'air plus ou moins fines ; parmi les polluants directement émis dans l'air, qualifiés de polluants *primaires*, on trouve des composés du soufre et de l'azote, des oxydes de carbone et des composés organiques volatils (solvants, hydrocarbures,...), des poussières fines et des métaux (plomb, mercure, cadmium, arsenic ...). L'atmosphère comprend également des polluants dits « *secondaires* » (ozone, certaines particules) qui résultent de l'évolution chimique dans l'air des polluants primaires (Hewitt & Jackson 2003). En particulier, la formation de l'ozone est due aux transformations subies par certains composés (NO_x et COV essentiellement) sous l'action du rayonnement UV solaire, par réactions photochimiques.

Les agences françaises chargées de la surveillance de la qualité de l'air ont défini certains polluants comme indicateurs de pollution. Ce sont :

- le dioxyde de soufre (SO_2),
- le monoxyde d'azote (NO),
- le dioxyde d'azote (NO_2),
- l'ozone troposphérique (O_3),
- les particules ou aérosols de différentes tailles ($\text{PM}_{2,5}$, PM_{10} , PM_{13}), c'est-à-dire dont le diamètre est inférieur à, respectivement, 2,5, 10 et 13 microns.

Ce sont aussi les principaux polluants de l'air ambiant de nos agglomérations et ceux dont les effets sont les plus importants et les mieux connus sur la santé de l'homme.

Les polluants de l'air extérieur constituent une part importante de la pollution mais il existe d'autres polluants de l'air entraînant pourtant des effets sur la santé tout aussi graves : les polluants de l'air intérieur (Kirchner 2001, 2003 ; Kotzias 2005). Dans certains cas et pour certaines substances, la pollution peut être plus importante à l'intérieur des locaux qu'à l'extérieur. Aux polluants de l'air ambiant qui trouvent leur source à l'intérieur des bâtiments, s'ajoutent les polluants de l'air extérieur qui pénètrent par les ouvertures ou le système de ventilation (Mosqueron & Nedellec 2001). Lorsqu'on parle de qualité de l'air en milieu urbain, il est donc important de considérer également cet aspect. Dans la littérature (Pernelet & Roth 2001), la prise en compte de la pollution de l'air intérieur se fait par le biais d'un rapport ; la concentration du polluant à l'intérieur s'exprime comme un

pourcentage de la pollution de ce polluant à l'extérieur. Le problème de la qualité de l'air à l'intérieur des bâtiments se ramène donc à un problème global de qualité de l'air extérieur.

Dans cette thèse, je m'intéresserai en particulier aux particules de type $PM_{2,5}$, PM_{10} et PM_{13} que j'engloberai sous le terme PM. Les particules contribuent majoritairement à la dégradation de la qualité de l'air en ville (smog) et leur quantité dans l'air est en nette augmentation, contrairement aux autres indicateurs de pollution, notamment pour les particules fines. Les particules sont, de plus, considérées comme des polluants très nocifs pour la santé car elles transportent des métaux lourds ou des composés organiques (Hewitt & Jackson 2003).

Au niveau européen et national, des normes de suivi des polluants sont mises en place.

La directive-cadre européenne du 27 septembre 1996 (JO L 296 du 21.11.1996, modification JO L 284 du 31.10.2003) établit les principes de base d'une stratégie commune visant à définir et fixer des objectifs concernant la qualité de l'air ambiant, afin d'éviter, de prévenir ou de réduire les effets nocifs pour la santé humaine et pour l'environnement, à évaluer la qualité de l'air ambiant dans les États membres, à informer le public, entre autres par des seuils d'alerte, ainsi qu'à améliorer la qualité de l'air lorsque celle-ci n'est pas satisfaisante.

Plus particulièrement, dans le cadre de la loi française sur l'air de décembre 1996 (n° 96-1236) dite Loi sur l'Air et l'Utilisation Rationnelle de l'Energie (LAURE 1997) dont « l'objectif est la mise en oeuvre du droit reconnu à chacun à respirer un air qui ne nuise pas à sa santé », les actions menées consistent « à prévenir, à surveiller, à réduire ou à supprimer les pollutions atmosphériques ». Cette loi a permis de donner les fondements et les moyens financiers du développement des réseaux de mesures de surveillance de la qualité de l'air pour les villes du territoire français.

Depuis cette loi, les AASQA (Agences Agréées de Surveillance de la Qualité de l'Air) se sont multipliées. Chaque ville française de plus de 100.000 habitants possède maintenant son propre réseau urbain de stations de mesure répondant ainsi à des objectifs bien identifiés des récents textes réglementaires comme la mesure de la qualité de l'air, le déclenchement d'alerte, la diffusion permanente de l'information, la surveillance en référence aux normes et la prévision des pics de pollution.

En outre, le Plan Régional pour la Qualité de l'Air (PRQA), prévu par la LAURE, et dont les modalités de mise en oeuvre sont précisées par le décret du 6 mai 1998 et dans la circulaire de la Direction Générale de la Santé (DGS) du 24 mars 1998, fixe des orientations visant à réduire les effets de la pollution atmosphérique.

L'application de la LAURE a conduit à un développement de la surveillance de la qualité de l'air (ADEME 2000). Les réseaux de stations de mesure permettent de donner rapidement l'alarme en cas de pics de pollution et d'aider les autorités à prendre les mesures d'urgence qui s'imposent pour la santé publique. En dépit de cela, du fait de la concentration des populations et des activités en ville, à cause du flux croissant des échanges et de l'accroissement constant du parc et de la circulation automobile, la pollution atmosphérique constitue toujours un enjeu particulièrement sensible et une préoccupation populaire (IPTS 2000). En effet, depuis 1950, le parc automobile n'a cessé de s'accroître jusqu'à devenir l'une des principales sources de pollution des villes européennes (Académie des sciences 1999).

2 Problématique

Aujourd'hui de multiples questions relatives à l'exposition des populations et à la représentation spatiale la plus réaliste ou proche possible des mesures collectées restent en suspens. Le Haut Comité de Santé Publique (HCSP 2000) a établi un rapport sur les moyens d'améliorer la connaissance et la prévention des phénomènes liés à la pollution atmosphérique et de renforcer l'efficacité des politiques de santé publique dans ce domaine. Ce rapport exprime le besoin d'outils permettant, notamment, de renforcer les connaissances en matière d'exposition réelle de la population à la pollution atmosphérique. L'évaluation de cette exposition consiste à croiser des données de répartition géographique de la population et les données de pollution. Weber *et al.* (2002) indiquent qu'il est nécessaire d'améliorer les connaissances des comportements de déplacements des citoyens, et aussi d'estimer précisément la pollution atmosphérique dans l'espace et dans le temps. Par conséquent, pour évaluer l'exposition, il faut d'une part et principalement, des cartes de répartition de la pollution, les informations ponctuelles délivrées par un réseau de mesures urbain ne représentant pas la variabilité spatiale de la pollution, et d'autre part, des modèles de déplacements des citoyens.

Les travaux de cette thèse s'inscrivent dans ce contexte global en s'orientant plus particulièrement vers l'estimation et la cartographie de la pollution atmosphérique, spécifiquement de la concentration en polluants particuliers.

Plusieurs outils existent aujourd'hui qui permettent de dériver des informations sur la qualité de l'air sous forme de cartes. Ces cartes sont généralement obtenues par :

- des méthodes d'interpolation spatiale des mesures ponctuelles,
- des modèles numériques utilisant la mécanique des fluides pour modéliser les phénomènes de transport et les transformations chimiques des polluants dans l'atmosphère.

Ces deux approches ont des limitations qui contraignent leur utilisation routinière par les AASQA. Pour la première approche, la principale limitation est l'insuffisance de stations de mesure de la qualité de l'air (Stalker & Dickerson 1962). Les villes ne peuvent pas se permettre d'équiper tous les quartiers, les rues et les places de dispositif de surveillance. Leur coût élevé limite le nombre de stations à une dizaine par ville, ce qui est insuffisant pour une cartographie précise par interpolation. Pour la seconde approche, l'emploi de ces méthodes est limité par la complexité des calculs et la méconnaissance de certains paramètres du modèle (conditions aux limites). Ceci entraîne que nous avons soit des modèles fonctionnant à grande échelle spatiale avec une grande couverture géographique : ils fournissent des cartes de pollution peu précises sur de grandes zones, soit des modèles fonctionnant seulement sur une partie de l'agglomération : les cartes sont très précises mais sur de très petites zones.

3 Objectifs de la thèse

Pour remédier à ces problèmes, les ministères de l'Environnement et de la Recherche ont financé des projets de recherche. En particulier au cours du programme (1999) de l'Action Concertée Incitative « Ville » (Weber *et al.* 2002), le groupe de travail constitué du laboratoire Image et Ville de l'Université Louis Pasteur de Strasbourg, de l'ASPA (Association pour la Surveillance et l'étude de la Pollution atmosphérique en Alsace) et du Centre Energétique et Procédés de l'Ecole des Mines de Paris a cherché à établir une stratégie de cartographie. Il a proposé des méthodes alternatives aux méthodes classiques de cartographie que sont l'interpolation des mesures et l'utilisation des modèles numériques. Ma thèse s'inscrit dans ce cadre d'établissement d'une stratégie de cartographie complète et précise de la pollution atmosphérique sur l'ensemble d'une agglomération à l'échelle de la rue (100 m). Plus précisément, j'ai pour objectif de mieux définir mathématiquement les méthodes alternatives proposées dans les travaux antérieurs, de cerner leur portée, leur applicabilité et en particulier leurs échelles spatiales et temporelles.

On se limite ici à la pollution due aux particules. Un cas d'étude a été choisi, la Communauté Urbaine de Strasbourg (C.U.S) pour laquelle on dispose de plusieurs jeux de données (mesures au sol, cadastre des émissions...). La C.U.S est le cas pratique sur lequel seront testées les différentes méthodes de cartographie. Toutefois notre travail devra être transposable à d'autres villes européennes disposant d'un réseau de surveillance de la pollution.

Deux méthodes alternatives ont été proposées. Ce sont la méthode des *champs typiques* et la méthode de densification du réseau de mesures, dont les originalités sont pour la première, l'utilisation de la fusion de données et pour la seconde, l'utilisation de *stations virtuelles*. Ces méthodes proviennent des travaux de Basly (2000), Ung (2003) et Ung *et al.* (2002) qui ont initié les concepts de *champs typique* et *stations virtuelles*. La stratégie de cartographie établie dans cette thèse s'inspire de ces concepts.

La méthode des *champs typiques* vise à améliorer la représentation de la pollution en appliquant des méthodes de fusion de données à des données ou cartes de pollution déjà existantes, issues de modèles numériques par exemple. Cette méthode s'applique dans le cas où l'on dispose déjà de données de pollution sous forme de cartes mais dont les résolutions sont différentes. La fusion de données constitue « *un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance de données provenant de sources diverses* ». Elle vise à « *l'obtention d'information de plus grande qualité* » (Wald 2002). En particulier, la fusion de données permet de synthétiser des cartes à des échelles très fines à partir de données ou de cartes à des échelles plus grossières. Ceci permet de pallier le manque de données disponibles à une échelle donnée. Les méthodes de fusion de données appliquées à la cartographie de la qualité de l'air n'ont encore jamais été testées sur un cas pratique. L'enjeu consiste à étudier la faisabilité, la validité ainsi que l'applicabilité d'une telle méthode et de son domaine d'application.

La méthode de densification du réseau de mesure vise à créer des cartes de pollution par interpolation. Elle s'applique dans le cas où l'on dispose seulement de relevés de mesures *in situ*. Les *stations virtuelles* s'ajoutent aux stations réelles du réseau de mesure pour le rendre plus dense et améliorer ainsi l'interpolation. La méthode de détermination des stations virtuelles telle qu'elle est présentée ici diffère de celle proposée par Ung (2003), Ung *et al.* (2002), et représente une innovation. L'enjeu est de montrer l'apport des stations virtuelles pour la réalisation de cartes de pollution à l'échelle locale par méthode d'interpolation, d'estimer le domaine de validité et d'étudier son applicabilité.

Pour valider la définition et la localisation des stations virtuelles, il est envisagé d'utiliser l'imagerie satellitaire en analysant son rôle possible dans l'estimation des concentrations en particules. Ce dernier point constituerait une innovation supplémentaire.

Un point critique de cette stratégie de cartographie concerne les échelles d'espace et de temps pour lesquelles cette stratégie est valide. La thèse s'attache donc à bien définir ces échelles en fonction des différentes hypothèses sous-tendant nos méthodes.

4 Démarche entreprise

La démarche mise en œuvre dans cette thèse est présentée chapitre par chapitre.

- Le chapitre 2 est un état de l'art des méthodes généralement employées pour cartographier la pollution atmosphérique. La première partie présente une étude critique des méthodes classiquement utilisées par les AASQA afin d'en montrer les principales limitations pour la cartographie de la pollution de l'air sur l'ensemble d'une agglomération à l'échelle de la rue. Cette étude met l'accent sur les échelles spatiales et temporelles impliquées. Pour pallier ces

insuffisances, des méthodes originales ont été proposées. L'existence des capteurs satellites haute résolution a poussé certains auteurs à exploiter les potentialités de l'imagerie satellitaire pour la cartographie à des échelles locales. D'autres ont étudié l'intérêt de la fusion de données dans ce domaine.

- Le chapitre 3 présente une méthode ayant recours aux outils de la fusion de données : la « méthode des champs typiques ». Le principe de cette méthode repose sur le fait que l'on suppose qu'il existe un nombre fini de formes de champs de pollution et que ceux-ci sont caractéristiques d'une situation liée à la météorologie et aux activités humaines. Le concept est décrit en détail en montrant les différences et améliorations par rapport au concept d'origine proposé par Ung (2003). En particulier le formalisme mathématique est bien établi. Une première mise en œuvre pratique de la méthode est présentée. Celle-ci permet d'évaluer la faisabilité de la méthode, son applicabilité et d'étudier son domaine d'application.
- L'interpolation des mesures ponctuelles est la manière la plus usitée par les AASQA pour obtenir des cartes de pollution. La limitation majeure est l'insuffisance de stations de mesures. Le faible nombre de mesures entraîne de graves imprécisions sur les cartes de pollution. Le chapitre 4 présente une méthode qui permet d'augmenter de façon virtuelle le nombre de mesures. Le principe de densification virtuelle du réseau de stations de mesures est de déterminer des endroits de la ville semblables, par leurs propriétés aérauliques et conditions d'émissions, à ceux où sont installées les stations de mesures constituant le réseau. Ces points de la ville sont appelés des *stations virtuelles*. Une classification de données multi-sources permet de dénombrer et localiser ces stations. La représentativité des stations virtuelles pour la cartographie de la pollution sera vérifiée. Puis, on montrera l'apport de ces stations pour la cartographie par interpolation des mesures in situ. L'amélioration de la qualité des estimations sera mise en évidence. Les résultats des interpolations seront évalués.
- Le chapitre 5 est une étude de l'apport de l'imagerie satellitaire haute et moyenne résolution pour la cartographie de la pollution particulaire et plus précisément, pour l'estimation des concentrations en particules. L'utilisation d'un modèle de transfert radiatif combiné à une approche par plans d'expériences numériques a permis de quantifier l'influence des variations de concentrations sur la radiométrie des images, la concentration minimale détectable a été déterminée. Les canaux les plus sensibles à ces variations pour une gamme de concentrations réalistes ont été précisés. En comparant les simulations et les observations (mesures au sol et données satellitaires), un modèle empirique permettant de passer de la donnée satellitaire à la concentration en particules a été déterminé. Grâce à ce modèle, le concept des stations virtuelles sera validé.

Chapitre 1 - Introduction

- Le chapitre 6 propose des conclusions aux travaux réalisés et offre une perspective sur l'application de cette carte de concentration en particules. Son emploi dans l'estimation de l'exposition des personnes aux polluants en ville est tout à fait adapté de par sa résolution spatiale et temporelle. Nous avons couplé des cartes de pollution à un modèle de déplacement d'un citoyen. En fait, nous avons construit une maquette de simulation de l'exposition individuelle au cours d'un trajet en ville qui montre qu'il est impératif de disposer d'une telle représentation pour, à terme, estimer l'exposition aux polluants urbains.

Bibliographie

Académie des sciences, 1999. Evaluer les effets des transports sur l'environnement, le cas des nuisances sonores. Rapport du CADAS n°16, novembre 1999.

ADEME, 2000. Classification et critères d'implantation des stations de surveillance de la qualité de l'air, recommandation du groupe de travail « caractérisation des sites ». Janvier 2000, 70 p.

Basly L., 2000. Télédétection pour la qualité de l'air en milieu urbain. Thèse de doctorat Sciences des technologies de l'information et de la communication, Université de Nice-Sophia Antipolis, Nice, France, 182 p.

HCSP (Haut Comité de la Santé Publique), 2000. Politiques publiques, pollution atmosphérique et santé: poursuivre la réduction des risques". <http://www.sante.gouv.fr/htm/actu/pollution/sommaire.htm>

Hewitt C.N., Jackson A., 2003. *Handbook of Atmospheric Science. Principles and Applications*. Blackwell Publishing, Oxford, UK, 633 p, 149.

IPTS, 2000. Report n°47, qualité de l'air dans les villes européennes: la question des transports urbains, septembre 2000. <http://www.jrc.es/pages/f-report.fr.html>

Kirchner S., 2003. Given metrological of indoor air. 41st congress of the Clinical Company of Toxicology, Nancy, France, Scientific and Technical Center of Building industry (CSTB), June 26-27, 15p., http://www.stc-congrès.org/Congrès_2003/Conférences.htm.

Kirchner S., Bailloit O., Collignan B., Flori J. P., Garret O., Laurent A. M., Le Moullec Y., O'Kelly P., Ramalho O., Sauvaget M., Villenave J. G., Vedel C., 2001. Experimental study of transfer conditions for atmospheric pollution from local origin inside residential buildings. Final report, Primequal program, Scientific and Technical Center of Building industry (CSTB), April 2001, 269 p.

Kotzias D., 2005. Indoor air and human exposure assessment – needs and approaches. *Experimental and Toxicologic Pathology*, 57, 1, 22 July, pp. 5-7.

LAURE Loi sur l'Air de l'Utilisation Rationnelle de l'Energie, 1996. Journal Officiel de la République Française. 1 Janvier 1997, 11 p.

Mosqueron L., Nedellec V., 2001. Inventory of French data on air quality. Observatory of Indoor Air Quality, Marne la Vallée, France, December 2001, 173 p.

Chapitre 1 - Introduction

Pernelet V., Roth C., 2001. Evaluation of the exposure of the Ile-de-France citizens to air pollution during the day. Rapport AIRPARIF, Paris, France, December 2001, 106 p.

Stalker W.W., Dickerson R.C., 1962. Sampling station and time requirements for urban air pollution surveys, II, suspended particulate matter and soiling index. Journal of Air Pollution Control Association, vol. 12, n°3, 111-128.

Ung A., Wald L., Ranchin T., Weber C., Hirsch J., Perron G., Kleinpeter J., 2002. Air pollution mapping : a new approach based on remote sensing and geographical databases. Application to the city of Strasbourg. Photo-Interpretation 2000 / 3-4, 38, p 53-63.

Ung A., 2003. *Cartographie de la pollution atmosphérique en milieu urbain à l'aide de données multisources*. Thèse de doctorat Sciences de l'environnement, Université Paris 7 Denis Diderot, Paris, France , 96 p.

Wald L., 2002. *Data Fusion. Definitions and Architectures - Fusion of Images of Different Spatial Resolutions*. Presses de l'Ecole, Ecole des Mines de Paris, Paris, France, ISBN 2-911762-38-X, 200 p.

Weber C., Hirsch J., Puissant A., Durrenberger M., Cerciat M., Schnell L., Bronner A.C., 2002. Morphologie urbaine et répartition spatiale des polluants : de la rue à l'agglomération. Application à la Communauté Urbaine de Strasbourg (CUS). Rapport de recherche, programme ACI "Ville" du MENRT, Strasbourg , 153 p.

Chapitre 2

Méthodes actuelles de cartographie –
Apports connus de l'imagerie satellitaire

Actuellement, la cartographie de la pollution atmosphérique en ville se fait par interpolation de mesures au sol dans la plupart des AASQA (Rouïl 2001) ou par modélisation CFD (Computational Fluids Dynamics) (Mestayer et al. 1997). Quelques travaux font état d'apports possibles des images satellitaires.

1 Méthodes classiques de cartographie

1.1. Interpolation des mesures ponctuelles

Le problème de l'interpolation des mesures ponctuelles est le suivant : je ne connais les concentrations en polluants qu'en un certain nombre de points disséminés sur une surface et je veux connaître les concentrations en polluants en tout point de cette surface. Ici, on va négliger les différences de hauteur entre stations de mesure et se ramener à un cas plan.

Je suppose qu'il existe une fonction F , liant les points selon leur emplacement sur la surface. La solution de notre problème est de trouver une fonction d'interpolation G approchant au mieux la fonction F . Dans notre cas F et G sont bidimensionnelles, et les méthodes permettant d'approcher F sont des méthodes de reconstruction de surface (Basly 2000). L'extrapolation consiste à déterminer la fonction G en dehors du domaine ayant formé par les points de mesure.

Il existe une grande variété de méthodes d'interpolation / extrapolation des données *in situ* de pollution. Les méthodes présentées ici sont les plus usitées pour la cartographie des polluants de l'air.

Les méthodes polynômiales de base. L'espace des surfaces est l'ensemble des polynômes à deux variables et de degré d . Chaque polynôme dispose de M degrés de liberté avec :

$$M = \frac{(d+1)(d+2)}{2} \quad (2.1)$$

Le degré du polynôme, d , est généralement choisi afin que M soit inférieur au nombre de points donnés, N . Dans le cas où $N \gg M$, la surface construite est dite sur-contrainte, *i.e* elle n'assure qu'une reconstruction des basses fréquences c'est-à-dire des variations lentes des données ; il y a donc un effet de lissage. Si N est voisin de M , la surface respecte mieux les données et le lissage disparaît. Pour s'adapter parfaitement, il faudrait que $M=N$. Cependant, l'adaptation peut entraîner un comportement erratique loin des points de données, ce qui conduit à des écarts incontrôlés (Basly 2000).

La méthode des « corrections successives », conçue par Bergthorsson & Doos (1955), est une méthode itérative qui converge vers un équilibre. Dans l'esprit de ces derniers, il s'agissait de combler les lacunes du réseau d'observation. Une pondération empirique des deux sources d'information

disponibles (réseau d'observation et relevé sur le terrain) permet de créer une ébauche initiale Φ_0 du champ de pollution Φ , laquelle est corrigée localement pour être adaptée aux valeurs observées aux stations. La méthode des corrections successives permet de lisser le champ des corrections $\varepsilon = \Phi - \Phi_0$. La méthode de Cressman-Chardonnal, présentée par Delfiner (1973), est fondée sur ce principe. La correction ε_g aux points de grille Γ est une combinaison linéaire des écarts ε_i entre l'observation et l'ébauche Φ_0 aux stations i , situées dans une certaine « zone d'influence » de Γ (cercle de centre Γ et de rayon D) :

$$\varepsilon_g = \sum_{i=1}^n \lambda_i \varepsilon_i \quad (2.2)$$

avec ε_i la correction et λ_i le coefficient de pondération à la station S_i ($i=1, \dots, N$). Ainsi, ε_i va influencer sur la valeur du champ en Γ , dans une proportion μ d'autant plus faible que la distance d' de Γ à la station i est plus forte. Cressman (1959) a préconisé la fonction de poids suivante :

$$\begin{cases} \mu = \frac{D^2 - d'^2}{D^2 + d'^2} & \text{avec } d' \leq D \\ \mu = 0 & \text{avec } d' \geq D \end{cases} \quad (2.3)$$

On peut choisir d'autres fonctions positives décroissantes avec D , conservant plus ou moins la particularité d'une mesure. Ensuite, cette opération est réitérée deux à quatre fois en pratique avec pour nouvelle ébauche $\Phi_1 = \Phi_0 + \varepsilon_g$, en diminuant le rayon d'influence (Puybonnieux-Textier 1996).

Cette méthode est utilisée en pratique pour la présentation des données. Sa mise en œuvre est facilitée par le fait qu'elle est simple à programmer et peu coûteuse en temps de calcul. Le paramètre D doit être adapté en fonction de la nature du champ. Le principal problème est l'estimation des coefficients de pondération empirique.

La méthode des « plaques minces » ou méthode « thin-plate spline » est préconisée par Ionescu *et al.* (1996, 2000). Cette méthode a été appelée ainsi par analogie avec la flexion d'une fine feuille de métal ; elle génère une surface de rigidité contrôlable. Dans le principe physique, la déflexion se fait dans la direction z orthogonale au plan. La « thin-plate spline » est l'équivalent bidimensionnelle de la spline cubique unidimensionnelle. Elle est la solution fondamentale à l'équation bi-harmonique et a la forme suivante :

$$U(r) = r^2 \ln r \quad (2.4)$$

Soit un ensemble de points de données, une combinaison linéaire de thin-plate splines centrée sur chaque point de donnée donne une fonction d'interpolation qui passe exactement par les points en minimisant « l'énergie de flexion ». L'énergie de flexion est définie ici par :

$$I[G(x, y)] = \iint_{R^2} (G_{xx}^2 + G_{xy}^2 + G_{yy}^2) dx dy \quad (2.5)$$

avec G la fonction d'interpolation et G_{xx} , G_{xy} , G_{yy} désignent les dérivées secondes.

La génération de surfaces de rigidité contrôlable permet de suivre les variations hautes fréquences sans pour autant avoir des effets de bords et des écarts importants loin des points de données. Cela évite également de trop lisser la surface (Basly 2000).

La méthode des « moindres carrés » consiste à ajuster le champ Φ à une combinaison linéaire de k fonctions de base f^l choisies à l'avance, en minimisant la somme ∂^2 des écarts quadratiques aux N points expérimentaux. L'estimation du champ s'écrit de la façon suivante :

$$\hat{\Phi}(x, y) = \sum_{l=0}^k a_l f^l(x, y) \quad (2.6)$$

La somme ∂^2 des écarts quadratiques aux N points expérimentaux est :

$$\partial^2 = \sum_{i=1}^N [\Phi(x_i, y_i) - \hat{\Phi}(x_i, y_i)]^2 \quad (2.7)$$

Delfiner (1973) préconise de se limiter pour une interpolation locale, à une quadratique de la forme :

$$\hat{\Phi}(x, y) = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2 \quad (2.8)$$

La méthode des moindres carrés débouche sur une méthode « filtrant » les variations de petite échelle. Dans tous les cas, les moindres carrés minimisent les écarts aux points expérimentaux. Ils ne garantissent en rien que l'erreur d'estimation sera faible en des points où les valeurs ne sont pas connues. La meilleure preuve est que l'on peut toujours trouver un ajustement passant exactement par les points expérimentaux, donc annulant ∂^2 . Ceci ne signifie pas que la fonction interpolée représente le champ en tout point (Puybonnieux-Textier 1996). Cette méthode se contente de reconstituer les champs non pas de façon exacte, mais en rendant compte de leurs traits structurants.

Toute la difficulté des méthodes pré-citées réside dans le choix de l'algorithme qui permet de construire en tous points les valeurs des concentrations. Il existe de nombreuses possibilités qui peuvent aboutir à des résultats sensiblement différents. Leur manque de fiabilité est l'une des principales critiques adressées à ces outils. La géostatistique (ou statistiques spatiales) correspond à

une technique plus élaborée qui, en complément des données ponctuelles mesurées, intègre des caractéristiques spatiales du phénomène de façon à assurer la continuité et la cohérence de la représentation. On parle parfois « d'interpolation intelligente » pour qualifier cette approche qui est supposée éviter les dérives des méthodes citées, au moyen des contraintes imposées par les données auxiliaires, représentatives des caractéristiques du problème. De plus, la géostatistique permet de quantifier la précision de l'estimation et offre une solution pour apprécier la qualité des représentations par rapport à un modèle (Rouïl 2001). Les méthodes géostatistiques et en particulier le krigeage sont les plus exploitées par les AASQA.

Le krigeage (Matheron 1963) est la méthode optimale d'estimation, au sens statistique du terme. Le problème qui a donné naissance au krigeage est la recherche de l'estimateur linéaire optimal de la valeur inconnue $Y(x_0)$ au point x_0 à partir des valeurs $Y(x_i)$ aux points expérimentaux voisins x_i . On cherche donc un estimateur Y^* de la forme (Chiles *et al.* 1975) :

$$Y^* = \sum_{i=1}^n \lambda_i Y(x_i) \quad (2.9)$$

Le problème consiste à déterminer la pondération c'est-à-dire les λ_i de chacun des points environnants. Le krigeage choisit les poids à partir du degré de similarité entre les valeurs de Y , i.e. à partir de la covariance entre les points en fonction de la distance entre ces points. Ici, je dirais simplement que la seule condition indispensable pour utiliser le krigeage est que les moments statistiques de la fonction d'accroissement $Y(x+h) - Y(x)$ soient stationnaires, c'est-à-dire ne dépendent pas du point x courant, seulement de la distance entre les points, h . Le krigeage utilise le semi-variogramme pour déterminer les poids de la première équation ; ce semi-variogramme $\gamma(h)$ est défini par :

$$\gamma(h) = \frac{1}{2} E[Y(x+h) - Y(x)]^2 \quad (2.10)$$

où x est le point courant sur lequel s'applique l'opérateur moyenne E . Le semi-variogramme est simplement la moitié de variance totale moins la covariance, en fonction de la distance entre les points. Pour appliquer le krigeage, on doit approcher le semi-variogramme expérimental, calculé pour quelques valeurs de h , par une fonction analytique de h . Dans la pratique, il existe trois modèles de base : linéaire, sphérique et exponentiel.

Le krigeage consiste ensuite à calculer les λ_i de la première équation à l'aide des valeurs de la fonction $\gamma(h)$, une fois N choisi. Il existe trois types de krigeage univarié, i.e. à une seule variable : le krigeage simple, le krigeage ordinaire ou le krigeage universel. La différence entre ces types d'estimation réside dans la connaissance de la statistique de la variable à interpoler :

- krigeage simple : variable stationnaire de moyenne connue ;
- krigeage ordinaire : variable stationnaire de moyenne inconnue ;
- krigeage universel : variable non-stationnaire (qui contient une tendance).

Le krigeage ordinaire, aussi appelé krigeage ponctuel, est le plus fréquemment utilisé.

Le cokrigeage désigne l'extension de cette méthode au cas où on étudie simultanément plusieurs variables liées entre elles. En effet, le cokrigeage se distingue du krigeage universel par la possibilité d'utiliser pour l'analyse objective non seulement les valeurs du champ étudié, mais encore d'autres sources complémentaires d'information.

Contrairement aux autres méthodes d'interpolation, les méthodes de krigeage et cokrigeage résolvent le problème de l'estimation des pondérations. Elles permettent également de calculer à la fois les estimateurs optimaux de ces coefficients et l'incertitude commise par rapport au modèle choisi de semi-variogramme. Selon Gratton (2002), c'est la méthode d'estimation la plus précise et est souvent préconisée (Frangi *et al.* 1996, Carletti *et al.* 2000).

1.2. Limite principale des méthodes d'interpolation

Toutefois, le principal problème dans ces méthodes, y compris pour les méthodes de krigeage et cokrigeage, concerne le manque de mesures fournies par les réseaux d'observation de la qualité de l'air. Les résultats sont extrêmement dépendants du nombre et de la répartition spatiale des mesures, ainsi que de la méthode d'interpolation.

La figure 2.1 montre trois cartes représentant la concentration en PM journalière, pour la journée du 15 août 2001. Trois méthodes d'interpolation différentes sont utilisées. On remarque que les représentations du champ de concentration sont très différentes alors qu'il s'agit de la même date. Il existe des différences assez fortes même si les zones de faibles et fortes concentrations sont situées aux mêmes endroits. Les différences sont plus marquées avec la méthode des plaques minces car cette méthode laisse moins la place aux petites échelles. Finalement, ces représentations mettent en évidence le fait que les mesures sont en nombre trop faible pour que la représentation du champ de concentration soit réaliste.

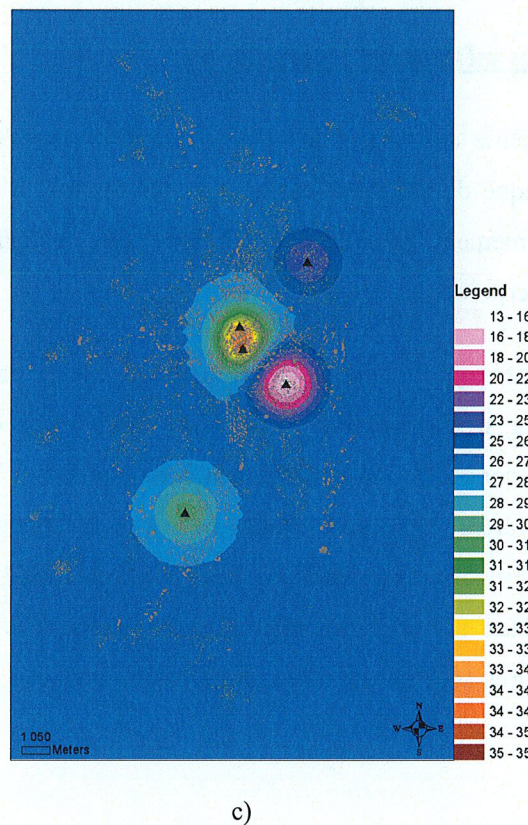
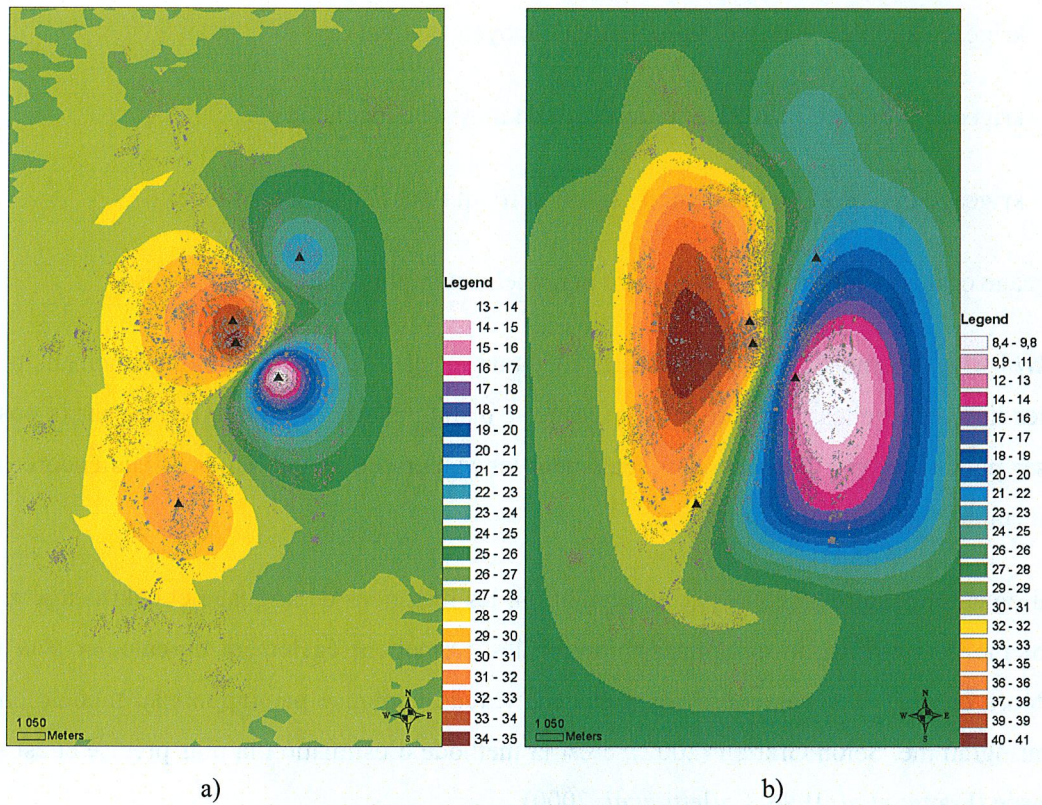


Fig. 2.1: Exemples de cartes de pollution particulaire obtenues par 3 méthodes d' interpolation différentes à partir des mesures issues des 5 stations (triangles noirs) mesurant les particules dans la ville de Strasbourg. En fond, en gris, l'ensemble du bâti de Strasbourg est représenté, l'échelle de couleur donne les valeurs de la concentration en PM en $\mu\text{g}/\text{m}^3$. a) IDW (Inverse Distance Weighted), b) plaques minces, c) krigeage simple.

1.3. Modèles numériques de dispersion/transport des polluants statistiques et déterministes

Le premier point caractérisant un modèle de dispersion de polluants est sa méthode de résolution. Ainsi, Descamps *et al.* (2000) proposent une classification selon deux catégories.

Les modèles statistiques (Bel *et al.* 2001) (NCFD : « No Computational Fluid Dynamics ») sont basés sur la recherche de relations simples entre différents paramètres le plus souvent observés ou prédits. Un exemple classique est donné par les modèles établissant un lien entre les concentrations atmosphériques (immissions) et les émissions soumises aux paramètres météorologiques. Leur principe général repose sur la recherche de relations simples entre concentrations et paramètres caractérisant la pollution (émissions, météorologie et immissions). Les variables explicatives les plus souvent rencontrées sont des paramètres météorologiques (température, vent, nébulosité, humidité relative) observés ou prévus. Les modèles prennent également en compte les niveaux des concentrations en polluants mesurés la veille ou le jour même. Les valeurs maximales sont en général utilisées. Dans certains cas, des caractéristiques locales telles que des informations météorologiques fines (flux turbulents et radiatifs), ou des données de sites (occupation du sol,...) peuvent compléter le jeu des entrées.

Ces modèles sont utilisés par les AASQA pour la représentation de la qualité de l'air sur un site donné. Ils sont largement répandus et presque exclusivement dédiés à la prévision des concentrations et des pics de pollution, notamment des dépassements en ozone (Rouil 2001, Blond 2002).

Les modèles déterministes (CFD) reconstituent l'information en tout point d'une grille maillant le domaine d'étude, grâce à la résolution des équations physiques et chimiques qui conditionnent les phénomènes. Dans le domaine de la qualité de l'air, il s'agit typiquement des équations de la mécanique des fluides (équations de Navier-Stokes ; Holton 1972) couplées avec des équations de transport et de transformation chimique des polluants. Un maillage tridimensionnel est appliqué au domaine d'étude. Une solution numérique du problème est recherchée en chaque point de la grille. Les équations dans les modèles varient selon les échelles d'espace et donc de temps puisque ces variables sont intrinsèquement liées en météorologie. Pour des échelles de l'ordre de 1000 km, dites régionales, les phénomènes prédominants relèvent de l'advection. A 100 km, les mouvements verticaux et les phénomènes de diffusion doivent être pris en compte. A 10 km, l'importance relative de ces derniers croît. A 1 km, l'influence du sol et du bâti se fait déjà sentir. A 100 m, il faut prendre en compte les bâtiments, leur orientation, leur hauteur, la nature du sol et sa rugosité.

Pour être performants, les modèles déterministes demandent une information complète sur :

- la météorologie (vent, température, humidité, ensoleillement) : dans le cas d'une étude du niveau moyen de pollution sur une journée, une caractérisation préalable de la climatologie du

site est nécessaire car les modèles de chimie-transport ne produisent pas leurs propres données météorologiques, nécessaires à l'évaluation des flux liés aux processus de dispersion, mélange, dépôt, aux réactions chimiques, etc... Par conséquent, il est nécessaire de leur fournir en entrée des données météorologiques, sur le domaine de modélisation et la période que l'on souhaite simuler,

- les émissions : la réalisation d'un inventaire spatialisé des émissions constitue l'une des étapes les plus lourdes. Il permet de définir les flux de polluants rejetés pour différents types de sources (trafic, industries, chauffage urbain, végétation...). Plus le domaine de calcul est important, plus l'exercice est difficile, notamment lors de la réalisation de calculs photochimiques,
- les conditions aux limites : elles qualifient les données météorologiques et de concentrations aux bornes du domaine de calcul. Ces informations sont indispensables pour assurer la cohérence du résultat et la fiabilité des résultats. Cela s'applique quelle que soit l'échelle de calcul.

Il existe bon nombre de codes numériques modélisant la pollution atmosphérique, et ce pour de nombreuses échelles de calcul. Certains codes numériques sont développés par des sociétés comme Aria Technologies, Acri... Ils modélisent les phénomènes de dispersion principalement, à des échelles très locales (rue canyon, sortie de tunnel, industrie...). Pour représenter de plus grandes zones avec une haute résolution spatiale (et donc temporelle), plusieurs modèles sont couplés. Les résultats du modèle à grande échelle servent de contraintes et conditions aux limites pour les modèles aux échelles plus petites. Aria Technologies, par exemple, utilise le modèle très connu : CHIMERE©, pour contraindre ses propres modèles.

Il existe des modèles très utilisés dans le domaine de la qualité de l'air. Ce sont les modèles CHIMERE©, STREET et ADMS. Leurs échelles spatiales de travail sont très différentes.

Le modèle CHIMERE©, développé par l'Institut Pierre Simon Laplace (IPSL / CNRS), est un modèle de « chimie-transport », c'est-à-dire qu'il modélise à la fois les réactions chimiques entre polluants et les phénomènes de transport. Il s'agit d'un modèle multi-échelle qui est *a priori* conçu pour faire des prévisions quotidiennes de concentration en ozone, aérosols ou autres polluants et pour faire des simulations à long-termes d'émissions pour l'analyse de scénarios. CHIMERE© fonctionne avec des échelles spatiales allant de l'échelle régionale (plusieurs milliers de kilomètres) à l'échelle d'une ville (100-200 km) avec un pas de grille allant de 1-2 km à 100 km. Il offre la possibilité d'intégrer des mesures pour ajuster les résultats de calcul, ce qui augmente la précision des cartographies obtenues.

La figure 2.2 est un exemple de carte de concentrations en particules (PM₁₀) obtenue par modélisation avec CHIMERE© associé au code de Météo France MOCAGE.

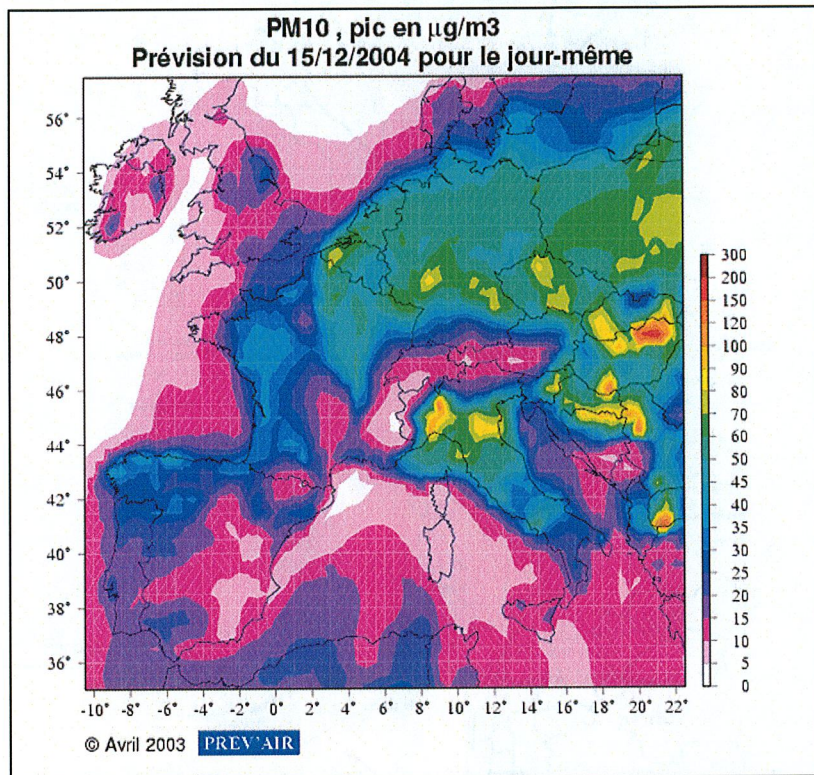


Fig. 2.2 : exemple de carte de concentrations en PM₁₀ sur l'Europe obtenue avec CHIMERE©, source www.prevoir.org.

STREET est un logiciel d'évaluation de la concentration dans l'air ambiant de polluants liés à la circulation automobile à l'échelle de la rue. Il s'appuie sur la typologie du réseau routier. Il a été conçu par la société allemande TÜV Ecoplan Umwelt GmbH. C'est un modèle simple et peu coûteux qui a séduit bon nombre d'AASQA pour établir des cartographies de pollution à l'échelle de la rue. La figure 2.3 est un exemple de carte des immissions en particules obtenue avec ce logiciel.

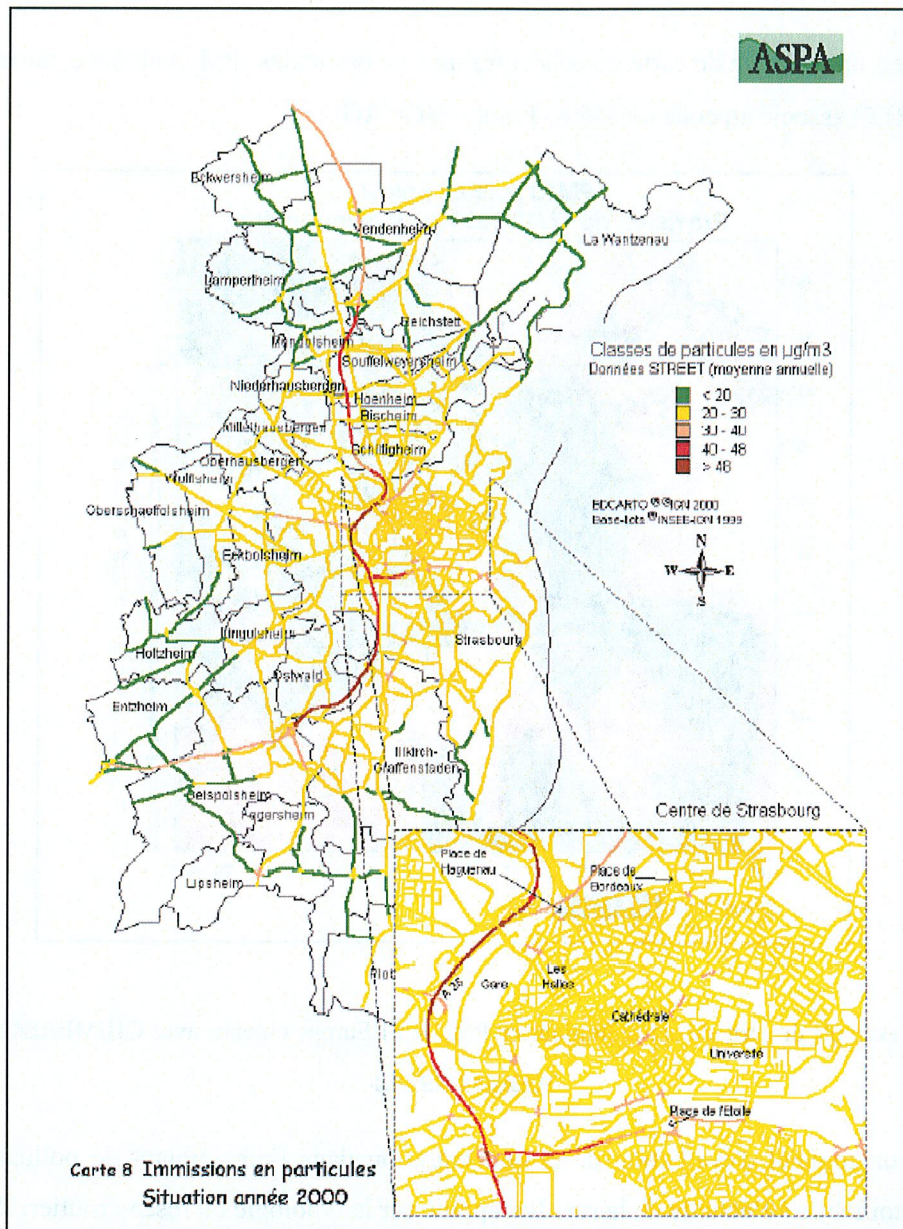


Fig. 2.3 : carte des immissions en particules sur la CUS obtenue par le logiciel STREET. Source : ASPA

Assez similaire, le modèle ADMS (Atmospheric Dispersion Modeling System), créé en 1998 par le CERC (Cambridge Environmental Research Consultants Ltd) est un modèle de dispersion qui permet d'évaluer et de prévoir la qualité de l'air urbain en tenant compte du trafic routier, des émissions domestiques et industrielles. Il est également possible de faire intervenir dans la modélisation les mesures issues des stations. Son domaine d'application va de la rue à l'agglomération avec un pas de grille de l'ordre du mètre. Les simulations peuvent se faire sur du long terme, en général elles vont de l'heure à l'année avec une résolution temporelle horaire. Ce modèle est utilisé dans de grandes métropoles européennes comme Londres et Budapest. Dans une étude sur les modèles de dispersion, Descamps *et al.* (2000) rapporte qu'il existe peu de codes permettant la modélisation correcte de la dispersion des particules en zone urbaine. Les modèles d'approche gaussienne comme ADMS,

simulent bien la dispersion des particules en zone urbaine mais donnent des estimations peu précises. L'approche gaussienne ne considère aucune hétérogénéité que ce soit pour l'écoulement, la turbulence ou le sol. Les résultats donnés par ce type de code sont souvent éloignés des concentrations réelles mesurées. La figure 2.4 est un exemple de carte de concentrations en dioxyde d'azote obtenue avec ADMS.

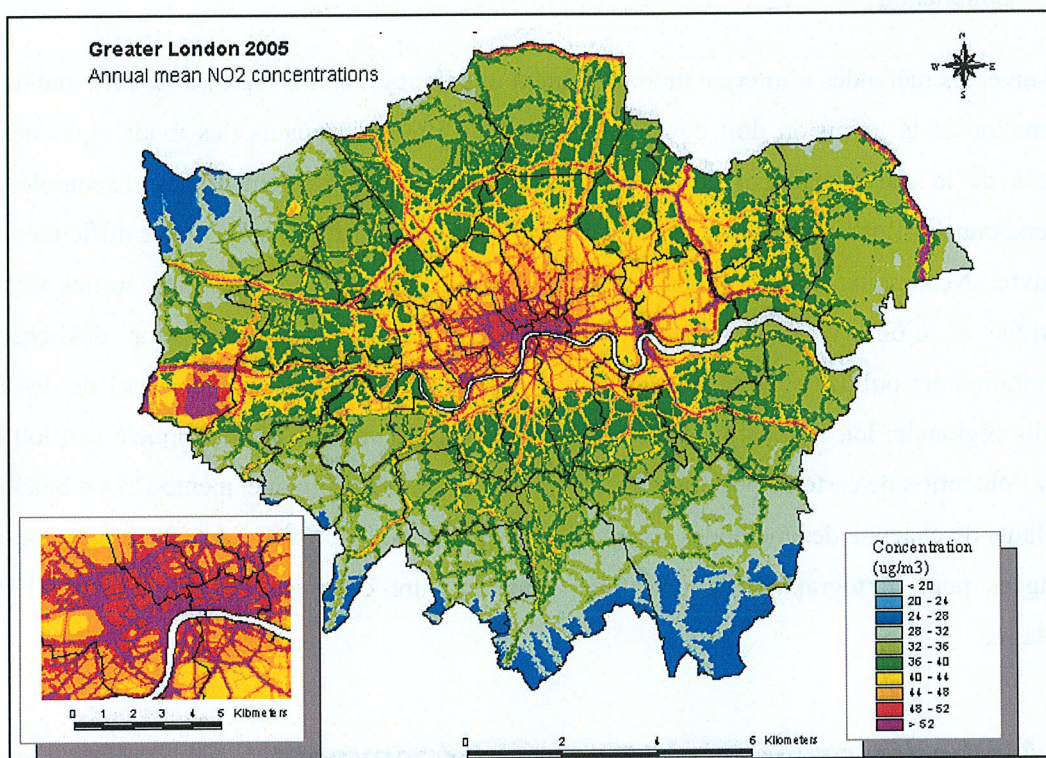


Fig. 2.4 : carte de concentrations en NO₂ obtenue avec ADMS sur l'agglomération de Londres, source <http://www.numtech.fr>

De manière générale, ces modèles donnent des résultats corrects mais leur usage par les AASQA est limité. Les AASQA privilégient l'interprétation des mesures qu'elles effectuent et donc l'approche par interpolation. L'utilisation des modèles déterministes pour la représentation de la qualité de l'air à l'échelle régionale reste donc assez exceptionnelle (Rouïl 2001). Ces outils sont plus utilisés pour mener des études de scénario ou des analyses d'épisodes de pollution. Les principaux obstacles à la représentation cartographique sont l'acquisition de l'inventaire d'émissions approprié et la compilation de données météorologiques appropriées, représentatives par exemple de la climatologie d'une région. Météo France mène cependant ce type d'analyse. L'emploi limité de ces méthodes par les AASQA s'explique également par le fait qu'en réalité, disposer d'une plate-forme opérationnelle de modélisation déterministe coûte cher et demande du personnel supplémentaire pour la mettre en œuvre.

Un autre inconvénient à l'utilisation de ces modèles est le fait qu'ils nécessitent l'intégration d'un nombre relativement élevé de paramètres d'entrée et que la modélisation reste très coûteuse en temps

de calculs. Les domaines d'applicabilité inhérents aux modèles posent également problème. Par exemple, pour STREET, la modélisation optimale est obtenue pour un domaine très peu étendu : une rue ; pour CHIMERE, sa résolution est trop basse (1-2 km) pour représenter de manière détaillée la variabilité de la concentration en polluants dans une ville.

1.4. Conclusion

En résumé, les méthodes d'interpolation présentent une limitation non négligeable en terme de qualité d'estimation et la précision doit être améliorée. Pour être performants, les modèles numériques à l'échelle de la rue demandent des paramètres qui ne sont généralement pas disponibles. Prises indépendamment l'une de l'autre, ces deux méthodes présentent des défauts et des difficultés de mise en œuvre. Néanmoins, Blond (2002) indique que l'utilisation combinée des sorties de modèle numérique et d'observations de surface permet d'améliorer la représentation des champs de concentration des polluants. Des cartes de concentrations en ozone sont obtenues sur l'Ile-de-France, à l'échelle régionale. Ici, il apparaît que la résolution spatiale imposée par le modèle soit toujours un frein à l'obtention de carte de pollution à l'échelle de la rue. Il semble que, même s'il y a amélioration, les défauts de chacune des méthodes soient toujours présents. Par conséquent, d'autres solutions sont envisagées pour cartographier la pollution. L'une d'entre elles est l'exploitation de l'imagerie satellitaire.

2 **Méthodes originales proposées récemment**

2.1. Utilisation de l'imagerie satellitaire

L'imagerie satellitaire est largement utilisée pour l'analyse et la surveillance de la pollution atmosphérique aux échelles globale ou régionale, c'est-à-dire pour des zones couvrant le globe, le continent ou un pays. Les concentrations en polluants sont mesurées directement par le capteur ou déduites des valeurs d'épaisseurs optiques estimées à partir des images satellites. Des satellites d'observation de la Terre sont entièrement dédiés à la surveillance de polluants principalement l'ozone, les aérosols et les gaz à l'état de traces. Les capteurs SCIAMACHY, GOME, et ModIS et ATSR-2 des satellites Envisat, Terra et Aqua, et ERS-2 en font partie. Le premier est un spectromètre d'absorption, le deuxième est un spectro-radiomètre et les derniers sont des radiomètres imageurs.

Certains auteurs utilisent les valeurs d'épaisseurs optiques des aérosols mesurées par MODIS pour déterminer la charge en aérosols troposphériques (Kaufman & Tanré 1998) ou les concentrations en PM_{2,5} et PM₁₀ (Engel-Cox *et al.* 2004). En première approximation, on fait souvent l'hypothèse que l'épaisseur optique des aérosols est une fonction linéaire de la concentration en aérosols (Retalis *et al.*

2003). La validation se fait par comparaison avec les données mesurées au sol. De la même façon, Holzer-Popp *et al.* (2002) utilisent simultanément GOME et ATSR-2 pour déduire les valeurs d'épaisseur optique des aérosols puis la concentration en aérosols. Robles-Gonzalez *et al.* (2003) utilisent ATSR-2 pour déduire des caractéristiques des aérosols : l'épaisseur optique et le coefficient d'Angström qui indique la taille des particules. Hashim *et al.* 2004 ont étudié l'influence des PM₁₀ sur la radiométrie des pixels d'une image NOAA-14 AVHRR. De cette étude, ils déduisent directement des images satellitaires la concentration en PM₁₀, puis par corrélation celles d'autres polluants, sur la péninsule malaisienne. La validation se fait généralement par comparaison avec des mesures au sol, des résultats de modèles numériques et des mesures issues d'autres capteurs spatio-portés. Pour ces auteurs, les cartes obtenues sont à petite échelle cartographique : elles couvrent de grandes zones, comme l'Europe ou l'Amérique du Nord, même si Engel-Cox *et al.* (2004) indiquent que les données de MODIS peuvent être utiles pour la surveillance de la qualité de l'air urbaine.

Les capteurs cités ci-dessus ont des résolutions spatiales qui limitent leur exploitation pour réaliser des cartes de pollution suffisamment précises à l'échelle qui nous intéresse, *i.e.* à l'échelle de la rue, même si les méthodes employées restent potentiellement transposables à nos échelles. Les travaux mentionnant l'emploi de l'imagerie satellitaire pour la cartographie de la pollution aux échelles de la ville sont beaucoup plus rares.

Toutefois, Retalis *et al.* (2003) expliquent que pour utiliser les satellites d'observation de la Terre et pouvoir estimer et cartographier la pollution atmosphérique sur les villes, ils ont développé des techniques *ad hoc* sur des données de moyenne résolution spatiale (NOAA-15 AVHRR). Ils obtiennent ainsi des cartes représentant l'épaisseur optique des aérosols qu'ils relient aux concentrations en PM₁₀ sur l'agglomération d'Athènes.

Dans la littérature, les satellites utilisés pour une cartographie à l'échelle de la rue ont de ce fait, plus souvent, des résolutions spatiales plus élevées. Le capteur TM de Landsat a une résolution de 30 m pour les canaux du visible et de 120 m pour l'infrarouge thermique. Les satellites SPOT-1 à -4 ont une résolution de 10 m en panchromatique et 20 m en multispectral.

Sifakis (1992) et Sifakis *et al.* (1992) utilisent le capteur TM pour localiser les voiles de pollution en dioxyde de soufre (SO₂) et en particules au-dessus de la ville d'Athènes. L'approche se fait par comparaison de deux images satellites. La première coïncide avec un épisode de pollution et la seconde à un épisode de non pollution. La méthode consiste en une comparaison réalisée localement entre les contrastes apparents, contenus dans des données relatives à l'image satellite de référence (*i.e.* date de non pollution) et ceux relatifs à des images représentatives de divers stades de pollution. En bref, la méthode consiste à trouver des anomalies de contrastes c'est-à-dire des modifications abruptes des gradients de l'image. En considérant que le laps de temps qui sépare les deux images n'est pas trop

important et que les changements d'origine humaine sont, par conséquent, inexistant, alors ces anomalies sont attribuées à la pollution atmosphérique. Cette méthode est appliquée aux longueurs d'onde du visible et de l'infrarouge thermique.

Dans le visible, les anomalies de contraste permettent de distinguer les variations de concentration en SO_2 et PM. La présence en quantité importante de ces polluants entraîne une diminution du contraste. Des cartes sont ainsi créées permettant la localisation du voile de pollution, les valeurs de concentrations en polluants restant inconnues. Sifakis *et al.* (1995) ont à nouveau appliqué cette méthode pour la détermination de la concentration en particules à partir d'images SPOT sur l'agglomération d'Athènes.

En ce qui concerne le canal infrarouge thermique TM6 (à $11 \mu\text{m}$), deux cartes de températures de surface (en réalité des cartes de luminance) sont créées puis comparées. Les baisses importantes de température sont attribuées au voile de pollution. L'apport de ce canal pour la représentation de la pollution (SO_2 et PM) a également été abordé par Finzi et Lechi (1991) sur la ville de Milan, par Poli *et al.* (1994) sur la ville de Rome ainsi que par Wald & Baleynaud (1999) sur Nantes.

Sifakis *et al.* (1999) ont amélioré la méthode, en combinant deux procédures appliquées conjointement aux données satellites. L'épaisseur optique de diffusion des aérosols dans le visible est ainsi évaluée. Les données satellite utilisées dans cet article sont des images TM et la zone étudiée est l'agglomération d'Athènes. Ces deux procédures mettent en évidence des effets atmosphériques physiquement indépendants. Ces effets s'expriment de façon précise sur des images HRS (Haute Résolution Spatiale) :

- « l'effet de flou » est un effet des particules sur les courtes longueurs d'ondes. Il se traduit par une réduction du contraste local engendrée par les particules diffusantes. Cette procédure s'inspire de la méthode appelée : « contrast-reduction method » proposée par Tanré *et al.* (1988).
- « l'effet de noircissement » est un effet des particules sur les grandes longueurs d'ondes. Il se traduit par un voilage des images engendré par l'atténuation des radiations due majoritairement à l'absorption par les particules.

A l'issue de cette méthode, des cartes représentant les variations d'épaisseurs optiques des aérosols sur Athènes sont produites. Puis des corrélations sont calculées pour relier ces valeurs aux concentrations de polluants.

Kanaroglou *et al.* (2002) combinent même à la fois des méthodes d'interpolation géostatistiques et des données satellitaires pour obtenir des cartes de pollution.

D'autres méthodes originales de cartographie ont été étudiées à l'Ecole des Mines de Paris, faisant souvent appel à l'imagerie satellitaire, et sont détaillées ci-après.

2.2. Méthode des «champs typiques» originelle (Ung 2003)

2.2.1. Principe

Le concept proposé par Ung (2003) est le suivant : il s'agit d'utiliser la fusion de données pour synthétiser des cartes de pollution représentant le champ de concentration sur l'ensemble de la ville à l'échelle de la rue, à partir de données à différentes résolutions spatiales et temporelles issues de sources différentes.

Ung définit le « champ typique » comme « l'ensemble des structures spatiales existantes d'un champ de concentration de polluant pour un type de situation météorologique ». Son idée est de créer un champ typique pour chaque situation météorologique, pour une zone donnée, afin de constituer une base de données réutilisable : la bibliothèque de champs typiques. A un instant donné, sur cette zone, estimer le champ de concentration devient plus facile puisque, connaissant la situation météorologique, on connaît également le champ typique correspondant, i.e. l'ensemble des structures de ce champ de concentration.

Les données que Ung propose d'utiliser sont :

- les observations de pollution aux stations de mesures,
- les cartes représentant le champ de concentration à l'échelle régionale fournies par les modèles numériques,
- les sorties des modèles numériques représentant la pollution à l'échelle d'une rue.

L'hypothèse de départ qui sous-tend le concept de « champ typique » est la suivante : une répétition des situations météorologiques, une faible variation temporelle de la rugosité de la ville et une répétition ou une continuité des sources d'émission des polluants engendrent nécessairement une répétition des épisodes de pollution. Les observations des AASQA ont confirmé cette hypothèse. L'idée est de caractériser la répartition spatiale des polluants pour toute situation météorologique définie.

La bibliothèque des champs typiques est un ensemble de champs typiques. Ces champs typiques de pollution peuvent être classés en plusieurs classes et sous-classes. Cette classification prend en compte la variabilité saisonnière et trois situations météorologiques pour chaque saison. Ung dénombre 12 champs typiques pour chacun des quatre polluants considérés. Si l'on suppose qu'une bibliothèque de

champs typiques a été établie pour l'ensemble des situations météorologiques connues pour la zone considérée, les étapes de mise en œuvre de la cartographie sont les suivantes :

- sélection du champ typique correspondant à la situation météorologique pour la date étudiée,
- relevé du niveau moyen de pollution de la zone en utilisant l'image du champ de concentration à basse résolution spatiale issue d'un modèle numérique,
- relevé des mesures au sol de concentrations,
- fusion du champ typique, de la sortie de modèle numérique et des mesures in situ dans un processus de cartographie assurant une cohérence des échelles spatiales.

A l'issue de la fusion, une cartographie du champ de concentration sur l'ensemble de la zone considérée à une échelle assez fine est réalisée. La figure 2.5 donne le schéma de principe de la méthode.

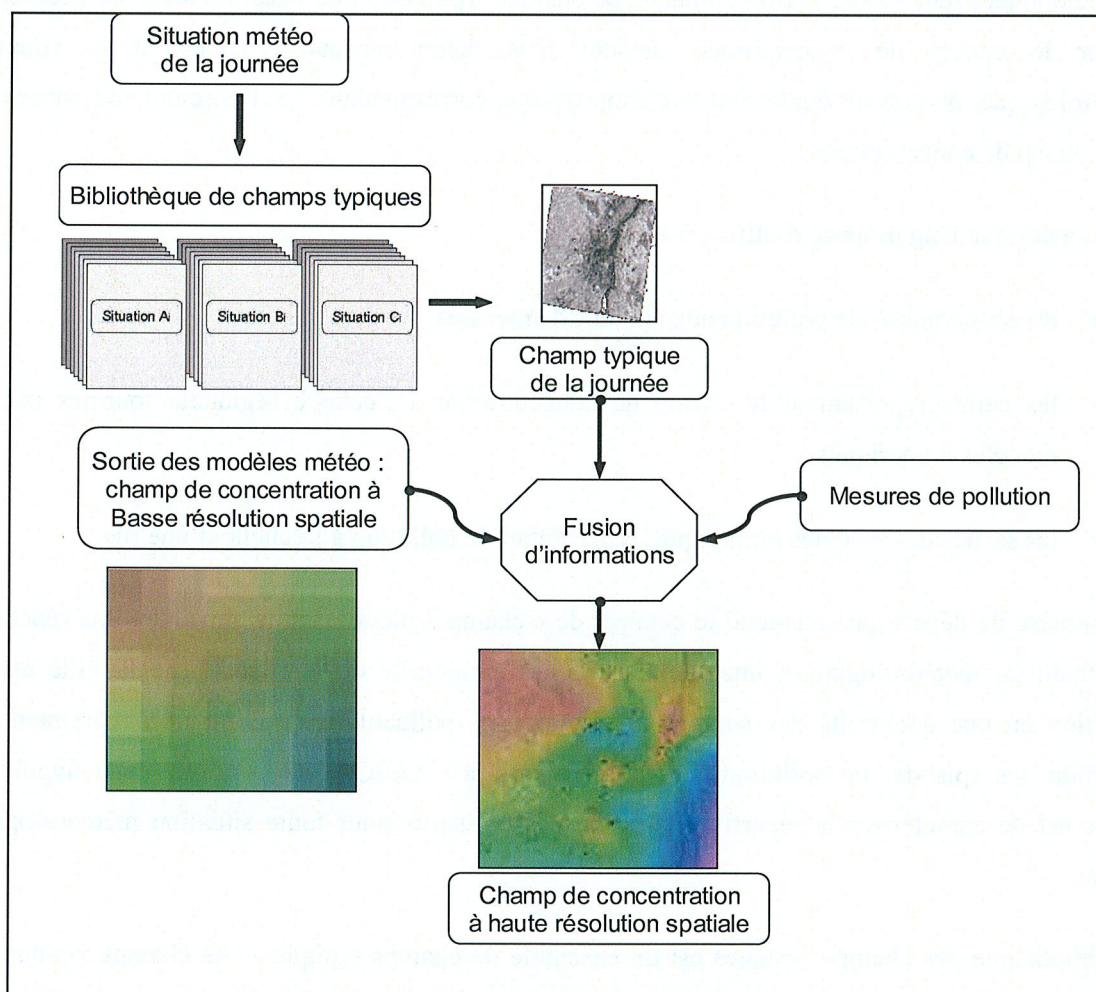


Fig. 2.5 : schéma de principe de la méthode des champs typiques (Ung 2003).

2.2.2. Définition et construction d'un champ typique

Un champ typique est défini comme étant l'ensemble des structures existantes d'un champ de concentration de polluants pour un type de situation météorologique. Les outils mathématiques utilisés pour définir un champ typique et fusionner les informations à des échelles différentes sont la transformée en ondelette et l'analyse multirésolution.

L'analyse multirésolution appliquée au domaine de la cartographie de la pollution de l'air permet d'établir une base de représentation des données de sources différentes. En d'autres termes, les données de sources différentes sont vues comme étant des approximations d'une fonction théorique désirée à une échelle beaucoup plus fine. La carte issue du modèle numérique à l'échelle régionale apparaît comme une approximation du champ de concentration à l'échelle de la rue.

Les informations sur les structures spatiales existantes en terme d'approximation du champ de concentration désirée sont les coefficients en ondelettes : ce sont des mesures locales de la structure du champ de concentration à différentes échelles spatiales. Autrement dit, le champ de concentration à haute résolution peut se résumer à un ensemble de coefficients en ondelettes, constitué par les sous-ensembles des coefficients en ondelettes présents à chaque échelle, et une approximation à basse résolution.

Pour assurer la cohérence entre toutes les échelles spatiales et la réalité terrain, Ung (2003) impose à certaines cellules du champ (celles correspondant aux stations de mesures) la valeur de concentration mesurée par les stations de surveillance.

La construction du champ de concentration à haute résolution nécessite la connaissance de l'ensemble des coefficients en ondelettes mais surtout de ceux à cette résolution. Cependant ces coefficients en ondelettes ne sont pas accessibles avec les outils actuels de cartographie : toutes les petites structures du champ de concentration à haute résolution ne sont pas connues. Ung propose alors une méthode pour les modéliser.

Cette modélisation est basée sur la connaissance des épisodes de pollution passés. Pour une situation météorologique passée identique à celle qu'on étudie, on suppose connu le champ de concentration à l'échelle la plus fine désirée ainsi que toutes les estimations de ce champ à toutes les échelles. Comme la situation météorologique est la même, les structures spatiales présentes vont être identiques à celles que l'on cherche. Seuls les niveaux de pollution varient mais ils sont connus grâce à l'image basse résolution. En normalisant, on extrait seulement les structures spatiales : on obtient un champ normalisé. Ung envisageait de faire cela pour l'ensemble des configurations possibles du champ de concentration existantes pour une situation météorologique donnée. Il voulait faire une sorte de « moyenne » afin de déterminer les tendances dans la distribution de la pollution.

Une analyse multirésolution appliquée au champ normalisé permet d'extraire l'ensemble des coefficients en ondelettes. Puis le processus inverse permet de créer le champ de concentration à haute résolution. Partant de l'image basse résolution qui impose les niveaux moyens de pollution, on introduit les coefficients en ondelettes et on synthétise ainsi le champ de concentration à l'échelle d'une rue.

Le problème majeur est la création des estimations de ces champs de concentrations passés. Ung propose d'utiliser les méthodes d'interpolation / extrapolation des données. Comme le peu de mesures à interpoler donne souvent une mauvaise représentation du champ de concentration, il propose de densifier artificiellement le réseau de mesures. La figure 2.6 schématise la construction d'un champ typique et celle de la carte haute résolution.

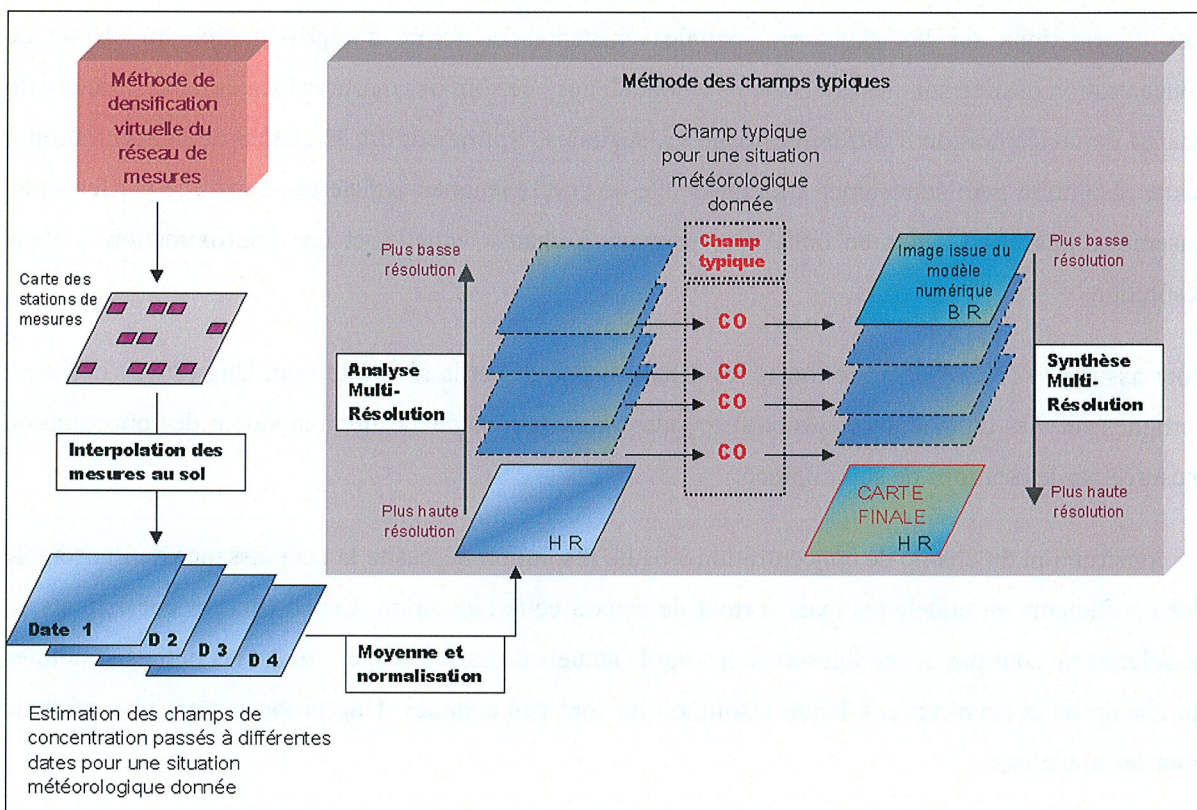


Fig. 2.6 : schéma de construction d'un champ typique et de la carte de pollution haute résolution.

2.2.3. Discussion

Le concept est relativement innovant du fait de peu de références dans le domaine mentionnant l'application de la fusion de données à la qualité de l'air (Sarigiannis *et al.* 2005, Wald & Baleynaud 1999). L'emploi de la fusion de données est attractif, par rapport aux méthodes classiques de cartographie, puisqu'elle permet de manipuler des échelles spatiales différentes ce qui est très utile

lorsque l'on dispose de peu de données à une même échelle. Cependant, Ung n'a abordé que la partie théorique de la méthode. Le concept a été décrit mais le formalisme mathématique n'a été qu'esquissé et la mise en œuvre est inexistante. Toutefois, la méthode semble prometteuse et une étude approfondie est nécessaire pour le confirmer.

La méthode d'estimation du champ de concentration, proposée par Ung, peut se présenter comme étant une injection des structures du champ de concentration pour améliorer les cartes de concentrations existantes à l'échelle régionale. Cette méthode est couramment employée dans le domaine du traitement d'image. Elle se rapproche du concept ARSIS « Amélioration de la Résolution Spatiale par Injection de Structures » (Ranchin & Wald 2000). Pour améliorer une image multispectrale haute résolution sans altérer l'information, les auteurs proposent de lui ajouter uniquement les structures manquantes à une échelle caractéristique voulue.

Ung propose certaines simplifications de mise en œuvre par rapport à la description théorique initiale de la méthode. Les modèles numériques à l'échelle de la rue fournissent une approximation du champ de concentration mais celle-ci est incomplète spatialement, ce qui rend ces données difficilement exploitables. Ung simplifie le problème et n'envisage pas d'utiliser les sorties de modèles numériques à l'échelle de la rue. Le nombre de données réellement utilisables pour cette méthode est très réduit. Il y a seulement les sorties de modèles numériques à l'échelle régionale et les mesures de concentrations relevées par le réseau de surveillance de la qualité de l'air. Ensuite, il écarte le problème des résolutions temporelles différentes, non gérées par les outils de fusion de données en proposant de créer des cartes instantanées. Les données qu'il propose d'utiliser ont alors la même résolution temporelle. La méthode n'a été qu'ébauchée et aucune mise en œuvre pratique n'a été faite, ce qui incite à vérifier la faisabilité de la méthode et son applicabilité (chapitre 3).

Finalement, la méthode des champs typiques telle qu'elle est présentée par Ung n'est que la première partie d'une méthode globale de cartographie associant la méthode des champs typiques à celle de densification virtuelle du réseau de mesures. L'inconvénient principal de cette approche est que pour obtenir une carte de pollution à haute résolution, on est obligé de disposer déjà de cartes de pollution à haute résolution. En effet, l'objectif est d'obtenir une carte couvrant l'ensemble d'une ville qui représente le champ de concentration à l'échelle de la rue parce que les méthodes actuelles de cartographie ne donnent pas de résultats satisfaisants à cette échelle. Cependant pour obtenir cette carte, Ung indique qu'il est nécessaire de disposer d'estimations du champ de concentration à l'échelle de la rue. Pour faire cela, il propose d'utiliser les méthodes classiques d'interpolation combinée à une nouvelle méthode basée sur la densification du réseau de mesures. Ceci constitue un inconvénient majeur.

La méthode présente également une faille au niveau de la détermination du nombre de champs typiques. En effet, les paramètres intervenant dans le classement des champs typiques en plusieurs classes et sous-classes ne concernent que la météorologie. Or les variabilités temporelles de la pollution sont également liées aux activités humaines, par exemple les émissions en polluants sont supérieures en semaine à celles du week-end. Cet aspect lié aux activités humaines devrait être pris en compte dans la détermination du nombre de champs typiques existant sur une zone donnée.

2.3. Méthode de densification du réseau de mesures

2.3.1. *Les premiers travaux (Basly 2000)*

La méthode de densification virtuelle a été initiée par Basly (2000) pour la cartographie des fumées noires sur la ville de Nantes. Dans sa thèse, Basly a trouvé une relation liant la concentration en fumées noires (FN) au valeur numérique des pixels du canal infrarouge thermique (TM6) du satellite Landsat 5. En admettant que la présence de FN soit la principale cause de modification de la valeur numérique des pixels, la relation peut s'écrire :

$$[FN] = aTM6 + b \quad (2.11)$$

avec $[FN]$, la concentration en FN, $TM6$, la mesure effectuée par le canal thermique, a , la pente et b , l'ordonnée à l'origine. Bien sûr, la relation liant le contenu en aérosols à la mesure satellitaire n'est pas linéaire. Cependant, l'intervalle des mesures, satellitaires et terrain, étant relativement faible, Basly a opté pour une approximation linéaire. Cette relation a été déterminée en corrélant les mesures relevées au niveau des stations à celle mesurées sur les pixels de l'image satellite. Pour cartographier les FN, il pense tout d'abord l'appliquer à l'ensemble des pixels de l'image TM6. Cependant, des considérations liées aux paramètres influant sur la radiométrie des pixels (par exemple : comment distinguer les pixels dont la radiométrie est influencée par la pollution et ceux dont le compte numérique est principalement dû à l'émissivité des objets ?) le pousse à sélectionner les pixels sur lesquels il va appliquer la relation.

Les endroits où sont localisées les stations de mesure répondent à certaines caractéristiques physiques d'occupation du sol. La relation (2.11) ne doit être appliquée qu'aux zones de l'image répondant aux mêmes caractéristiques. Basly va donc rechercher les pixels de l'image TM6 qui sont spectralement similaires aux pixels correspondants aux stations. Cette sélection de pixels est basée sur l'information multibande fournie par les six autres bandes du capteur TM (TM1, 2, 3, 4, 5, 7). Les pixels retenus sont appelés des pseudo-stations. A ces pixels, on affecte une valeur de concentration en FN calculée par application de la relation liant les luminances aux mesures au sol. Basly augmente ainsi artificiellement le nombre de points de mesure. La cartographie se fait ensuite par interpolation /

extrapolation des valeurs des pseudo-stations. Pour affiner sa méthode de cartographie, Basly a testé différentes combinaisons de canaux pour la détermination des pseudo-stations. Finalement, la combinaison donnant les meilleurs résultats cartographiques est la combinaison des canaux TM1, 2, 3 et 4. Basly a comparé ces résultats à ceux obtenus par des méthodes classiques d'interpolation des mesures au sol. Il constate que la corrélation est augmentée, le biais divisé par deux et la RMSE (Root Mean Square Error) est diminuée de près de moitié. Il en déduit que cette méthode de densification artificielle du réseau de mesures améliore la cartographie de la distribution spatiale des FN.

2.3.2. Les apports de Ung (2003)

Ung a repris le concept de « pseudo-stations » et l'a enrichi de deux nouvelles notions : la notion de « station virtuelle » et celle de « carte d'identité ».

Chaque élément de la ville est identifié par une carte d'identité. Elle contient un ensemble d'éléments permettant la caractérisation du lieu considéré :

- la position géographique du lieu considéré,
- le type d'occupation du sol,
- la proximité par rapport aux sources d'émissions,
- la morphologie de l'espace entourant la zone,
- les conditions climatiques et météorologiques du lieu.

Ung définit les pseudo-stations comme des lieux de la ville présentant des propriétés morphologiques, climatiques, de pollution et une occupation du sol similaires à celles des stations de mesures. Il suppose que les pseudo-stations, ayant des propriétés similaires à celles des stations de mesures, ont un comportement semblable à celui des stations vis-à-vis de la pollution. La localisation des pseudo-stations relève d'une classification basée sur les cartes d'identité.

Ung a identifié les données répondant aux objectifs de caractérisation du lieu. Pour la position géographique, il utilise une base de données topographiques fournie par l'IGN (Institut Géographique National). Il propose d'acquérir des informations sur le type d'occupation du sol grâce aux images satellites (Landsat 5 TM). Pour connaître la proximité par rapport aux sources d'émissions, il propose le cadastre des émissions. Enfin, la morphologie de l'espace entourant la zone est caractérisée par des indicateurs morphologiques, une surface de visibilité et un périmètre de visibilité.

Finalement, pour cette application, Ung se limite uniquement aux données satellitaires provenant des huit dates de prises de vue satellite. La carte d'identité contient alors uniquement la signature spectrale

multidate des stations de mesure. La localisation des pixels ayant une même signature spectrale est effectuée station de mesure par station de mesure et date par date. Ung effectue une analyse en composante principale (ACP) pour chacune des huit date de prises de vue satellitaire, prise séparément, et à chaque fois sur toutes les bandes spectrales à l'exception du canal thermique TM6. Ce canal est exclu de la classification puisqu'il ne renseigne pas sur l'occupation du sol. L'ACP permet de déterminer les canaux qui à eux seuls décrivent le mieux l'image. Il trouve une combinaison de trois canaux. L'ensemble des données est projeté sur trois axes principaux.

Ensuite, une classification est effectuée qui regroupe les pixels semblables aux pixels-stations. Elle prend en compte la dynamique du signal projeté. Un pixel est classé comme étant semblable à un pixel-station si les valeurs numériques des pixels dans chacun des trois canaux sont identiques à celles des pixels-stations. Pour chacune des prises de vue satellitaires et à partir des 11 stations de mesures, il obtient onze classes de pixels semblables dont la population (nombre de pixels localisés) varie suivant la classe. Pour déterminer les pixels qui sont réellement des pseudo-stations *i.e.* qui présentent une signature spectrale stable dans le temps et identique aux pixels-stations, une intersection multidate des pixels est effectuée. Ung dénombre 38 pseudo-stations c'est-à-dire beaucoup moins qu'en ne considérant qu'une date de prise de vue mais nettement plus que les 11 stations de mesures de départ.

A partir de ces pseudo-stations, Ung va déterminer les stations virtuelles. Il appelle « stations virtuelles » le sous-ensemble des pseudo-stations, sur lequel une estimation de la pollution est possible à l'aide des images satellitaires et des observations de concentrations de polluants. L'estimation de la pollution à partir des images satellites nécessite l'établissement d'une loi reliant les mesures satellites aux concentrations des polluants via la transmittance de l'atmosphère. En première approximation, la réflectance mesurée au niveau du satellite s'écrit :

$$\rho_{sat} = \rho_{atm} + \rho_{surf} \tau_{atm} \quad (2.12)$$

avec ρ_{atm} , la réflectance de l'atmosphère, ρ_{surf} , la réflectance de la surface terrestre, ρ_{sat} , la réflectance mesurée par le satellite et τ_{atm} , la transmittance de l'atmosphère.

Le principe retenu dans son approche, est que la différence de réflectance d'une pseudo-station par rapport à la station de la même classe reflète une variation d'épaisseur optique et donc de concentration de polluants par rapports à cette station. Les difficultés consistent à éliminer les effets purement atmosphériques puis à estimer un modèle convertissant les variations de réflectance en concentration. Ung précise qu'il est vraisemblable que le principe ne peut s'appliquer réellement qu'à un sous-ensemble des pseudo-stations, d'où cette notion de station virtuelle.

Finalement, Ung a calculé les corrélations entre les réflectances corrigées des effets de l'atmosphère et les concentrations en polluants mais n'a pas déterminé de relation les liant. Il n'a donc pas déterminé le sous-ensemble des stations virtuelles et testé la cartographie par interpolation comme l'a fait Basly.

2.3.3. Discussion

La principale différence entre le travail de Basly (2000) et Ung (2003) est le domaine spectral. Basly a cartographié des fumées noires à partir des données de l'infrarouge thermique et Ung s'est attaché à étudier la pollution quelle soit particulaire ou moléculaire à partir des données satellites issues du domaine visible.

L'introduction de la notion de « carte d'identité » est innovante : elle a une réalité physique et un aspect pratique. Toutefois elle n'a été exploitée qu'en partie. En effet, la seule utilisation des images satellites n'a permis de satisfaire que quelques éléments de la carte d'identité : l'occupation du sol, le niveau de pollution et la localisation. La morphologie des sites décrite par les indicateurs, surface et périmètre de visibilité n'a pas été prise en compte alors qu'elle joue un rôle majeur dans la circulation des masses d'air polluantes.

L'application de l'ACP sur la totalité de l'image peut se discuter. En effet, pour ne pas traiter la totalité des canaux de l'image et utiliser suffisamment d'informations, Ung effectue une ACP. Il détermine ainsi les canaux qui décrivent le mieux l'image dans son ensemble. Cependant, à un endroit particulier de l'image, pour un pixel donné, par exemple celui correspondant à une station, la combinaison des trois canaux déterminés par l'ACP ne décrira peut-être pas entièrement ce lieu. La détermination des pseudo-stations sera biaisée. Les pixels correspondant à la station n'auront pas exactement les mêmes caractéristiques.

La méthode proposée par Ung est intéressante en particulier la notion de « carte d'identité ». Mais la difficulté de modélisation de la relation liant les concentrations aux réflectances n'a pas permis le passage des pseudo-stations aux stations virtuelles. Finalement, le nombre de mesures n'a pas été augmenté ce qui n'a pas permis de réaliser la cartographie et d'évaluer ainsi l'intérêt de la méthode. Toutefois, les éléments proposées par Ung : carte d'identité, pseudo-station et station virtuelle, sont innovants et méritent d'être exploités.

Bibliographie

- Basly L., 2000. Télédétection pour la qualité de l'air en milieu urbain. Thèse de doctorat Sciences des technologies de l'information et de la communication, Université de Nice-Sophia Antipolis, Nice, France, 182 p.
- Bel L., Poggi J. M., Portier B., Vautard R., 2001. Statistical forecasting of ozone peaks over Paris area with a nonlinear additive model. Impact of imported ozone. *Applied Stochastic Models and Data Analysis*, Eds G. Govaert, J. Janssen, N. Limnios, Vol 1, 166-171.
- Bergthorsson P., Doos B., 1955. Numerical weather map analysis. *Tellus*, 7, 329-340.
- Blond, N., 2002. Assimilation de données photochimiques et prévision de la pollution troposphérique, Ecole Polytechnique, Palaiseau, France, 2002, 202 p., p. 161-162.
- Carletti R., Picci M., Romano D., 2000. Kriging and bilinear methods for estimating spatial pattern of atmospheric pollutants. *Environmental Monitoring and Assessment*, Vol. 63, pp. 341-359.
- Chiles J. P., Chauvet P., Delfiner P., 1975. *Analyse des champs météorologiques par krigeage*. Centre de Morphologie Mathématique de Fontainebleau. Ecole des Mines de Paris, juillet 1975, 115 p.
- Cressman, G. P., 1959. *An operational objective analysis system*. Monthly Weather Review, 87, pp. 367-374.
- Descamps I., Meunier M. A., Harion J. L., 2000. Modèles de dispersion de particules en milieu urbain. Etude n° 3, Département Energétique Industrielle, Ecole des Mines de Douai, novembre 2000, 61 p.
- Delfiner P., 1973. Analyse objective du géopotential et du vent géostrophique par krigeage universel. Note EERM n° 321.
- Engel-Cox J.A., Holloman C.H., Coutant B.W., Hoff R.M., 2004. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmospheric Environment*, 38, 2495-2590.
- Finzi G., Lechi G.M., 1991. Landsat images of urban air pollution in stable meteorological conditions. *Il Nuovo Cimento*, 14C, 433-443.

Frangi J.P., Jacquemoud S., Puybonnieux-Textier V., Lazard H., 1996. Suivi spatio-temporel de la concentration en NO_x en île de France. *Compte Rendu Académie des Sciences, Paris* 323, série II a, 373-379.

Gratton Y., 2002. Le krigeage : la méthode optimale d'interpolation spatiale. Les articles de l'Institut d'Analyse Géographique, juin 2002, 4 p., <http://www.iag.asso.fr>

Hashim M., Kanniah K. D., Ahmad A., Rasib A. W., Ibrahim A. L., 2004. The use of AVHRR data to determine the concentration of visible and invisible tropospheric pollutants originating from a 1997 forest fire in Southeast Asia. *International Journal of Remote Sensing*, 25 (21), 4781-4794.

Holton J. R., 1972. An introduction to dynamic meteorology. *International Geophysics Series*, Academic Press, London, UK, 319 p, p. 28.

Holzer-Popp T., Schroedter M., Gesell G., 2002. Retrieving aerosol optical depth and type in the boundary layer over land and ocean from simultaneous GOME spectrometer and ATSR-2 radiometer measurements, 1, Method description. *Journal of Geophysical Research*, 107, D21, AAC16, 1-17.

Holzer-Popp T., Schroedter M., Gesell G., 2002. Retrieving aerosol optical depth and type in the boundary layer over land and ocean from simultaneous GOME spectrometer and ATSR-2 radiometer measurements, 2, Case study application and validation. *Journal of Geophysical Research*, 107, D24, AAC10, 1-8.

Ionescu A., Mayer E., Colda I., 1996. Méthodes mathématiques pour estimer le champ de concentration d'un polluant gazeux à partir des valeurs mesurées aux points dispersés. *Pollution Atmosphérique*, 149, 78-89.

Ionescu A., Candau Y., Mayer E., Colda I., 2000. Analytical determination and classification of pollutant concentration fields using air pollution monitoring network data. Methodology and application in the Paris area, during episodes with peak nitrogen dioxide levels. *Journal of Environmental Modelling & Software* 15, 565-573.

Kanaroglou P. S., Soulakellis N. A., Sifakis N. I., 2002. Improvement of Satellite Derived Pollution Maps with the Use of a Geostatistical Interpolation Method. *Journal of Geographical Systems*, 4, 193-208.

Kaufman Y. J., Tanré D., 1998. Algorithm for Remote Sensing of Tropospheric Aerosol from MODIS. http://tpwww.gsfc.nasa.gov/MODIS-Atmosphere/docs/atbd_mod02.pdf (date de l'accès : 12 juillet 2006)

Matheron, G., 1963, Principles of Geostatistics: Economic Geology, v. 58, p. 1246-1266.

Mestayer P.G., Bottema M., Costes J. P., Sini J. F., 1997. Modelling urban canopy and terrain for transport-diffusion simulations at sub-mesoscales. *International Journal of Environment and Pollution*, 8, n° 3-6, pp 675-682.

Poli U., Pignatoro F. Rocchi V. and Bracco L., 1994. Study of the heat island over the city of Rome from Landsat-TM satellite in relation with urban air pollution. In: Proceedings of 13th EARSeL Symposium, Remote sensing – From research to operational applications in the new Europe, Dundee, Scotland, UK, Springer Hungarica, pp. 413-422.

Puybonnieux-Texier V., 1996. Cartographie de la pollution atmosphérique, interpolation spatiale. Mémoire de DEA "Sciences de l'Information Géographique", Université de Marne-la-Vallée.

Retalis, A., Sifakis, N., Grosso, N., Paronis, D. and Sarigiannis, D., 2003. Aerosol optical thickness retrieval from AVHRR images over the Athens urban area. In: Proceedings of IGARSS 2003. http://sat2.space.noa.gr/rsensing/documents/IGARSS2003_AVHRR_Retalisetal_web.pdf

Robles-Gonzalez C., Schaap M., de Leeuw G., Builtjes P.J.H., van Loon M., 2003. Spatial variation of aerosol properties derived from satellite observations and comparison with model calculations. *Atmospheric Chemistry and Physics*, 3, pp. 521-533.

Rouïl L., 2001. *Représentation et prévision de la qualité de l'air en France : bilan et perspectives*. Rapport INERIS n° 56, 13 juin 2001, 11 p.

Sarigiannis D. A., Gotti A., Sifakis N. I., Tombrou M., Dandou A., Schaefer K., Emeis S., Soulakellis N., 2005. *High-resolution estimation of urban aerosol from fusion of satellite and ground data with numerical modelling results*. In Proceedings of the 5th international conference on Urban Air Quality, Valencia, Spain, mars 2005.

Sifakis N. I., 1992. *Potentialités de l'imagerie à haute résolution spatiale pour le suivi de la répartition de pollutions atmosphériques dans la basse troposphère. Etude du cas d'Athènes*. Thèse de doctorat, chimie de la pollution atmosphérique et physique de l'environnement, université Paris 7 Denis Diderot, Paris. 292 p.

Sifakis N., Bildgen P., Gilg J.P., 1992. *Utilisation du canal 6 (thermique) de Themactic Mapper pour la localisation de nuages de pollution atmosphérique. Application à la région d'Athènes (Grèce)*. Pollution atmosphérique, Vol. 34, pp. 96-107.

Sifakis N., 1995. *La télédétection des voiles de pollution atmosphérique et de la dégradation de l'environnement dans la région d'Athènes*. Photo-Interprétation, Vol. 4, pp. 220-225.

Sifakis N., Soulakellis N. and Paronis D., 1999. *Quantitative mapping of air pollution density using Earth observations: A new processing method and application on an urban area*. International Journal of Remote Sensing, January 1999, vol. 19, no. 17, pp. 3289-3300.

Tanré D., Deschamps P. Y., Devaux C., Herman M., 1988. *Estimation of Saharan aerosol optical thickness from blurring effects in Thematic Mapper data*. Journal of Geophysical Research, 93, pp. 15955-15964.

Ung A., 2003. *Cartographie de la pollution atmosphérique en milieu urbain à l'aide de données multisources*. Thèse de doctorat Sciences de l'environnement, Université Paris 7 Denis Diderot, Paris, France , 96 p.

Wald L. ,Baleynaud J.M., 1999. *Observing air quality over the city of Nantes by means of Landsat thermal infrared data*. International Journal of Remote Sensing, Vol. 20, 5, pp. 947-959.

Chapitre 3

Méthode des « champs typiques »

1 Contexte et objectifs

La première partie de l'état de l'art (chapitre 2) a montré les limitations des méthodes "classiques" de cartographie. Pour remédier au problème de précision des méthodes d'interpolation / extrapolation des mesures au sol et pour éviter la complexité des modèles numériques, Basly (2000), Ung (2003) et Ung et al. (2002) ont proposé un nouveau concept de cartographie de la qualité de l'air basé sur la fusion de données. Dans sa thèse, Ung a abordé la partie théorique de la méthode. Le concept a été décrit mais le formalisme mathématique n'a été qu'esquissé et la mise en œuvre est inexistante (chapitre 2, seconde partie). Le concept est innovant. Il existe peu de références dans le domaine mentionnant l'application de la fusion de données à la qualité de l'air (Sarigiannis *et al.* 2005, Wald & Baleynaud 1999). Par rapport aux méthodes classiques utilisées pour cartographier la qualité de l'air, la fusion de données permet de manipuler des échelles spatiales différentes ce qui est très utile lorsque l'on dispose de peu de données à une même échelle. Elle permet également de synthétiser des cartes à des échelles très fines à partir de données à des échelles plus grossières. Il nous paraissait indispensable d'explorer cette piste pour la cartographie de la qualité de l'air à des échelles très locales, d'autant plus que la plupart des données dont nous disposons sont à des échelles assez grandes.

Dans ce chapitre, je reprends le concept initié par Basly (2000), Ung (2003) et Ung *et al.* (2002) pour proposer une version plus aboutie. Je m'attache à décrire le formalisme mathématique de façon formelle. J'en fais une application à notre cas d'étude : la ville de Strasbourg. Enfin, je discute des difficultés de mise en œuvre et des aspects les plus prometteurs.

La méthode des *champs typiques* vise à améliorer la représentation de la pollution en appliquant des méthodes de fusion de données à des données ou cartes de pollution déjà existantes, issues de modèles numériques par exemple. Cette méthode s'applique dans le cas où l'on dispose déjà de données de pollution sous forme de cartes mais dont les résolutions sont différentes. La fusion de données vise à « l'obtention d'information de plus grande qualité » (Wald 2002). Les méthodes de fusion de données appliquées à la cartographie de la qualité de l'air n'ont été que rarement testées sur un cas pratique (Sarigiannis *et al.* 2005). L'enjeu immédiat consiste à étudier la faisabilité, la validité ainsi que l'applicabilité d'une telle méthode et son domaine d'application. Les objectifs à plus long terme seraient de créer des cartes haute résolution de la pollution urbaine locale pour chaque situation type bien identifiée sur une zone donnée. Une base de données cartographiques sur la pollution de l'air serait ainsi créée pour cette zone. Une fois construite, cette « bibliothèque » serait réutilisable *a posteriori* et à tout instant. En effet, on a fait l'hypothèse que, pour une situation type, la répartition spatiale des polluants est la même. Seul le niveau de pollution et l'amplitude moyenne varient. Ainsi pour chaque situation type, la répartition spatiale des polluants est caractérisée : on a construit un champ typique. Nous disposerions alors d'une bibliothèque de champs typiques correspondants à chaque type de situation. A un instant donné, grâce à cette bibliothèque, il serait possible de construire

un champ de concentration, beaucoup plus rapidement qu'avec les méthodes actuelles. Le facteur temps est important dans une optique de dissémination de l'information en temps réel vers des décideurs ou le public.

2 Hypothèse fondatrice

La méthode des champs typiques est fondée sur une hypothèse. On suppose que pour un polluant particulier, pour une situation météorologique précise et pour une situation où il y a constance des sources d'émissions, la forme des structures spatiales du champ de pollution reste la même. Cette approche par scénarios s'appuie sur des observations faites par les AASQA et est classiquement utilisée lors de la création de modèles physiques numériques dédiés à la qualité de l'air (Gaillard & Batton-Hubert 2003) ou pour la prévision des épisodes de pollution (Kim Oanh *et al.* 2005).

La distribution spatiale des polluants suit un cycle saisonnier et diurne différent selon le polluant considéré (Kuang-Ling Yang, 2002). En effet, la pollution atmosphérique n'est pas la même en toutes saisons. En Europe et en été, la chaleur, l'ensoleillement et l'absence de vent favorisent la formation de polluants secondaires. Les pics de pollution sont essentiellement identifiés comme des hausses de l'ozone troposphérique : on parle alors de pollution estivale ou photochimique pouvant s'étendre sur de longues distances. En automne-hiver, la pollution est caractérisée par la présence d'oxyde d'azote, d'oxyde de soufre et de particules. Aux sources de polluants classiques (transport, industrie) s'ajoute la pollution due au chauffage domestique. De plus, les conditions météorologiques (vent, froid, pluie) peuvent favoriser le transport des polluants ou sont à l'origine d'inversion thermique piégeant les polluants sous la couche d'inversion : on parle alors de pollution hivernale. Ces notions de « pollution estivale » et de « pollution hivernale » montrent que la pollution atmosphérique peut être différenciée suivant les saisons et les conditions météorologiques (Ung 2003).

Le Plan Régional pour la Qualité de l'Air (PRQA 2000) en Alsace, distingue quatre saisons dans le climat de Strasbourg du point de vue de la qualité de l'air et parle de types de temps. Les différents types de temps identifiés peuvent être associés à des configurations caractéristiques sur les cartes météorologiques quotidiennes. En Alsace, la durée moyenne d'un type de temps est de deux à trois jours. Une persistance plus longue favorise des conditions critiques, dans le domaine de la qualité de l'air. Les types de temps dépendent des saisons. Les quatre saisons identifiées sont :

- la saison froide (novembre à février), caractérisée par des conditions généralement favorables à la dispersion des polluants ou des particules. Cette dispersion est extrêmement dépendante des conditions de vent. Les inversions de températures surviennent fréquemment à cette période ;

- le printemps (mars à mai). A cette saison, l'atmosphère est fréquemment instable, ce qui favorise la dispersion des polluants. Les journées anticycloniques, chaudes et ensoleillées, se caractérisent par la présence de convection thermique dans la couche de mélange ce qui assure de bonnes conditions à la dispersion des polluants. Toutefois, lorsque une vaste zone anticyclonique persiste, cela entraîne l'apparition de fortes pollutions photochimiques à l'ozone ;
- la saison chaude. Elle est caractérisée par l'installation de beaux temps anticycloniques qui favorisent les processus photochimiques et conduisent à des augmentations importantes des concentrations en ozone ;
- l'automne (septembre à octobre) caractérisé par l'affaiblissement de l'insolation et la diminution de la durée du jour qui accentuent la persistance des inversions thermiques pendant les types de temps anticycloniques fréquents. Ces conditions sont défavorables à une bonne qualité de l'air.

On observe donc une répétition des situations météorologiques qui entraîne une répétition des épisodes de pollution. De la même façon, le caractère cyclique des activités humaines joue un rôle important dans les répétitions des épisodes de pollution, à cette échelle. A cette échelle locale, les pointes de pollution sont liées aux variations d'émissions périodiques (trafic pendulaire en semaine, régime de fonctionnement des industries différent en semaine et le week-end) ou épisodiques (opérations industrielles particulières) (PRQA 2000). Il existe donc un nombre limité de configurations spatiales de la pollution. On suppose alors qu'il existe un nombre fini de formes de champs de concentrations en polluant.

Soit un type de situation météorologique, caractérisé par des conditions aérauliques et d'insolation précises : direction et force du vent, stabilité de l'air, ennuagement, ensoleillement. Parce que la rugosité aérodynamique de la ville est constante temporellement pour cette situation et qu'il y a constance des sources d'émissions – ou, du moins, constance saisonnière ou journalière, on peut faire l'hypothèse que la forme des structures spatiales du champ de concentration reste la même pour ce type de situation ; seules la valeur moyenne et l'intensité (amplitude du gradient) des formes varient. Cet ensemble de formes ou structures, d'un champ de concentration de polluants, caractéristique d'une situation type liée à la météorologie et aux activités humaines est appelé un *champ typique*.

A partir de ces hypothèses, j'ai dénombré les situations typiques à l'origine des champs typiques de pollution en particules (PM) sur le ville de Strasbourg. Ces situations typiques sont classées en classes et sous-classes prenant en compte la variabilité temporelle des sources d'émissions et les situations météorologiques. La variabilité temporelle des sources d'émissions est principalement liée à l'utilisation ou non du chauffage domestique (saisons) ainsi qu'au régime des activités industrielles et

du trafic que l'on soit en semaine ou le week-end (Almeida *et al.* 2006). En outre, les conditions météorologiques différentes suivant les saisons en liaison notamment avec la température et l'humidité relative entraîne des réactions chimiques différentes et donc des formations de polluants différents. Par exemple, la réaction suivante NH_4NO_3 (solide) \rightleftharpoons NO_3 (gaz) + NH_3 (gaz), se fait plus facilement dans le sens droite-gauche en automne / hiver, créant ainsi des aérosols secondaires alors que le sens gauche-droite est favorisé au printemps / été, privilégiant la formation de polluants gazeux (Almeida *et al.* 2006). J'ai choisi de décrire les situations météorologiques par trois paramètres caractérisant la dispersion des PM (PRQA 2000) : la pluviométrie, le vent, l'humidité de l'air ou le brouillard. Le paramètre température est pris en compte de façon implicite à travers la variabilité saisonnière. Le vent a deux directions privilégiées sur la ville de Strasbourg (Rekkip 1995) du fait de la présence de la vallée du Rhin. Je dénombre ainsi (figure 3.1) 48 situations typiques différentes soit 48 champs typiques caractérisant la distribution spatiale des particules.

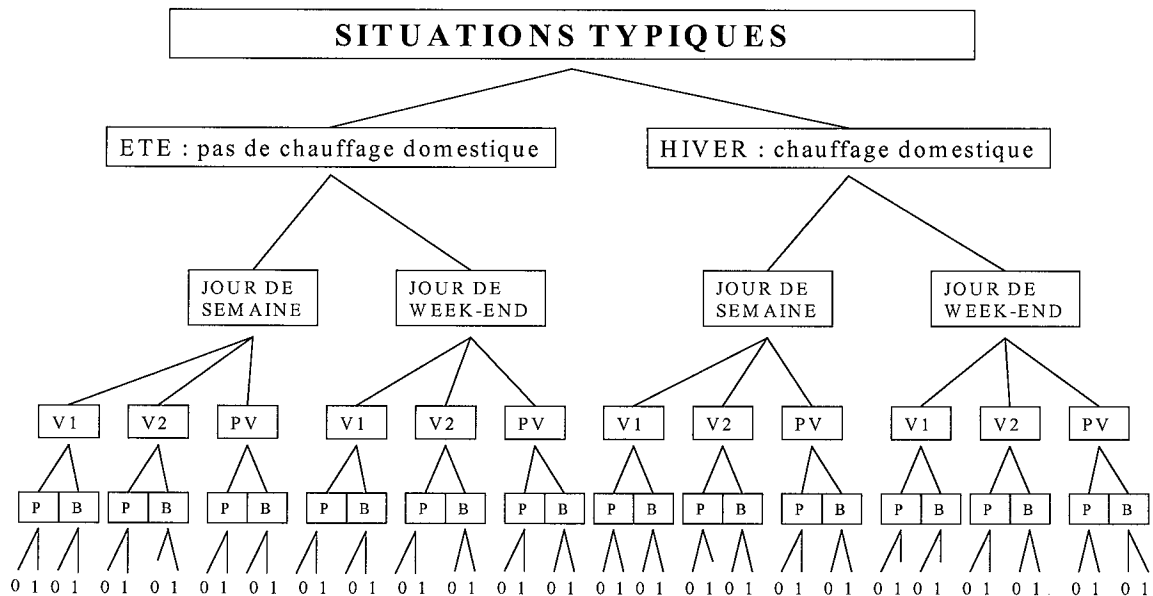


Fig. 3.1 : schéma de dénombrement des champs typiques. V1 signifie vent dans la direction 1 ; V2, vent dans la direction 2 ; PV, absence de vent ; P, pluie ; B, brouillard.

Avant d'aller plus loin et de donner une définition formelle du champ typique et de décrire la méthode des champs typiques, je dois donner quelques notions sur les échelles et présenter l'outil mathématique qu'est l'analyse multirésolution.

3 Outils mathématiques

3.1. Notions de résolution et d'échelle

J'appelle « résolution », la taille du pixel équivalent, c'est-à-dire la dimension du plus petit objet discernable dans l'image (ou le champ). J'utilise la notion d'échelle à la manière des physiciens et météorologues, c'est-à-dire définissant la taille typique des phénomènes. J'écrirai « résolution » et « échelle » indifféremment.

La méthode des champs typiques fait appel à une modélisation de la distribution spatiale du champ. Cette modélisation est effectuée à l'aide de deux outils mathématiques que sont la transformée en ondelettes et l'analyse multirésolution.

3.2. Transformée en ondelettes

La théorie des ondelettes ne sera pas développée en détail dans cette thèse. Une version plus détaillée figure en annexe B (paragraphe 2). Pour encore plus de précision, le lecteur peut se reporter à Daubechies (1988) ou Ranchin (1993) ainsi qu'à Meyer (1992) qui présente dans son ouvrage une perspective historique des ondelettes.

On considère le champ de pollution comme une fonction $d(x, y)$ caractérisée par des structures bi-dimensionnelles à différentes échelles spatiales imbriquées.

La principale propriété de la transformée en ondelettes est d'adapter la fenêtre d'analyse au phénomène étudié, ce qui permet de fournir une information locale. La transformée en ondelettes permet de décomposer la fonction d sur une base de fonctions orthogonales élémentaires $\{\psi_{a, b}\}$: les ondelettes. Cette base ou famille est générée par translations et dilatations d'une unique fonction ψ appelée *ondelette-mère*.

La décomposition en ondelettes ou analyse d'une fonction d dans le cas continu s'écrit de la façon suivante :

$$C_{a,b} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} d(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx \quad (1)$$

a et $b \in \mathbf{R}$, $a \neq 0$. Le paramètre a est appelé "pas de dilatation" et b "pas de translation". $\psi_{a, b}$ est l'ondelette-mère, $\overline{\psi_{a,b}}$ est le complexe conjugué de ψ et $C_{a,b}$ est appelée le coefficient d'ondelette. Pour une échelle a et une position b , $C_{a,b}$ représente l'information contenue dans $d(x)$ pour cette échelle et à cette position. Ce coefficient représente une mesure de l'intensité des variations locales du signal.

Son amplitude sera grande si l'échelle de variation de l'ondelette $\psi_{a,b}$ est proche de celle de la singularité ou de la structure locale du signal. Si le signal présente peu ou pas de variations pour cette échelle, la valeur du coefficient d'ondelette sera faible. $C_{a,b}$ est donc une caractérisation des structures présentes dans l'image pour cette échelle et à cette position.

Les capteurs utilisés délivrent des signaux qui sont discrétisés : les images numériques. C'est pour cette raison que j'utiliserai la version discrète de la transformée en ondelettes. A partir de ces ondelettes et de la fonction étudiée d , je construis les coefficients en ondelettes.

Les coefficients en ondelettes caractérisent l'ensemble des structures spatiales existantes d'un champ de concentration en polluant. A chaque échelle, est associé un champ de coefficients en ondelettes. L'ensemble de ces champs de coefficients pris à toutes les échelles caractérise un champ typique.

3.3. Analyse multirésolution et algorithme « à trous »

Le concept d'analyse multirésolution a été introduit par Mallat (1988, 1989). Dans le cas des algorithmes pyramidaux, on utilise souvent la représentation de la figure 3.2.

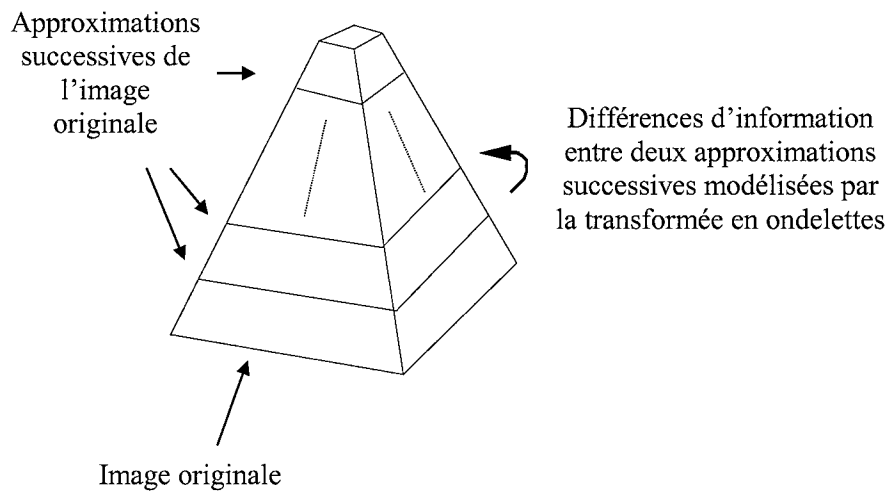


Fig 3.2 : Représentation des approximations successives d'une même image à l'aide d'un algorithme pyramidal.

La figure précédente (figure 3.2) est la représentation d'un algorithme pyramidal. Les algorithmes pyramidaux permettent de "fabriquer" des approximations successives et de plus en plus grossières de l'image originale. La base de la pyramide correspond à l'image originale. Quand on "monte" dans la pyramide, les différents étages représentent les approximations successives de l'image originale. La limite théorique de ces algorithmes est le sommet de la pyramide qui correspond à un pixel unique. L'analyse multirésolution est un cas particulier des algorithmes pyramidaux.

D'un point de vue pratique, l'application de l'analyse multirésolution est réalisée par le biais de filtres passe-bas et la différence d'informations existant entre deux approximations successives par un filtre passe-haut. Pour plus de détails, on peut se référer à l'ouvrage très complet de Daubechies (1992).

La mise en œuvre de l'analyse multirésolution se décompose en deux étapes : l'analyse (quand on « monte » la pyramide de l'image originale jusqu'au pixel unique) et la synthèse (quand on « descend » la pyramide). Ces deux opérations permettent une reconstruction exacte du signal original.

Il est important de préciser que je n'utiliserai pas d'algorithme pyramidal dans la méthode des champs typiques. La théorie et les figures ont été données à des fins pédagogiques et d'illustration. J'utiliserai l'algorithme « à trous » (Dutilleux 1987). Ce n'est pas un algorithme pyramidal car la taille des images est identique si on « monte » dans les échelles. Il n'y a donc pas de rééchantillonnage avec le changement d'échelle. C'est cette caractéristique ainsi que le fait qu'il n'y a qu'un plan d'ondelettes non-directionnel au lieu de trois directionnels (vertical, horizontal, diagonal) dans l'implantation classique d'un algorithme pyramidal, qui m'a fait choisir cet algorithme, de ce fait, plus facile à mettre en œuvre.

L'association de l'analyse multirésolution (annexe, paragraphe 3) et de la transformée en ondelettes permet une analyse complète et un traitement efficace des images. On utilise le terme « approximation d'une image » pour désigner une représentation d'une image originale dans laquelle les structures de petites tailles ont été filtrées, c'est-à-dire éliminées. Par itération, on obtient une série d'approximations présentant de moins en moins de détails. La différence d'informations existant entre deux approximations successives (j) et ($j-1$) d'une même image construites par l'analyse multirésolution est décrite par les coefficients en ondelettes, $C_j(x,y)$.

L'approximation de l'image originale est obtenue par filtrage et l'image des coefficients d'ondelettes par soustraction de l'approximation de l'image originale pixel par pixel. D'où :

$$C_j(x,y) = d_{j-1}(x,y) - d_j(x,y) \quad (2)$$

Cet algorithme produit, à chaque itération, une approximation et une image non-directionnelle de coefficients d'ondelettes (figure 3.3). Le filtre utilisé est symétrique, de taille 5x5 et ses coefficients sont :

$$1/256 \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

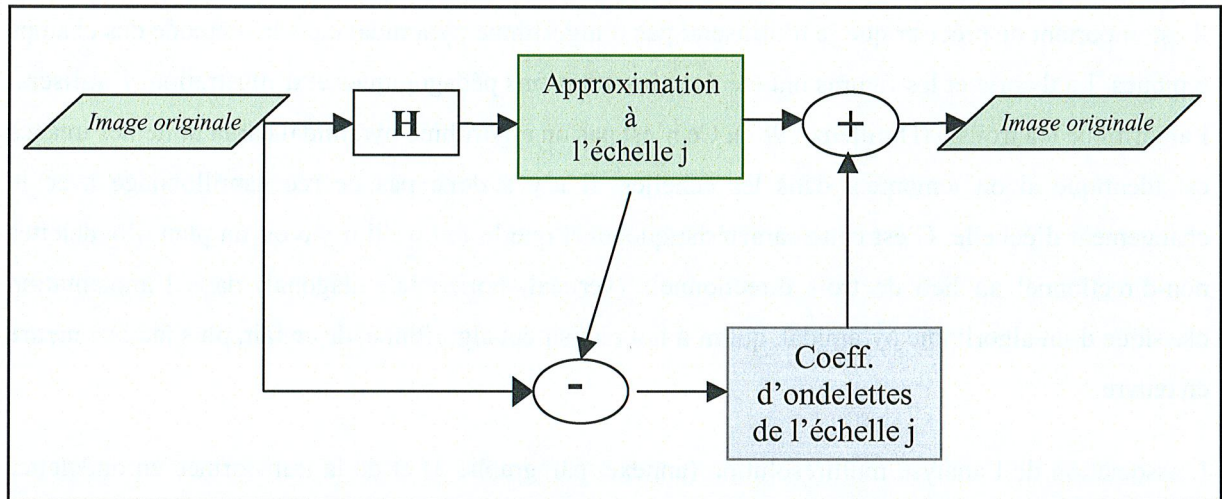


Fig 3.3 : Schéma de principe de l'algorithme « à trous ».

Il n'y a aucun sous-échantillonnage de l'image originale et toutes les images gardent la même taille. La reconstruction s'effectue suivant l'équation :

$$d(x, y) = d_J(x, y) + \sum_{j=1}^J C_j(x, y) \quad (3)$$

où $d_J(x, y)$ est la dernière approximation calculée par l'algorithme et $C_j(x, y)$ les coefficients d'ondelettes de $d(x, y)$. Ainsi le signal échantillonné original est reconstruit de manière exacte, sans perte d'informations.

4 Présentation de la méthode

Je précise maintenant la définition d'un champ typique au sens mathématique : le champ typique est l'ensemble des structures existantes d'un champ de concentration de polluants pour un type de situation, et plus exactement, l'ensemble des coefficients d'ondelette pour les structures comprises entre 10 m et 10 km et indépendants du niveau de pollution.

La méthode des champs typiques que je propose et implante, diffère de celle de Ung (2003) sur un point fondamental : les données utilisées. Je rappelle que chez Ung la construction des champs typiques s'appuie uniquement sur une carte de concentration à l'échelle locale. Celle-ci est obtenue par interpolation/extrapolation des données de concentrations issues à la fois des mesures in situ et des estimations effectuées au niveau des « stations virtuelles » i.e. des stations supplémentaires qu'il détermine au préalable. Une analyse multirésolution est appliquée à cette image à haute résolution, ce qui permet d'extraire les structures caractéristiques du champ de pollution à chaque échelle. On a, cependant, une sorte de contradiction, dans la mesure où les cartes interpolées sont à l'échelle locale qui est justement celle pour laquelle on cherche à obtenir une cartographie précise avec cette nouvelle méthode.

La méthode que je propose résout cette contradiction. Elle fait appel à d'autres informations susceptibles d'aider à la modélisation de la distribution spatiale de la pollution. Ces informations existent à différentes échelles spatiales. Ce sont des informations sur les polluants directement ou sur la structure du bâti ayant une influence sur la dispersion des polluants. Ces dernières informations se présentent sous la forme de sectorisations. Une *sectorisation* est un découpage en secteurs géographiques au sein desquels un paramètre donné est identique sur toute la zone.

Je dispose par exemple d'une sectorisation du cadastre des émissions (fig. 3.4, à gauche). Ces densités annuelles d'émissions sont exprimées en tonnes/an/km² et les valeurs sont ici réparties en dix classes. Chaque secteur délimite une zone pour laquelle la quantité de polluants émis est la même en moyenne sur la zone. Sur cette même figure est présentée une autre sectorisation, celle des niveaux de pollution en PM. Ces sectorisations sont discutées plus loin.

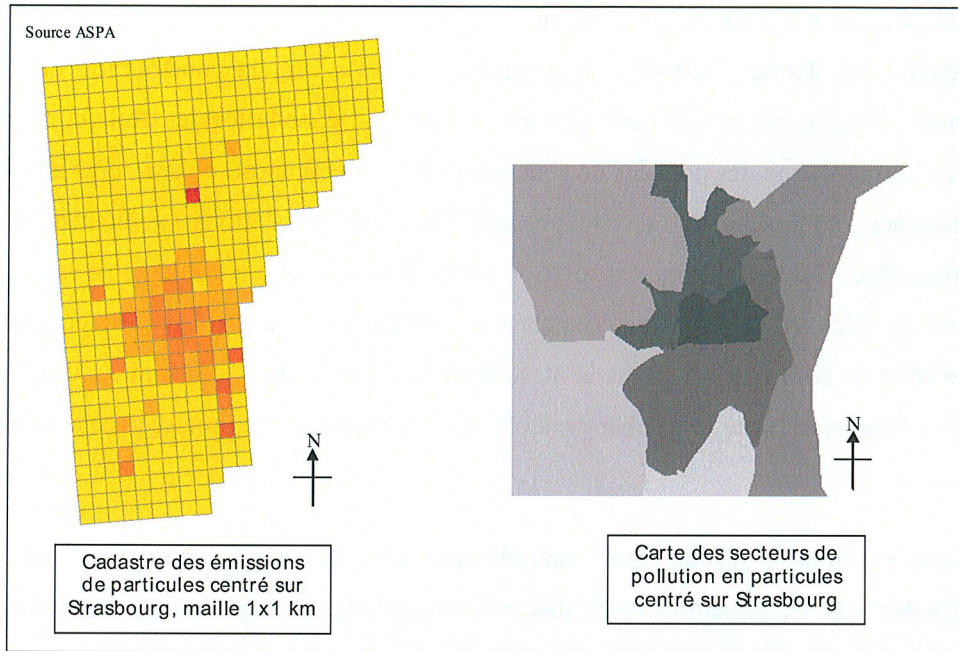


Figure 3.4 : exemples de sectorisations ; à gauche, plus on tend vers le rouge plus les émissions de particules sont élevées, à droite, plus on tend vers le gris foncé plus les émissions de PM sont élevées.

A partir de ces secteurs, on extrait les coefficients en ondelettes. Puis on normalise les coefficients par la moyenne de l'approximation de l'image de cette sectorisation à cette échelle. L'ensemble de ces coefficients normalisés constituent le champ typique. Le champ typique pour la situation météorologique j suit l'équation (4) :

$$\text{Champ typique} = CT_j = \left\{ C_0(x, y); C_1(x, y); \dots ; C_8(x, y) \right\} \quad (4)$$

Le champ typique est utilisé dans la reconstruction comme dans l'équation 3 mais de manière un peu plus complexe, car une même échelle peut être représentée par plusieurs coefficients normalisés provenant de sectorisations différentes. Ces sectorisations vont influencer de manière différente la représentation du champ de concentration à chaque échelle spatiale. En effet, on suppose qu'une information pertinente à une échelle peut l'être moins à une autre échelle. Cette idée impose de pondérer les coefficients en ondelettes $C_j^i(x, y)$ normalisés extraits des sectorisations i , à chaque échelle. Le champ typique s'exprime en valeur relative. L'amplitude du niveau de pollution doit être donné par une autre source d'information. J'ai choisi ici l'approximation à l'échelle régionale d_j (fig. 3.5), et connue par les modèles de prévision ou ré-analyse de Météo-France. En multipliant ces coefficients en ondelettes par la moyenne d_{moy} de cette carte de concentration, nous obtenons une modélisation des coefficients en ondelettes pour une situation recherchée.

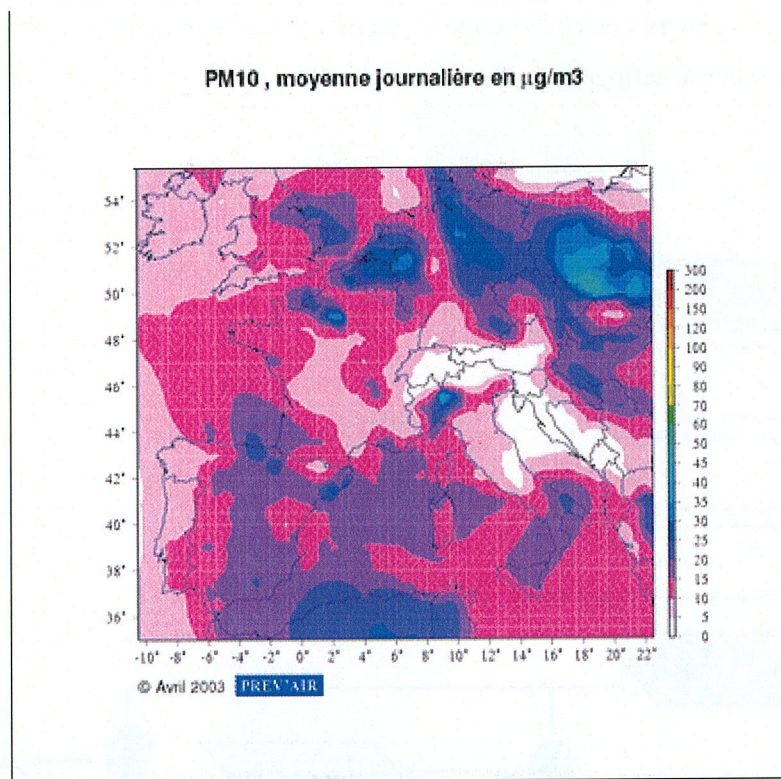


Figure 3.5 : carte du champ de concentration en PM à l'échelle régionale.

Pour établir l'équation (5), j'ai fait l'hypothèse que l'importance des informations est dépendante de l'échelle spatiale. L'introduction des w_j permet de pondérer ces informations. L'obtention des w_j doit être étudiée. Ces poids peuvent être des coefficients de corrélation ou bien les valeurs propres issues d'une analyse en composantes principales (ACP) effectuée sur l'ensemble des paramètres à une échelle donnée.

Cette synthèse s'exprime avec le champ typique de la façon suivante :

$$d(x, y) = d_j(x, y) + d_{moy} \sum_{j=0}^{j=J-1} \left[\frac{1}{\sum_{i=1}^{N_j} w_j^i} \sum_{i=1}^{N_j} w_j^i C_j^i(x, y) \right] \quad (5)$$

où les w_j^i représentent les pondérations et N_j le nombre de sectorisations à cette échelle j .

4.1. Principe général de la méthode

La méthode des champs typiques permet à la fois de créer un champ typique et d'obtenir une première cartographie du champ de concentration à l'échelle locale. La cartographie finale s'appuie, en amont,

sur la carte météorologique à l'échelle régionale et en aval, sur les données des stations de mesure à l'échelle locale. Aux échelles intermédiaires, les coefficients en ondelettes constituant le champ typique sont introduits à chaque échelle selon l'équation 5. La figure suivante (figure 3.6) schématise la méthode pour obtenir la cartographie du champ de concentration.

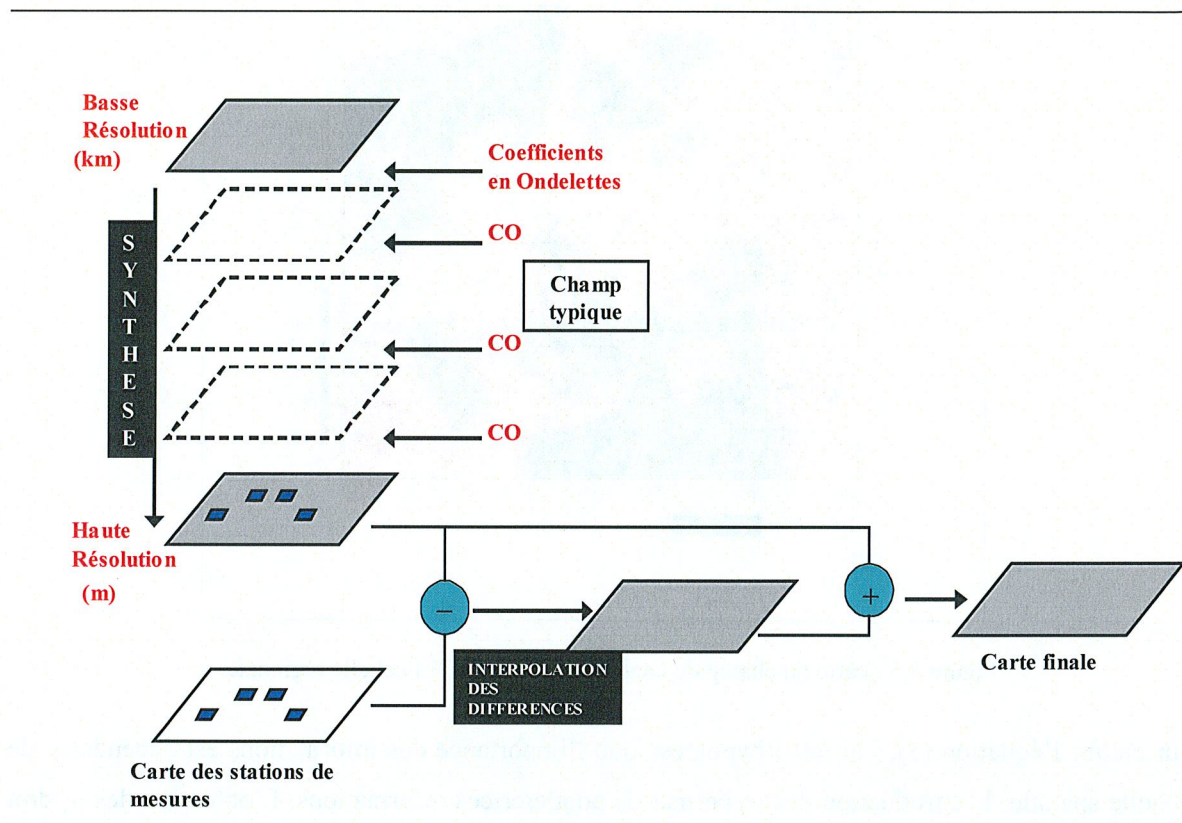


Figure 3.6 : schéma de la méthode des champs typiques proposée pour parvenir à une cartographie de la distribution de la pollution à l'échelle de la ville.

En effectuant la synthèse (multirésolution) à partir de la carte météorologique représentant les concentration en polluant ($R \sim 15$ km) et le champ typique correspondant, on construit une approximation du champ de concentrations à haute résolution ($R \sim 60$ m). Afin de réajuster les valeurs de cette approximation à celles effectivement mesurées au sol, on calcule la différence entre les mesures et les estimations. Ces différences sont ensuite interpolées avec une méthode classique de type « pondération par l'inverse de la distance » pour obtenir une carte des différences. Cette dernière est ajoutée à l'approximation pour obtenir la carte finale.

Cette méthode est inspirée du travail de Beyer *et al.* (1997) qui ont créé des cartes pour le nouvel atlas européen du rayonnement solaire (ESRA) et surtout du concept ARSIS « Amélioration de la Résolution Spatiale par Injection de Structures » (Ranchin et Wald 2000), couramment utilisée dans le domaine du traitement d'image. Ce concept permet d'améliorer une image multispectrale haute

résolution sans altérer l'information en lui ajoutant uniquement les structures manquantes à une échelle caractéristique voulue. Ce n'est pas très éloigné de la méthode d'estimation du champ proposé ici, qui peut se présenter comme étant une injection des structures du champ de concentration pour améliorer les cartes de concentration existant à l'échelle régionale.

Je présente ici une première mise en oeuvre de la méthode pour les PM afin de montrer la faisabilité de la méthode sur la ville de Strasbourg et d'en cerner les avantages et les limitations.

5 Mise en oeuvre et résultats

5.1. Inventaire des données disponibles et création d'un champ typique

A la plus basse résolution, je dispose d'une carte de concentrations en PM_{10} de résolution 15 km environ (fig.3.7, à gauche). Elle est issue du modèle météorologique CHIMERE. Cette carte est fournie par une organisation de prévision et d'observation de la qualité de l'air (<http://www.prevoir.org>, INERIS) et montre des champs de concentrations en particules au-dessus de l'Europe. Elle représente de manière continue la distribution des particules. Cette carte est le terme $d_j(x, y)$ de l'équation 5.

L'image à l'échelle régionale, issue du modèle météorologique, représente une fonction continue sur laquelle on peut facilement appliquer la transformée en ondelettes pour extraire les coefficients en ondelettes. Cependant, les autres cartes présentent de fortes discontinuités puisqu'il s'agit d'une description par classes ou par secteurs. Pour extraire les coefficients en ondelettes, il y a deux cas à considérer.

Premièrement, j'ai une description par classe et il existe une relation d'ordre entre ces classes. Chaque classe est recodée de manière à avoir une moyenne non nulle sur l'ensemble des classes. Cette moyenne me permettra par la suite de normaliser les coefficients en ondelettes. Deuxièmement, j'ai une description par secteur et il n'existe pas de relation d'ordre. En recodant, je crée une relation d'ordre et je me retrouve dans les conditions du premier cas. La traduction d'une information qualitative en coefficients en ondelettes et l'introduction des coefficients correspondants reste à approfondir et à mettre en oeuvre.

Par manque de données à toutes les échelles d'analyses, j'ai créé des données supplémentaires aux échelles manquantes à partir de données quantitatives (densité du bâti). J'aurai également pu le faire à partir de données qualitatives (type de bâtiments). J'obtiens des cartes de sectorisations correspondant à un découpage du bâti suivant la densité (ou bien le type de bâtiments, l'homogénéité...). Un exemple est fourni figure 3.7. Les données sont fournies par Azadeh Kallhori, doctorante en géographie au

Laboratoire Image et Ville, à Strasbourg. Les données qualitatives nécessitent d'être recodées pour obtenir des données quantitatives qui sont gérées par l'analyse multi-échelles. Par exemple, à partir de la sectorisation suivant les densités de bâti j'obtiens une sectorisation des niveaux de pollution en PM (fig. 3.4, droite). Ceci a été possible en faisant l'hypothèse, à partir de Hewitt & Jackson (2003), que la dispersion des PM était beaucoup moins facile dans un environnement de bâtiments denses. Plus la densité de bâtiments est importante, plus l'air a des difficultés à s'écouler, créant ainsi des zones d'accumulation de polluants.

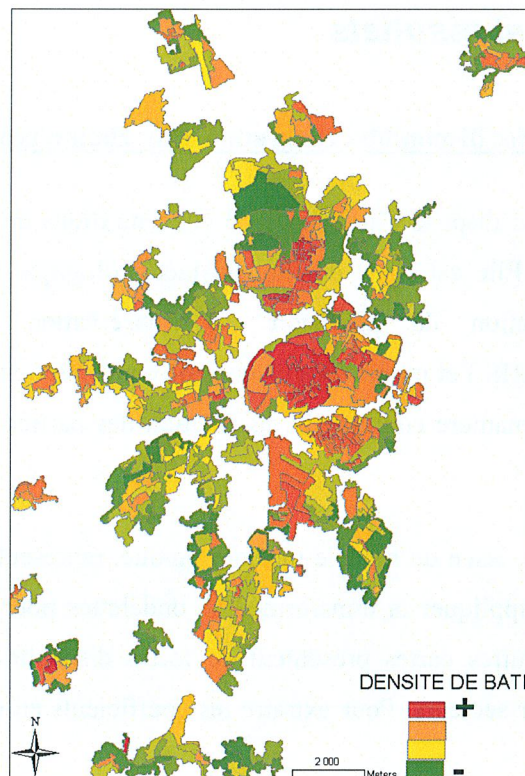


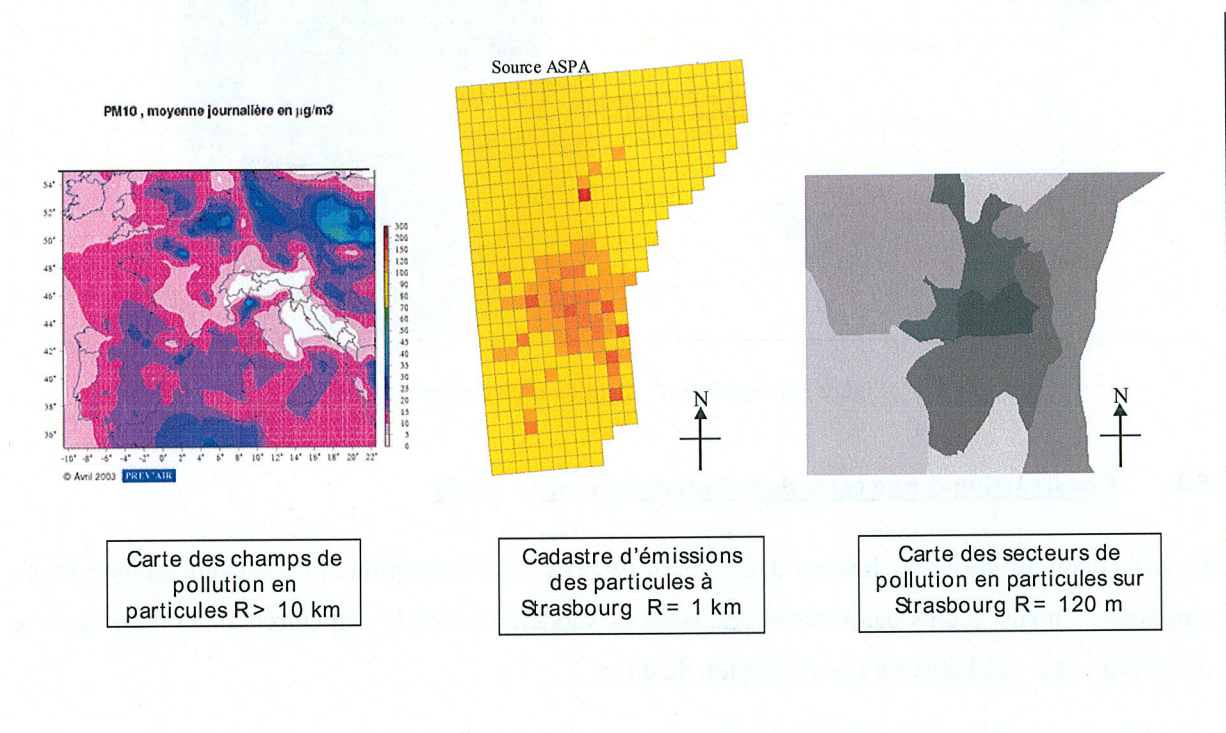
Fig. 3.7 : Densité de bâti de l'agglomération strasbourgeoise.

A une résolution plus élevée, je dispose du cadastre des émissions en polluants (fig.3.7 au centre). Ce cadastre fourni par l'ASPA (association pour la surveillance et l'étude de la pollution atmosphérique en Alsace) a une résolution de 1 kilomètre et décrit la distribution des émissions en particules fines pour l'agglomération de Strasbourg. Il donne en fait la quantité de particules émises par zones de 1 km², pour un an. La forme du cadastre et l'absence de données sur la partie droite sont dues au fait que je ne dispose pas des données d'émissions pour le territoire allemand. Les valeurs s'étendent de 0 à 310 000. Pour créer une carte, il est nécessaire de procéder à une classification pour ne représenter qu'un nombre limité de classes, par exemple 10. Pour déterminer les classes d'émission, j'ai procédé à un découpage de l'histogramme d'après un schéma standard de classification de ArcGIS 9. La construction des classes est basée sur des groupements naturels inhérents aux données. Le logiciel identifie les points d'arrêt en sélectionnant les coupures de classe qui groupent le mieux les valeurs

semblables et maximisent les différences entre les classes (ESRI® ArcMap™ 9.0). En d'autres termes, cet histogramme présente des seuils naturels c'est-à-dire des creux et des pics. Chaque creux va délimiter une nouvelle classe. On dénombre alors 10 classes d'émissions. D'après Hewitt & Jackson (2003), une source de pollution en particules influence son environnement sur un rayon de 100 m, avec une décroissance exponentielle. Pour tenir compte de ce fait, j'ai appliqué au cadastre un filtre de type gaussien. L'image résultante est de fait plus une carte représentant la pollution de fond en PM annuelle qu'un registre des émissions.

A une résolution encore plus élevée, des cartes de sectorisations sont créées puis exploitées. Ces sectorisations portent sur différents thèmes : l'homogénéité du bâti, la densité des bâtiments etc. Elles sont fondées sur des critères morphologiques essentiellement. Pour pouvoir être utilisées, ces cartes subissent un recodage décrit dans le paragraphe précédent.

La carte des secteurs de pollution en PM sur Strasbourg (fig. 3.8, à droite) dérive de ces sectorisations.



5.2. Création d'un champ typique pour une situation donnée

L'algorithme « à trous » est appliqué à chaque image dont la résolution est inférieure à la résolution de l'image d_J , c'est-à-dire 15 km, pour extraire les structures du champ (les coefficients en ondelettes) à chaque résolution (fig. 3.9). A chaque échelle, l'information est pondérée selon son importance puis sommée comme dans la deuxième partie de l'équation (5). Pour simplifier, j'ai choisi ici des poids w_i

tous égaux à 1. L'ensemble des images des structures spatiales de toutes les échelles d'analyse constitue le champ typique.

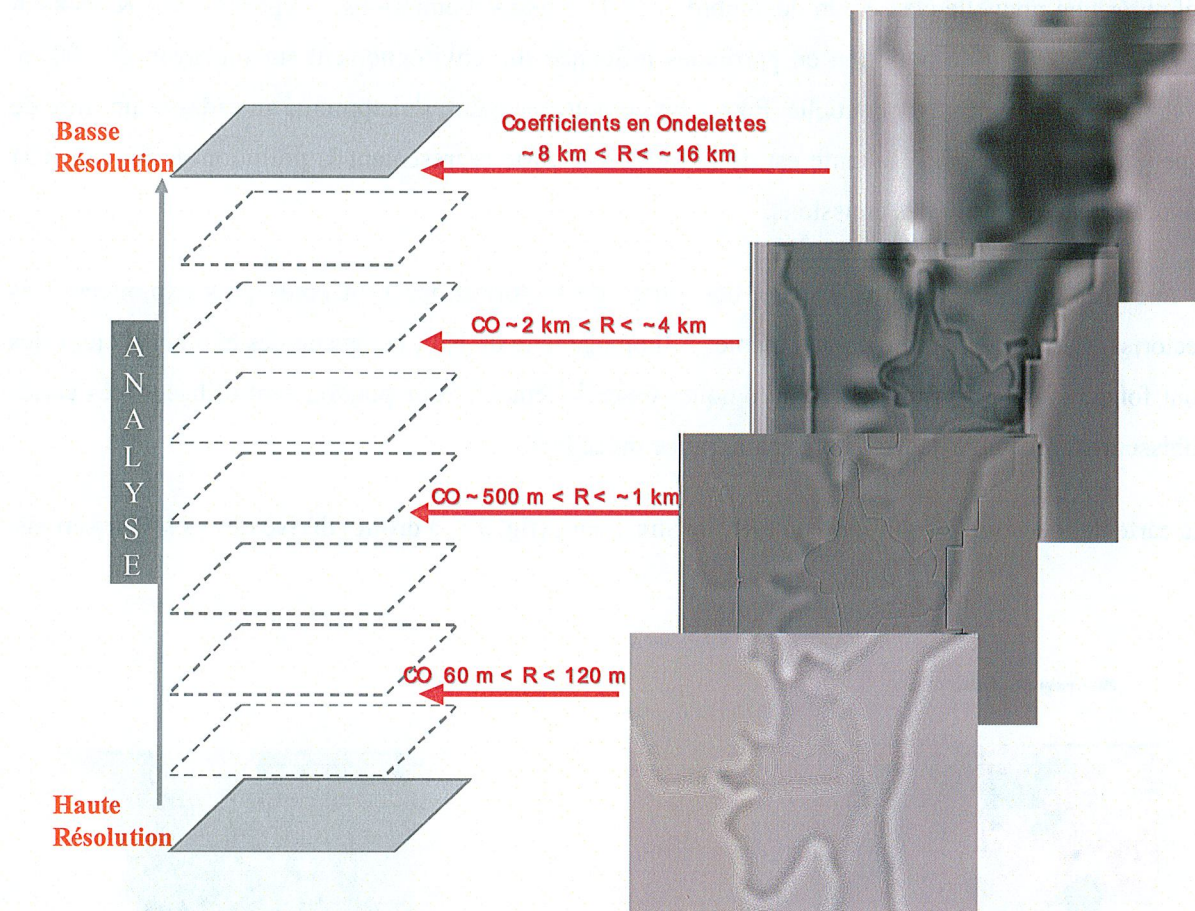


Figure 3.9 : schéma d'obtention d'un champ typique

5.3. Construction d'une carte de pollution à l'échelle locale

En effectuant l'algorithme inverse à celui utilisé pour la construction d'un champ typique et en employant l'image à plus basse résolution issue du modèle CHIMERE en entrée, une première carte de pollution est synthétisée à une résolution de 60 m.

À partir de la basse résolution, les structures appropriées du champ typique sont ajoutées à l'information brute à chaque échelle. Le processus itératif est effectué jusqu'à ce que la plus petite échelle soit atteinte. La carte obtenue (fig. 3.10, à gauche) représente une approximation du champ de concentration en particules au-dessus de la ville de Strasbourg avec une résolution de 60 mètres (en haut, à gauche). On calcule ensuite la différence entre les valeurs estimées de cette manière et celles mesurées au niveau des stations de mesures de Strasbourg en moyenne sur la période considérée. Une carte des différences est obtenue par interpolation (en haut, au centre). L'interpolation utilisée ici est une méthode classique de type « pondération par l'inverse de la distance ». Plus la distance entre la

valeur à interpoler et la mesure est grande, moins la mesure à d'importance. On ajoute ensuite cette carte des différences à la carte de résolution 60 m, selon le schéma de la figure 3.10. Les mesures au sol permettent d'améliorer la dernière carte en ajustant ses niveaux sur ces mêmes mesures. On obtient ainsi une carte finale à la résolution de 60 m.

De manière évidente, le résultat de l'interpolation dépend du nombre de stations et de leur répartition. Plus on a de stations bien réparties, meilleur est le résultat. Ici je ne dispose que de 5 stations pour les PM. Nous verrons au chapitre suivant comment augmenter artificiellement le nombre de stations pour l'interpolation.

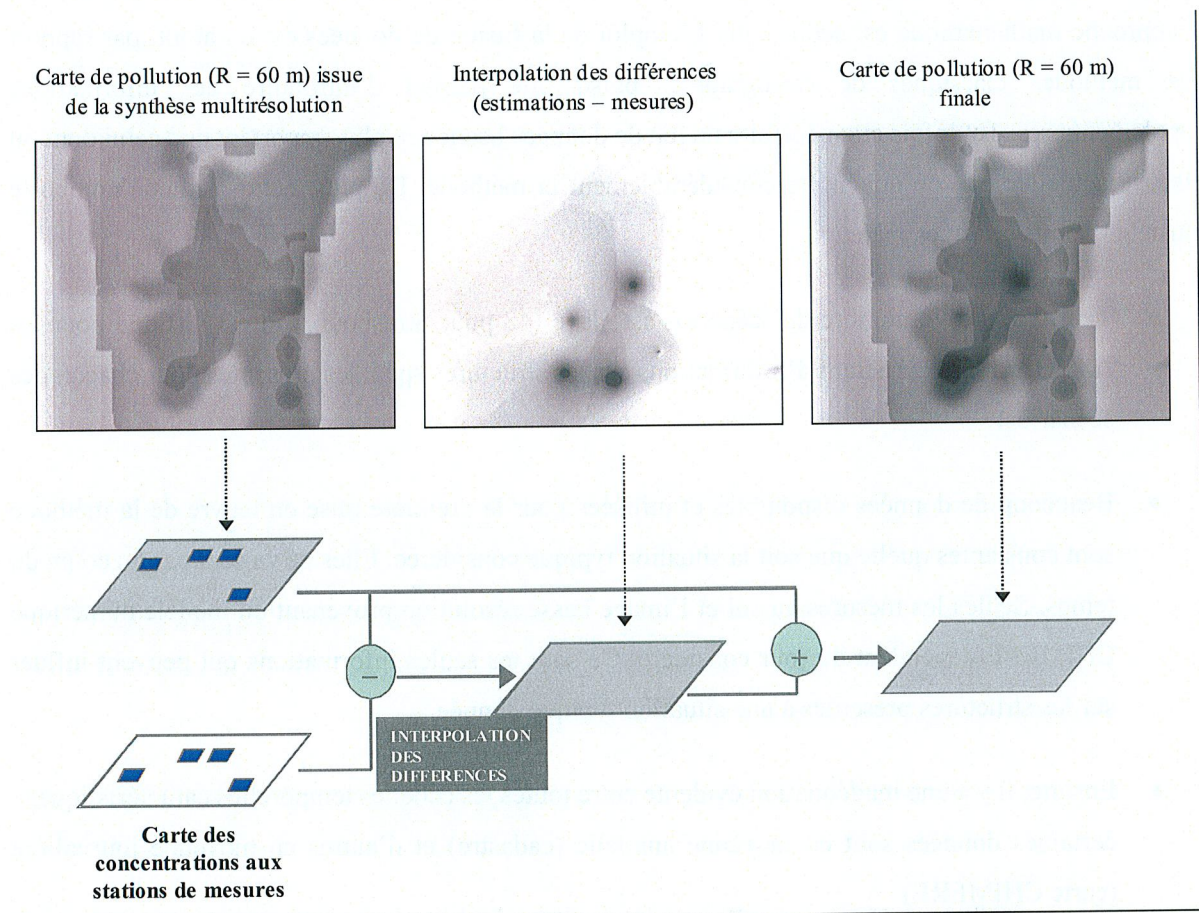


Figure 3.10 : schéma de construction de la carte de pollution finale, à l'échelle locale.

Finalement, j'ai créé une carte représentant la répartition de la pollution en particules sur la ville de Strasbourg à une résolution de 60 m.

6 Conclusion et discussion

L'objectif du travail sur la méthode des champs typiques était d'étudier la faisabilité et l'applicabilité d'une telle méthode au domaine de la qualité de l'air.

Premièrement, je me suis attachée à donner un formalisme mathématique à cette méthode, jusqu'alors inexistant. Puis, j'ai corrigé les défauts de la méthode originale et la théorie élaborée a été améliorée et complétée. Ensuite, pour mieux appréhender la faisabilité de la méthode, j'ai fait une première mise en œuvre à partir des données dont je dispose. Par manque de données à toutes les échelles d'analyses, j'ai créé des données supplémentaires aux échelles manquantes à partir de données qualitatives, ce que les géographes appellent les *sectorisations*. Les données qualitatives sont recodées pour obtenir des données quantitatives qui sont gérées par l'analyse multi-échelles. Pour ce premier test de faisabilité, le cadastre des émissions a été retraité de manière à prendre en compte l'influence des sources d'émissions sur leur voisinage.

L'approche mathématique est séduisante. L'emploi de la fusion de données est un atout, par rapport aux méthodes classiques de cartographie, puisqu'elle permet d'introduire des informations supplémentaires. Toutefois elle nécessite un jeu de données beaucoup plus important que celui dont on dispose actuellement, ce qui limite considérablement la méthode. De plus, cette méthode comporte plusieurs points critiques majeurs.

- Tout d'abord, le nombre de scénarios est élevé (48 pour Strasbourg). Il est difficile pour les AASQA d'obtenir autant d'informations sur les structures spatiales présentes pour chacun des scénarios.
- Beaucoup de données disponibles et utilisées pour la première mise en œuvre de la méthode sont constantes quelle que soit la situation typique considérée. Elles ne varient pas au cours du temps. Seules les mesures au sol et l'image basse résolution provenant du modèle numérique CHIMERE dépendent du jour considéré. Ce sont les seules informations qui peuvent influencer sur les structures présentes à une situation typique donnée.
- Ensuite, il y a une inadéquation évidente entre toutes les échelles temporelles caractéristiques : certaines données sont en moyenne annuelle (cadastre) et d'autres en moyenne journalière (carte CHIMERE).

De ce fait, le résultat est peu satisfaisant. La méthode est intéressante parce qu'elle permet de manipuler des échelles spatiales différentes ce qui est très utile lorsque l'on dispose de peu de données à une même échelle. Par contre, la méthode ne gère pas les échelles temporelles, ce qui implique qu'il est impossible de manipuler des données avec des échelles temporelles différentes.

Certains aspects de la méthode sont intéressants et méritent d'être mentionnés :

- Tout d'abord, la méthode des champs typiques permet d'introduire de l'information supplémentaire : on introduit de manière quantitative des informations qualitatives. Ces

informations qualitatives peuvent porter sur des sectorisations fonction du type de tissu urbain ou de la présence ou non de surfaces végétales (Puissant 2003). De la même façon, le réseau de rues ou les données épidémiologiques peuvent servir de base à la création de cartes de sectorisation de la pollution qui, une fois introduites, vont servir à contraindre l'interpolation. En fusionnant de multiples informations d'origines diverses, qualitatives : épidémiologie, urbanisme, ou quantitatives, provenant d'autres échelles spatiales, cela permet d'affiner l'interpolation et de représenter la pollution de façon plus précise.

- Ensuite, l'idée de créer une base de données disponible et exploitable facilement sans calcul supplémentaire : la bibliothèque des champs typiques, est intéressante, même si toute la difficulté réside dans la détermination du nombre de situations caractéristiques, dans l'adéquation entre la situation actuelle et une situation type.

Certes, la méthode est faisable et applicable c'est-à-dire que l'algorithme proposé est utilisable et applicable aux diverses données dont on dispose, ce que montre le test réalisé. Mais il manque des données pour parvenir à des résultats satisfaisants et surtout la méthode présente les défauts majeurs cités ci-dessus. Pour ces raisons, je ne la recommande pas pour l'établissement de la stratégie de cartographie.

Bibliographie

- Almeida S.M., Pio C.A., Freitas M.C., Reis M.A., Trancoso M.A., 2006. Source apportionment of atmospheric urban aerosol based on weekdays/weekend variability: evaluation of road re-suspended dust contribution. *Atmospheric Environment*, 40, 11, April 2006, 2058-2067.
- Basly L., 2000. Télédétection pour la qualité de l'air en milieu urbain. Thèse de doctorat Sciences des technologies de l'information et de la communication, Université de Nice-Sophia Antipolis, Nice, France, 182 p.
- Beyer H.-G., Czeplak G., Terzenbach U., Wald L., 1997. Assessment of the method used to construct clearness index maps for the new European solar radiation atlas (ESRA). *Solar Energy*, 61, 6, 389-397
- Daubechies I., 1988. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, vol. XLI, 909-906, 1988.
- Daubechies I., 1992. Ten lectures on wavelets. CBMS-NSF regional conference series in applied mathematics 61, SIAM, Philadelphia, USA, 357 p., 1992.
- Dutilleux P., 1989. An implementation of the « algorithme à trous » to compute the Wavelet Transform. In: Combes, J.M., Grossman, A., Tchamitchian, Ph. (Eds.), *Wavelets: Time-Frequency Methods and Phase Space*. Springer, Berlin, pp. 298-304.
- Gaillard E., Batton-Hubert M., 2003. Classification de données météorologiques pour l'étude de la concentration en ozone. Actes du 3ème Colloque STIC et Environnement, INSA , 19-20 juin 2003, Rouen, France.
- Hewitt C.N., Jackson A., 2003. *Handbook of Atmospheric Science. Principles and Applications*. Blackwell Publishing, Oxford, UK, 633 p, 149.
- Kim Oanh N.T., Chutimon P., Ekbordin W., Supat W., 2005. Meteorological pattern classification and application for forecasting air pollution episode potential in a mountain-valley area. *Atmospheric Environment*, 39, 7, March 2005, 1211-1225.
- Kuang-Ling Yang, 2002. Spatial and seasonal variation of PM10 mass concentrations in Taiwan. *Atmospheric Environment*, 36, 3403-3411.

Chapitre 3 – Méthodes des « champs typiques »

- Mallat S. G., 1989. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674-693, 1989.
- Mallat S. G., 1988. Multiresolution representations and wavelets. Ph. D., U.M.I., Ann Arbor, USA, 147 p.
- Meyer Y., 1992. *Les ondelettes – Algorithmes et applications*. Editeur : Armand Colin, Paris. 172 p.
- PRQA (plan régional pour la qualité de l'air) en Alsace 2000. Disponible sur le site de la direction régionale de l'industrie, de la recherche et de l'environnement en Alsace : <http://www.alsace.drire.gouv.fr/environnement/>.
- Puissant A., 2003. Information géographique et image à très haute résolution. Thèse de Doctorat, Laboratoire Image et ville, Université Louis Pasteur de Strasbourg, 349 p.
- Ranchin T., 1993. Applications de la transformée en ondelettes et de l'analyse multirésolution au traitement des images de télédétection. Thèse de doctorat, Ecole doctorale sciences pour l'ingénieur, université Nice-Sophia Antipolis, France, 110 p.
- Ranchin T. & Wald L., 2000. Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogrammetric Engineering & Remote Sensing*, 66, 1, 49-61.
- REKLIP Regio-Klima-Projekt, 1995. Communauté trinationale REKLIP sous la responsabilité scientifique de F. FIEDLER – Atlas climatique du Fossé Rhénan méridional. Vdf Zurich, IFG Offenbach, Ed. Coprur Strasbourg. Vol. de 78 cartes au format A2, vol. de textes, 240 p.
- Sarigiannis D. A., Gotti A., Sifakis N. I., Tombrou M., Dandou A., Schaefer K., Emeis S., Soulakellis N., 2005. *High-resolution estimation of urban aerosol from fusion of satellite and ground data with numerical modelling results*. In Proceedings of the 5th international conference on Urban Air Quality, Valencia, Spain, mars 2005.
- Ung A., Wald L., Ranchin T., Weber C., Hirsch J., Perron G., Kleinpeter J., 2002. Air pollution mapping : a new approach based on remote sensing and geographical databases. Application to the city of Strasbourg. *Photo-Interpretation 2000 / 3-4*, 38, p 53-63.
- Ung A., 2003. Cartographie de la pollution atmosphérique en milieu urbain à l'aide de données multisources. Thèse de doctorat Sciences de l'environnement, Université Paris 7 Denis Diderot, Paris, France , 96 p.
- Wald L. ,Baleynaud J.M., 1999. Observing air quality over the city of Nantes by means of Landsat thermal infrared data. *International Journal of Remote Sensing*, 20, 5, 947-959.

Chapitre 3 – Méthodes des « champs typiques »

Wald L., 2002. *Data Fusion. Définitions and Architectures - Fusion of Images of Different Spatial Resolutions*. Presses de l'Ecole, Ecole des Mines de Paris, Paris, France, ISBN 2-911762-38-X, 200 p.

Chapitre 4

Méthode de densification virtuelle du
réseau de mesures

1 Introduction

Ce chapitre est dédié à la présentation de la deuxième méthode alternative : la méthode de densification virtuelle du réseau de mesures. Le concept de densification virtuelle du réseau de mesures a été initié partiellement par Basly (2000) puis surtout par Ung (2003). Je le reprends et je donne une nouvelle version de cette méthode. L'objectif est de remédier au manque de mesures issues des réseaux urbains de surveillance qui sont à l'origine des représentations peu réalistes du champ de pollution par interpolation (Chap. 2, fig. 2.1). Il s'agit donc d'améliorer la qualité de l'interpolation en augmentant le nombre de données de concentrations en un certain polluant.

Notre travail a fait l'objet d'une proposition d'article en anglais, soumis à la revue « Environment International ». J'ai intégré cette proposition à ce chapitre et le lecteur y trouvera tous les détails de la méthode. Cependant afin « d'en faciliter la lecture », je présente brièvement auparavant les éléments principaux de la méthode et les résultats. J'ouvre ensuite une discussion sur les conclusions et perspectives. Puis la proposition d'article est reproduite.

La méthode consiste à densifier virtuellement le réseau de mesures conventionnel en y ajoutant des stations supplémentaires appelées « stations virtuelles ». Le point de départ du concept de « station virtuelle » repose sur des observations faites par les agences de surveillance de la qualité de l'air. Les AASQA ont observé qu'il existait des endroits en ville présentant des propriétés similaires à celles des lieux d'implantations des stations de mesure du réseau (Weber *et al.* 2002). De là, j'ai fait l'hypothèse suivante. Un lieu ayant les mêmes caractéristiques aérodynamiques qu'un lieu où existe déjà une station de mesure, appartenant à la même classe d'émission que cette station et présentant un environnement semblable à celui entourant la station, aura le même comportement vis-à-vis de la circulation des polluants et par conséquent la même concentration. J'appellerai ce lieu ponctuel de la ville une « station virtuelle ».

Les éléments caractérisant une station de mesure sont les émissions annuelles de particules, le type d'occupation du sol, la distance aux voies majeures de communication et les caractéristiques morphologiques du bâti entourant la station. Le cadastre des émissions renseigne sur la quantité de particules émises (en unité de masse) par an et par km². Le type de sol joue un rôle important dans la dispersion des polluants. Des changements dans le type d'occupation du sol (par exemple des bâtiments en comparaison avec des arbres) engendrent des changements de la longueur de rugosité aérodynamique, ce qui modifie l'écoulement de l'air en générant des turbulences (Tennekes & Lumley 1972). Cette longueur est un paramètre clé dans l'estimation de la pollution de l'air aux échelles locales (Briggs *et al.* 2000) et est souvent utilisée en entrée des modèles de dispersion (Hasager & Thykier-Nielsen 2001). Comme les classes de rugosité aérodynamiques sont calculées à partir des

classes d'occupation des sols, on peut directement utiliser ces dernières dans la méthode. La distance aux voies majeures de communication permet de prendre en compte la pollution due au trafic automobile. Elle sert souvent d'indicateur de l'exposition au trafic (Hoek *et al.* 2002a, b). Les indicateurs morphologiques (Miller et Gravélius) permettent de décrire la forme de l'espace entourant la station. Ils renseignent sur la façon dont s'écoule l'air à cet endroit.

Ces cinq éléments (les indices de Miller et Gravélius, la classe d'émission, le type d'occupation du sol, la distance aux routes principales) constituent les éléments discriminants du processus de classification. J'ai créé une image de la zone étudiée dont chaque pixel à une résolution de 10 m. Cette image contient cinq modalités (une par élément) et pour chacun des pixels de l'image, tous les éléments sont connus. Chaque pixel correspondant à une station de mesure réelle est comparé à tous les autres. Seuls les pixels dont les cinq caractéristiques sont identiques à celles de la station de mesure sont conservés. Ils constituent le jeu de stations virtuelles associées à cette station réelle. La méthode fait appel à des outils d'analyse spectrale tel que le hashcoding (Albuisson 1995). Les détails sont présentés dans l'article.

2 Cartes et résultats

A partir des cinq stations de mesures réelles composant le réseau de surveillance de Strasbourg, je dénombre 635 stations virtuelles réparties sur l'ensemble de la ville. Le tableau 4.1 donne le nombre de stations virtuelles pour chaque station réelle.

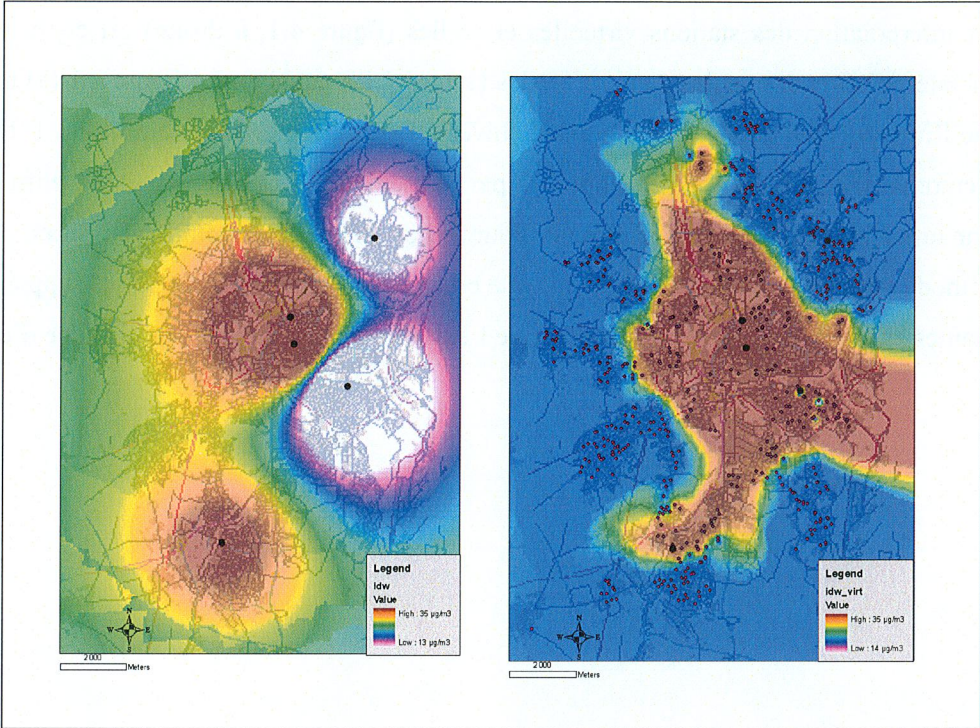
Stations réelles	Nombre de stations virtuelles
STG Illkirch	71
STG Centre	104
STG Nord	450
STG Clemenceau	6
STG Est	4

Tableau 4.1 : nombre de stations virtuelles par station.

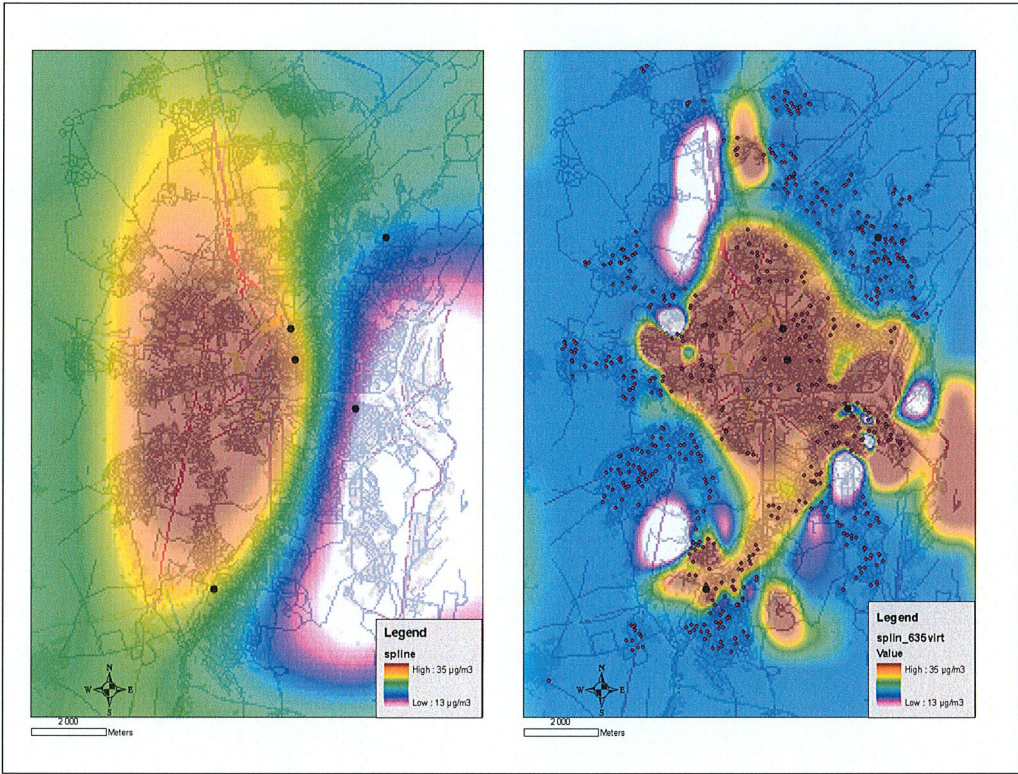
La station de mesures STG Nord possède un nombre très important de stations virtuelles, tandis que les stations STG Clémenceau et STG Est possèdent très peu de stations virtuelles associées. Cela signifie que les éléments sélectionnés pour la carte d'identité ne sont pas assez discriminants dans le premier cas et peut-être trop dans le second cas. D'autres exemples seraient nécessaires pour approfondir ce point.

Pour évaluer de manière qualitative le bénéfice de la densification virtuelle, j'ai construit des cartes représentant la concentration en particules sur la ville de Strasbourg pour la journée du 14 août 2001. Trois méthodes d'interpolation différentes sont utilisées. Pour chaque méthode, le résultat

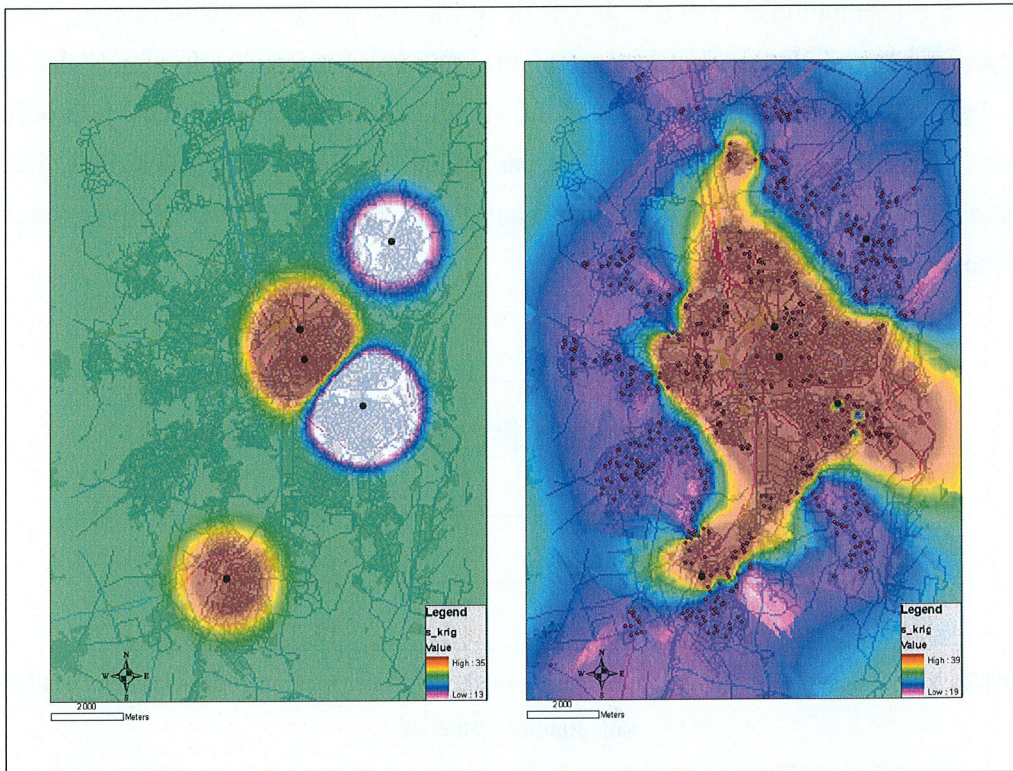
obtenu par interpolation des stations virtuelles et réelles (figure 4.1, à droite) est comparé à celui obtenu par interpolation des seules stations réelles (figure 4.1 à gauche). La figure 4.1 (a) montre les résultats de l'interpolation par la méthode IDW (Inverse Distance Weighted) qui utilise l'inverse des distances comme pondération, (b) ceux obtenus par la méthode des plaques minces, enfin (c) ceux obtenus par une méthode de Krigeage simple. Pour toutes les méthodes employées et en particulier pour la méthode de Krigeage qui nécessite plus que cinq points de mesure, des points supplémentaires ont été rajoutés aux angles et sur les bords afin que l'extrapolation ne dépasse pas la zone d'étude.



(a)



(b)



(c)

Fig. 4.1 : Cartes de pollution sur la ville de Strasbourg représentant les concentrations en PM₁₀ obtenues par interpolation des mesures issues des stations réelles (à gauche, les points noirs correspondent aux stations) et par interpolation des mesures issues des stations réelles et des estimations des stations virtuelles (à droite, les petits points rouges correspondent aux stations virtuelles). La méthode d'interpolation utilisée est (a) la méthode IDW, (b) la méthode des « plaques minces » et (c) une méthode de krigeage simple.

Les cartes obtenues à l'aide des stations virtuelles sont plus réalistes en particulier dans la zone délimitée par les stations. On notera que les cartes issues des stations réelles présentent une dichotomie est-ouest de la CUS, totalement infondée. Cette dichotomie disparaît avec l'usage des stations virtuelles. Des variations fines sont maintenant visibles, alors que sur les cartes de gauche, les variations sont plus grossières et la distribution de la pollution est d'avantage dictée par les fonctions mathématiques utilisées dans la méthode d'interpolation que par la réalité. C'est la représentation issue de la méthode des « plaques minces » qui présente le plus de variations fines et qui, de fait semble la plus précise et la plus réaliste.

Pour confirmer nos observations, j'ai cherché à évaluer nos résultats de manière quantitative. Pour évaluer l'apport de la densification virtuelle du réseau, j'ai calculé les RMSE (Root Mean Square Error ou écart quadratique moyen) par rapport aux mesures relevées au niveau des stations réelles, pour chaque méthode d'interpolation. Ce calcul se fait de la façon suivante : dans le cas où seules les stations réelles sont utilisées, quatre stations sur cinq servent à l'interpolation. La cinquième sert de

référence et sa concentration est estimée. Je fais de même pour les autres stations, cinq fois au total. Chacune des stations réelles devient à son tour une station de référence. Je prends finalement la moyenne des erreurs. Dans le cas où les stations virtuelles sont en plus utilisées, je procède de la même façon. Je réitère le processus cinq fois par groupe de stations virtuelles. Puis, j'estime la concentration à la station réelle du groupe manquant. Pour chaque méthode d'interpolation testée, je prends la moyenne des erreurs.

RMSE relative	5 stations réelles	Stations réelles + virtuelles
IDW	34.8 %	14.9 %
Plaques minces	49.0 %	14.9 %
Krigeage simple	33.5 %	15.4 %

Tableau 4.2 : Valeur des RMSE relatives pour chaque méthode d'interpolation et dans les deux cas : avec ou sans stations virtuelles.

Les valeurs des RMSE relatives confirment les observations faites sur les cartes. La méthode des « plaques minces » donne la plus faible valeur de RMSE et est extrêmement proche de celle correspondant à la méthode IDW. L'amélioration du résultat de l'interpolation est visible et quantifiable. La qualité de l'interpolation est améliorée de près de 55 % pour la méthode de krigeage et jusqu'à près de 70 % pour la méthode des « plaques minces ».

Ces résultats confirment le bénéfice d'ajouter des stations virtuelles à un réseau de mesures déjà existant et sont en accord avec les résultats déjà obtenus par Basly (2000). Dans une étude sur la ville de Nantes, en utilisant une approche similaire à la nôtre, Basly a trouvé que les valeurs des RMSE relatives diminuaient de 70 % (pour des interpolations sur des stations réelles uniquement) à 50 % (pour des interpolations utilisant en plus des stations virtuelles), dans le cas des « fumées noires ». Autrement dit, il trouve que la qualité de la représentation peut être améliorée d'environ 30 %. Néanmoins, il est important de préciser que le manque de référence pour comparer mes résultats, me contraint à estimer l'erreur qu'en un nombre très faible de points (5).

Nos résultats sont très encourageants. Toutefois l'hypothèse fondatrice des stations virtuelles reste à valider. Il faudrait également vérifier la qualité de la discrimination pour la création des stations virtuelles. Il faut noter également que les méthodes n'ont pas été optimisées. Par exemple, pour le krigeage simple, la méthode n'a pas été ajustée avec précision. En effet, on peut affiner encore l'interpolation en analysant en détail le variogramme pour déterminer des effets de structures (pépites, palier, portée). Les méthodes que j'ai utilisées sont ajustées automatiquement. Ce sont celles dont on dispose par défaut sur le logiciel SIG Arcview 9 qui sont configurées de manière à être facilement

utilisables tout en donnant des résultats satisfaisants. Néanmoins, cela n'enlève rien à la démonstration de l'apport des stations virtuelles pour la cartographie de la pollution urbaine quelle que soit la méthode d'interpolation choisie. Je n'ai pas cherché ici à préconiser une méthode d'interpolation plutôt qu'une autre. J'ai cherché à montrer que l'interpolation était améliorée en utilisant les stations virtuelles.

3 Conclusion et perspectives

L'objectif de ce travail était d'améliorer la représentation de la pollution par interpolation. La méthode proposée consiste à augmenter le nombre de mesures pour l'interpolation en densifiant artificiellement le réseau conventionnel de stations. Cette densification se fait au moyen de stations virtuelles. Celles-ci sont déterminées à partir d'éléments caractéristiques décrivant une station réelle. Ces éléments sont liés aux sources de pollution (émissions, trafic) « alimentant » la station et aux propriétés aérodynamiques (occupation du sol, indicateurs morphologiques) du lieu où se trouve la station. Chaque lieu ponctuel de la ville dont les éléments sont identiques à ceux d'une station, devient une station virtuelle associée à cette station réelle. La détermination des stations virtuelles et leur localisation se fait grâce à une classification de données multi-source. La mise en oeuvre de la méthode pour la pollution particulaire sur la ville de Strasbourg a été concluante. A partir de cinq stations de mesure au sol, 635 stations virtuelles ont été trouvées, réparties de façon homogène sur l'ensemble de l'agglomération. Les cartes de pollution en particules obtenues par interpolation des stations réelles et virtuelles ont confirmé le bénéfice d'une telle méthode pour la cartographie. Des variations fines de concentrations sont visibles sur les cartes et l'erreur sur l'estimation des concentrations est fortement diminuée.

Bien que ces résultats soient en accord avec ceux trouvés lors des études précédentes, d'autres cas doivent être analysés et testés pour que la méthode soit opérationnelle. Ce travail offre certainement une bonne base pour la sélection d'éléments relatifs à la circulation des polluants permettant de discriminer avec pertinence des lieux de la ville. De plus, la mise en oeuvre de la méthode est relativement aisée pour les AASQA notamment, puisque les outils employés sont bien connus et certaines données utilisées sont fournies par ces agences.

Une perspective de travail à court terme est bien sûr la validation des stations virtuelles, plus précisément la validation de l'hypothèse fondatrice du concept, selon laquelle la station et les stations virtuelles associées ont la même concentration. L'ASPA et les géographes du laboratoire Image et Ville de l'université Louis Pasteur à Strasbourg semblent convaincus par l'emploi des stations virtuelles suite à une campagne de mesure qui se déroula en juin 2002 dont le but était de valider le concept de stations virtuelles en comparant les mesures de concentrations au niveau des stations réelles et virtuelles. Pour cela, des stations mobiles avaient été disposées à l'emplacement de certaines

stations virtuelles désignées par Ung (2000). Cependant, je reste plus réservée quant à la représentativité de cette campagne de mesures. En effet, aux dates de la campagne, les niveaux de pollution étaient relativement bas. Aucun épisode de pollution n'est survenu. Ceci implique que les valeurs des concentrations étaient toutes relativement proches. Il était donc aisé de trouver de fortes corrélations entre différents points de la ville.

Je vais donc chercher à valider la méthode de détermination des stations virtuelles telle qu'elle est expliquée dans cette thèse. Soit j'ai recours à une nouvelle campagne de mesures mais cela nécessite des moyens matériels et financiers très importants. De plus, le choix des dates est très contraignant puisqu'il faut que la campagne coïncide, à la fois, avec des épisodes de pollution et avec des périodes de faibles pollution pour être représentative. Soit j'ai recours à l'imagerie satellitaire. C'est une solution alternative, plus simple à mettre en œuvre et surtout moins onéreuse. Ce travail fera l'objet du chapitre suivant.

4 Démarche de validation des stations virtuelles

L'hypothèse de départ des stations virtuelles suppose que la concentration en particules à la station virtuelle est la même qu'à la station réelle associée. Pour que l'hypothèse se justifie, les stations virtuelles sont choisies de manière à ce qu'elles présentent des propriétés liées à l'écoulement de l'air et à la proximité de sources de particules identiques à celles des stations réelles.

Pour valider l'hypothèse :

$$[PM]_{station} = [PM]_{station_virtuelle} \quad (1)$$

Je vais chercher une relation entre ce qui est observé par le capteur satellitaire (la réflectance ρ_{sat}) et la concentration en particules $[PM]$. Si je trouve une relation de type :

$$\rho_{sat} = f([PM]) \quad (2)$$

je peux adopter la démarche de validation suivante.

Le capteur satellitaire observe une réflectance ρ_{sat} qui s'exprime de la façon suivante :

$$\rho_{sat} = \rho_{atm} + T \rho_{sol} \quad (3)$$

avec ρ_{atm} , la réflectance de l'atmosphère sans polluant, ρ_{sol} , l'albedo de surface et T , la transmittance globale de l'atmosphère.

Comme d'après la carte d'identité, la station et la station virtuelle ont le même type de sol puisqu'elles appartiennent à la même classe d'occupation du sol, je peux écrire :

$$\rho_{sol_station} = \rho_{sol_station_virtuelle} \quad (4)$$

et comme, je suppose que la réflectance de l'atmosphère sans polluant ρ_{atm} est constante sur toute la zone considérée, du fait de la faible variabilité de l'atmosphère à des échelles très locales pour un instant donné, alors si :

$$\rho_{sat_station} = \rho_{sat_station_virtuelle} \quad (5)$$

alors, j'ai montré l'équation (1) avec une erreur au premier ordre.

Si je parviens à prouver cela à l'aide de l'imagerie satellitaire (chapitre 5 suivant), j'aurai validé l'hypothèse de départ et ainsi le concept de stations virtuelles (chapitre 5, partie 6).

5 Article soumis à *Environment International*

ADDING VIRTUAL MEASURING STATIONS TO A NETWORK FOR URBAN AIR POLLUTION MAPPING

A. L. Beaulant^{1*}, G. Perron², J. Kleinpeter², C. Weber³, T. Ranchin¹, L. Wald¹

^{1*}*Centre Energétique et Procédés, Ecole des Mines de Paris, B.P 207, 06904 Sophia Antipolis Cedex, France, anne-lise.beaulant@ensmp.fr.*

²*ASPA (Association pour la Surveillance et l'étude de la Pollution atmosphérique en Alsace), 5 rue de Madrid, 67300 Schiltigheim, France*

³*Laboratoire Image et Ville, UMR 7011, Université Louis Pasteur, 3 rue de l'Argonne, 67000 Strasbourg, France*

ABSTRACT

Maps of pollutants concentration are usually generated by means of interpolation and extrapolation methods. The quality of the results depends mainly of the number of permanent or temporary measuring stations. This paper deals with a method for the virtual densification of the network of stations. The method creates "virtual measuring stations". It aims at improving the quality of interpolation by increasing the number of data on pollutant concentration. The virtual stations are determined by the means of a classification method applied to each pixel of the area under concern. Discriminating elements are pollutants emission classes, land cover types, urban morphological indicators created to this purpose and distance to major roads. A first implementation was made for particulate matter (PM) for the city of Strasbourg (France) using thin-plates spline interpolation method in Arcview 9 GIS. The relative Root Mean Square Error decreases from 49% for five input stations down to 15% for the virtual stations.

Key Words: pollution map, virtual measuring station, morphological indicators, thin plates spline, RMSE

1. INTRODUCTION

Many large cities in Europe have acquired a measuring network operated by ad hoc agencies in order to monitor and analyze air quality. Air quality agencies inform local authorities, which rely on these studies to take decisions and to inform the population in order to reduce health impact caused by air pollution. European policies in this domain request accurate space-time knowledge of individual or collective exposure to pollutants. Therefore, it is becoming more and more essential to know or model the spatial distribution of pollutants concentration at any time and any place in the city in order to map pollution at very local scale.

To answer this need, agencies in charge of studying air pollution are using two different types of methods, with possible combination of both types. Either they map pollutants concentration with numerical models, either they interpolate concentrations values issued from permanent or temporary measuring ground network. The combination of the two methods is a methodology, which have been already tested in the frame of the AirProche project, giving encouraging results. Modeling pollution variability at local scale involves to solve very complex phenomena and needs many inputs such as meteorological parameters, emissions parameters, boundary conditions. Several models are used to map pollutants concentrations at large scale, i.e. 100-1000 km (Schmidt et al. 2001, Vautard et al. 2003, Monteiro et al. 2005). Others such as STREET and ADMS model pollution at very local scale, i.e. street and are usually of limited geographical areas (McHug et al. 1997a, b). Models require many input parameters to provide accurate results and these parameters are seldom available. For these reasons, practitioners presently generate maps of concentrations by means of interpolation and extrapolation methods. Among those, the most used are the kriging method (Carletti et al. 2000) or co-kriging and kriging with trend method using explained co-factors (ex. : emissions, land use) and the "thin plates spline" method (Ionescu et al. 1996, 2000).

The accuracy of the results of the interpolation depends mainly on the number of known measurements input to the method. Majority of cities have a monitoring network composed by a small number of measuring stations. According to Stalker & Dickerson (1962), a network of conventional ground measurements requires at least one sampling station of PM per surface of 2.5 km² to estimate the actual concentration at $\pm 20\%$. For our study area, the city of Strasbourg, whose surface is 306 km², a network of approximately 122 stations dedicated to PM would be needed instead of the 5 current ones. But, increasing the number of ground measuring stations requires important material means and is extremely expensive. Furthermore, maintenance has a considerable cost.

To overcome this problem, we propose a method for increasing the number of measurements input to the interpolator. It consists in a virtual densification of the network. Our study is based on the work of Ung (2003), Ung et al. (2001, 2002), who introduced the concept of "virtual stations". The first part of this paper presents the method. Then, we apply the method to the city of Strasbourg and discuss the results. Finally, future work is sketched.

2. METHOD

The starting point of the concept of "virtual station" comes from observations made by agencies in charge of air quality monitoring. These agencies (such as ASPA in Strasbourg) observed that there are places in the city, which present properties similar to those of the measuring stations (Weber et al. 2002). We made the assumption that a place having the same aerodynamic features than an actual measuring station, belonging to the same emission class than this station and with similar surroundings, will have the same behavior regarding pollutants circulation and therefore, the same concentration. Such a place is a virtual station.

The major attributes, or features, of a station are annual emission of pollutants, land cover, distance to major roads and morphological characteristics of the surrounding buildings that affect the aerodynamic flow.

The emission cadaster informs about the quantity of pollutants (in weight unit) emitted per year per squared kilometer. The land cover plays an important role in the pollutants dispersion. Changes in land cover (e.g., buildings compared to trees) imply changes in aerodynamic roughness length, which in turn modify air flow by creating turbulences (Tennekes & Lumley 1972). It is a key parameter to assess air pollution at small scale (Briggs et al. 2000) and as input in dispersion model (Hasager & Thykier-Nielsen 2001). The distance to major roads allows us to take traffic pollution into account in the classification. It is often used as exposure indicator to traffic intensity (Hoek et al. 2002a, b).

Morphological indicators describe the shape of the space surrounding the station. Individual buildings, their height, orientation and their arrangement influence wind flows and thus spatial distribution of the pollutants (Turbelin 2000). For example, in a very large avenue, air flows disperse pollutants, which decreases pollutants concentrations. Inversely pollutants accumulate in a very confined space. We have selected two indicators to characterize the space surrounding the station.

The Miller indicator I_M characterizes the circularity of this space. It varies from 0 for a linear shape to 1 for a perfectly circular shape (eq. 1):

$$I_M = \frac{4\pi S}{P^2} \quad (1)$$

where S is the visibility surface of the area and P its perimeter.

The Gravelius indicator I_G is an indicator of compacity. It compares the perimeter of the visibility surface to a perimeter of a disk having the same area (eq. 2):

$$I_G = \frac{P}{2\sqrt{\pi S}} \quad (2)$$

The visibility surface S characterizes the open space around the station. It is computed by a rays throw technique by taking into account the buildings position around the station. It is expressed as follows (eq. 3):

$$S = \sum_{i=1}^{N_{ray}} S_{angle}(i) \quad (3)$$

where $S_{angle}(i)$, is a sector in the direction i (eq. 4):

$$S_{angle}(i) = \frac{\pi D_{angle}^2(i)}{N_{ray}} \quad (4)$$

with $Dangle(i)$, the visibility distance computed in the direction i , and $Nray$ is the number of thrown rays. Here, $Nray$ is set to 360.

Prior to the determination of virtual stations and their localization, the urban area is divided into cells by a regular grid. The cell size defines the spatial resolution. In our application, the cell size is set to 10 m. For each cell, the Miller and the Gravelius indicators, the emission class, the land cover type and the distance to major roads are computed or already known. Each cell containing a measuring station is compared to all others. Only are kept those cells whose features are similar to those of the station; they are "virtual stations" for this measuring station and they are precisely localized in the city. In other words, a classification is performed: each pixel is allotted to a measuring station if similar, or rejected. The method uses spectral analysis tools, namely hashcoding tools (Albuisson 1995).

3. APPLICATION TO THE CITY OF STRASBOURG

3.1 Study area

The city of Strasbourg is located in Eastern France, separated from Germany by the Rhine river. Geographical coordinates are: 48.33° and 7.38° . ASPA is the local agency in charge of the air quality measuring network in the city of Strasbourg and vicinity. Fig. 1 displays a map of the measuring network in the city of Strasbourg and vicinity. ASPA operates 11 stations; 5 of them measure ($PM_{10, 13, 2.5}$). Atmospheric pollution in Strasbourg is mostly due to motor vehicles since there is no heavily polluting industries in this area except a refinery located in Reichstett in the North. Wind is blowing mostly from South-West or North-East because it is channeled in the Rhin valley (REKLIP 1995). It can bring significant pollution clouds from the industrial region of Rhur in Germany.

3.2 Data used

The morphological indicators Miller and Gravelius were computed for the whole city of Strasbourg (fig. 2). We developed scripts in Arcview 3.2, Geographical Information System (GIS) software to that aim and run them on the geographical database BD TOPO® - © IGN of the French Institute of Geography. The database is georeferenced and contains a volumetric 3D description of the city for buildings. Each building is described by a polygon and each polygon has attributes such as height (minimum, maximum, mean), area and perimeter... The computation provides two images of Strasbourg whose pixels contain the values of the Miller and Gravelius indicators. The computing time requested was very large (several days). All data used are stored into a GIS.

The land cover map is the CORINE Land Cover (2000) with an initial resolution of 100 m. It contains 44 different types of land cover. It was oversampled to a grid cell of 10 m using a nearest neighbor technique (Fig.3, left).

The emission cadaster is a map in a vector format; each polygon represents 1 km² area. The map was oversampled to 10 m using a bilinear technique. The quantity of pollutants was then recoded into a limited number of classes (9). The number of classes is

determined so that real stations belong to different classes as possible. A polluting source has an influence till 100 m over a surrounding area (Hewitt & Jackson 2003). A spatial filter of Gaussian type was applied to take this influence into account. The resulting image is no longer the emission register but more a map of annual background pollution (Fig. 3, right).

Distance to major roads is computed for both virtual and real stations thanks to a Arcview 9 GIS tool. The distance to traffic pollution is no more than 50 m for real stations. As Hewitt & Jackson (2003) indicate that a polluting source at ground level has an influence till 100 m, we used this 100 m limit to discriminate virtual stations. All cells whose distance to roads exceeds 100 m are rejected. On the figure 4, black triangles are cells that will be removed.

4. RESULTS AND DISCUSSION

Table 1 reports on the number of virtual stations found for each of the five stations measuring PM.

The total number of virtual stations is 635. They are scattered all over the city, which is a true advantage for interpolation method. Compared to the initial 5 stations, one may easily understand the benefit of the virtual densification.

However, we are puzzled in this very first application of the method by the large number of virtual stations for the station STG Nord. Conversely, the number of virtual stations is quite low for two others STG Clemenceau and STG Est. It means that the selected features are sufficiently discriminating in the last cases and maybe not in the first one. More examples are necessary to better understand the creation of the virtual stations.

Mapping the PM concentration from this set of virtual stations has been performed using Arcview 9 GIS tools. Two maps of PM_{10} concentration were obtained by performing the "thin plates spline" method of the Arcview GIS (Fig. 5). The area is approximately 18 x 28 km² with a spatial resolution of 10 m. Buildings and roads of Strasbourg were laid in the background for better readability. Fig. 5 (left) exhibits the concentration obtained from the five actual measuring stations (black dots). The low number of measuring points yields to a very uniform pollution map, which is not realistic. The spatial variation of the pollution is closer to mathematical functions used in this method than to reality. Figure 5 (right) is a pollution map obtained from the virtual stations for the same date. Compared to the figure on the left, the map based on virtual stations shows a more realistic pollution distribution. Into the area delimited by measurements, the representation is very realistic : local variations are visible. However, outside this area, the pollution distribution is less realistic: pollution is uniform, concentrations values are identical over the area. This map has not been validated; nevertheless a measuring campaign made in June 2003 indicates that several virtual stations preliminary spotted by Ung *et al.* (2002) were actually behaving like the real stations as expected (Ung 2003; Puissant 2003).

A quantitative assessment was made as follows. Only four stations over five were used for interpolation. The result for the missing station was compared to the actual value. By repeating this procedure, we obtain five errors and compute a root mean square error (RMSE) in both case: actual stations only and virtual stations.

Table 2 reports on the relative RMSE. It shows that using virtual stations for the mapping improves mapping quality : the relative RMSE decreases from around 49% down to 15%. In other words, the improvement of the quality of the interpolation is around 70% when using the “thin plates spline” method. Fairly similar results are obtained with other interpolation methods (simple kriging, Inverse Distance Weighted). Those results confirm the benefit of adding virtual stations and are in agreement with those obtained by Basly (2000). Using a similar approach, in a study on the city of Nantes (France), Basly found that the relative RMSE decreased from 70% (only real stations) to 50% (virtual stations), for the case of black smokes.

However, this method has a limitation. Each real station has a different number of virtual station. The higher the number is, the more it will influence the mapping. Indeed, many places will have the same concentrations. The number of virtual stations acts as a weight for the interpolation, which tends to smooth very local structures.

3. CONCLUSION

In this paper, we present a method to make denser the network measuring pollution and thus to increase the concentration measurements in order to improve the pollution mapping by interpolation. Making denser consists in virtually adding new measuring stations. Virtual stations are places of a city exhibiting features similar to those of measuring stations. Features are linked to pollution sources (emission register, traffic) and aerodynamical properties of the considered place (morphological indicators, land cover). The approach is not new by itself; this paper offers a strong basis for selecting discriminating features with strong relevance to pollutants circulation. Previous studies by Basly (2000) or Ung (2003) exploited series of satellite data as discriminating features and proved to be sensitive to selected images and to the characteristics of spaceborne sensor.

The first implementation of the method was done for particulates (PM) and for the case study of the city of Strasbourg. From the 5 PM measuring stations initially, we obtained 635 virtual stations homogeneously spread over the area. The benefit of virtual stations to the interpolation is proved in a qualitative way by the pollution map and in a quantitative way by the reduction of the RMSE. Though previous studies support these findings, more cases should be analyzed. Efforts are underway to better analyze the created virtual stations and to validate them. The latter point is not obvious and requests expensive fields campaigns. Technically, the method is easily feasible by practitioners since it involves well-known tools and it uses data they already have. It offers promises to map pollution with high accuracy and thus to represent the pollution distribution variability. This method also appears as complementary to those already existing as virtual stations can constrain air pollution model.

4. ACKNOWLEDGEMENTS

The authors would like to thank Giath Do'Un - Laboratoire Génie Urbain, Environnement et Habitat (LGUEH), Université Marne la Vallée, Pôle Ville - for computing morphological indicators.

REFERENCES

Albuisson M., 1995. Optimisation des traitements des images multi-sources. In Comptes-Rendu du 1er colloque international "l'imagerie scientifique et le traitement d'image", Cannes, 4-6 avril 1995, 8-17, Association Aéronautique et Astronomique de France.

Basly L., Wald L., Baleynaud J.-M., 1999. Satellite mapping for urban air pollution. In : Proceedings, EARSeL Symposium 1998 "operational remote sensing for sustainable development", Enschede, The Netherlands, Nieuwenhuis G., Vaughan R., Molenaar M. eds, Balkema, Rotterdam, 133-139.

Briggs D., de Hoogh C., Gulliver J., Wills J., Elliott P., Kingham S., Smallbone K., 2000. A regression-based method for mapping traffic-related air pollution: application and testing in four contrasting urban environments, *The Science of The Total Environment*, Volume 253, Issues 1-3, 15 May 2000, Pages 151-167.

Carletti R., Picci M., Romano D., 2000. Kriging and bilinear methods for estimating spatial pattern of atmospheric pollutants. *Environmental Monitoring and Assessment*, Vol. 63, 341-359.

Hasager C. B., Thykier-Nielsen S., 2001. IRS-1C LISS III land cover maps at different spatial resolutions used in real-time accidental air pollution deposition modelling, *Remote Sensing of Environment*, Volume 76, Issue 3, June 2001, 326-336.

Hewitt C.N., Jackson A., 2003. *Handbook of Atmospheric Science. Principles and Applications*. Blackwell Publishing, Oxford, UK, 149.

Hoek G., Meliefste K., Cyrys J., Lewné M., Bellander T., Brauer M., Fischer P., Gehring U., Heinrich J., van Vliet P., Brunekreef B., 2002. Spatial variability of fine particle concentrations in three European areas, *Atmospheric Environment*, Volume 36, Issue 25, September 2002, 4077-4088.

Hoek G., Brunekreef B., Goldbohm S., Fisher P., van den Brandt P., 2002. Association between mortality and indicators of traffic-related air pollution in the Netherlands: a cohort study. *The Lancet*, Vol. 360, October 19, 2002, 1203-1209.

Ionescu A., Mayer E., Colda I., 1996. Méthodes mathématiques pour estimer le champ de concentration d'un polluant gazeux à partir des valeurs mesurées aux points dispersés. *Pollution Atmosphérique*, janvier-mars 1996, pp. 78-89.

Ionescu A., Candau Y., Mayer E., Colda I., 2000. Analytical determination and classification of pollutant concentration fields using air pollution monitoring network data. Methodology and application in the Paris area, during episodes with peak nitrogen dioxide levels. *Journal of Environmental Modelling & Software* 15, 565-573.

Monteiro A., Vautard R., Borrego C., Miranda A.I., 2005. Long-term simulations of photo oxidant pollution over Portugal using the CHIMERE model. *Atmospheric Environment*, Vol. 39, Issue 17, June 2005, 3089-3101.

McHugh C. A., Carruthers D. J., Edmunds H. A., 1997. ADMS-Urban: an air quality management system for traffic, domestic and industrial pollution. *International Journal of Environment and Pollution*, Vol. 8, Nos.3-6, 666-675.

McHugh C. A., Carruthers D. J., Edmunds H. A., 1997. ADMS and ADMS-Urban. *International Journal of Environment and Pollution*, Vol. 8, Nos. 306, 437-440.

Puissant A., 2003. Information géographique et images à très haute résolution : utilité et applications en milieu urbain. Thèse de doctorat en géographie, Université Louis Pasteur, Strasbourg I, France, 349 p.

REKLIP Regio-Klima-Projekt, 1995. Communauté trinationale REKLIP sous la responsabilité scientifique de F. FIEDLER – Atlas climatique du Fossé Rhénan méridional. Vdf Zurich, IFG Offenbach, Ed. Coprur, Strasbourg, France, 240 p.

Schmidt H., Derognat C., Vautard R., Beekmann M., 2001. A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. *Atmospheric Environment*, Vol. 35, 2449–2461.

Stalker W.W., Dickerson R.C., 1962. Sampling station and time requirements for urban air pollution surveys, II, suspended particulate matter and soiling index. *Journal of Air Pollution Control Association*, vol. 12, n°3, 111-128.

Tennekes H., Lumley J.L., 1972. *A First Course in Turbulence*. The Massachusetts Institute of Technology Press, Cambridge, MA, USA.

Turbelin G., 2000. Modélisation de la turbulence atmosphérique en vue de l'étude du chargement aérodynamique des structures soumises aux effets du vent. Thèse de doctorat Sciences Mécaniques. Université d'Evry Val d'Essonne, France, 183 p.

Ung A., Weber C., Perron G., Hirsch J., Kleinpeter J., Wald L., Ranchin T., 2001. Air pollution mapping over a city - virtual stations and morphological indicators. *Proceedings of 10th International Symposium "Transport and Air Pollution"*, Boulder, Colorado, USA.

Ung A., Wald, L., Ranchin T., Weber C., Hirsch J., Perron G., Kleinpeter J., 2002. Satellite data for the air pollution mapping over a city – The use of virtual station. In: Gérard Begni (Ed.), *Proceedings of the 21th EARSeL Symposium*, Paris, France, A. A. Balkema, Lisse, Abingdon, Exton (PA), Tokyo, 147-151.

Ung A., 2003. Cartographie de la pollution atmosphérique en milieu urbain à l'aide de données multisources. Thèse de doctorat Sciences de l'environnement. Université Paris 7 Denis Diderot, Paris, France , 96 p.

Vautard R., Martin D., Beekmann M., Drobinski P., Friedrich R., Jaubertie A., Kley D., Lattuati M., Moral P., Neininger B., Theloke J., 2003. Paris emission inventory diagnostics from ESQUIF airborne measurements and a chemistry transport model. *Journal of Geophysical Research*, Vol. 108, 8564.

Weber C., Hirsch J., Puissant A., Durrenberger M., Cerciat M., Schnell L., Bronner A.C., 2002. Morphologie urbaine et répartition spatiale des polluants : de la rue à l'agglomération. Application à la Communauté Urbaine de Strasbourg (CUS). Rapport de recherche, programme ACI "Ville" du MENRT, Strasbourg, 153 p.

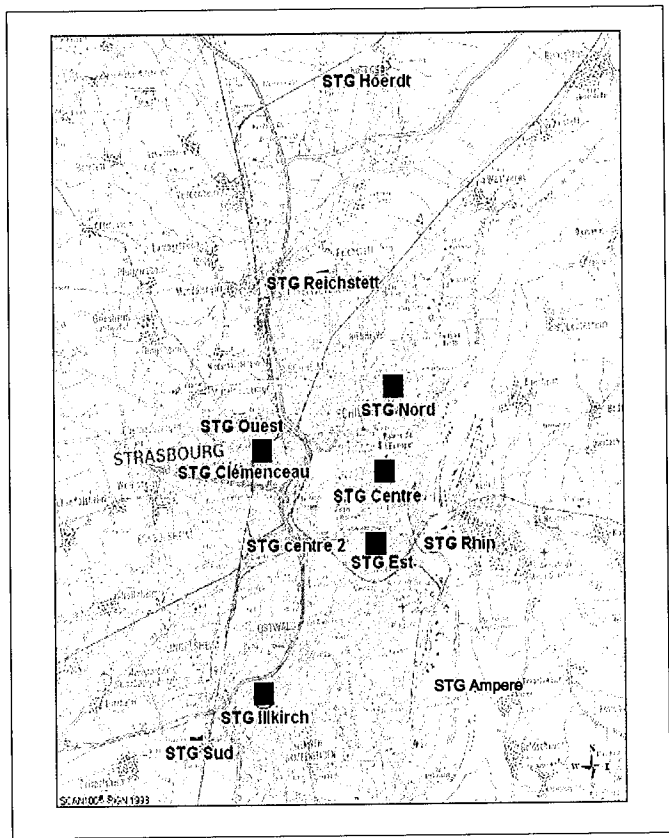


Fig. 1: Stations of the network measuring air quality in the city of Strasbourg. Squares denote stations measuring PM. The width of the map is approximately 20 km.

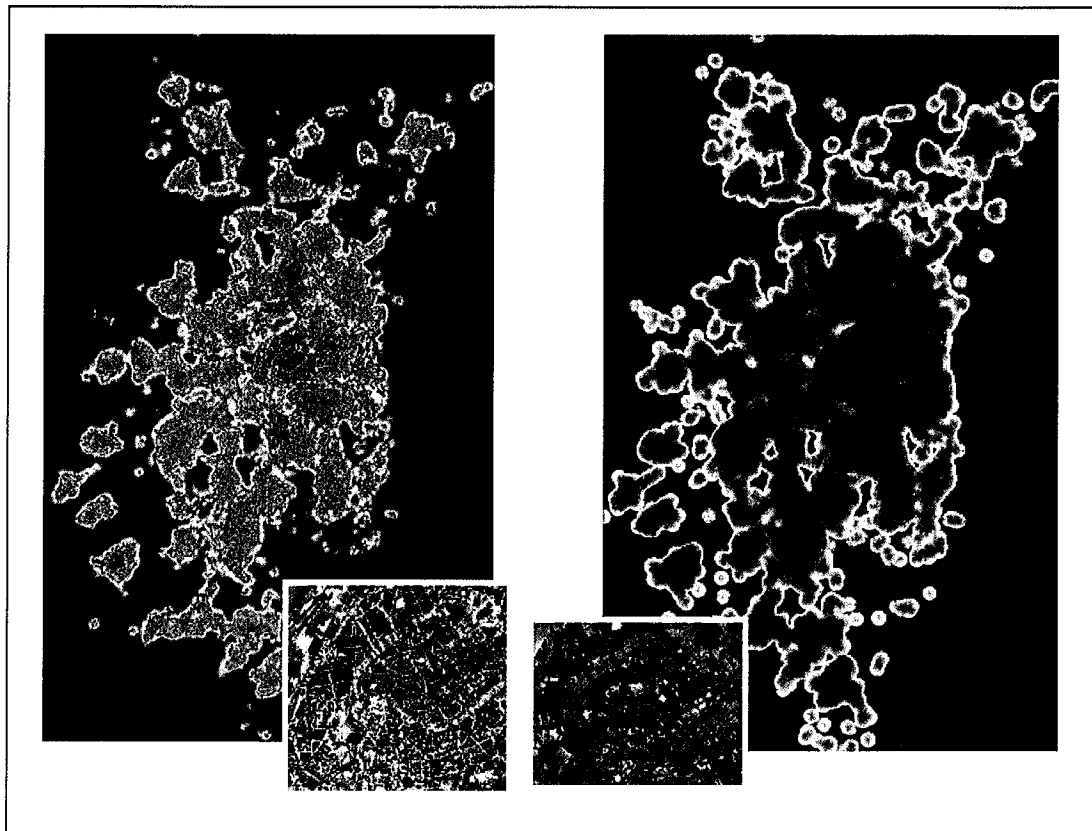


Fig. 2: Images of the Gravelius (left) and Miller (right) indicators for the city of Strasbourg. Images at bottom display zooms of the city center for each indicator.

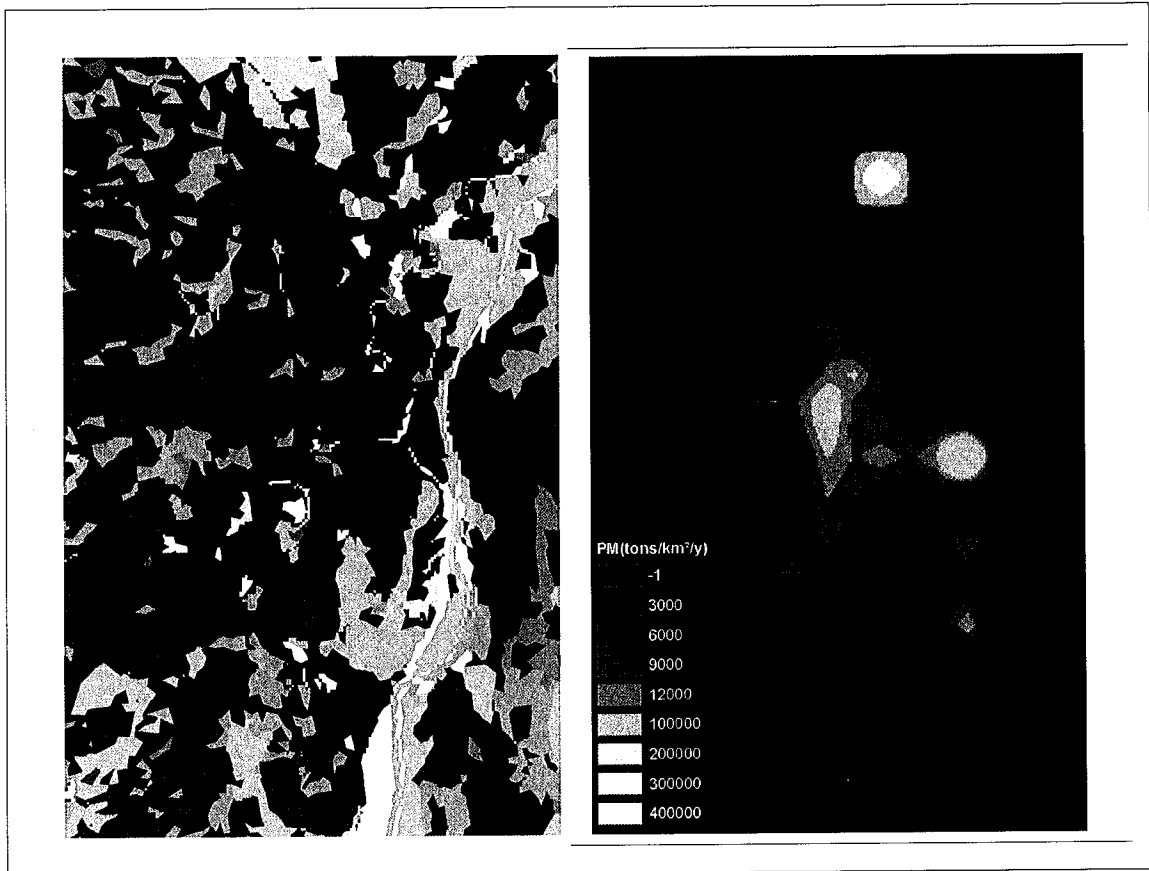


Fig. 3: Images of the land cover map (left) and the emission register for PM for 2004 (right). On the land cover map, the 44 different types of land cover are in different gray levels. For example, urbanized lands are in black.

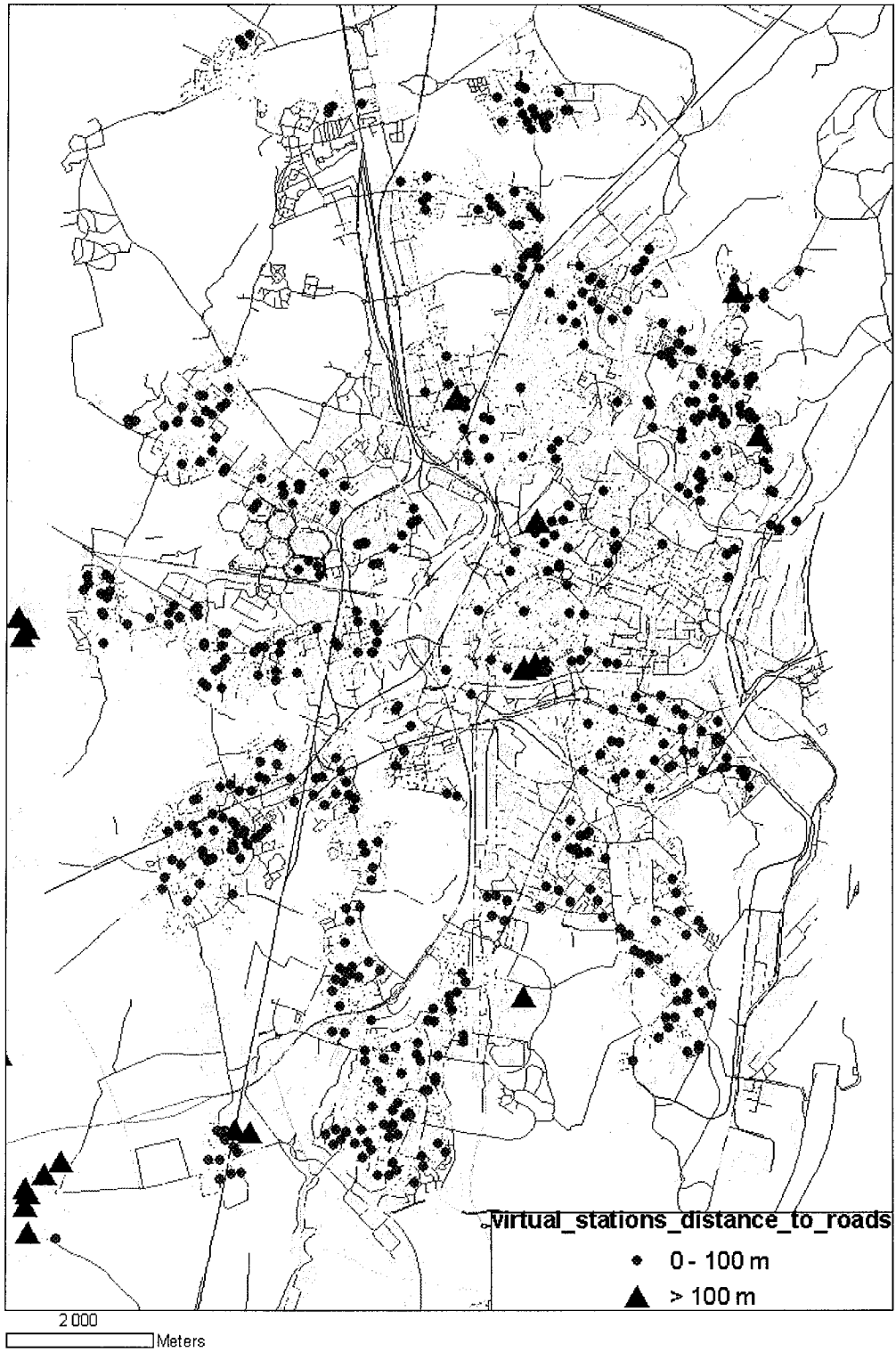


Fig 4 : virtual stations in Strasbourg. Triangles will be rejected.

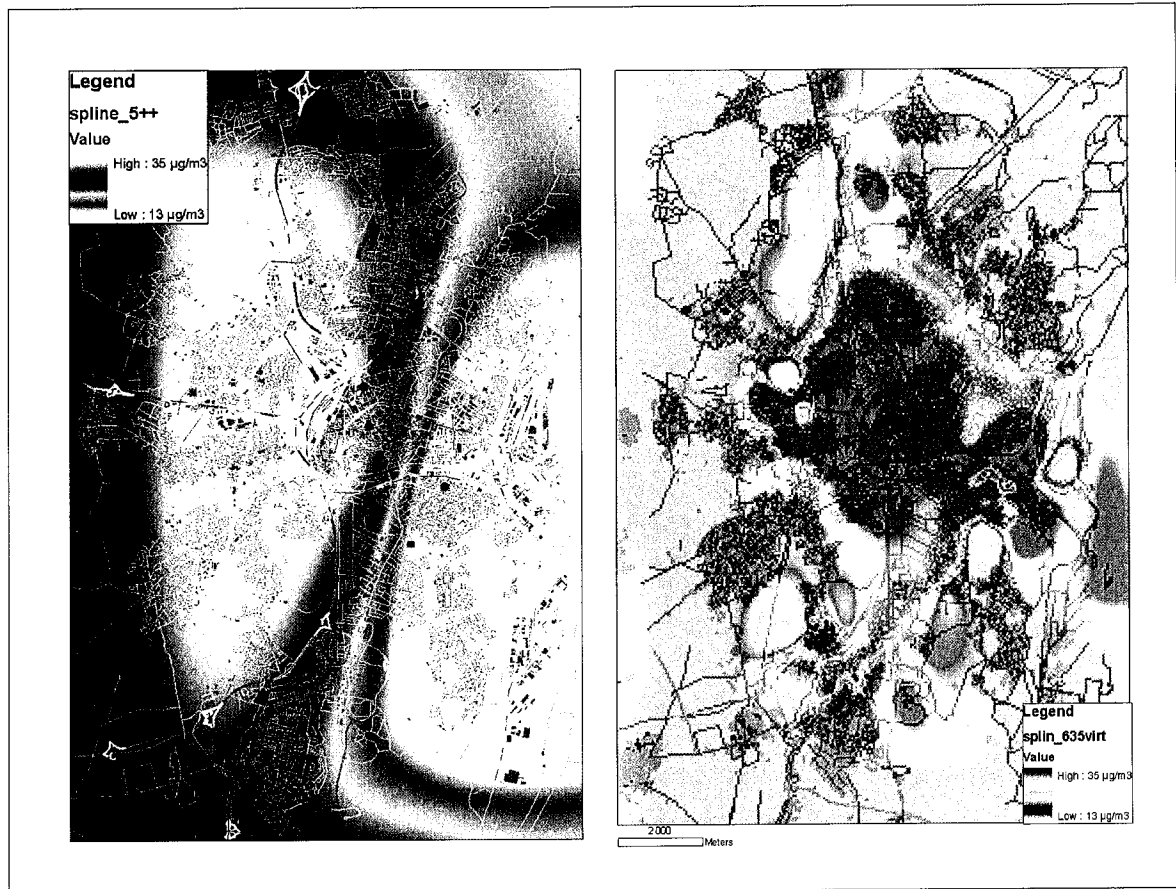


Fig. 5: PM10 concentration maps over Strasbourg obtained by interpolation of real measuring stations only (left, black dots) and of both real and virtual measuring stations (right).

Table 1: Number of virtual stations for each real measuring station

Real station name	Number of virtual stations
STG Illkirch	71
STG Centre	104
STG Nord	450
STG Clemenceau	6
STG Est	4

Table 2: Relative RMSE values for an interpolation with and without virtual measuring stations

“Thin plates spline” method	5 real measuring stations	Real + virtual measuring stations
Relative RMSE	49%	15%

Chapitre 5

Détection des aérosols urbains (PM) à
l'aide d'images satellitaires

1 Démarche

Ce chapitre concerne la validation du concept de « station virtuelle ». Il passe tout d'abord par une étude de l'apport de l'imagerie satellitaire haute et moyenne résolution pour la cartographie de la pollution particulaire et plus précisément, pour l'estimation des concentrations en PM, dans le but ultime de valider le concept de « station virtuelle ». Dans ce chapitre, je cherche à savoir s'il est possible de détecter la pollution particulaire directement sur les images satellites. Les nouveaux capteurs offrent des avantages pour la détection des polluants : de bonnes résolutions spatiale et temporelle, une couverture de la gamme de longueurs d'ondes assez large et des largeurs de bandes spectrales assez fines. Toutefois, ces avantages peuvent aussi devenir des inconvénients en plus de l'inconvénient majeur que constitue la couverture nuageuse pour l'imagerie satellitaire. Il n'existe pas de capteurs à haute résolution spatiale dédiés à la détection de la pollution particulaire. Pour déterminer le capteur idéal qui va me permettre de faire mon étude, il s'agit donc de faire un compromis entre les trois paramètres suivants : la résolution spatiale, le temps de revisite ou résolution temporelle et l'enneuagement. Le tableau 5.1 résume les différentes caractéristiques des capteurs pressentis pour cette étude.

Capteurs spatiaux	Landsat 5 - TM	Envisat - MeRIS	SPOT 1, 2, 3 - HRV
Résolution spatiale	Vis : 30 m IRT [10,4 μm -12,5 μm] : 120 m	300 m	20 m Pan : 10 m
Résolution temporelle	18 jours	3 jours	3 à 26 jours
Nombre de bandes spectrales	7	15	3 + Pan
Intervalle spectral couvert (μm)	[0,450 – 12,50]	[0,41 – 1,00]	[0,50 à 0,89]
Largeurs des bandes spectrales (μm)	0,06 à 2,1	0,0037 à 0,0200	0,07 à 0,10

Tableau 5.1 : Caractéristiques des capteurs satellite Landsat-TM, Envisat-MeRIS et SPOT-HRV. (Sources ESA et Lhomme 2005)

Pour cette étude, j'ai choisi les capteurs TM du satellite Landsat et MeRIS de Envisat. Malgré un temps de revisite assez long, le capteur TM offre une résolution spatiale intéressante (30 m) pour une étude à l'échelle de la rue. Pour le capteur MeRIS, malgré une résolution spatiale relativement

grossière, il offre une bonne résolution temporelle et des largeurs de bandes spectrales très fines. Je dispose de huit images Landsat et de six images MeRIS de la zone d'étude.

Le chapitre 2 a montré que l'imagerie satellitaire est déjà utilisée pour la détection de la pollution atmosphérique (Hashim *et al.* 2004, Retalis *et al.* 2003, Sifakis *et al.* 1992, 1995, 1999). La plupart des auteurs procèdent par comparaison de deux images dont l'une sert de référence. La texture est un des éléments important de comparaison. Des cartes de pollution sont produites. Elles permettent de rendre compte de l'étendue du voile de pollution, de ses variations de concentrations mais ne permettent pas d'estimer directement la concentration en polluant aux échelles spatiales qui nous intéressent (100 m). Notre démarche complète celles déjà existantes en ce sens où il s'agit d'une part, de confirmer la détection de PM sur des images satellitaires et d'autre part, d'en estimer la concentration en certains points de la ville. Ces points de la ville correspondent aux stations virtuelles déterminées dans le chapitre 4. Ce travail doit servir finalement à la validation de ces stations virtuelles. Ce chapitre a donc un double objectif :

- 1) confirmer qu'il est possible de détecter des aérosols urbains grâce à des images satellite à des échelles locales,
- 2) valider le concept des stations virtuelles.

La physique du phénomène est complexe car de nombreux paramètres entrent en jeu, comme nous le verrons plus loin. Pour résoudre ce problème, je vais faire appel à des outils impliquant des principes liés au transfert radiatif dans l'atmosphère et à la physique du capteur ou de la mesure. Je vais utiliser le modèle de transfert radiatif 6S combinée à une approche par plans d'expériences numériques. Le modèle 6S simule la réponse d'un capteur satellite qui reçoit un signal ayant traversé l'atmosphère. Quant aux plans d'expériences numériques, ils permettent de prendre en compte plusieurs paramètres atmosphériques et, surtout, de réduire le nombre de simulations nécessaires.

Ces simulations vont servir à deux choses :

- déterminer les contributions relatives de chaque paramètre pour savoir si la concentration en aérosols influence de manière conséquente le signal reçu par le satellite et s'il est possible de détecter ce phénomène sur une image,
- déterminer le ou les canaux optimaux pour l'observation, la détection et la quantification de la pollution particulaire.

L'étape suivante consiste à comparer les résultats issus des simulations avec les observations faites au sol et par le satellite. Cette étape doit conduire à l'établissement d'un modèle empirique qui permettra

le passage de la donnée satellitaire à la concentration en PM. De cette façon, les concentrations des stations réelles et virtuelles pourront être comparées.

Notons ici, que les termes « PM », « aérosols urbains » et « particules » désignent la même chose et seront donc employés indifféremment.

2 Approche simplifiée du problème et hypothèses

2.1. Effets atmosphériques

L'atmosphère modifie le rayonnement électromagnétique qui la traverse (dans notre étude, la source de rayonnement est le soleil), par des effets d'absorption et de diffusion dus :

- aux molécules gazeuses entrant dans sa composition ou présentes dans l'air : azote, oxygène, argon, dioxyde de carbone, ozone, vapeur d'eau et polluants gazeux,
- aux particules qu'elle contient : aérosols, gouttelettes d'eau, poussières...

Les gaz et les particules, mais essentiellement les gaz, sont responsables de l'absorption sélective de certaines longueurs d'onde. Pour les gaz de l'atmosphère les plus importants, le domaine de longueur d'onde dans lequel il y a absorption est connu (courbes d'absorption). Ainsi, les rayonnements dans le visible et dans le proche infrarouge sont peu absorbés par les gaz (Vermote *et al.* 1994).

Quant à la diffusion (figure 5.1), elle est due aux particules, molécules, en suspension dans l'atmosphère et dépend de leur taille par rapport à la longueur du rayonnement électromagnétique.

- Les molécules gazeuses, comme N₂ et O₂, et les particules de petites tailles dont la dimension est nettement inférieure à la longueur d'onde du rayonnement diffusent selon la loi de Rayleigh. Cette diffusion est proportionnelle à λ^{-4} .
- Pour les particules, dont la taille est proche de la longueur d'onde du rayonnement ou supérieure, c'est la diffusion de Mie qui est prédominante. Pour les aérosols, la diffusion est proportionnelle à $\lambda^{-\alpha}$ avec $0 < \alpha < 4$ (Wald 2006). Elle est beaucoup moins sélective que la diffusion Rayleigh. Sur les gouttes d'eau, les poussières, la diffusion est importante pour toutes les longueurs d'onde du spectre visible et infrarouge.

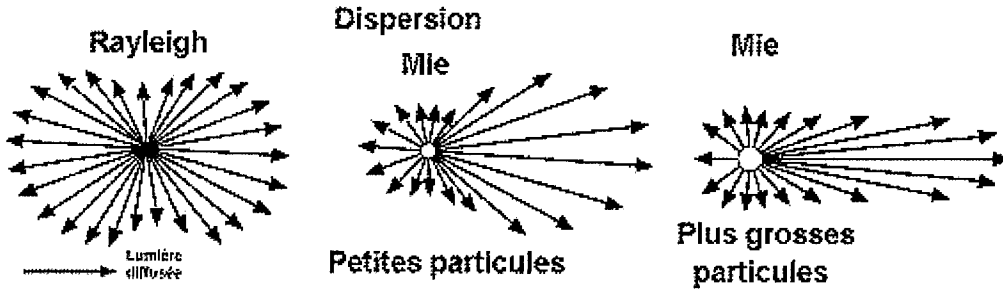


fig. 5.1 : Illustration des différents type de diffusion. Source Wikipédia.

Suivant la composition de l'aérosol, les propriétés absorbantes et diffusantes sont différentes. Le tableau 5.2 donne les valeurs des coefficients d'absorption et de diffusion pour diverses espèces d'aérosols. Le coefficient de diffusion augmente avec l'humidité relative de l'atmosphère.

Propriétés optiques	Sulfate	Carbone organique	Poussière minérale	Suie	Aérosol fin
$k_{SP} (m^2 g^{-1})$	5 (3.6 – 7)	5 (3.0 – 7)	$k_{EXT} \sim 0.7$	3	3 (2 – 4)
$k_{AP} (m^2 g^{-1})$	0	0		10 (8 – 12)	(0 – 10)

Tab. 5.2 : Coefficients de diffusion (k_{SP}) et d'absorption (k_{AP}) pour différentes espèces d'aérosols. k_{EXT} est le coefficient d'extinction et $k_{EXT} = k_{AP} + k_{SP}$. Source Hewitt & Jackson 2003.

Pour prendre en compte la diffusion ou l'absorption par les molécules ou les aérosols, il existe un paramètre qui permet d'évaluer le pouvoir absorbant et diffusant (opacité) de l'atmosphère : l'épaisseur optique. Celle-ci peut parfois être estimée par la visibilité. Pour les gaz, il est facile d'obtenir l'épaisseur optique en fonction de la concentration. Pour une longueur d'onde λ donnée, la transmittance monochromatique dans le cas de molécules gazeuses est reliée à la concentration de la façon suivante (Liou 1980) :

$$\tau_{\lambda} = e^{-\frac{k_{\lambda} C}{\mu}} \quad (1)$$

où k_{λ} est le coefficient d'absorption en $cm^2 \text{ molécule}^{-1}$, C est la concentration du gaz absorbant et μ est une fonction de l'angle zénithal local du soleil.

L'équation (5) relie la transmittance τ_{λ} à l'épaisseur optique δ_{λ} :

$$\tau_{\lambda} = e^{-\frac{\delta_{\lambda}}{\mu}} \quad (2)$$

Finalement, pour les molécules gazeuses, la concentration des gaz et l'épaisseur optique sont liées selon l'équation (3) :

$$\delta_{\lambda} = k_{\lambda} C \quad (3)$$

Par contre pour les aérosols, il n'en est pas de même. Selon Angström, la transmittance globale due à la diffusion par les aérosols (secs et humides) est fonction de la taille des particules et de la longueur d'onde, comme nous l'avons vu précédemment mais également de leur nombre et de leur masse. Il n'existe donc pas de relation mathématique connue entre l'épaisseur optique des aérosols et la concentration en aérosols. Plusieurs travaux (Pancrati 2003, Kacenelenbogen 2004) faisant appel à la base de données AERONET (<http://www.aeronet.gsfc.nasa.gov>) et utilisant des données sol ou issues de spectro-radiomètres embarqués ont étudié le lien entre ces deux grandeurs. Ces études expérimentales permettent d'aboutir à une approximation linéaire souvent utilisée en première approximation (Retalis *et al.* 2003). En me basant sur les travaux de Pancrati (2003), j'ai établi la relation (4) entre la concentration en particules du type « aérosol urbain » et l'épaisseur optique à 550 nm.

$$\delta_{550} = 0,0082 C_{PM} + 0,0022 \quad (4)$$

Les valeurs d'épaisseur optique que l'on obtient avec cette relation en prenant des valeurs de concentration réalistes sont proches de celles mesurées expérimentalement (Bellouin *et al.* 2003, Kacenelenbogen 2004, Thieuleux *et al.* 2004).

2.2. Mesures par satellite

Pour s'affranchir en partie de ces effets atmosphériques qui vont perturber le signal, avant tout de l'absorption moléculaire, les capteurs-imageurs à bord des satellites observent la Terre dans des fenêtres atmosphériques où les phénomènes d'absorption sont moindres et où l'intensité du rayonnement émis par la source est suffisamment fort pour être détecté. Le capteur satellite mesure une valeur de luminance. La luminance est l'intensité d'une source de rayonnement dans une direction donnée, divisée par l'aire apparente de cette source dans cette même direction. De la luminance, on peut facilement en déduire la valeur de réflectance ρ , i.e. le rapport entre l'énergie réfléchie et l'énergie incidente, $0 < \rho < 1$.

La réflectance mesurée par le satellite s'exprime, en première approximation, de la façon suivante :

$$\rho_{sat} = \rho_{atm} + T\rho_{sol} \quad (5)$$

avec ρ_{sat} , la réflectance mesurée par le satellite, ρ_{atm} , la réflectance de l'atmosphère, ρ_{sol} , la réflectance du sol (appelée ici, albedo), et T , la transmittance globale de l'atmosphère.

T s'écrit :

$$T = e^{-\frac{\delta}{\mu}} \quad (6)$$

δ est l'épaisseur optique globale de l'atmosphère. Elle peut s'écrire comme la somme de l'épaisseur optique de l'atmosphère sans polluant (ou atmosphère claire), δ_{atm} , et de l'épaisseur optique de la pollution, δ_{pol} (équation 7).

$$\delta = \delta_{atm} + \delta_{pol} \quad (7)$$

alors T s'écrit :

$$T = T_{atm} T_{pol} \text{ avec } T_{pol} = e^{-\frac{0,0082 C_{PM} + 0,0022}{\mu}} \quad (8)$$

La transmittance T et donc la réflectance observée par le capteur satellite ρ_{sat} sont influencées par les éléments qui constituent l'atmosphère et notamment par la pollution en particule. D'après ces éléments théoriques qui font appel à la physique de l'atmosphère et de la mesure, une variation de concentration en particules devrait théoriquement entraîner une variation de la réflectance ρ_{sat} . Je vais vérifier la théorie à l'aide de simulations numériques. Plus précisément, l'utilisation combinée de plans d'expériences numériques et d'un modèle de transfert radiatif va permettre de quantifier l'influence des variations de concentrations en PM sur la radiométrie des images satellites.

2.3. Outils expérimentaux

2.3.1. *Plans d'expériences numériques*

La concentration en particules n'est pas le seul facteur influençant le signal reçu au niveau du capteur satellite. D'autres paramètres qu'ils soient atmosphériques ou de surface, ont une influence sur ce que mesure le capteur. Prendre en compte tous ces paramètres reste complexe. Néanmoins, il est possible d'obtenir une modélisation réaliste en ne faisant varier que les paramètres les plus importants. Ainsi, j'ai choisi de faire varier : le contenu en vapeur d'eau (UH_2O), le contenu en ozone (UO_3), l'épaisseur optique des aérosols urbains (AOTp) et celle de l'atmosphère claire (OTatm) et enfin l'albedo de surface (ρ_{sol}).

Le terme « atmosphère claire » signifie l'atmosphère sans les polluants considérés qui sont les aérosols urbains. Son épaisseur optique est variable car outre les aérosols urbains, d'autres polluants, molécules ou aérosols, dont la concentration varie peuvent avoir une influence sur son épaisseur optique. Pour avoir l'épaisseur optique de l'atmosphère considérée, il faut additionner OTatm et AOTp.

Les effets et les contributions des paramètres atmosphériques et de surface sont étudiés pour chaque bande spectrale grâce à l'utilisation de plans d'expériences numériques. L'échantillonnage statistique des valeurs des variables par un plan d'expériences permet d'explorer leur intervalle de définition plus rigoureusement, en un nombre limité de simulations. Autrement dit, l'approche par plans d'expériences numériques permet :

- de prendre en compte les variations conjointes de plusieurs paramètres atmosphériques,
- de définir leur domaine de variation,
- et de réduire le nombre de simulation nécessaires pour modéliser le phénomène.

Les plans d'expériences sont destinés à maximiser l'information accessible à partir d'un nombre limité d'expériences. La pratique des plans d'expériences en télédétection est assez récente (Dechambre & Le Gac 2001). Ces auteurs ont validé et comparé des modèles de diffusion du rayonnement micro-ondes dans la végétation.

Considérons un modèle dont on cherche à étudier la dépendance vis-à-vis de ses variables d'entrée. La réponse ρ du modèle, qui correspond ici à la réflectance simulée, est fonction de plusieurs variables v_i (appelées paramètres ou facteurs dans la terminologie des plans d'expériences) : $\rho = F(v_1, v_2, \dots, v_p)$. Ces facteurs peuvent être liés à la chimie (UH₂O) ou à la physique (OTatm) de l'atmosphère, ou au type de surface (ρ_{sol}) sur laquelle le rayonnement est réfléchi.

L'approche par plans d'expériences est souvent employée pour des études de sensibilité de modèle (Bacour 2001). L'approche classique de l'analyse de sensibilité du modèle à ses variables consiste, après avoir déterminé une gamme de variation pour chaque facteur, à fixer tous les facteurs sauf un, v_i , et à tracer $\rho = F(v_i)$. En général, les facteurs figés prennent la valeur moyenne de leur intervalle de définition.

La méthode des plans d'expériences consiste à choisir un plan d'expérimentation comportant un minimum d'essais compte tenu des résultats souhaités, tout en apportant un maximum de précision dans les résultats. Au cours des essais successifs, les niveaux de tous les facteurs varient simultanément, de manière programmée et raisonnée. Ainsi, par rapport à l'approche classique et pour l'étude d'un même modèle, une telle démarche permet :

- de diminuer le nombre d'essais,
- d'étudier plus de facteurs,
- de détecter les interactions entre facteurs.

Les méthodes de construction des plans d'expériences sortent du cadre de l'étude. Le lecteur peut se référer à l'ouvrage de Benoist et al. (1994) pour une description exhaustive.

Parmi toutes les variables d'entrée du modèle 6S, l'analyse de sensibilité n'est effectuée que pour cinq d'entre eux, à savoir : le contenu en vapeur d'eau (U_{H_2O}), le contenu en ozone (U_{O_3}), l'épaisseur optique des aérosols urbains (AOTp) et celle de l'atmosphère claire (OTatm) et enfin l'albedo de surface (ρ_{sol}). Les variations de l'épaisseur optique des aérosols urbains (AOTp) sont choisies de manière à correspondre aux variations réalistes de la concentration en PM en ville, grâce à l'emploi de l'équation (4).

Les résultats des simulations sont collectés à partir d'un plan appelé Hyper Gréco Latin Géométrique pour lequel tous les facteurs ont le même nombre de niveaux. Le tableau d'expérience utilisé, selon la notation de Taguchi, est $L_{343}7^{57}$. Il permet de prendre en compte jusqu'à 57 facteurs à 7 niveaux et comprend 343 simulations. La construction de ce plan d'expériences a été effectuée par Bacour (2001).

Pour chacun des 5 facteurs que j'ai choisis, a été définie une gamme de variations réalistes. Les valeurs prises par les variables au cours des simulations sont équiréparties (aux arrondis près), de telle sorte que la valeur minimale (maximale) correspond à la borne inférieure (supérieure) incrémentée de 5 % de la taille de l'intervalle. Ces valeurs sont rassemblées dans le tableau 5.3.

paramètre	variation	niveaux
AOT polluant	0,1579 - 1,2077	0,1662 ; 0,3302 ; 0,4942 ; 0,6582 ; 0,8222 ; 0,9862 ; 1,1502
OT atm	0,144 - 0,819	0,152 ; 0,257 ; 0,361 ; 0,466 ; 0,570 ; 0,675 ; 0,780
UH2O	0,33 - 3,31	0,35 ; 0,82 ; 1,28 ; 1,75 ; 2,21 ; 2,68 ; 3,15
UO3	0,234 - 0,367	0,246 ; 0,263 ; 0,281 ; 0,298 ; 0,315 ; 0,332 ; 0,350
ρ_{sol}	0,057 - 0,630	0,06 ; 0,15 ; 0,24 ; 0,33 ; 0,42 ; 0,51 ; 0,60

Tab. 5.3 : Gammes de variations et niveaux des différents paramètres ou facteurs utilisés dans le plan d'expériences.

Les valeurs de l'albedo sont celles que l'on trouve en milieu urbain (albedo du béton, bitume, végétation...) (Lachéradé *et al.* 2006). Les valeurs du contenu en ozone proviennent des données du satellite GOME (<http://www.wdc.dlr.de>) et celles du contenu en vapeur d'eau proviennent des mesures du spectroradiomètre MODIS (<http://modis-atmos.gsfc.nasa.gov>).

2.3.2. Modèle de transfert radiatif 6S

Le modèle 6S (Second Simulation of the Satellite Signal in the Solar Spectrum) permet de simuler des observations effectuées depuis l'espace, de prendre en compte les conditions aux limites du sol, la

transmission par les gaz et les particules dans l'atmosphère (absorption, diffusion Rayleigh ou Mie) (Vermote et al. 1994).

Avec ce modèle, je vais simuler la réflectance observée par le capteur TM de Landsat 5 pour chacun des essais déterminés par le plan d'expériences et pour chacune des bandes spectrales. Au cours de chaque simulation, les variations conjointes des paramètres atmosphériques cités ci-dessus sont prises en compte.

Ainsi un total de 2058 (343 x 6 bandes spectrales) réflectances ont été calculées pour une date.

3 Effets moyens et contributions des paramètres atmosphériques

L'étude des effets moyens et des contributions de chaque paramètres doit permettre de renforcer l'idée qu'il est possible de détecter la pollution particulaire par satellite.

3.1. Effets moyens

Les effets moyens de chaque facteurs sont calculés pour chaque bande spectrale. Le niveau m du facteur v_f apparaît m_f fois, alors la moyenne des résultats d'essais pour cette modalité s'écrit :

$$\bar{\rho}_{f m} = \frac{\sum \rho_{f m}}{m_f}.$$

L'effet moyen de chaque facteur, qui représente les écarts entre les valeurs moyennes de la réponse de la modalité m et de la moyenne générale $\bar{\rho}$, s'écrit comme :

$$\hat{e}_{f m} = \bar{\rho}_{f m} - \bar{\rho} \quad (9)$$

Soit $\hat{e}_{f m}(\lambda)$ l'effet du facteur f lorsqu'il prend la valeur m . Il s'exprime par rapport à la réflectance moyenne estimée à partir des simulations, dans une configuration spectrale donnée : $\bar{\rho}(\lambda)$. L'effet est maintenant exprimé en pourcentage par rapport à la réflectance moyenne, tel que :

$$E_{f m}(\lambda) = \frac{\bar{\rho}_{f m}(\lambda) - \bar{\rho}(\lambda)}{\bar{\rho}(\lambda)} 100 \quad (10)$$

Les effets des variables UH_2O , UO_3 , O_{atm} , $AOTp$, ρ_{sols} , sur la réflectance moyenne simulée par 6S sont présentés sur les figures 5.2, 5.3 et 5.4. Les résultats mettent clairement en évidence les facteurs spectralement prédominants.

Dans le visible (TM1, 2, 3, 4, 5) et de façon moins flagrante dans le proche infra-rouge (TM7), c'est l'épaisseur optique des aérosols urbains et donc la concentration en PM qui affecte majoritairement les niveaux de réflectance. Sur son intervalle de définition, l'amplitude totale de variation de l'effet moyen de AOTp est de l'ordre de 215 % (TM1) puis 180 % (TM2) et 145 % (TM3), son effet devenant moins important à mesure que l'on s'approche des longueurs d'ondes de l'infra-rouge. Le facteur qui affecte le plus la réflectance après AOTp est O_{atm} suivi de α , U_{H_2O} et U_{O_3} dont les effets sont à peu près similaires. Pour comparer avec AOTp, l'amplitude totale de variation de l'effet moyen de O_{atm} est de l'ordre de 115 % (TM1) puis 90 % (TM2) et 75 % (TM3).

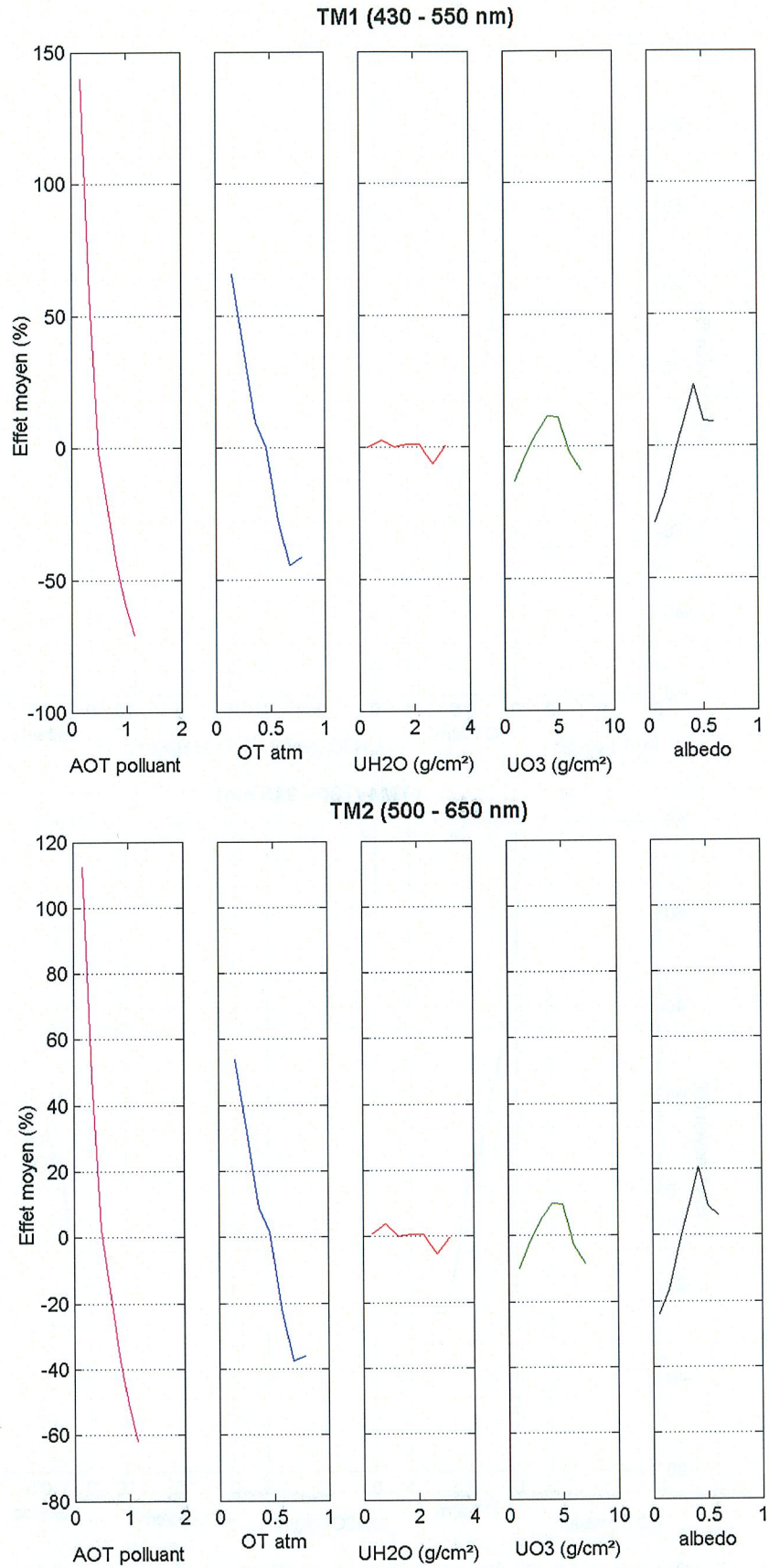


Fig. 5.2 : effet moyen (en %) des paramètres sur la réflectance des canaux TM1 et TM2.

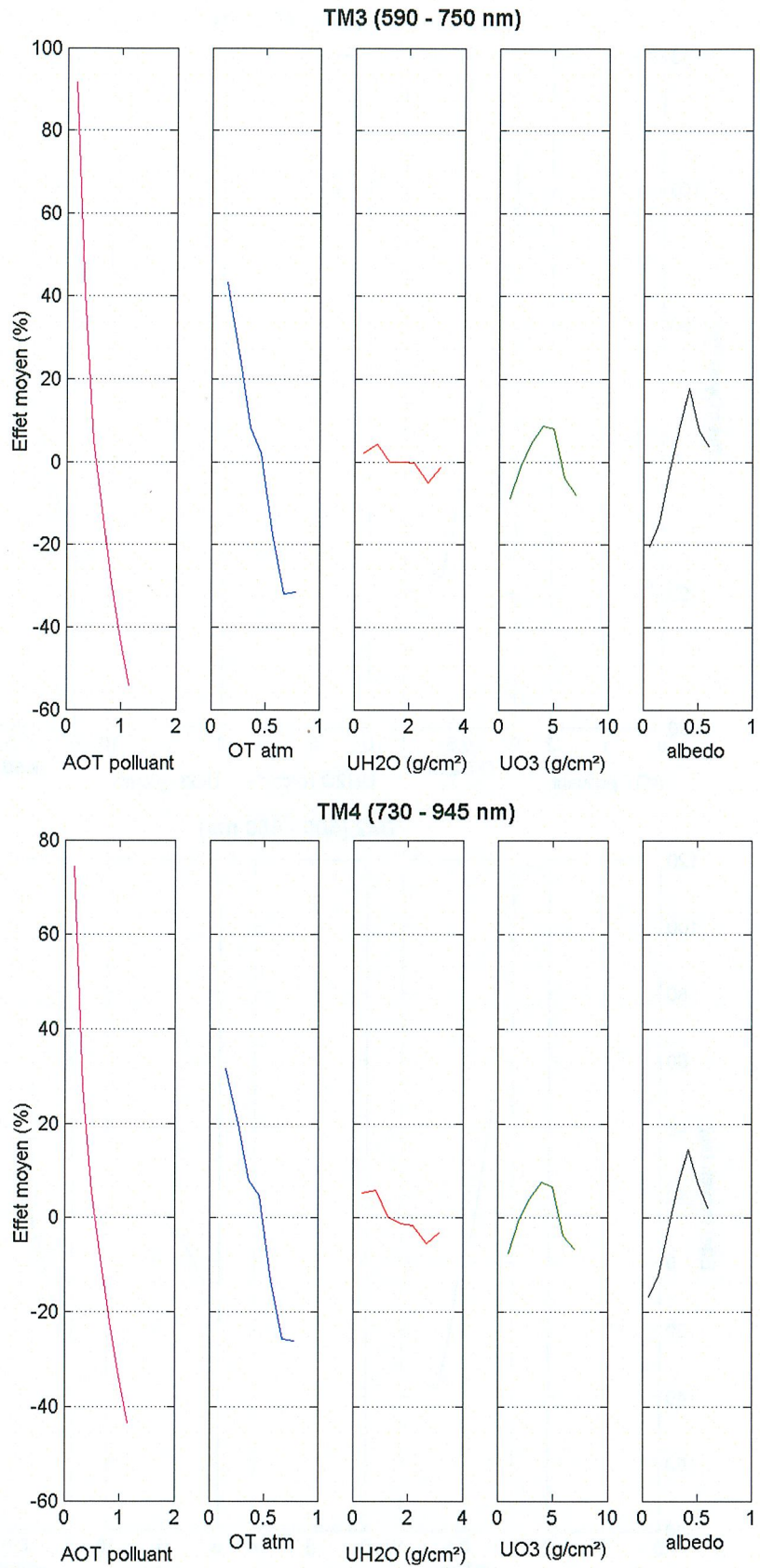


Fig. 5.3 : effet moyen (en %) des paramètres sur la réflectance des canaux TM3 et TM4.

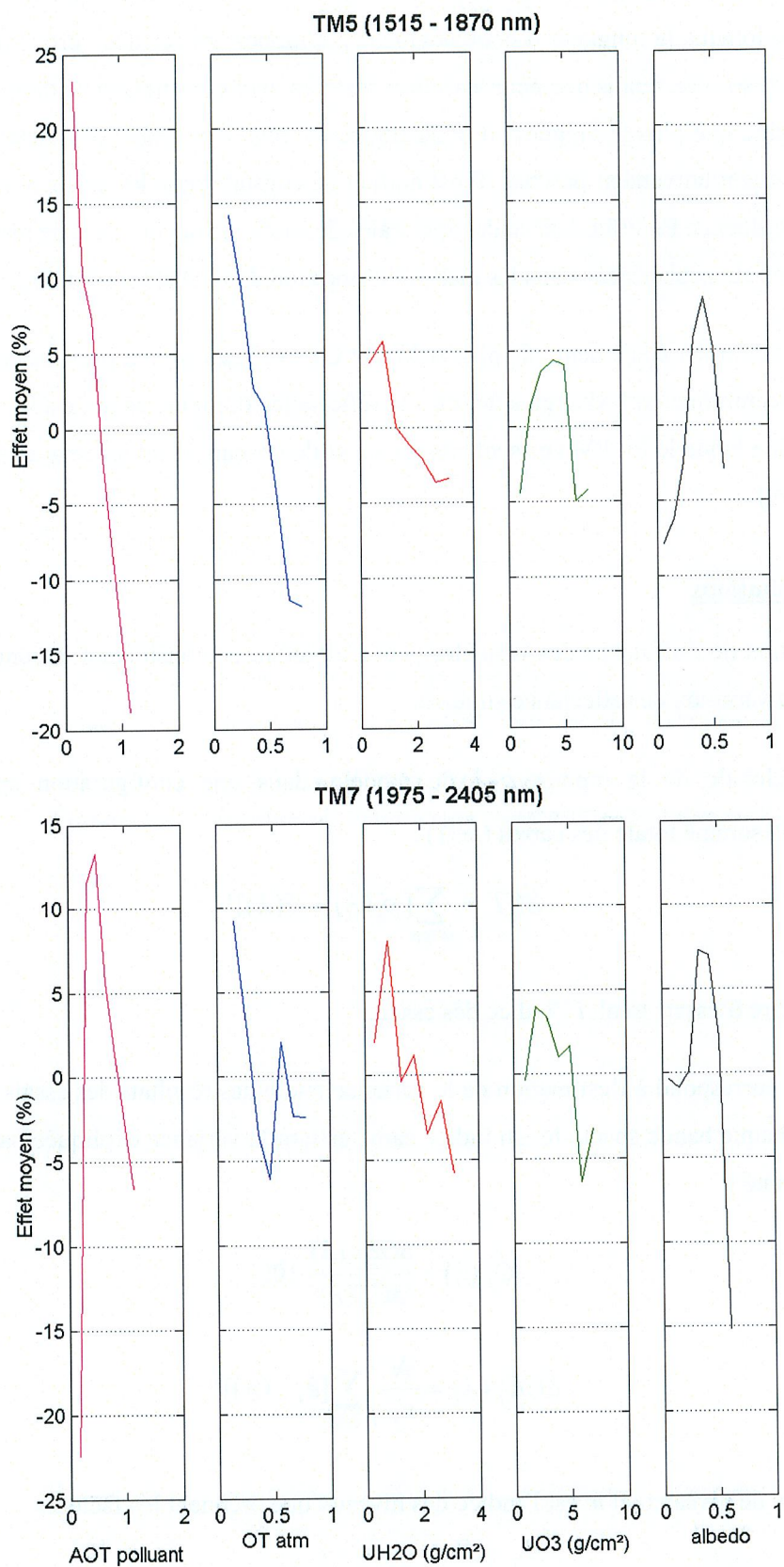


Fig. 5.4 : effet moyen (en %) des paramètres sur la réflectance des canaux TM5 et TM7.

Du visible vers le proche infra-rouge, les effets moyens de AOTp et OTatm diminuent de façon proportionnelle lorsque la longueur d'onde augmente, le rapport de ces deux effets moyens restant le même sur tout le spectre, tandis que pour les autres facteurs, leur effet moyen ne diminue pratiquement pas. Ceci entraîne que plus la longueur d'onde augmente, plus l'ensemble des facteurs tend vers des effets moyens quantitativement proches. Il est normal de constater que les effets moyens de UH₂O et UO₃ sont peu influents. En effet, les bandes spectrales du satellite sont choisies de manière à coïncider avec les fenêtres atmosphériques dans lesquelles l'absorption de H₂O et O₃ est faible.

Il est important de noter également que plus AOTp et OTatm augmentent, plus leur effet diminue. En fait plus la concentration en PM augmente, plus la réflectance diminue, ce qui est bien en accord avec l'hypothèse selon laquelle les PM agissent comme un voile obscurcissant les images satellite (Wald & Baleynaud 1999).

3.2. Contributions

La hiérarchisation de l'influence des variables peut s'effectuer en considérant la contribution relative de chacune à la variation de réflectance simulée.

La variation globale de la réponse $\rho(\lambda)$ du modèle, dans une configuration spectrale donnée, s'exprime par la somme totale des carrés (SCT) :

$$SCT = \sum_{1 \leq i \leq N} [\rho_i(\lambda) - \bar{\rho}(\lambda)]^2 \quad (11)$$

avec N , le nombre d'essais total, i l'indice des essais.

Cette grandeur correspond à l'estimation de la variance totale des résultats des essais au facteur $I/(N-I)$ près. Pour chaque bande spectrale, un indice caractérisant la variance expliquée par chaque facteur est ainsi déterminé :

$$C_f(\lambda) = \frac{SCE_f(\lambda)}{SCT(\lambda)} 100 \quad (12)$$

avec

$$SCE_f(\lambda) = \frac{N}{n_f} \sum_{1 \leq m \leq n_f} [\hat{e}_{f,m}(\lambda)]^2 \quad (13)$$

n_f est le nombre de niveaux, et m est l'indice des niveaux que prennent les facteurs.

La figure 5.5 représente la contribution relative de chacune des variables en fonction de la bande spectrale. On notera que la somme des contributions de chaque facteur, dans une configuration spectrale donnée, est inférieure à 100 % (tableau 5.4). Cela est dû aux erreurs résiduelles et au fait que

les interactions entre variables ne sont pas prises en compte. En effet, l'utilisation des plans d'expériences numériques permet en outre de quantifier l'interaction entre différentes variables d'un même modèle de réflectance, mais cela n'a pas été étudié dans cette thèse.

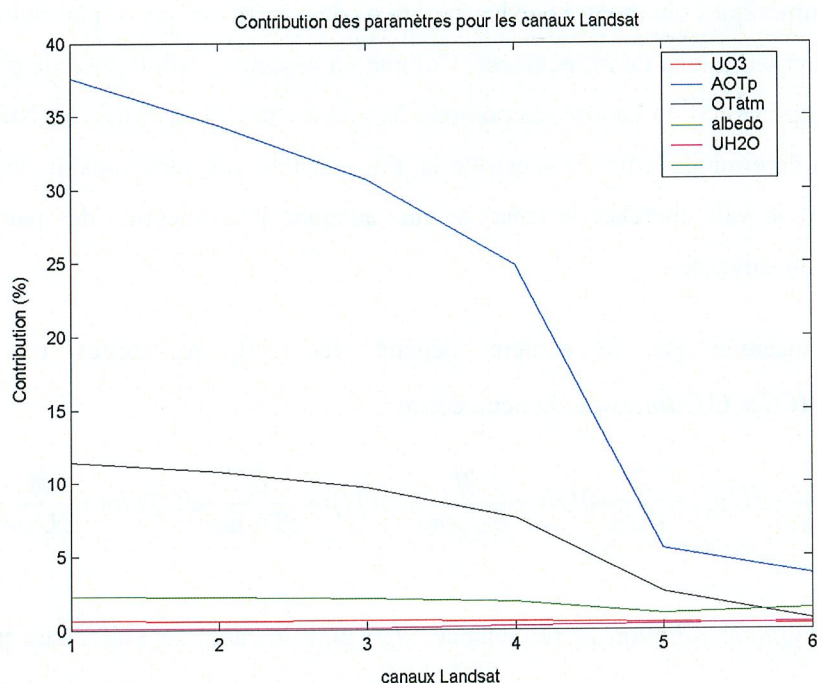


Fig. 5.5 : contribution relative des paramètres aux variations de réflectances du satellite Landsat-TM.

Canaux	TM1	TM2	TM3	TM4	TM5	TM7
Caotp	37.62	34.42	30.67	24.88	5.49	3.76
Cotatm	11.43	10.77	9.66	7.64	2.52	0.67
CUh2o	0.06	0.07	0.10	0.26	0.35	0.49
CUo3	0.60	0.58	0.61	0.58	0.49	0.35
Calbedo	2.26	2.21	2.10	1.89	1.06	1.41

Tab. 5.4 : calcul des contributions (%) de chacun des paramètres

La figure 5.5 et le tableau 5.4 renforcent les observations précédentes (effets moyens) sur la hiérarchisation de l'influence des paramètres. Dans le visible (jusqu'à 38 %) mais beaucoup moins dans le proche infra-rouge (~ 4 %), c'est principalement l'absorption et la diffusion dues aux particules (AOTp) qui déterminent la variation de la réflectance. Les contributions des paramètres caractéristiques de la composition de l'atmosphère (UH₂O, UO₃) et de la surface urbaine (ρ_{sol}) sont assez similaires et plutôt faibles.

D'après ces simulations numériques, les variations de concentrations en PM sont visibles par le biais des variations de réflectances des images du satellite Landsat-TM.

4 Etude de la sensibilité des réflectances aux PM par simulations numériques

4.1. Détermination de la sensibilité spectrale du capteur

Les simulations numériques ont montré que les variations de concentrations en particules contribuaient majoritairement aux variations de réflectances. Comme un capteur satellite possède plusieurs bandes spectrales (ici le capteur TM de Landsat en possède 7), pour ne pas traiter toutes les bandes spectrales, je vais chercher à déterminer la bande spectrale la plus sensible aux variations de concentrations en particules. En fait, je vais chercher le canal le plus adéquat à la détection des particules. Je vais procéder de la façon suivante.

La réflectance mesurée par le capteur dépend des cinq paramètres que j'ai choisi : $\rho = f(U_{H_2O}, U_{O_3}, AOTp, OTatm, \rho_{sol})$. Je peux écrire :

$$d\rho = \frac{\partial f}{\partial U_{H_2O}} dU_{H_2O} + \frac{\partial f}{\partial U_{O_3}} dU_{O_3} + \frac{\partial f}{\partial U_{AOTp}} dAOTp + \frac{\partial f}{\partial U_{OTatm}} dOTatm + \frac{\partial f}{\partial U_{\rho_{sol}}} d\rho_{sol} \quad (14)$$

Je cherche la plus grande variation de réflectance, $d\rho$, pour le canal considéré dès que $\frac{\partial f}{\partial U_{AOTp}}$ est

grand et que $\frac{\partial f}{\partial U_{H_2O}}$, $\frac{\partial f}{\partial U_{O_3}}$, $\frac{\partial f}{\partial U_{OTatm}}$, $\frac{\partial f}{\partial U_{\rho_{sol}}}$ sont petits. En fait, je cherche la bande spectrale

pour laquelle il y a la plus grande variation de réflectance et pour laquelle la variation de réflectance est principalement due aux variations de l'épaisseur optique en aérosols. Cette condition élimine de fait les canaux TM5 et TM7. Ensuite, sur la figure 5.6, je vais chercher le canal pour lequel il y a le plus grand écart entre la réflectance minimale (courbe rose) et la réflectance maximale (courbe bleue). Plus l'écart sera grand, plus il sera facile de détecter les variations de réflectances et donc les variations de pollution, sur une image de ce canal.

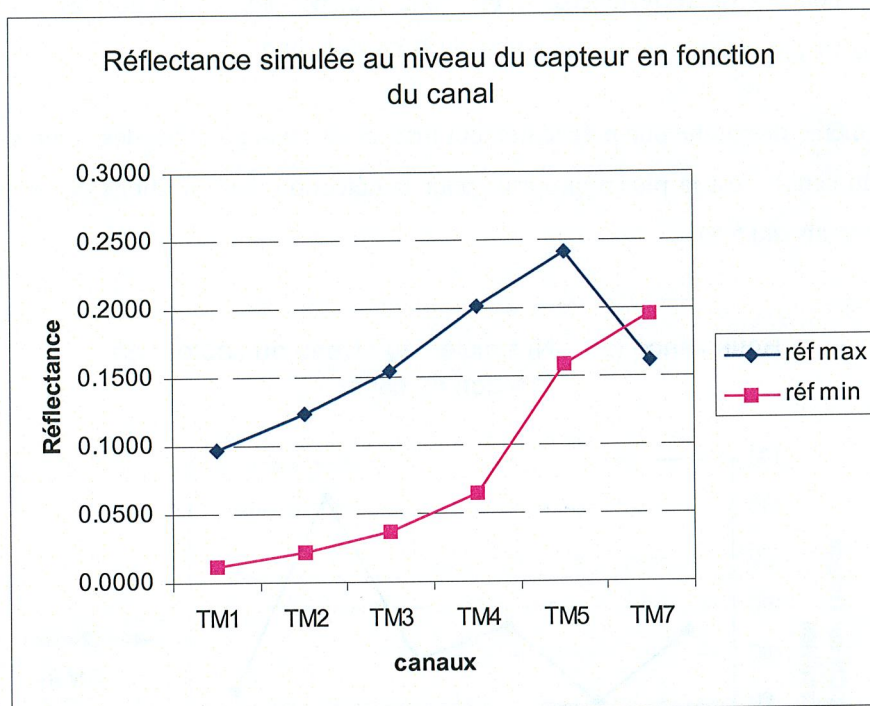


Fig. 5.6 : réflectances minimale et maximale simulées en fonction des canaux de Landsat-TM.

	TM1	TM2	TM3	TM4	TM5	TM7
réflectance	0.0858	0.1015	0.1169	0.1358	0.0821	-0.0331
CN	79	50	71	54	52	-10

Tab. 5.5 : valeurs des écarts entre les réflectances minimales et maximales et entre les valeurs de comptes numériques correspondants (CN).

Remarque : le modèle 6S fournit en sortie des valeurs normalisées de réflectance.

D'après la figure 5.6 et le tableau 5.5 résultant des simulations, le canal 4 du capteur Landsat-TM est le canal le plus sensible aux variations de concentrations en PM. Ce canal correspond aux longueurs d'ondes comprises entre 730 et 900 nm. Le capteur TM de Landsat a, d'après le modèle, une sensibilité spectrale aux aérosols plus importante dans cette gamme de longueur d'ondes.

Cette analyse de sensibilité doit servir à déterminer le canal adéquat pour la détection de particules par satellite. Sur l'image de ce canal, il doit être possible de distinguer les variations de réflectances i.e. les variations de concentrations en PM. Cependant sur une image satellite, les réflectances sont codées ; elles s'expriment en valeur de comptes numériques ou niveaux de gris. Ce sont ces valeurs que l'on « lit » sur une image satellite. De plus, les comptes numériques (CN) ne sont pas issus directement des luminances reçues par le capteur du satellite. Le gain du capteur peut modifier sensiblement la position de l'histogramme. Le gain du capteur est fonction du canal et de la date de prise de vue. Les positions respectives des histogrammes des divers canaux ne sont pas forcément celles qu'on pourrait déduire

simplement des courbes de réflectances. Il est donc indispensable d'étudier les images satellites directement à partir des valeurs numériques (Girard & Girard 1997).

En adoptant la même démarche que précédemment mais cette fois-ci à partir des comptes numériques, l'image issue du canal 1 est la plus appropriée pour la détection des variations de concentrations en PM (figure 5.7 et tableau 5.5).

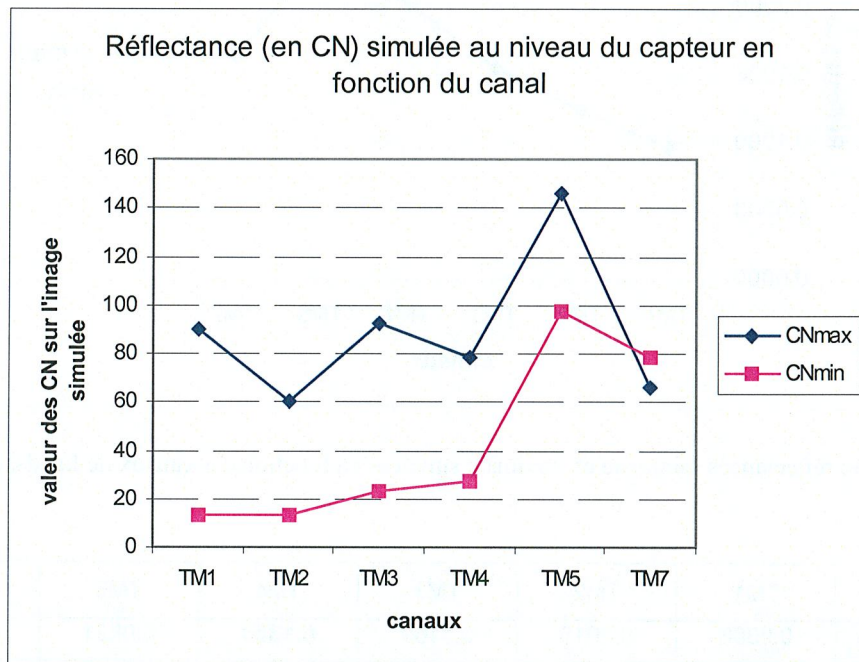


Fig. 5.7 : réflectances minimales et maximales exprimées en comptes numériques (CN), simulées pour chaque canal de Landsat-TM.

D'après le tableau 5.5, il faut 79 comptes numériques pour décrire la dynamique du polluant dans le canal TM1 et 54 dans TM4. Comme la gamme de variation des épaisseurs optiques des aérosols qui a servi aux simulations correspond à une gamme de concentrations en PM allant de 20 à 140 $\mu\text{g}/\text{m}^3$, on dispose alors de 79 valeurs pour décrire une variation de 120 $\mu\text{g}/\text{m}^3$. Pour une variation réaliste de concentration en PM, je vais pouvoir discriminer un nombre suffisant de concentrations différentes avec TM1.

4.2. Quantification de la sensibilité

La sensibilité du canal TM1 de Landsat dépend de la date de prise de vue du satellite (figure 5.8). Six dates de prises de vues ont été utilisées au lieu de huit car pour les deux dates manquantes, les données issues des stations de mesures au sol sont indisponibles. Pour que la comparaison entre les simulations et les observations soit cohérente et que la détection des concentrations en PM soit suffisamment précise, j'ai déterminé la sensibilité du canal TM1 en termes de PM pour chacune des six dates.

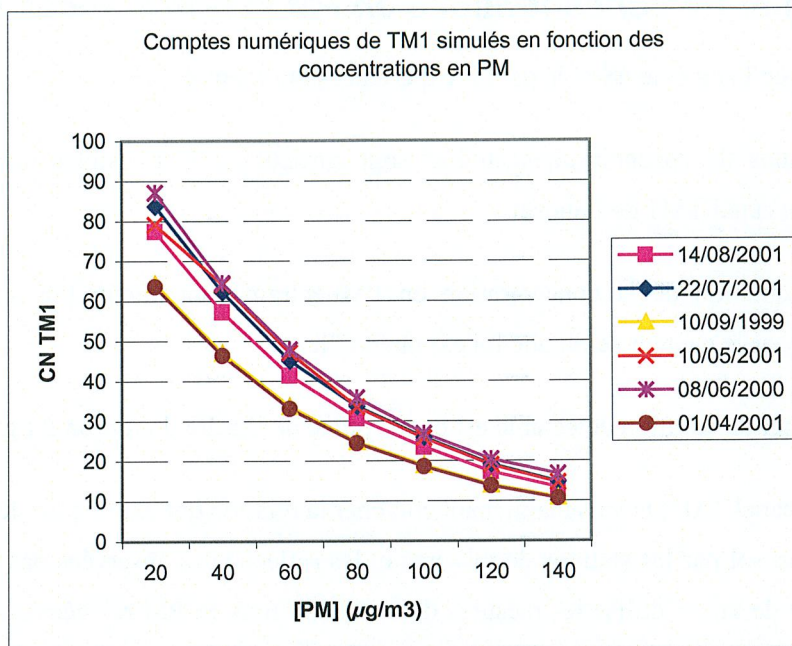


Fig 5.8 : réflectances de TM1 simulées, traduites en comptes numériques, pour chaque dates et en fonction de la concentration en PM.

La figure 5.8 montre que le canal TM1 réagit différemment selon la date mais que l'influence de la concentration en PM sur la réflectance garde la même allure. En effet, quand la concentration en PM augmente, la réflectance diminue selon une exponentielle décroissante. Cette observation est conforme à la théorie explicitée dans la partie 2.2 de ce chapitre, selon les équations (5) à (8).

Pour chaque date, j'ai calculé la sensibilité du canal TM1 (tableau 5.6) c'est-à-dire la plus petite valeur de concentration que je vais pouvoir distinguer sur l'image de ce canal.

dates	14/08/2001	22/07/2001	10/09/1999	10/05/2001	08/06/2000	01/04/2001
sensibilité ($\mu\text{g}/\text{m}^3$) : 1CN \Leftrightarrow	1.9	1.7	2.2	1.9	1.7	2.3

Tab. 5.6 : sensibilité du canal TM1 aux PM pour chaque date.

La sensibilité du canal TM1 est en moyenne de $1,9 \mu\text{g}/\text{m}^3$ sur l'ensemble des dates. Par exemple, une variation classique des niveaux de pollution en particules de 15 à $40 \mu\text{g}/\text{m}^3$ va entraîner une variation d'environ 13 CN sur une image du canal 1. Un tel changement dans la radiométrie de l'image est facilement perceptible.

Finalement, d'après les simulations numériques, il est possible de détecter la pollution en PM sur une image satellite. A partir de ce constat, je vais utiliser les images satellitaires du canal TM1, pour chercher une relation entre les concentrations en PM mesurées au sol et les réflectances observées par Landsat.

5 Relation entre concentrations en PM et réflectances observées

Les simulations avec le modèle 6S et le plan d'expériences ont montré que :

- les variations de concentrations en PM sont particulièrement visibles pour les longueurs d'ondes du canal TM1 de Landsat,
- une augmentation de la concentration en PM entraîne une diminution de la réflectance observées par le capteur selon une loi exponentielle,
- et la plus petite variation détectable est en moyenne de l'ordre d'environ $2 \mu\text{g}/\text{m}^3$.

Sur une image du canal TM1, je vais maintenant chercher la relation qui existe entre les concentrations en PM mesurées au sol par les stations de mesures et les réflectances observées par le satellite. Pour six dates de prises de vue, j'utilise les mesures de concentrations en PM relevées au niveau des cinq stations au sol et les réflectances issues de l'image du canal TM1. En traçant, pour chaque date de prise de vue, les mesures de réflectances (en CN) en fonction des concentrations en PM (figure 5.9), j'obtiens une répartition des points que je peux modéliser par une exponentielle dont l'équation dépend des dates de prise de vue (tableau 5.7).

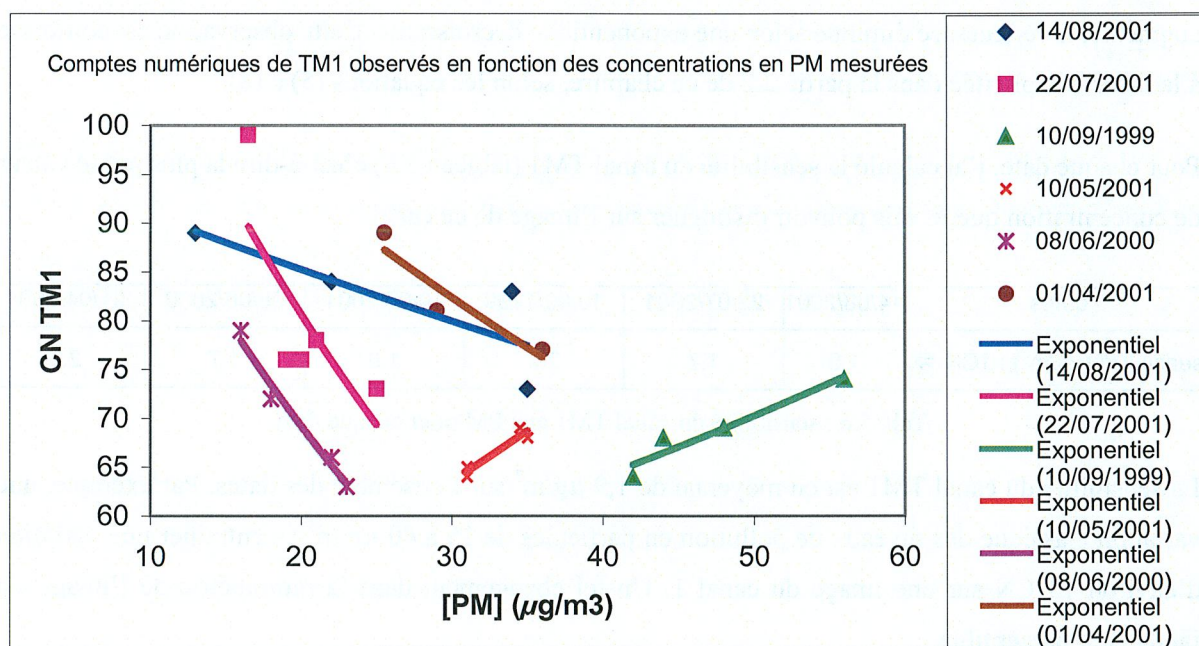


Fig. 5.9 : comptes numériques relevés sur l'image de TM1 tracés en fonction des concentrations en PM mesurées au niveau des stations de mesures au sol, pour les six dates de prise de vue. Les courbes de couleur correspondent aux courbes de tendance exponentielles.

dates	Observations			Simulations
	équation de la courbe de tendance	Corrélation R ²	Sensibilité 1CN ⇔	Sensibilité 1CN ⇔
14/08/2001	y = 96.881e-0.0064x	0.64	2.63	1.87
22/07/2001	y = 147.81e-0.0303x	0.60	1.53	1.74
10/09/1999	y = 44.324e0.0092x	0.91	/	2.24
10/05/2001	y = 39.738e0.0156x	0.90	/	1.86
08/06/2000	y = 125.75e-0.0299x	0.97	1.78	1.7
01/04/2001	y = 121.12e-0.0129x	0.87	1.63	2.26

Tab. 5.7 : modèles empiriques déduits des courbes précédentes, leur corrélation avec les points de mesure et leur sensibilité pour les différentes dates de prise de vue.

J'observe que j'ai bien une décroissance des valeurs de réflectances (en CN) quand la concentration en PM augmente, ce qui est conforme aux résultats des simulations, pour quatre dates sur six. Cette décroissance peut être approchée assez justement (coefficient de corrélation moyen : $\bar{R}^2 > 0,8$) par une exponentielle. Je peux alors déduire, de ces courbes, des modèles empiriques avec lesquels je peux estimer la concentration à partir des CN, pour chaque date. Avec ces modèles exponentiels, je confirme, à $\pm 20\%$ en moyenne, pour quatre dates, la sensibilité déterminée avec les simulations (voir les deux dernières colonnes du tableau 5.7). Pour ces quatre dates, j'ai un modèle empirique avec lequel je suis capable de prédire la concentration en PM à partir d'une valeur de compte numérique d'un pixel d'une image du canal TM1 de Landsat.

On observe des différences de comportement entre les dates, notamment au niveau de la décroissance des courbes. Je les explique de la façon suivante. Il est important de signaler, avant d'entamer l'explication, que les stations de mesure au sol sont toutes placées sur des surfaces de même type. L'albédo de surface de ces stations ne joue donc aucun rôle dans ces différences de comportement.

L'épaisseur optique des aérosols, paramètre d'entrée de la concentration en PM dans le modèle 6S (AOTp), est l'intégration du coefficient d'extinction sur toute la colonne atmosphérique. Dans le modèle 6S, les valeurs de ce coefficient suivent un profil vertical connu (Mc Clatchey *et al.* 1971). Par translation, on peut dire que les valeurs de concentrations en PM suivent ce profil vertical, dans les simulations. Cependant, les observations par les stations de mesures se font au niveau du sol. En fait, dans la réalité, la répartition du polluant ne suit pas forcément ce profil. La concentration mesurée au sol, [PM], n'est pas égale à la concentration globale de ce polluant dans l'atmosphère, [PM_{globale}]. Par contre, le capteur du satellite mesure la concentration globale.

- Par conséquent, quand il y a une décroissance des valeurs de CN en fonction de [PM], c'est que l'on se trouve dans la configuration du modèle 6S pour toutes les stations. [PM] mesurée au sol correspond à une partie de la concentration totale de PM dans l'atmosphère qui suit une

répartition verticale comme celle du modèle : $[PM] < [PM_{\text{globale}}]$ pour toutes les stations. A la station 1 : $[PM]_1 < [PM_{\text{globale}}]_1$, à la station 2 : $[PM]_2 < [PM_{\text{globale}}]_2$ avec $[PM]_1 < [PM]_2$ et $[PM_{\text{globale}}]_1 < [PM_{\text{globale}}]_2$. Ceci entraîne que l'on observe bien $CN_1 > CN_2$ quand $[PM]_1 < [PM]_2$. Et ainsi de suite pour les autres stations.

- Quand il y a une croissance des valeurs de CN en fonction de [PM], c'est que l'on ne se trouve pas dans la configuration du modèle 6S pour toutes les stations. A la station 1, $[PM]_1$ mesurée au sol correspond à une partie de la concentration totale de PM dans l'atmosphère qui suit une répartition verticale comme celle du modèle : $[PM]_1 < [PM_{\text{globale}}]_1$. Par contre, à la station 2, la répartition de la pollution ne suit pas le profil type et $[PM]_2 \sim [PM_{\text{globale}}]_2$. Toute la pollution est concentrée au niveau du sol, il n'y a pas de polluants en altitude. On a toujours $[PM]_1 < [PM]_2$ mais $[PM_{\text{globale}}]_1 > [PM_{\text{globale}}]_2$: la concentration globale dans l'atmosphère a diminué d'une station à l'autre. Ceci entraîne que l'on observe $CN_1 < CN_2$ quand $[PM]_1 < [PM]_2$.

En fait, dans le cas le plus fréquent, les variations de la concentration en PM mesurée au sol sont corrélées positivement avec celles de la concentration globale en PM dans l'atmosphère. Dans le cas le moins fréquent, les variations de la concentration mesurée au sol sont corrélées négativement avec la concentration globale d'aérosols dans l'atmosphère.

Pour expliquer les différences de comportement des courbes de la figure 5.9, j'ai cherché d'autres explications liées à la taille des particules ou à leur composition. En effet, comme le diamètre des particules a un impact sur le type de diffusion et sur sa configuration et comme la composition des particules influe sur leur pouvoir absorbant ou diffusant (partie 2.1 de ce chapitre), j'ai envisagé une explication de ce type. Cependant, il semble que l'explication la plus vraisemblable soit celle que j'ai proposée sur l'importance de la distribution verticale des aérosols de pollution (Mallet *et al.* 2003).

Il est important de mentionner ici que ces explications sont à mettre en coïncidence avec le nombre relativement faible de points expérimentaux, ce qui conduit à des résultats à interpréter avec précaution.

A partir des modèles empiriques que j'ai déterminés pour chaque date, je vais chercher à valider le concept de stations virtuelles. En particulier je vais chercher à confirmer l'hypothèse selon laquelle la concentration en PM aux stations virtuelles est la même qu'à la station réelle.

6 Validation des stations virtuelles

Pour valider le concept de stations virtuelles, la démarche est celle expliquée à la fin du chapitre 4. Valider les stations virtuelles signifie confirmer l'hypothèse qui sous-tend la méthode des stations virtuelles. Il s'agit de prouver que la concentration en PM à la station virtuelle est la même qu'à la

station réelle associée. La méthode de détermination des stations virtuelles étant indépendante du temps, la validation des stations virtuelles ne dépend pas de la date considérée.

J'ai relevé les valeurs de réflectances (en CN) correspondant aux stations virtuelles sur les images du canal TM1 pour chaque date de prise de vue. Pour relever les CN sur les images du canal TM1, j'ai utilisé le SIG pour extraire de façon automatique les valeurs de CN au niveau des stations virtuelles et réelles. Les images Landsat du canal TM1 de résolution 30 m ont été recalées par Ung (2003) à partir d'une image SPOT de résolution 10 m. Les images dont je dispose, ont alors une résolution de 10 m. Ung explique que l'erreur sur le recalage est de l'ordre de 30 m. Quand le SIG extrait les valeurs de CN, il fait une moyenne des CN sur un voisinage de 3 x 3 pixels soit sur une zone de 30 m x 30 m, ce qui permet ainsi de prendre en compte l'erreur sur le recalage. Par ailleurs, les auteurs qui cherchent à extraire de l'information de pollution à partir des pixels de l'image, prennent en général un voisinage de cette taille. Le choix de ce nombre de pixels est un compromis entre le besoin d'avoir un nombre suffisamment grand pour améliorer le sens statistique du résultat et de prendre en compte la variabilité spatiale et naturelle des données (Hashim *et al.* 2004).

L'histogramme des valeurs de réflectances des stations virtuelles associées à une station réelle montre une répartition telle que 90 % des valeurs de réflectances sont comprises entre $\pm 20\%$ de la valeur de la réflectance de la station réelle correspondante. Autrement dit, pour une station réelle, il y a 90 % des stations virtuelles qui ont une réflectance proche à $\pm 20\%$ de celle de leur station réelle. La figure 5.10 donne un exemple d'histogramme obtenu pour une station en particulier.

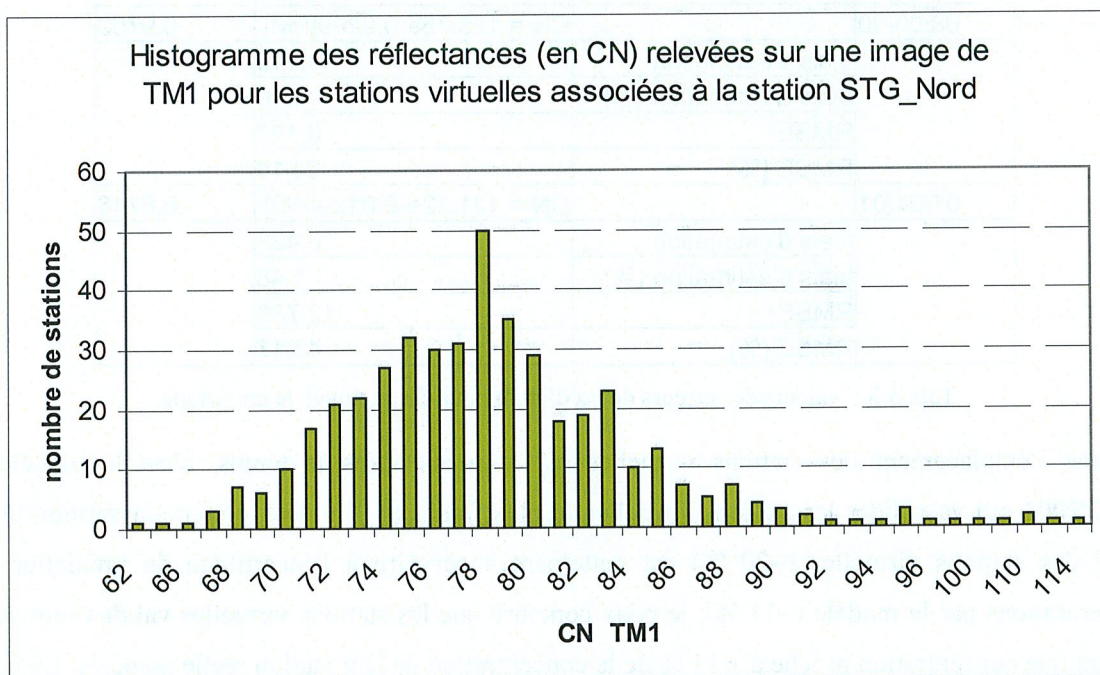


Fig. 5.10 : histogramme des valeurs de CN d'une image de TM1 correspondant aux stations virtuelles associées à la station réelle STG-Nord.

La répartition des CN relevés au niveau des stations virtuelles associés à une station réelle montre une allure de gaussienne centrée sur la valeur du CN de la station réelle. L'application du modèle empirique pour chaque date permet le passage de la réflectance en comptes numériques aux concentrations en particules. La relation qui lie ces deux grandeurs est exponentielle. J'observe alors une répartition des valeurs de concentrations en PM des stations virtuelles quasiment identique à celle des comptes numériques. En fait, il y a une légère dissymétrie de l'histogramme du fait de la non linéarité de la relation mais comme la gamme de valeur est peu large, l'allure de histogramme est peu affectée. Finalement, j'en déduis que la répartition des valeurs de concentrations suit une loi gaussienne centrée sur la valeur de concentration à la station réelle.

Le modèle empirique est plus ou moins précis selon la date considérée (Tableau 5.8). L'incertitude de prédiction la plus faible correspond à la date du 22/07/2001.

date	statistiques	Modèle	R ²
14/08/01		CN = 96.881e-0.0064[PM]	0.6443
	biais d'estimation	-0.016	
	biais d'estimation (%)	-0.06	
	RMSE	6.033	
	RMSE (%)	23.20	
22/07/01		CN = 147.81e-0.0303[PM]	0.5969
	biais d'estimation	0.003	
	biais d'estimation (%)	0.01	
	RMSE	2.289	
	RMSE (%)	11.28	
08/06/00		CN = 125.75e-0.0299[PM]	0.9752
	biais d'estimation	-3.116	
	biais d'estimation (%)	-14.16	
	RMSE	6.182	
	RMSE (%)	28.10	
01/04/01		CN = 121.12e-0.0129[PM]	0.8746
	biais d'estimation	0.438	
	biais d'estimation (%)	1.48	
	RMSE	12.789	
	RMSE (%)	43.17	

Tab. 5.8 : valeurs des erreurs de prédiction pour chaque modèle empirique.

Comme l'emplacement des stations virtuelles est le même dans le temps, c'est le modèle du 22/07/2001 qui va valider les stations virtuelles le plus finement. En effet, comme la variabilité des [PM] des stations virtuelles (~20 %) est nettement supérieure à l'incertitude de prédiction des concentrations par le modèle (~11 %), je peux conclure que les stations virtuelles valides sont celles qui ont une concentration proche à ± 11 % de la concentration de leur station réelle associée. La figure 5.11 illustre ceci. L'incertitude de prédiction que je fais avec le modèle est représentée par la gaussienne en bleue. La répartition des valeurs de concentrations des stations virtuelles suit une loi

gaussienne (en rouge). Les stations virtuelles sont validées à l'intérieur du volume délimité par les deux courbes.

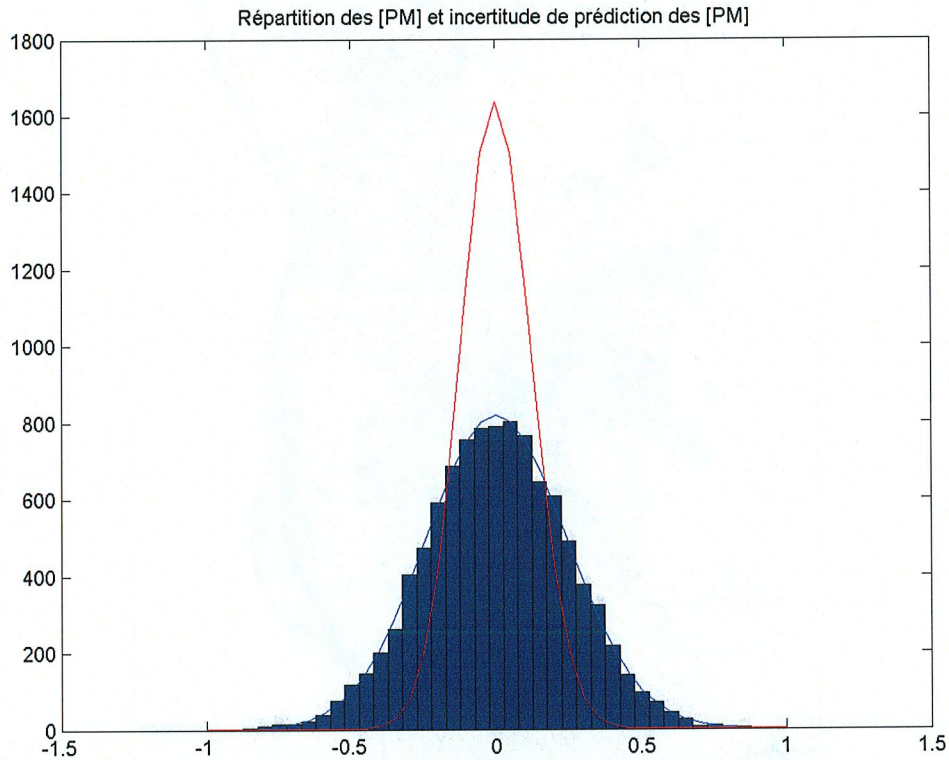


Fig. 5.11 : répartition des concentrations en PM des stations virtuelles (histogramme et gaussienne bleue) et incertitude de prédiction des concentrations en PM (gaussienne rouge).

Avec cette méthode de validation, 71 % des stations virtuelles déterminées avec la méthode de densification virtuelle du réseau sont validées. Sur 635 stations, 182 sont éliminées. La figure 5.12 montre la nouvelle répartition des stations virtuelles sur la ville de Strasbourg. Cette figure est à mettre en regard de la figure 4 de l'article du chapitre 4.

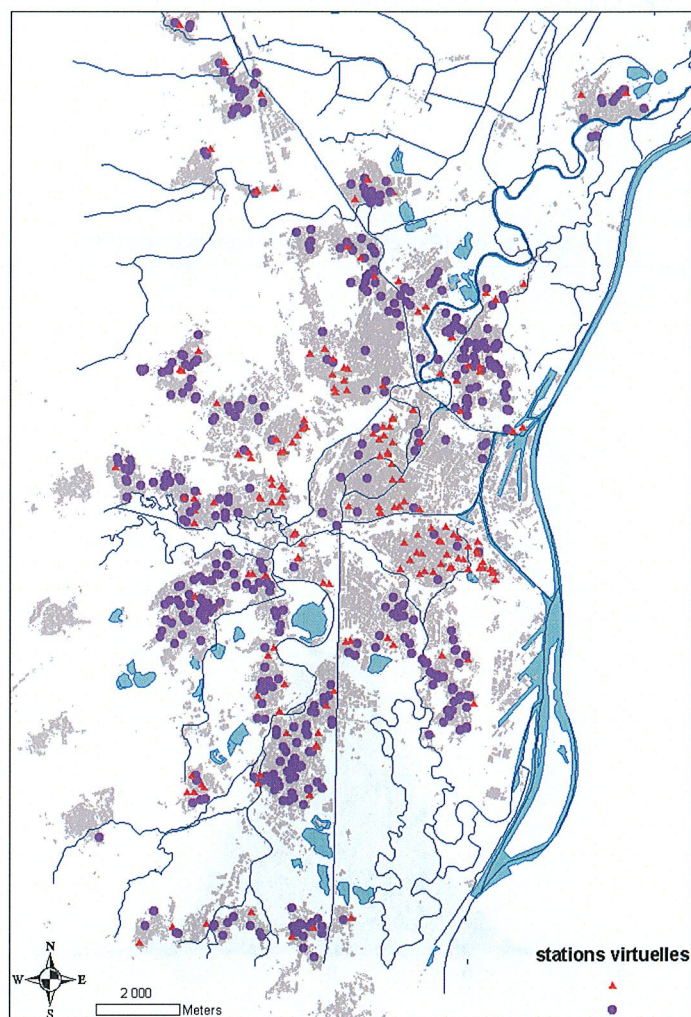


Fig. 5.12 : répartition des stations virtuelles sur la ville de Strasbourg après validation. Les points mauves sont les stations virtuelles validées par l'imagerie satellitaire. Les triangles rouges sont les stations non validées par cette méthode qu'il faut éliminer.

435 stations virtuelles, soit 71 % des stations déterminées par la méthode de densification du réseau de mesures proposée au chapitre 4, ont été validées grâce à l'imagerie satellitaire.

D'après la figure 5.12, il semble qu'il soit plus difficile d'estimer, de façon précise, la concentration en PM en centre ville avec l'imagerie satellitaire. En centre ville, les sources de pollution se concentrent et la variabilité spatiale de la pollution est très locale, le gradient peut être très grand sur des zones très peu étendues. Peut-être peut-on voir ici une limite de l'imagerie satellitaire en tout cas, du capteur TM de Landsat pour la validation des stations virtuelles et donc pour l'estimation des concentrations ? La résolution de 30 m apparaît comme encore trop grossière pour détecter finement la pollution. En effet, même si cette méthode ne valide pas ces stations virtuelles, il est possible que localement, la station réelle et la station virtuelle aient la même concentration.

7 Conclusion

Cette étude préalable sur les potentialités de l'imagerie satellitaire comme outil de détection des particules polluantes à l'échelle locale répond à son premier objectif. Elle confirme les travaux antérieurs sur ce sujet et réaffirme donc qu'il est possible de détecter la pollution particulaire par satellite. Elle montre que le capteur TM de Landsat peut servir à la détection de cette pollution et en particulier, que cette détection est facilitée si on emploie le canal TM1. Les simulations avec le modèle 6S combiné à une approche par plans d'expériences, ont montré que les aérosols urbains affectent particulièrement les longueurs d'ondes autour de 815 nm. Le canal TM4 (730 – 900 nm) est le canal le plus sensible aux variations de concentrations en particules. Néanmoins, quand le gain du capteur est pris en compte, le canal TM1 est le canal adéquat pour détecter les variations de concentrations en particules sur l'image. De plus, ce canal est suffisamment sensible pour qu'une variation moyenne d'environ $2 \mu\text{g}/\text{m}^3$ soit discernable sur une image de ce canal. Les simulations ont également montré qu'une augmentation de la concentration en PM entraîne une diminution de la réflectance mesurée par le capteur, selon une loi exponentielle.

Les observations ont confirmé les simulations et ont montré que, malgré une relation a priori complexe entre la concentration en PM et la réflectance, il est possible dans certains cas de modéliser cette relation par une exponentielle décroissante avec une assez bonne justesse ($\bar{R}^2 > 0,8$) et de confirmer la valeur de la sensibilité proche de $2 \mu\text{g}/\text{m}^3$. Cette cohérence entre les simulations et les observations constitue déjà un résultat réellement encourageant et confirme l'importance d'étudier plus en profondeur encore les apports possibles de la télédétection spatiale.

Cette étude répond également à son second et principal objectif qui est la validation du concept de « station virtuelle » mis en avant dans le chapitre 4. A partir des modèles empiriques déduits des observations, j'ai pu estimer les valeurs de concentrations au niveau des stations virtuelles et valider 71 % des stations virtuelles déterminées par la méthode décrite au chapitre 4. 453 stations virtuelles ont bien la même concentration en PM que leur station réelle associée à environ 10 % près. Cette méthode de validation montre qu'il est utile d'employer l'imagerie satellitaire mais que toutefois, celle-ci présente des limites. La répartition de l'ensemble des stations virtuelles non validées par cette méthode (fig. 5.12) révèle que la méthode est moins fiable en centre ville, c'est-à-dire pour des zones où la variabilité spatiale de la pollution particulaire est extrêmement locale. Ce constat implique que, au delà du fait d'employer des capteurs satellitaires avec une meilleure résolution spatiale pour améliorer la validation, recourir à une campagne de mesures au sol peut constituer une solution à ce problème.

La réponse apportée par cette étude au second objectif reste partielle et nécessite d'être complétée. En effet, l'erreur (RMSE) relative de 11 % n'est obtenue que pour une date de prise de vue. Sans l'image

prise à cette date, et en utilisant l'un des trois autres modèles obtenus à partir des images prises aux autres dates, on déduirait que 90 % des stations virtuelles ont la même concentration que leur station réelle associée à 20 % près (fig. 5.11) et que l'imagerie satellitaire ne permet pas une meilleure précision dans cette validation. Ce constat suscite alors de nombreuses questions : est-ce que cette image du 22 juillet 2001 est un cas isolé ? Existe-t-il d'autres images donnant des modèles avec une RMSE encore plus faible ? Est-ce qu'une image donnant un modèle avec une RMSE équivalente donnerait les mêmes stations virtuelles validées ? Autant de questions auxquelles il faut répondre avant d'accorder un crédit opérationnel à cette méthode de cartographie.

Une perspective à court terme de ce travail serait d'utiliser ces 453 stations virtuelles pour réaliser une interpolation, comme cela a été fait dans le chapitre 4, pour vérifier qualitativement et quantitativement dans quelle mesure la cartographie s'en trouve améliorée. Les résultats permettraient d'appuyer ou non le choix de cette méthode de validation par imagerie satellitaire.

Le même travail sur le capteur MeRIS d'Envisat n'a pas été effectué dans cette thèse. Il serait cependant intéressant de vérifier si l'on retrouve les mêmes résultats et d'étudier l'influence de la résolution spatiale (300 m) et de la largeur spectrale (Tab. 5.1) sur la détection des aérosols urbains.

Bibliographie

- Bacour C., 2001. Contribution à la détermination des paramètres biophysiques des couverts végétaux par inversion de modèles de réflectance : analyses de sensibilité comparatives et configurations optimales. Thèse de Doctorat "Méthodes Physiques en Télédétection", Université Paris 7, 228 p.
- Bellouin C. Boucher O., Tanré D., Dubovic O., 2003. Aerosol absorption over the clear-sky oceans, deduced from POLDER-1 and AERONET observations. *Geophysical Research Letters*, 30(14), 1748.
- Benoist D., Tourbier Y., Germain-Tourbier S., 1994. *Plans d'expériences : construction et analyse*. coll. Tec & Doc, Lavoisier, Paris, 693 p.
- Dechambre M., le Gac C., 2001. Statistical sensitivity analysis of a vegetation canopy microwave scattering model based on an experiment design method. In: Proceedings of the Geoscience and Remote Sensing symposium, IGARSS 2001, vol. 3, 1070-1072, 10.1109/IGARSS.2001.976749.
- Girard M.C., Girard C.M., 1999. *Traitement des données de télédétection*. Edition Dunod, Paris, France, 529 p.
- Hashim M., Kanniah K. D., Ahmad A., Rasib A. W., Ibrahim A. L., 2004. The use of AVHRR data to determine the concentration of visible and invisible tropospheric pollutants originating from a 1997 forest fire in Southeast Asia. *International Journal of Remote Sensing*, 25 (21), 4781-4794.
- Hewitt C.N., Jackson A., 2003. *Handbook of Atmospheric Science. Principles and Applications*. Blackwell Publishing, Oxford, UK, p. 244, 633 p.
- Kacenenbogen M., 2004. Application de la télédétection spatiale pour la surveillance de la pollution de l'atmosphère en aérosols. Rapport de DEA Méthodes physiques en télédétection, Université Pierre et Marie Curie (Paris VI), Laboratoire d'Optique Atmosphérique, UMR CNRS 8518, Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, 31.
- Lachérade S., Miesch C., Boldo D., Briottet X., Le Men H., Valorge C., 2006. An inverse radiative transfer model to extract ground spectral reflectance of urban areas. 1st workshop of the EARSeI Special Interest Group Urban Remote Sensing : Urban remote sensing, Challenges & Solutions, 2-3 March, Berlin Adlershof, Germany.
- Liou K.N., 1980. *An introduction to Atmospheric Radiation*. Academic Press, 392 p.

Mc Clatchey R. A., Fenn R. W., Selby J. E. A. , Volz F. E., Garing J. S., 1971. Optical properties of the atmosphere (revised), AFCRL 71-0279, *Environmental Research*, Paper 354, Bedford, Massachusetts, U.S.A., 1971.

Mallet, M., Roger, J.C., Despiaud, S., Dubovik, P., Putaud, J.P, 2003. Microphysical and optical properties of aerosol particles in urban zone during ESCOMPTE. *Atmospheric Research*, 69, 73-97.

Pancrati O., 2003. Télédétection de l'aérosol désertique depuis le sol par radiométrie infrarouge thermique multibande. Thèse de doctorat Lasers, molécules, rayonnement atmosphérique, Université des Sciences et Technologies de Lille, France, 200 p.

Retalis, A., Sifakis, N., Grosso, N., Paronis, D. and Sarigiannis, D., 2003. Aerosol optical thickness retrieval from AVHRR images over the Athens urban area. In: Proceedings of IGARSS 2003. http://sat2.space.noa.gr/rsensing/documents/IGARSS2003_AVHRR_Retalisetal_web.pdf

Sifakis N., Bildgen P., Gilg J.P., 1992. Utilisation du canal 6 (thermique) de Themactic Mapper pour la localisation de nuages de pollution atmosphérique. Application à la région d'Athènes (Grèce). *Pollution Atmosphérique*, Vol. 34, 96-107.

Sifakis N., 1995. La télédétection des voiles de pollution atmosphérique et de la dégradation de l'environnement dans la région d'Athènes. *Photo-Interprétation*, Vol. 4, 220-225.

Sifakis N., Soulakellis N. and Paronis D., 1999. Quantitative mapping of air pollution density using Earth observations: A new processing method and application on an urban area. *International Journal of Remote Sensing*, 19 (17), 3289-3300.

Thieuleux F., Moulin C., Breon F.M., Maignan F., Poitou J., Tanré D., 2004. Remote sensing of aerosols over the oceans using MSG/SEVIRI Imagery. Instrumental Notes of IPLS n° 51, novembre 2004, ISSN 1626-8334.

Vermote E.D., Tanré D., Deuzé J.L., Herman M., Morcette J.J., 1997. Second simulation of the satellite signal in the solar spectrum. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 35, 675-686.

Wald L., Baleynaud J.M., 1999. Observing air quality over the city of Nantes by means of Landsat thermal infrared data. *International Journal of Remote Sensing*, 20 (5), 947-959.

Wald L., 2006. Solar radiation energy (fundamentals and theory). UNESCO

Chapitre 6

Conclusion générale et perspectives

Cette partie résume la contribution de cette thèse à la représentation cartographique de la pollution particulaire en ville. Les perspectives offertes par ce travail sont ensuite discutées.

L'objectif majeur de cette thèse était de contribuer à l'établissement d'une stratégie de cartographie complète et précise de la pollution atmosphérique sur l'ensemble d'une agglomération à l'échelle de la rue (100 m). Plus précisément, j'avais pour objectif de mieux définir mathématiquement les méthodes alternatives proposées dans les travaux antérieurs, de cerner leur portée, leur applicabilité et en particulier leurs échelles spatiales et temporelles. Les méthodes alternatives incluent les deux méthodes de cartographie (champs typiques et densification du réseau de mesures) et l'étude de l'apport de l'imagerie satellitaire pour la détection de polluants et avant tout pour la validation des stations virtuelles.

Les conclusions de ce travail peuvent être synthétisées en trois parties.

Le peu de données disponibles aux échelles locales m'a incité à approfondir l'étude sur l'apport possible de la fusion de données pour le domaine de la qualité de l'air. La première partie a donc consisté à approfondir la méthode des « champs typiques ». Pour cela, je lui ai donné un formalisme mathématique et j'ai examiné sa faisabilité à travers une mise en œuvre pratique. L'emploi de la fusion de données s'avère intéressant pour l'exploitation de données à des échelles spatiales différentes. Cependant, un problème majeur lié à l'insuffisance de données et à la non-gestion des échelles temporelles par la méthode proposée, a conduit à ne pas la recommander pour la cartographie. Des améliorations possibles pourraient être de transposer les données dans un même référentiel temporel en appliquant des profils de variations temporelles pour passer des grandes aux petites échelles et de pallier le manque de données en créant d'avantage de sectorisations à partir des données qualitatives fournies par les géographes.

Le deuxième groupe de conclusions concerne l'étude de l'apport de l'imagerie satellitaire haute et moyenne résolution pour la cartographie de la pollution sur une agglomération i.e. à des échelles très locales. L'absence de mesures m'a conduit à m'appuyer sur les principes de la physique de l'atmosphère et, conjointement, à utiliser l'imagerie satellitaire pour valider les stations virtuelles. L'utilisation combinée du modèle de transfert radiatif 6S et de la méthodologie des plans d'expériences numériques a permis, en un nombre limité de simulations, d'explorer rigoureusement l'influence relative de chaque variable atmosphérique et la sensibilité spectrale de chaque capteur. Cette analyse a permis de confirmer qu'il était possible de détecter la pollution particulaire par satellite, en particulier dans la bande spectrale TM1 (450 – 515 nm) du satellite Landsat-TM. Les résultats de simulation ont été confirmés par les observations qui font intervenir les mesures au sol des concentrations en particules et la radiométrie des pixels de l'image satellitaire. En outre, ces résultats sont en accord avec ceux trouvés dans la littérature.

Cette confrontation des résultats a abouti à la création d'un modèle empirique qui permet de convertir les comptes numériques de l'image en concentrations en PM. Cependant, la comparaison avec plusieurs dates de prises de vue a montré que le modèle empirique est variable. Un modèle universel n'existe pas encore.

La dernière partie de cette conclusion est dédiée au développement de la méthode de densification virtuelle du réseau de mesure, dont le but est d'améliorer la cartographie de la pollution par interpolation spatiale. Mon choix s'est porté sur cette méthode en particulier car il s'avère que les stations virtuelles peuvent contraindre à la fois les méthodes d'interpolation et les modèles de dispersion. Le problème du peu de mesures à interpoler engendrant une représentation non réaliste et des incertitudes importantes sur les estimations a été résolu. En effet, grâce à cette méthode, les cartes représentant le champ de pollution particulaire ont été améliorées du point de vue de la représentation cartographique et de l'estimation des concentrations, et ce, quelle que soit la méthode d'interpolation choisie. De plus, leur résolution spatiale de 10 m est suffisamment fine pour décrire des variations spatiales à l'échelle de la rue. Les modèles empiriques déterminés à l'aide des images satellitaires ont servi à la validation des stations virtuelles par comparaison des réflectances mesurées aux stations réelles et virtuelles. Soixante et onze pour cent des stations virtuelles déterminées par la méthode de densification du réseau de mesures ont été validées. Cependant, cette méthode de validation présente une limitation. Elle est moins fiable en centre ville du fait que la résolution spatiale du capteur ne soit pas assez fine.

Les perspectives de travail qui font suite à cette thèse sont nombreuses. Voici quelques propositions d'amélioration et d'approfondissement à très court et à plus long terme.

Dans une perspective d'amélioration de la méthode de détermination des stations virtuelles par analyse discriminante (chap. 4), les images issues de la bande spectrale sensible à la pollution peuvent devenir des éléments discriminants.

La résolution spatiale du capteur TM semble être une limitation à l'estimation des concentrations dans des zones urbaines très denses. Les solutions à ce problème peuvent être soit d'utiliser un capteur avec une meilleure résolution spatiale (SPOT : 10 m) soit de recourir à une campagne de mesures au sol pour la validation des stations virtuelles. La télédétection spatiale a montré son apport pour l'estimation des concentrations et pour la validation. Il me semble intéressant d'approfondir cette étude pour mieux cerner les limites et possibilités de la télédétection appliquée à la cartographie de la pollution atmosphérique. Il serait judicieux d'effectuer ce genre d'étude sur une zone pour laquelle on dispose de nombreuses données au sol. On pourrait ainsi savoir si les méthodes proposées ici sont transposables à une autre agglomération, à d'autres capteurs spatiaux, à d'autres polluants et

éventuellement les améliorer. Les mesures terrain provenant de la campagne ESCOMPTE sur l'agglomération Marseillaise peuvent constituer une bonne base de données pour cette étude.

A plus long terme, une perspective d'application de ces travaux de thèse est d'intégrer ces cartes à un système opérationnel d'estimation du bilan espace-temps d'un citoyen en ville. Par bilan espace-temps, on entend évaluation de la quantité de polluants auxquels est soumis un citoyen au cours de ses déplacements et de ses activités. Ce bilan espace-temps, fourni par les autorités locales, est, à terme, une nécessité réglementaire. Un citoyen pourra exiger de connaître son bilan espace-temps au cours de la journée c'est-à-dire la concentration cumulée de polluants auxquels il est exposé. Pour établir ce bilan, il faudra coupler à une cartographie de la qualité de l'air, des données sur la mobilité. Une fois la connaissance spatio-temporelle établie, l'information d'exposition pourra être diffusée au moyen d'un indice approprié permettant d'évaluer l'exposition du citoyen. L'emploi d'un indice peut constituer une façon simple de fournir une information claire et accessible à tous. Les autorités locales, aidées des agences de surveillance de la qualité de l'air, pourront alors envisager de diffuser cette information par le biais de la téléphonie mobile de troisième génération qui s'auto-localise et s'auto-informe périodiquement, ou encore en proposant un service web. Ce service web permettrait à l'utilisateur de calculer son itinéraire en ville et de lui indiquer les zones polluées le long de son itinéraire. J'ai construit une maquette du bilan espace-temps pour amorcer le travail sur ce système opérationnel. En utilisant le module de gestion de données temporelles Tracking Analyst® du logiciel SIG ArcGIS™ 8.3, j'ai obtenu une représentation dynamique du bilan espace-temps. L'article donné en annexe A donne un aperçu de ce travail.

Annexe A

Publications

A GEOGRAPHICAL TOOL FOR ASSESSING PERSONAL EXPOSURE TO PARTICULATE MATTER (PM)

Anne-Lise Beaulant*¹, Thierry Ranchin¹, Christiane Weber², Gilles Perron³, Joseph Kleinpeter³

¹Groupe Télédétection & Modélisation, Ecole des Mines de Paris, B.P 207, 06904 Sophia Antipolis Cedex, France. Tel :

+ 33 (0)4.93.95.74.52 Fax : + 33 (0)4.93.95.75.35

E-mail : anne-lise.beaulant@ensmp.fr

²Laboratoire Image et Ville, UMR 7011, Université Louis Pasteur, 3 rue de l'Argonne, 67000 Strasbourg, France.

³ASPA (Association pour la Surveillance et l'étude de la Pollution atmosphérique en Alsace), 5 rue de Madrid, 67300

Schiltigheim, France.

ABSTRACT :

The European policy in urban atmospheric pollution aims at reducing its impact on human health. This particular problem of public health is closely related to the daily exposure of the population. Exposure assessment should be done through analysis of individuals activities. Sooner or later the local authorities in charge of air quality will have to provide to the population their individual exposure.

We have developed a tool that supplies this information. It is based on a multi-source approach. It exploits a Geographic Information System (GIS) gathering information such as individuals mobility, the topographic database, and concentrations of pollutants. We have obtained maps and dynamic representations of individual exposure to PM in urban area.

1. INTRODUCTION

Atmospheric pollution is becoming a critical factor of anticipated deaths (Filleul *et al.*, 2001). Pollutants are the cause of pulmonary and even cardiovascular diseases: babies, asthmatic and old people are most affected. By atmospheric pollution, we mean ambient air quality as well as air quality in houses and places of work. Atmospheric pollution, like all other environmental issues, rests on a pragmatic and sensitive relationship between man and his surrounding environment (Roussel & Raimbaud, 2005). It exists many examples for which health risks due to outdoor pollution are proved. More than 30 million people have been exposed to phases of photochemical pollution in France, during the scorching summer of 2003 that spread all over the West and Central Europe. According to the French Agency of the Environment and the Control of Energy

(ADEME, 2004), health consequences, still badly explained, were without precedent. In the other hand health risks due to indoor pollution are not yet clearly described. But, first results of studies (Nerriere *et al.*, 2005), which aims at collecting data inside buildings, show very high levels of PM pollution inside. These alarming results let predict a more important health impact since it is underhand and unsuspected (Kotzias, 2005). In large European cities, the policy is turned towards the reduction of the negative impact on health. Hence for the well being of humans, it is therefore necessary to evaluate the actual human exposure to ambient pollution.

One way to perform it, is to aboard the problem in a static way. One can exploit data issued from monitoring network. Indeed, most of the large cities in Europe have acquired a monitoring network for air quality. A network is composed of static measuring stations. It allows continuous monitoring of atmospheric pollution at a station location. Pollution data is collected in near real time and already used to compute a global atmospheric pollution indicator, called ATMO in France or AQI in Germany. Creating ATMO index was initiated by ASPA (http://w3.atmo-alsace.net/rubrique/mesures/explication_indices.htm). Permanent (stations) or temporary (passive sensors) measures are also used to create pollution maps with interpolation methods or to constrain dispersion models. Modellings, maps and indicators aim at informing local authorities, as well as the population, of the air quality. Public authorities may take restrictive measures with car traffic in particular and with certain air polluting companies, when high levels are indicated, in order to minimize the impact on public health.

An other way to estimate pollution exposure is to consider the problem dynamically. In this case, people are wearing sensors. In litterature, many studies are dedicated to such measurements technique (Kaur *et al.*, 2005 ; Adams *et al.*, 2001 ; Pernelet & Roth, 2001 ; Nerriere *et al.*, 2005). The so-called "sentinels of the air" which was held from summer 2001 to winter 2002, (<http://www.appanpc-asso.org>) aim to better evaluating the individual exposure to air pollution through series of national measurements. Another program called Genotox'ER, to which ASPA participated, was carried out in four French metropolitan areas. It aimed at describing urban population exposure to PM_{2,5}; PM₁₀ and NO₂ or benzene. Measurements are taken thanks to mobile sensors carried by people throughout the day. Such campaign allows for a better understanding of the time of exposure spent in contact with high levels of pollution and it brings also out data from indoor and outdoor pollution. They take into account displacements of the inhabitants within the agglomerations, as well as the contribution of indoor pollution in selected confined spaces (homes, schools, public buildings, etc.).

However, the static way to evaluate exposure does not take into account either mobility or activities of a person in a city. As concern the dynamic way, it is interesting and useful because it considers these two last points. But it still remains exploratory and the results could not be easily generalized to every citizen in the city. The purpose of the present paper is to propose a geographical tool, in order to overcome those problems. It combines pollution maps and mobility. The work presented here, relates to the development of a software tool aiming at establishing a space-time assessment of the personal exposure to a given pollutant, for a citizen and a given day. Our approach supplements those presented above since it consists in modelling the individual exposure of a particular citizen, during one day thanks to a geographical information system (GIS). GIS is very often the favourite tool use by local air quality agencies to manage data and air quality survey activities.

The first section describes the area studied, data sets and tools. In this section, the multi-sources approach is presented. Then a first implementation is shown followed by a discussion about the results obtained. Our last section deals with perspectives of research.

2. STUDY AREA, DATA AND TOOLS

STUDY AREA :

The study area is the city of Strasbourg, France. The study area is approximatively 18 x 28 km². It is located in Eastern France near the German frontier. Geographical coordinates are 48°35'04" latitude North and 07°44'55" longitude West. It counts 451 223 inhabitants according to the last national census in 1999 and has a temperate climate dominated by continental influences. Atmospheric pollution in Strasbourg is mostly due to motor vehicles since there is no heavily polluting industries in this area except a refinery located in Reichstett in the North. ASPA (Association for the monitoring and the study of air pollution in Alsace, <http://www.atmo-alsace.net>) is the local agency in charge of the air quality monitoring in the city of Strasbourg and vicinity.

DATA SET :

Our database is based on an urban topographic data base (BDTOPO®) provided by the French national geographic institute (IGN©). BD TOPO® is a reference database including the physical three-dimensional description of a territory. It is the necessary topographic component to build up localised information

systems. It has a metric precision. This base contains eight principal topics : roads, electrical supply network of transport, hydrography, buildings, vegetation, description of the relief, administrative limits and finally toponymy.

Pollutants concentrations measurements of ambient air are provided by ASPA. Data concerning outdoor pollution relates to specific measurements of particulate matter ($PM_{2,5}$, PM_{10} , PM_{13}) catches on the 5 catches fixed measuring stations distributed over our area of study. In fact stations measure at the same time the background and the local pollution of various pollutants (PM , SO_2 , NO_x , O_3). In addition to the concentrations of pollutants, they also provide weather data such as wind speed, relative humidity, air temperature or solar radiation... The pollutants concentrations are provided for each measuring station every $\frac{1}{4}$ hour during one day. This data is collected by the local measurement network of ASPA.

Indoor pollution data are provided by OQAI (Observatory of Indoor Air Quality, <http://www.air-interieur.org>). Indoor air pollution can be approximated as a ratio between the pollutant concentration inside and its concentration outside (Mosqueron & Nedellec, 2001, Pernelet & Roth, 2001 ; Kirchner *et al.*, 2001). This ratio I/O depends on the pollutant, the type of building, season, system of air renewal, presence or not of smokers and presence or not of sources of pollution inside. This ratio is around 1 for $PM_{2,5}$ and PM_{10} whatever buildings and season according to Mosqueron & Nedellec (2001). Kirchner *et al.* (2001) has found a ratio around 1 for $PM_{2,5}$ and between 0.89 and 0.93 for black smokes. In order to determine PM concentrations inside, we applied a ratio of 1 to pollutants concentrations observed outside.

TOOLS :

We have used ArcGIS™ 8.3, a geographical information system, to represent an individual typical exposure to atmospheric pollution during one day. Thanks to the Tracking Analyst® module, extension of ArcGIS™ 8.3, it is possible to manage, represent and analyse temporal phenomena. We used the Tracking Analyst module to manage data on mobility and the evolution of pollutants concentration in time. Then we have created a dynamical representation in order to display the rates of pollutants to which a citizen is subjected throughout his way. This tool works also with real time data supplied by a connection to the Internet.

In order to estimate the quantity of PM a person is in contact with during a day, we propose to use the following expression :

$$Cumulation = \sum_{i=1}^N C_i t_i \quad (1)$$

with i indicating an activity (working, going outside, staying at home...), C_i the hourly mean PM concentration during this activity, t_i the time, in hours, spent in contact with this concentration and N the number of activities. In fact, we do not know very well activities duration (t_i). It depends of everyone (Roussel & Rambaud 2005). For the simulation with GIS, we assume that a typical citizen spends about 8 hours at work, 2 hours outside and the rest of his time at home. It is then possible to compute the “cumulation”.

We have used European air quality norms and thresholds as references in order to quantify exposure in terms of dangerousity .

3. RESULTS AND DISCUSSION

Our work is based on the following hypothesis. A person belonging to the active population stays 8 hours at its place of work, spends 2 hours outside and the rest of its time in its residence.

First of all, we have created hourly PM concentration maps from measures (figure 1). The mapping method is the “ordinary kriging” method, which is a method of interpolation usually used by air quality agencies. This map is used for the computation of the indoor/outdoor ratio. Such maps allow us to estimate the concentration everywhere.

Then we have modelled a typical path in the city taken by a citizen to go from home to work (figure 2). This citizen goes from its residence located in the South-East of Strasbourg to its place of work located in the North-West. He thus walks in the city every day. Its way lasts 1 hour in the morning from 8h to 9h and 1 hour in the evening from 18h to 19h.

We have gathered all set of information in GIS. With Tracking Analyst® module, we have obtained a dynamic representation of individuals exposure during a day: colourful disks appear and show the quantity of PM the citizen is in contact with. When European norms or thresholds are exceeded, the colourful disks become bigger and a warning symbol appears. Figure 3 is a snapshot of the dynamic representation acquired with the Tracking Analyst® module.

Finally, we have computed the “cumulation” (figure 4). It shows the total quantity of PM, which a person is breathing during a day. It is compared with the daily European air quality norm which is $50 \mu\text{g}/\text{m}^3$.

The exposure to pollution during one day remains difficult to estimate. To improve it, more realistic data of mobility would be needed, more information on health risks. Actually many data concerning indoor air quality are missing. Only few documents are available on this subject. It would seem that the majority of the studies launched on the exposure to indoor air are still exploratory. It is only recently that the effects of this pollution are taken into account in calculation of the individual exposure. Consequently, neither thresholds of exposure, nor thresholds of recommendations can be defined. The principal problem is the absence of indoor pollution indicators (Kirchner, 2003). By studying the space-time budget of the working population to some air pollutants in six European cities, Expolis (Air Pollution Exposure Distributions of Adult Urban Populations in Europe), a study in progress, hopes to define indicators (<http://www.ktl.fi/expolis/expolis2000/42.html>). An other major problem is the relative lack of precise pollution map at city scale. With such a map, the exposure could be assessed at every places and every instants. The maps used for this study are obtained by simple interpolation method based on *in situ* measurements. But the low number of measurements yields to vague assessment. The use of interpolation methods with cofactors as cokriging could notably improve map. Nevertheless this work offers an original approach of the exposure assessment. It does not pretend to integrate all available data but to show a tool that could become a decision or information support system.

4. CONCLUSIONS

In this paper, we have shown the existence of tools combining indoor and outdoor pollution that can inform about exposure to such pollution. We have shown a way to represent this information and to estimate exposure (“cumulation”). But it exists many other way to represent exposure assessment. For example, we could create an index especially dedicated to individual exposure.

In other words, we have shown, in particular by the dynamical representation, that useful tools are available and are promising to model exposure. Moreover, digital maps and geographical data become accessible and cover the territory more and more largely. Thus the precise localization of a person becomes all the more possible (Jensen 1999). But the problem in the exposure assessment is the exposure itself. On the one hand, health risks are not yet well known and estimated. For example, is it more harmful to be in contact with high levels of PM concentrations during a short time or with low levels of PM concentrations during a long time ? On the other hand, it does not exist any consensual way to compute exposure.

Nevertheless, the tool presented here can have numerous applications. Many scenarios can be imagined. A person who wants to know his exposure to the atmospheric pollutants, i.e. health risks along a way downtown or for a particular day, could visualize on his computer or telephone screen, such a representation of the exposure combined with a route calculation system. The purpose of this modelling is to give individualized information on the air quality that a person breaths during a day, to calculate his exposure and thus to give him recommendations. These recommendations can relate to spot or hours to avoid such a day.

Our work is also addressed to the local authorities. By the means of this system, they can have a further tool to manage exposure data, to analyse them and to provide information or recommendations to the population. Thanks to the Tracking Analyst® module, it could become possible to receive real time data on pollution and then to diffuse information through the internet. We limited our study to the exploitation of a GIS tool for the individual pollution exposure assessment. We know that it could be a part of an operational software dedicated to air pollution risks prevention and that it could be helpful for local authorities.

Our work is demonstrated on the town of Strasbourg but can be generalized at any city having an urban topographic database as well as an air quality monitoring network.

6. REFERENCES

- ADEME (Agency of Environment and Control of Energy), 2004. Air pollution and heat wave 2003. Assessment and consequences. Report of the technical day of March 4, 2004, Paris, France.
- Filleul L., Vandentorren S., Baldi I., Tessier J.F., 2001. Daily respiratory mortality and PM10 pollution in Mexico City. *European Respiratory Journal*, 17, pp. 1055–1056.
- Jensen S. S., 1999. A Geographic Approach to Modelling Human Exposure to Traffic Air Pollution using GIS. PhD Thesis of the National Environment Research Institute, Ministry of Environment and Energy, Denmark, 165 p.
- Kirchner S., 2003. Given metrological of indoor air. 41st congress of the Clinical Company of Toxicology, Nancy, France, Scientific and Technical Center of Building industry (CSTB), June 26-27, 15p., http://www.stc-congrès.org/Congrès_2003/Conférences.htm.
- Kirchner S., Bailloit O., Collignan B., Flori J. P., Garret O., Laurent A. M., Le Moullec Y., O’Kelly P., Ramalho O., Sauvaget M., Villenave J. G., Vedel C., 2001. Experimental study of transfer conditions for atmospheric pollution from local origin inside residential buildings. Final report, Primequal program, Scientific and Technical Center of Building industry (CSTB), April 2001, 269 p.
- Kotzias D., 2005. Indoor air and human exposure assessment – needs and approaches. *Experimental and Toxicologic Pathology*, 57, 1, 22 july, pp. 5-7.

Mosqueron L., Nedellec V., 2001. Inventory of French data on air quality. Observatory of Indoor Air Quality, Marne la Vallée, France, December 2001, 173 p.

Nerriere V., Zmirou-Navier D., Blanchard O., Momas I., Ladner J., Le Moullec Y., Personnaz M. B., Lameloise P., Delmas V., Target A., Desqueyroux H., 2005. Can we use fixed ambient air monitors to estimate population long-term exposure to air pollutants ? The case of spatial variability in the Genotox ER study. *Environmental Research*, 97, pp. 32-42.

Pernelet V., Roth C., 2001. Evaluation of the exposure of the Ile-de-France citizens to air pollution during the day. AIRPARIF, Paris, France, December 2001, 106 p.

Roussel I., Rambaud J. M., 2005. The assessment of personal exposure : what for ? In proceedings of the 3rd international symposium on air quality management at urban, regional and global scales, pp. 1526-1532, 26-30 september 2005, Istanbul, Turkey.

WHO World Health Organization, 2000. Air Pollution. Memory Help n° 187, <http://www.who.int/inf-fs/fr/am187.html>.

WHO, 2000. Air Quality Guidelines for Europe. 2nd Edition. Regional WHO Publications, European Series 91, Geneva, Switzerland, 288 p.

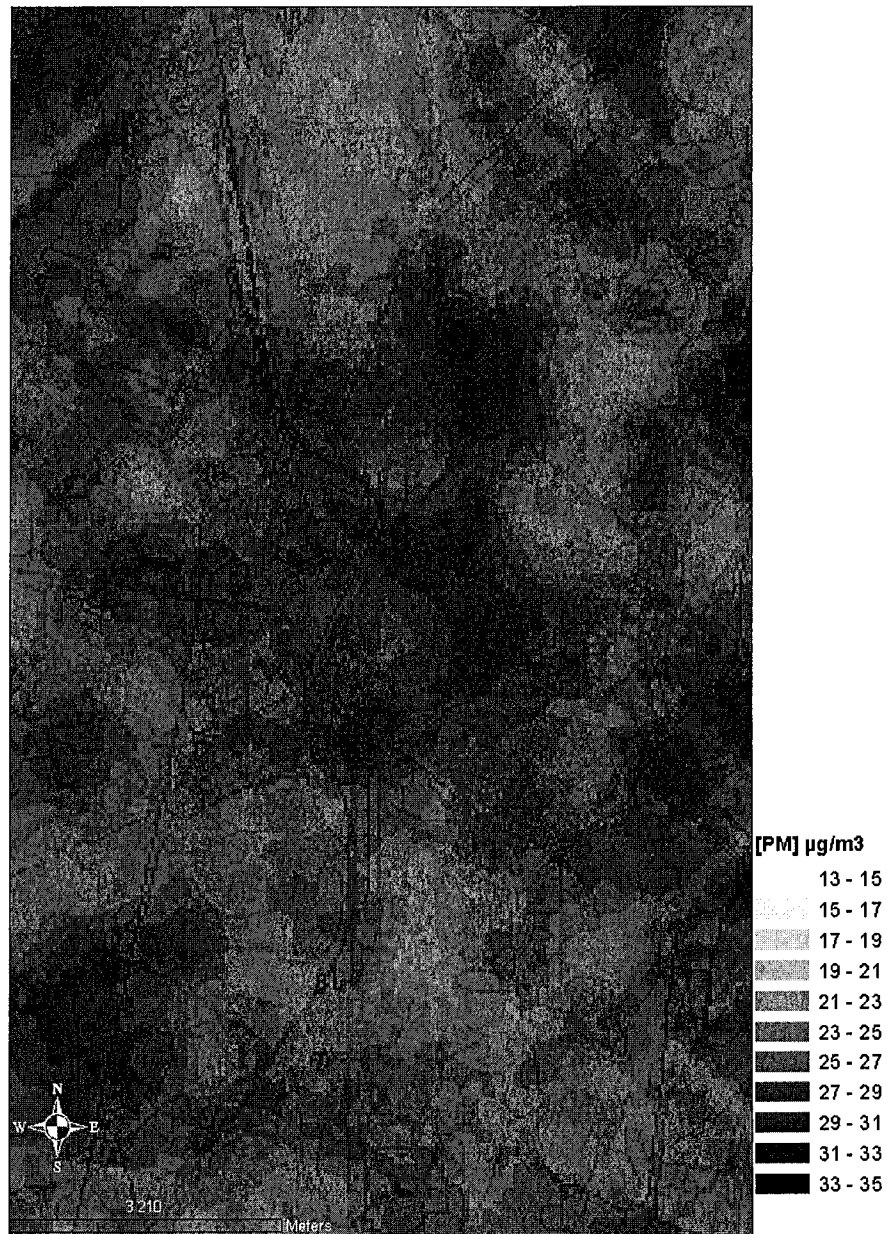


Fig 1 : Example of a hourly PM concentration map, on Strasbourg city. In the background, buildings, major roads and rails can be distinguished.

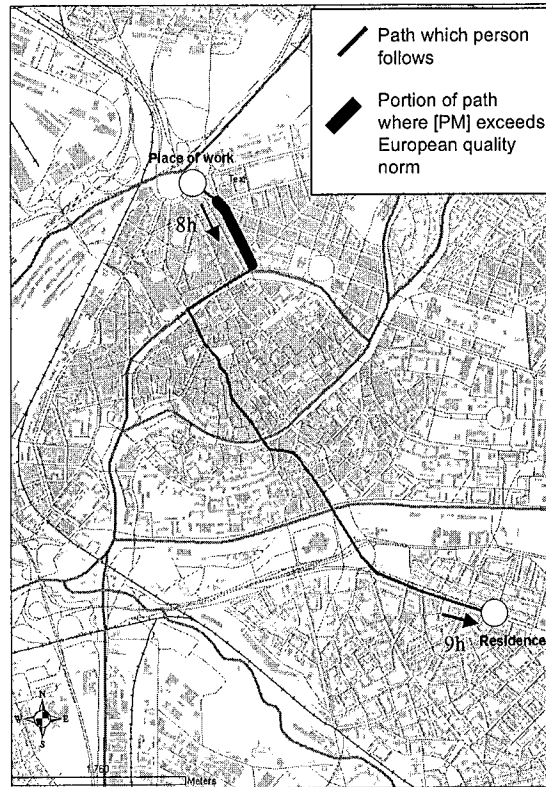


Fig 2 : Representation of exposure to PM along a path which person follows to go from home to work in the city center of Strasbourg. In the background, there are buildings, major roads and rails.

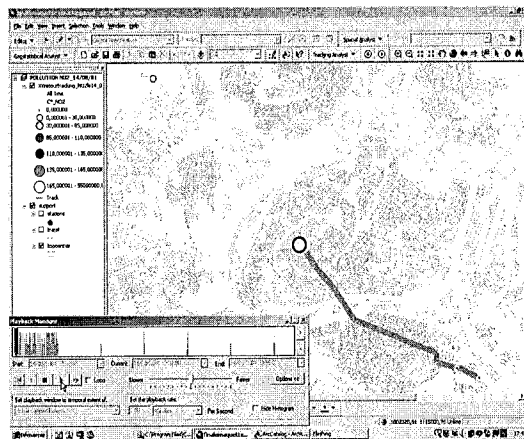


Fig 3 : Snapshot of the animation. The path and the colourful disks representing pollutants concentration appear progressively. In the bottom left corner, the progression bar allows to display the time and to manage the animation, to stop it at a precise place or moment to zoom in, for example.

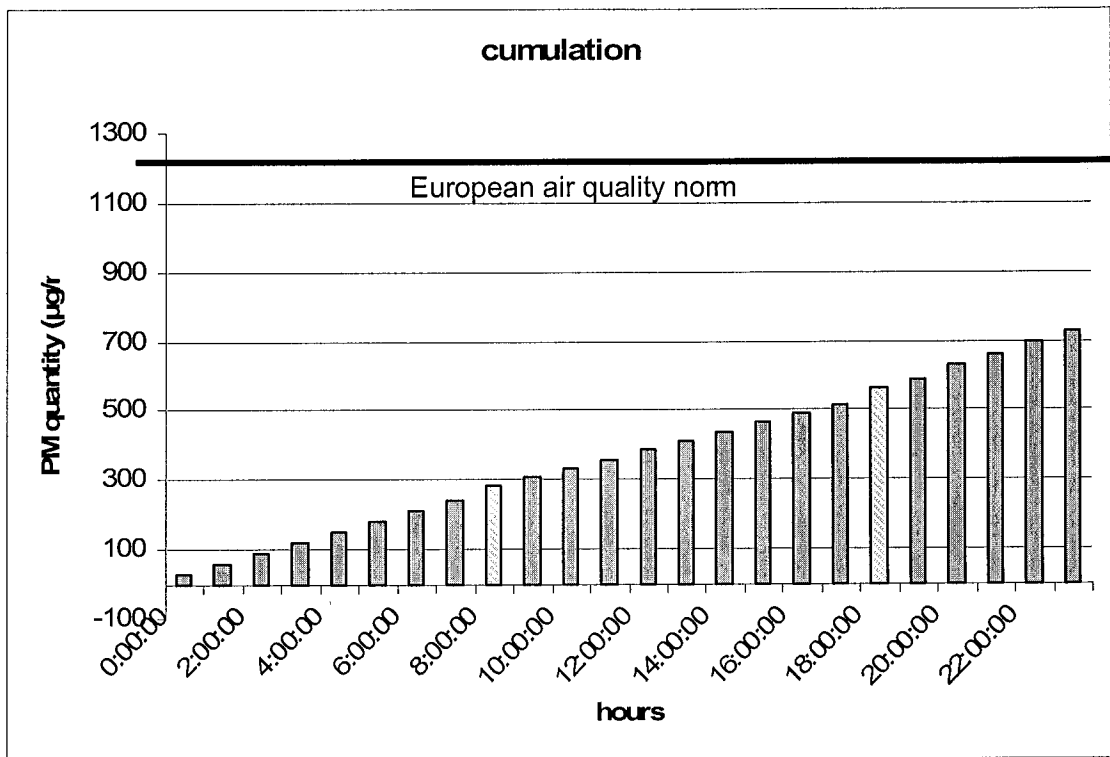


Fig 4 : Representation of the cumulation for a given day. Bars represent the cumulative quantity of PM a person is in contact with every hour during 24 hours. Hatchings correspond to the path outdoor.

Annexe B

Analyse multirésolution et transformée en
ondelettes

1 **Nomenclature**

d	fonction représentant le champ de concentration
t	temps
d^t	approximation du champ de concentration à l'instant t
$\psi_{a,b}$	ondelette-mère
a	pas de dilatation
b	pas de translation
$C_{a,b}$	coefficient en ondelette
j	résolution donnée
V_j	sous-espace vectoriel à la résolution j
ϕ	fonction d'échelle
h	filtre discret de la transformée de Fourier (filtre passe-bas)
g	filtre discret de la transformée de Fourier (filtre passe-haut)
m	situation météorologique donnée
D_m	champ de concentration normalisé pour la situation météorologique m
N	nombre d'estimations du champ de concentration
CT_m	champ typique pour une situation météorologique m
x	variable d'espace selon les longitudes
y	variable d'espace selon les latitudes

2 **Transformée en ondelettes**

Avant de définir les coefficients en ondelette, il est important d'expliquer la transformée en ondelettes. La théorie des ondelettes ne sera pas développée en détail dans cette thèse. Pour plus de précision,

j'invite le lecteur à se reporter à Daubechies (1988) ou Ranchin (1993) ainsi qu'à Meyer (1992) qui présente dans son ouvrage une perspective historique des ondelettes.

Pour représenter le champ de pollution, on considère une fonction caractérisée par des structures bi-dimensionnelles à différentes échelles spatiales imbriquées. Dans le cadre mathématique, on se place dans l'espace L^2 (ou espace de Hilbert des fonctions de carrés sommables). Soit $d(x) \in L^2(\mathbf{R})$, cette fonction.

Pour des raisons de simplicité, je vais décrire la transformée en ondelette dans le cas uni-dimensionnel.

La principale propriété de la transformée en ondelettes est d'adapter la fenêtre d'analyse au phénomène étudié, ce qui permet de fournir une information locale. La transformée en ondelette permet de décomposer la fonction d sur une base de fonctions orthogonales élémentaires $\{\psi_{a,b}\}$: les ondelettes. Cette base ou famille est générée par translations et dilatations d'une unique fonction ψ appelée *ondelette-mère*, ce qui s'écrit dans le cas uni-dimensionnel :

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) \quad (1)$$

où a et $b \in \mathbf{R}$, $a \neq 0$. Le paramètre a est appelé "pas de dilatation" et b "pas de translation".

Les capteurs que nous utilisons délivrent des signaux qui sont discrétisés : les images numériques. C'est pour cette raison que j'utiliserai la version discrète de la transformée en ondelettes. Les paramètres a et b prennent tous deux des valeurs discrètes et $d(x)$ est appelé « ligne image ».

La décomposition ou analyse d'une fonction d dans le cas discret s'écrit de la façon suivante :

$$C_{a,b} = \langle d, \psi_{a,b}(x) \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} d(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx \quad (2)$$

où $\psi_{a,b}$ est défini comme précédemment, $\overline{\psi_{a,b}}$ est le complexe conjugué de ψ , $\langle d, \psi_{a,b}(x) \rangle$ est le produit scalaire défini dans $L^2(\mathbf{R})$ et $C_{a,b}$ est appelée le coefficient d'ondelette. Pour une échelle a et une position b , $C_{a,b}$ représente l'information contenue par $d(x)$ pour cette échelle et à cette position. Ce coefficient représente une mesure de l'intensité des variations locales du signal. Sa valeur sera importante si l'échelle de variation de l'ondelette $\psi_{a,b}$ est proche de celle de la singularité ou de la structure locale du signal. Si le signal présente peu ou pas de variations pour cette échelle, la valeur du coefficient d'ondelette sera faible. $C_{a,b}$ est donc une caractérisation des structures présentes dans l'image pour cette échelle et à cette position.

Inversement, on peut reconstruire exactement $d(x)$ à partir des coefficients d'ondelettes et si la condition d'admissibilité est satisfaite, de la manière suivante :

$$d(x) = \frac{1}{K} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} C_{a,b} \psi\left(\frac{x-b}{a}\right) \frac{da db}{a^2} \quad \text{où} \quad K = \int_{-\infty}^{+\infty} \frac{|\psi(u)|^2}{|u|} du \quad (3)$$

Cette équation peut être interprétée de deux manières différentes :

- $d(x)$ peut être reconstruite exactement quand on connaît sa transformée en ondelettes,
- $d(x)$ est une superposition d'ondelettes.

Ces deux points de vue conduisent à des applications différentes qui sont, d'une part le traitement et, d'autre part, l'analyse des signaux.

3 Analyse multirésolution

3.1. Théorie

Le concept d'analyse multirésolution tel que l'a introduit Mallat (1988, 1989). Dans le cas des algorithmes pyramidaux, on utilise souvent la représentation de la figure 1.

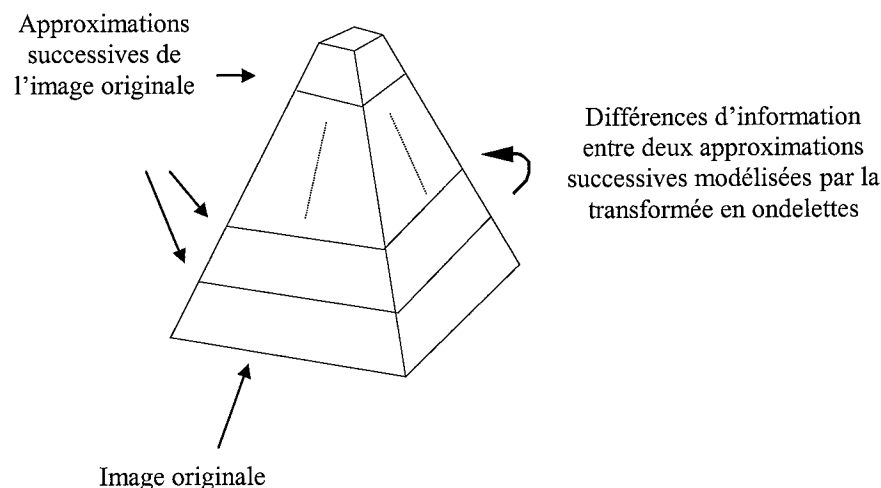


Fig 1 : Représentation des approximations successives d'une même image à l'aide d'un algorithme pyramidal.

La figure précédente (figure 1) est la représentation d'un algorithme pyramidal. Les algorithmes pyramidaux permettent de "fabriquer" des approximations successives et de plus en plus grossières de

l'image originale. La base de la pyramide correspond à l'image originale. Quand on "monte" dans la pyramide, les différents étages représentent les approximations successives de l'image originale. La limite théorique de ces algorithmes est le sommet de la pyramide qui correspond à un pixel unique. L'analyse multirésolution est un cas particulier des algorithmes pyramidaux.

L'analyse multirésolution d'une fonction d de $L^2(\mathbf{R})$ se traduit par l'obtention d'une suite d_n de fonctions, dont chacune est une approximation de d à une résolution bien déterminée. Chacune de ces approximations est la fonction $d(x)$ lissée par des fonctions de plus en plus concentrées. Par définition, une analyse multirésolution est donc une suite croissante $\{V_j\}$, $j \in \mathbf{Z}$ de sous-espaces vectoriels de $L^2(\mathbf{R})$, ayant certaines propriétés dont quatre en particulier sont utiles pour notre étude :

- Emboîtement des sous-espaces :

$$V_j \subset V_{j+1}. \text{ Si } d(x) \in V_j \text{ alors } d(x) \in V_{j+1}$$

- Propriété d'invariance par translation :

$$\text{Si } d_j(x) \in V_j \text{ alors } \forall k \in \mathbf{Z}, d_j(x - k) \in V_j$$

- Propriété d'invariance par dilatation :

$$\text{Si } d_j(x) \in V_j \text{ alors } d_j(2x) \in V_{j+1}$$

Le facteur 2 correspond à celui de l'analyse dyadique.

- Existence d'une base :

Il existe une fonction $\phi(x)$ telle que l'ensemble des fonctions $\{\phi(x - k), k \in \mathbf{Z}\}$ forme une base de V_0 .

Chaque V_j se trouve associé à une résolution j et l'approximation d'un signal $d(x)$ à cette résolution est obtenue par projection sur le sous-espace correspondant. Une base de V_j peut alors se déduire de la base V_0 à partir de la seule fonction $\phi(x)$ appelée « fonction d'échelle » ou « base d'interpolation ». La fonction d'échelle permet de définir des approximations successives de la fonction $d(x)$.

En se plaçant dans un algorithme pyramidal et dyadique (c'est-à-dire en utilisant des puissances de 2 pour la résolution spatiale : $a = 2^j$ et $b = k 2^j$, $k, j \in \mathbf{Z}$), Si l'ensemble des fonctions $\phi(x - k)$ est une base de V_0 alors l'ensemble des fonctions $2^{\frac{j}{2}} \phi_j(2^j x - k)$ (par translation et dilatation) est une base de V_j .

Soit $d_j, j \in \mathbb{Z}$, une approximation en multirésolution et $\phi(x)$ une base d'interpolation. Les fonctions $\{2^{-\frac{j}{2}}\phi_j(x-2^{-j}k)\}$ forment une base de V_j . Or les deux sous-espaces V_j et V_{j+1} étant emboîtés, une fonction quelconque de V_j peut être exprimée en fonction de la base de V_{j+1} , par projection. Après projection de la fonction sur la nouvelle base, on en déduit le coefficient d'interpolation, pour passer d'une résolution à une autre, tel que :

$$\begin{aligned} d_j(n) &= \langle d(x), \phi_j(x-2^{-j}n) \rangle \\ &= \sum_{l=-\infty}^{+\infty} h(l-2n) d_{j+1}(l) \\ &= \sum_{l=-\infty}^{+\infty} \bar{h}(2n-l) d_{j+1}(l) \end{aligned} \quad (4)$$

d_j est le coefficient d'interpolation à la résolution j et h est un filtre discret et $\bar{h}(l) = h(-l)$.

3.1. Décomposition ou Analyse : passage d'une résolution donnée à une résolution inférieure

Nous venons de voir que passer d'une résolution à une autre revient à projeter la fonction d_{j+1} appartenant à un espace V_{j+1} sur un espace V_j dont les fonctions $2^{-\frac{j}{2}}\phi_j(x-2^{-j}k)$ forment une base. L'information perdue lors de cette projection correspond à la projection C_j de d_{j+1} sur l'espace O_j complémentaire de V_j dans V_{j+1} ($V_{j+1} = V_j \oplus O_j$). C_j correspond alors au détail de d_{j+1} . Par analogie, l'ensemble $2^{-\frac{j}{2}}\psi_j(x-2^{-j}k)$ forme une base (d'ondelettes) de O_j . Il est alors possible de décomposer toute fonction de O_j sur V_{j+1} .

On en déduit alors l'expression des coefficients de détails :

$$\begin{aligned} C_j(n) &= \langle d(x), \psi_j(x-2^{-j}n) \rangle \\ &= \sum_{k=-\infty}^{+\infty} g(l-2n) d_{j+1}(l) \\ &= \sum_{k=-\infty}^{+\infty} \bar{g}(2n-l) d_{j+1}(l) \text{ avec } \bar{g}(l) = g(-l) \end{aligned} \quad (5)$$

g est un filtre discret.

On obtient donc l'approximation d_{j-1} du signal et le signal de détail C_{j-1} à partir de d_j en convoluant d_j respectivement par \bar{h} et \bar{g} et en retenant un échantillon sur deux (fig. 2).

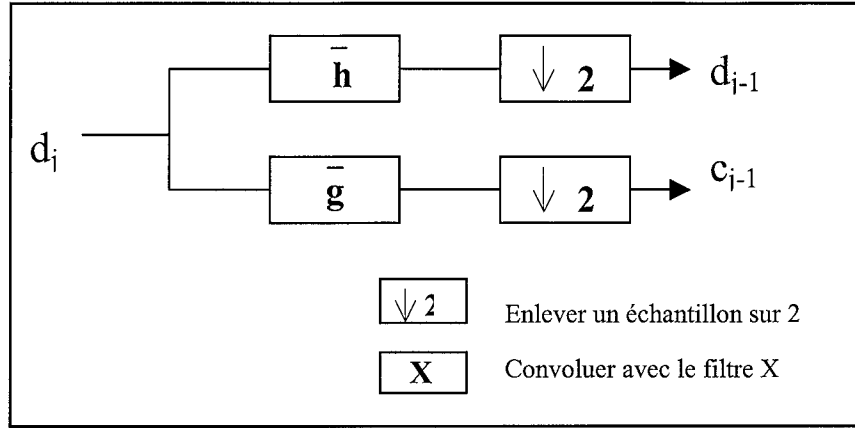


Fig 2 : Passage d'une résolution donnée à une résolution inférieure

3.2. Reconstruction ou Synthèse : passage d'une résolution donnée à une résolution supérieure

Le signal discret peut être reconstruit à partir de sa pyramide. Toute fonction $d_{j+1}(x) \in V_{j+1}$ est décomposable sur une base de ce sous-espace vectoriel, or comme $V_{j+1} = V_j \oplus O_j$, l'ensemble $\{2^{-\frac{j}{2}}\phi_j(x-2^{-j}k), 2^{-\frac{j}{2}}\psi_j(x-2^{-j}k)\}_{(k \in \mathbb{Z})}$ forme une base orthogonale de ce sous-espace vectoriel. En décomposant alors $\phi_{j+1}(x-2^{-j-1}k)$ sur cette base et en utilisant les filtres $h(l)$ et $g(n)$ définis précédemment, on passe d'une résolution donnée à une résolution plus basse de la façon suivante :

$$\langle d(x), \phi_{j+1}(x-2^{-j-1}n) \rangle = 2 \sum_{l=-\infty}^{+\infty} h(n-2l) \langle d(x), \phi_j(x-2^{-j}l) \rangle + 2 \sum_{l=-\infty}^{+\infty} g(n-2l) \langle d(x), \psi_j(x-2^{-j}l) \rangle \quad (6)$$

Ce qui s'écrit aussi (cf. Eq. 4 et 5) :

$$d_{j+1}(n) = 2 \sum_{l=-\infty}^{+\infty} [h(n-2l) d'_j(l) + g(n-2l) C_j(l)] \quad (7)$$

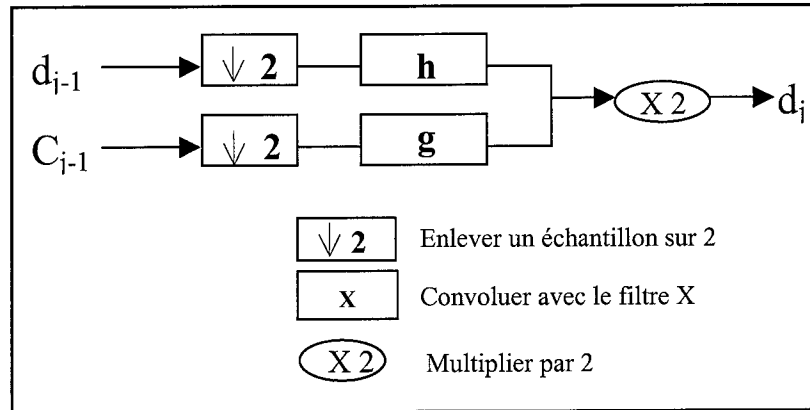


Fig 3 : Passage d'une résolution donnée à une résolution supérieure

On reconstruit la fonction en mettant des zéros entre deux échantillons de d_j et C_j , et en convoluant les signaux avec les filtres h et g . Ce qui revient à interpoler les approximations puis à les sommer.

3.3. Mise en œuvre et exemple

D'un point de vue pratique, l'application de l'analyse multirésolution est réalisée par le biais de filtres passe-bas et la différence d'informations existant entre deux approximations successives par un filtre passe-haut. Pour plus de détails, on peut se référer à l'ouvrage très complet de Daubechies (1992).

La mise en œuvre de l'analyse multirésolution se décompose en deux étapes : l'analyse (quand on « remonte » la pyramide de l'image originale jusqu'au pixel unique) et la synthèse (quand on « descend » la pyramide). Ces deux opérations permettent une reconstruction exacte du signal original.

En se plaçant toujours dans un algorithme pyramidal et dyadique ($a = 2^j$ et $b = k 2^j$, $k, j \in \mathbb{Z}$), l'approximation de l'image à la résolution inférieure est représentée par la fonction $d_{j+1}(x,y)$. La différence d'information entre l'approximation à la résolution 2^j et celle à la résolution 2^{j+1} est représentée par les images de coefficients d'ondelettes. Ces coefficients sont la représentation des irrégularités de l'image rencontrées à une résolution donnée dans une direction donnée.

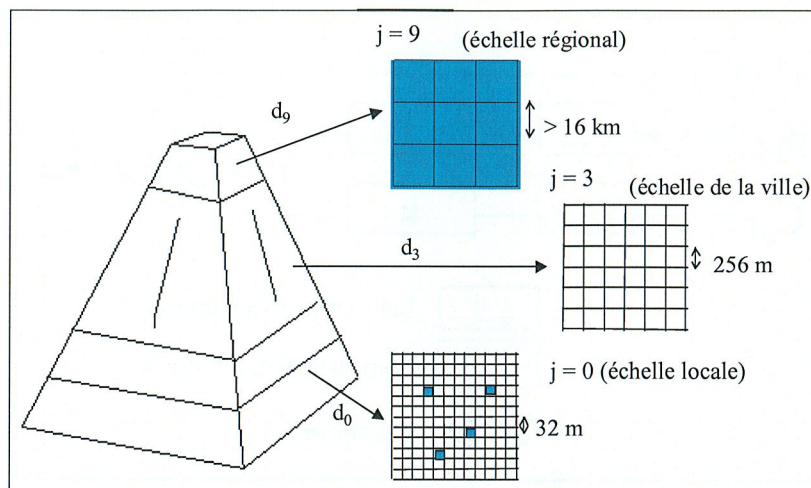


Fig 4 : Représentation des approximations successives d'une fonction à l'aide d'un algorithme pyramidal et dyadique.

La figure 4 représente les approximations successives d'une fonction $d(x)$ à l'aide d'un algorithme pyramidal et dyadique.

A la base de la pyramide (échelle locale), d_0 représente un champ de concentration théorique désiré. Prenons la résolution spatiale de l'image égale à 32 m. A cette échelle ($j = 0$, $a = 2^j = 1$), seules les structures locales ayant une longueur caractéristique supérieure à 32 m sont visibles. Lorsque l'échelle d'observation a augmente, la résolution diminue et les longueurs caractéristiques des structures visibles sur l'image « approximation » d_j augmentent. Pour $j = 3$, l'approximation d_3 contient toutes les structures ayant des longueurs caractéristiques supérieures à 256 ($= 32 \times 2^3$) m.

Au sommet de la pyramide, d_9 représente l'approximation de la fonction avec un paramètre d'échelle a égal à 2^9 , donc avec une résolution spatiale supérieure à 16 km. Cette approximation est à l'échelle régionale. Les structures ayant des longueurs caractéristiques inférieures à 16 km sont filtrées et ne sont plus représentées dans cette approximation.

Il est ainsi possible de représenter et d'étudier l'évolution de la fonction d suivant des échelles différentes, et donc à des résolutions différentes.

Il est important de préciser que je n'utiliserai pas d'algorithme pyramidal dans la méthode des champs typiques. La théorie et les figures ont été données à des fins pédagogiques et d'illustration. J'utiliserai l'algorithme « à trous ». Ce n'est pas tout à fait un algorithme pyramidal car la taille des images est identique si on « monte » dans les échelles. Il n'y a donc pas de rééchantillonnage avec le changement d'échelle. C'est cette caractéristique ainsi que le fait qu'il n'y a qu'un plan d'ondelettes non-directionnel au lieu de trois directionnels (vertical, horizontal, diagonal) dans l'implémentation classique d'un algorithme pyramidal, qui m'a fait choisir cet algorithme, de ce fait, plus facile à mettre en œuvre.

CARTOGRAPHIE DE LA POLLUTION PARTICULAIRE EN VILLE

Résumé

Le Haut Comité de Santé Publique donne des recommandations sur les moyens d'améliorer la connaissance des phénomènes liés à la pollution atmosphérique dans le but d'aider à l'évaluation de l'exposition des citoyens. Les travaux de cette thèse s'inscrivent dans ce contexte global en s'orientant vers la cartographie de la concentration en particules en ville. Plusieurs outils existent aujourd'hui qui permettent de dériver des informations sur la qualité de l'air. Ce sont des cartes obtenues par des méthodes d'interpolation spatiale des mesures ponctuelles ou des modèles numériques.

Ces deux approches ont des limitations. L'objectif de cette thèse est de contribuer à l'établissement d'une stratégie de cartographie de la pollution atmosphérique sur l'ensemble d'une agglomération à l'échelle de la rue (100 m) en s'appuyant notamment sur l'imagerie satellitaire.

La stratégie de cartographie comprend deux méthodes. La méthode des champs typiques vise à améliorer la représentation de la pollution en appliquant des méthodes de fusion de données à des cartes de pollution déjà existantes. Les méthodes de fusion de données sont appliquées à la cartographie de la qualité de l'air. Une formalisation de la méthode des champs typiques est donnée et un essai sur un cas pratique est réalisé. La méthode de densification du réseau de mesure vise à créer des cartes de pollution en interpolant les valeurs de concentration mesurées par les stations. Des stations virtuelles s'ajoutent aux stations réelles du réseau de mesure pour le rendre plus dense et améliorer ainsi l'interpolation. Ces stations virtuelles sont déterminées à partir d'une classification sur des éléments décrivant l'environnement des stations. Pour la ville de Strasbourg, 635 stations virtuelles sont ajoutées aux cinq stations réelles existantes. La cartographie par interpolation avec ces stations est améliorée qualitativement et quantitativement (jusqu'à 70 %).

Pour valider la définition des stations virtuelles, l'imagerie satellitaire est utilisée. Les longueurs d'ondes autour de 815 nm sont les plus sensibles aux particules. La bande spectrale TM1 du capteur TM de Landsat5 est appropriée pour la détection des particules et est utilisée pour valider les stations virtuelles. 70 % des stations virtuelles ont été validées avec cette approche.

Mots clés :

Cartographie, particules, champ typique, fusion de données, station virtuelle, interpolation, imagerie satellitaire, Landsat5-TM

CARTOGRAPHY OF PARTICULATE POLLUTION IN URBAN AREAS

Abstract

The French High Committee of Public Health gives recommendations on the means of improving knowledge of the phenomena related to air pollution in order to better evaluate citizens' exposure to pollution. The work presented in this thesis was conducted in this context and was particularly directed towards the cartography of particle concentrations. Several tools exist which make it possible to derive information on air quality. They are maps obtained by measurement interpolation methods or numerical models.

These two approaches have limitations. The objective of this thesis is to contribute to the establishment of a strategy for particulate pollution mapping over urban area at a local scale (100 m).

The mapping strategy includes two methods. The typical fields method aims at improving the representation of pollution by applying data fusion methods to existing pollution maps. The data fusion methods are applied in this thesis to map air pollution. A formalization of the typical fields method is given and tested on a case-study. The network densification method aims at creating pollution maps by interpolating pollution values provided by measurement stations. Virtual stations are added to the real stations belonging to the measurement network to make it denser thus improving the interpolation. The virtual stations are determined using a classification of elements describing the environment of the real stations. In a case-study of Strasbourg, 635 virtual stations were added to the five real stations located in this city. The mapping by interpolation with these stations was improved qualitatively and quantitatively (up to 70 %).

In order to validate the definition of the virtual stations, satellite images are used. Wavelengths around 815 nm are the most sensitive to particles. The spectral band (TM1) of the TM-Landsat5 sensor is appropriate for the detection of particles and was used to validate the virtual stations. 70 % of the virtual stations were found to be valid with this approach.

Key words :

Air pollution mapping, particles, typical field, data fusion, virtual station, interpolation, satellite images, Landsat5-TM

Laboratoire d'accueil : Centre Énergétique et Procédés - Ecole des Mines de Paris
Rue Claude Daunesse - B.P. 207 - F-06904 Sophia Antipolis Cedex

Thèse présentée par : BEAULANT Anne-Lise le : 27 septembre 2006

Discipline : Méthodes Physiques en Télédétection - Université Paris 7 Diderot
(Ecole Doctorale « Sciences de l'Environnement d'Ile de France »)
