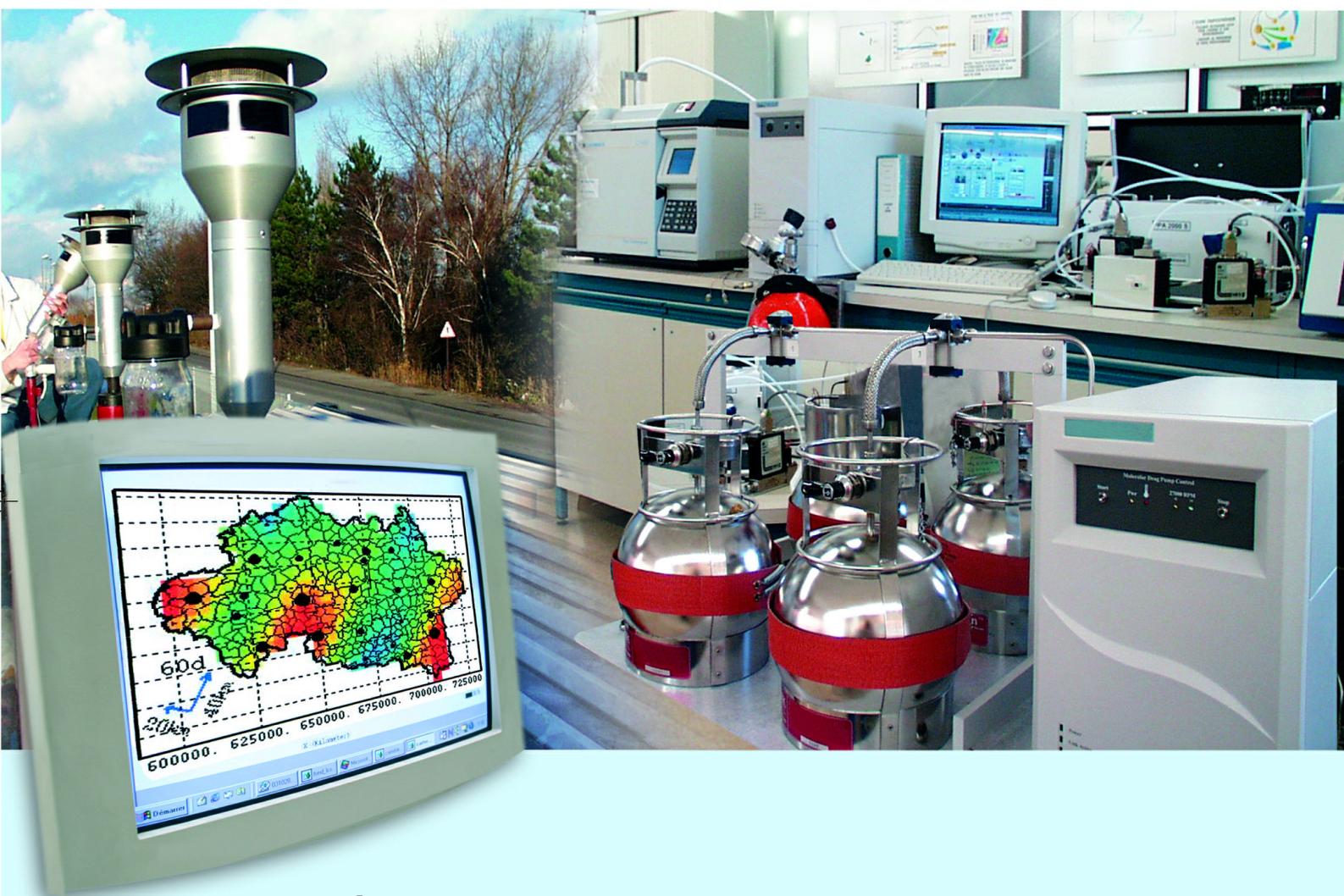




Laboratoire Central de Surveillance de la Qualité de l'Air



SYNTÈSE DES TECHNIQUES DE CARTOGRAPHIE DU DIOXYDE D'AZOTE : VALORISATION DES TRAVAUX 2003 – 2005

Laure Malherbe et André Wroblewski
Décembre 2006
Version finale

PREAMBULE

Le Laboratoire Central de Surveillance de la Qualité de l'Air

Le Laboratoire Central de Surveillance de la Qualité de l'Air est constitué de laboratoires de l'Ecole des Mines de Douai, de l'INERIS et du LNE. Il mène depuis 1991 des études et des recherches finalisées à la demande du Ministère chargé de l'environnement, sous la coordination technique de l'ADEME et en concertation avec les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA). Ces travaux en matière de pollution atmosphérique supportés financièrement par la Direction des Préventions des Pollutions et des Risques du Ministère de l'Ecologie et du Développement Durable sont réalisés avec le souci constant d'améliorer le dispositif de surveillance de la qualité de l'air en France en apportant un appui scientifique et technique aux AASQA.

L'objectif principal du LCSQA est de participer à l'amélioration de la qualité des mesures effectuées dans l'air ambiant, depuis le prélèvement des échantillons jusqu'au traitement des données issues des mesures. Cette action est menée dans le cadre des réglementations nationales et européennes mais aussi dans un cadre plus prospectif destiné à fournir aux AASQA de nouveaux outils permettant d'anticiper les évolutions futures.

Synthèse des techniques de cartographie du dioxyde d'azote : valorisation des travaux 2003-2005.

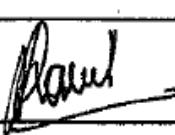
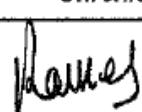
Laboratoire Central de Surveillance de la Qualité de l'Air

Thème : Traitements numériques

Programme financé par la Direction des Préventions des Pollutions et des Risques (DPPR)

Novembre 2006

Personnes ayant participé à l'étude : Giovanni CARDENAS

	Rédaction	Vérification	Approbation
NOM	Laure MALHERBE	Laurence ROUIL	M.RAMEL
Qualité	Ingénieur d'études Direction des Risques Chroniques	Ingénieur d'études Direction des Risques Chroniques	Responsable LCSQA/INERIS Direction des Risques Chroniques
Visa			

ECOLE DES MINES DE DOUAI

DEPARTEMENT CHIMIE ET ENVIRONNEMENT

**Synthèse des techniques de cartographie du dioxyde
d'azote : valorisation des travaux 2003-2005.**

André Wroblewski,
En collaboration avec Laure Malherbe : INERIS

DECEMBRE 2006

1. LCSQA 2006 / RESUME

Synthèse des techniques de cartographie de NO₂ : valorisation des travaux 2003-2005.

Etude conjointe EMD/INERIS

Pour l'information des autorités et du public, les AASQA produisent de nombreuses cartographies des polluants. Cette activité les a conduites à développer des compétences dans la mise en œuvre de campagnes d'échantillonnage et l'usage d'outils numériques plus ou moins sophistiqués :

- modélisation déterministe,
- interpolation entre les données expérimentales, généralement par géostatistique,
- méthodes paramétriques fondées sur des corrélations.

La diversité des techniques et outils de cartographie et la variété des situations dans lesquelles ils sont employés conduisent cependant à s'interroger sur le choix de la meilleure approche. La cartographie du dioxyde d'azote est tout particulièrement concernée. Ce polluant réglementé fait en effet l'objet de multiples études, aussi bien à l'échelle régionale, en réponse aux besoins de la surveillance, qu'à l'échelle d'une ville ou d'une route, du fait de son caractère local.

A partir des travaux conduits par l'EMD et l'INERIS ou des rapports d'AASQA, ce document dresse un bilan préliminaire des différentes méthodes de cartographie du NO₂. Plusieurs échelles d'espace : zones régionales/rurales, agglomérations, proximité routière, et différentes échelles de temps : année, jour, heure, sont considérées.

Sur un domaine rural influencé par des sources locales (routes, ville), l'étude montre en premier lieu le bon accord des trois catégories de méthodes avec les mesures. Un écart inférieur à 30 % sur la moyenne annuelle, conforme à la directive européenne de 1999, est ainsi enregistré. Les représentations cartographiques obtenues, qui toutes reposent sur l'inventaire des émissions de NO_x, sont de plus cohérentes entre elles.

La modélisation déterministe est la plus délicate à mettre en œuvre mais elle ouvre le plus grand champ d'application, aussi bien d'un point de vue spatial (modélisation de haute résolution) que d'un point de vue temporel (cartographie horaire par les modèles de chimie-transport). La technique des corrélations, plus aisée à utiliser, offre un résultat robuste. En revanche, il s'agit d'une estimation de grande échelle, et non d'une description fine des concentrations. La géostatistique constitue un compromis entre les deux approches. Avec des données de mesure suffisamment riches et des variables auxiliaires soigneusement choisies, elle permet de cartographier avec réalisme la pollution de NO₂. Quelle que soit la

méthode, la précision des résultats dépend grandement de la qualité des données d'entrée.

Aux plus petites échelles d'espace, avec un échantillonnage suffisamment dense, la géostatistique conserve sa pertinence, qu'il soit question de cartographie moyenne saisonnière ou annuelle ou, grâce à des développements récents, de cartographie journalière. En ce qui concerne la modélisation, la nécessité de disposer rapidement de représentations à fine échelle pour des études sur le long terme et/ou des cartographies heure par heure, conduit à l'usage d'outils relativement simples, fondés sur une formulation gaussienne ou empirique des concentrations.

TABLE DES MATIERES

1. LCSQA 2006 / RESUMÉ	5
2. INTRODUCTION.....	9
3. CARTOGRAPHIE DU NO₂ : MÉTHODES DISPONIBLES, CRITÈRES DE COMPARAISON	11
4. CARTOGRAPHIE A L'ECHELLE RÉGIONALE / EN ZONE RURALE (AVEC INFLUENCE ROUTIÈRE)	12
4.1 CARTES MOYENNES MENSUELLES, SAISONNIÈRES OU ANNUELLES	12
4.1.1 <i>Application des méthodes géostatistiques.....</i>	13
4.1.2 <i>Corrélations</i>	26
4.1.3 <i>Modélisation déterministe</i>	30
4.1.4 <i>Comparaison des approches.....</i>	33
4.2 CARTOGRAPHIE JOURNALIÈRE OU HORAIRE	45
4.2.1 <i>Application des méthodes géostatistiques ou statistiques.....</i>	45
4.2.2 <i>Modélisation déterministe</i>	45
4.2.3 <i>Perspective : couplage modélisation déterministe/géostatistique.....</i>	46
5. CARTOGRAPHIE EN ZONE URBAINE	46
5.1 CARTES MOYENNES SAISONNIÈRES OU ANNUELLES	46
5.1.1 <i>Application des méthodes géostatistiques.....</i>	46
5.1.2 <i>Modélisation déterministe ou semi-empirique</i>	52
5.1.3 <i>Comparaison des approches.....</i>	53
5.2 CARTES JOURNALIÈRES OU HORAIRES	55
5.2.1 <i>Application des méthodes géostatistiques (cartographie automatique)</i>	55
5.2.2 <i>Modélisation déterministe / semi-empirique</i>	57
6. CARTOGRAPHIE AU VOISINAGE DES ROUTES	58
6.1 CARTES MOYENNES SAISONNIÈRES OU ANNUELLES	58
6.1.1 <i>Application des méthodes géostatistiques.....</i>	58
6.1.2 <i>Modélisation empirique ou semi-empirique</i>	59
7. CONCLUSION	60
8. RÉFÉRENCES.....	62
9. LISTE DES ANNEXES	69
10. ANNEXE 1	70
11. ANNEXE 2 : LE KRIGEAGE	73
11.1 KRIGEAGE ORDINAIRE	74
11.2 VARIABLES AUXILIAIRES DISPONIBLES SUR UN ENSEMBLE IRRÉGULIER DE POINTS OU SUR UN MAILLAGE LÂCHE.....	75

11.3 VARIABLES AUXILIAIRES CONNUES PARTOUT (I.E. SUR UNE GRILLE RELATIVEMENT FINE)	75
12. ANNEXE 3 : VARIABLES AUXILIAIRES	78
13. ANNEXE 4 : ELABORATION DE CARTES SAISONNIÈRES – CAS D'UNE AGGLOMERATION	82
13.1 COKRIGEAGE TEMPOREL	82
13.1.1 <i>La méthode</i>	82
13.1.2 <i>Application à des données de LILLE</i>	83
13.2 CARTOGRAPHIE JOURNALIÈRE AUTOMATIQUE	92
14. ANNEXE 5 : APPLICATION DE LA GÉOSTATISTIQUE À L'EXPLOITATION DE MESURES CONJOINTES DE FOND ET DE PROXIMITÉ – RAPPEL DES RÉSULTATS	95
14.1 CARTOGRAPHIE DE LA POLLUTION DE FOND	95
14.1.1 <i>Relations avec les variables auxiliaires</i>	96
14.1.2 <i>Modélisation du variogramme</i>	98
14.1.3 <i>Evaluation des modèles</i>	100
14.1.4 <i>Estimation</i>	105
14.2 PRISE EN COMPTE DE LA POLLUTION DE PROXIMITÉ	109
14.2.1 <i>Comparaison avec les mesures de proximité</i>	109
14.2.2 <i>Modélisation de l'augmentation de concentration en zone de proximité</i>	110
14.2.3 <i>Incertitude</i>	112
14.2.4 <i>Représentation cartographique</i>	113

2. INTRODUCTION

Depuis quelques années, souhaitant d'une part, répondre aux exigences réglementaires et d'autre part, informer clairement le public, les AASQA ont engagé des moyens dans la réalisation de cartographies des polluants. Cette activité les a conduites à développer des compétences dans la mise en œuvre de campagnes d'échantillonnage et l'usage d'outils numériques plus ou moins sophistiqués. Deux approches se distinguent : soit l'utilisation de modèles déterministes, lorsqu'un inventaire des émissions est disponible sur la zone d'étude, soit une interpolation entre les données expérimentales, généralement par géostatistique, lorsque les sites de mesure sont en nombre suffisant. Dans ce contexte, le LCSQA a pris une part active au développement et à l'application de ces méthodes afin, notamment, de cartographier le dioxyde d'azote. Ce polluant réglementé fait régulièrement l'objet de campagnes d'échantillonnage, ce qui permet de disposer de nombreuses données pour la spatialisation. Les mesures réalisées portent sur de vastes domaines, de façon à couvrir les zones dépourvues de stations automatiques, mais à cause du caractère local des sources et des impacts du dioxyde d'azote, elles peuvent tout aussi bien concerner des domaines plus retrêints (villes, bords des routes,...).

La diversité des techniques et outils de cartographie et la variété des situations dans lesquelles ceux-ci sont employés (échelle régionale, urbaine, de proximité ; cartographie annuelle, journalière, horaire) conduisent cependant à s'interroger sur le choix de la meilleure approche.

A partir des études conduites par l'EMD et l'INERIS depuis près de cinq ans, avec l'appui de collaborations diverses (AASQA, Centre de Géostatistique de l'Ecole des Mines de Paris), et des informations extraites de travaux d'AASQA, ce rapport dresse un bilan préliminaire des différentes méthodes de cartographie du NO₂. Il a pour but d'en évaluer les avantages et les limites et, selon les données disponibles, l'environnement géographique et l'objectif de la cartographie, d'émettre des recommandations sur les choix techniques les plus pertinents. Fondé sur l'expérience acquise et sur l'état actuel des connaissances, ce document ne prétend pas à l'exhaustivité. Ainsi il n'aborde que partiellement la cartographie du NO₂ à haute résolution spatiale et/ou temporelle, qui demeure un sujet d'étude (voir les travaux AIR PROCHE (AFSSET) : Tognet, 2006 ; AIRMARAIX et al., 2006).

La question est traitée par échelle spatiale :

- cartographie régionale / rurale (avec influence routière)
- cartographie urbaine (grandes et petites agglomérations)
- cartographie en zone de proximité routière

puis par échelle temporelle :

- cartographie moyenne saisonnière ou annuelle
- cartographie quotidienne / horaire.

En préambule, un tableau mentionne les méthodes applicables à chaque situation caractérisée par une échelle d'espace et une échelle de temps. Le corps du guide détaille et compare les conditions de leur mise en œuvre et fournit des préconisations et avertissements. Les travaux du LCSQA exploités dans ce document sont répertoriés et, pour certains, synthétisés en annexe.

3. CARTOGRAPHIE DU NO₂: METHODES DISPONIBLES, CRITERES DE COMPARAISON

Le tableau suivant fait l'inventaire des méthodes de cartographie abordées dans ce guide. Certaines d'entre elles ne seront citées ou traitées que brièvement car elles concernent des travaux récents, en cours ou prévus, réalisés à l'initiative du LCSQA ou d'AASQA.

Echelle spatiale →	Régionale / Rurale (avec influence routière)	Urbaine	Proximité routière
Echelle temporelle ↓			
Cartographie moyenne saisonnière ou annuelle	Géostatistique Statistique (corrélations) Modélisation déterministe (eulérienne)	Géostatistique Modélisation déterministe (eulérienne ou gaussienne)	Géostatistique Modélisation empirique ou semi-empirique (modèles de rues)
Couplage géostatistique/modèles de rue			
Cartographie journalière ou horaire	Modélisation déterministe (eulérienne) Couplage modélisation déterministe/géostatistique (analyse)	Géostatistique Modélisation déterministe gaussienne (Urban'air à Strasbourg)	Modélisation semi-empirique (SIRANE à Lyon)

Pour chaque méthode on examinera les points suivants :

- les hypothèses et la théorie sous-jacentes : prise en compte ou non de la physique du problème, degré de simplification des phénomènes, etc.
- les données d'entrée requises/recommandées
- l'aspect des cartographies obtenues : résolution, réalisme, ... ; l'utilisation qu'on peut faire des cartes
- l'évaluation des incertitudes
- la facilité et la durée de mise en œuvre, les compétences nécessaires
- les outils informatiques.

Les commentaires et recommandations correspondants sont étayés d'exemples extraits des études LCSQA.

Les cartographies présentées sont issues :

- de TRANSCHIM, mis en œuvre à l'EMD, et de CHIMERE, mis en œuvre dans PREV'AIR, en ce qui concerne la modélisation déterministe ;
- du logiciel ISATIS (Géovariances) en ce qui concerne la géostatistique.

4. CARTOGRAPHIE A L'ECHELLE REGIONALE / EN ZONE RURALE (AVEC INFLUENCE ROUTIERE)

Ce chapitre a pour objet de recenser et de confronter différentes techniques permettant de cartographier la pollution par le dioxyde d'azote sur une **région** ou en **zone rurale influencée**.

Le terme de région désigne aussi bien une région administrative qu'un domaine de grande superficie (plusieurs milliers de km²) composé d'étendues rurales et de territoires urbanisés.

Une zone rurale est dite influencée si elle est soumise à des sources de pollution externes au domaine (comme un panache de pollution) et à des sources internes localisées (comme des autoroutes).

Il s'agit donc d'évaluer les concentrations de NO₂ à l'échelle de **plusieurs dizaines, voire centaines de kilomètres**.

Pour ces domaines composites, qui englobent des territoires naturels ou agricoles, des axes routiers, des sources ponctuelles de pollution et éventuellement des villes, l'enjeu est de fournir une représentation cohérente avec la physique du phénomène, suffisamment détaillée dans le rendu des variations spatiales de la concentration, et conforme aux exigences de qualité réglementaires.

4.1 CARTES MOYENNES MENSUELLES, SAISONNIERES OU ANNUELLES

Cette partie est plus particulièrement illustrée par les travaux LCSQA-EMD de 2003 (Wroblewski et al., 2003) et LCSQA-ILERIS de 2005 (Malherbe et Cárdenas, 2005) réalisés en collaboration avec LIG'Air.

Dans le premier cas, le domaine étudié, compris entre Orléans (région Centre) et le sud-ouest de la région parisienne, représente une superficie de 4356 km² (66x66 km²). Il a été choisi parce qu'il correspond à une zone rurale influencée par la région parisienne et par des autoroutes. Trois méthodes de cartographie ont été appliquées : géostatistique, statistique, déterministe.

Dans le second cas, le domaine d'étude couvre la région Centre. Il s'étend sur environ 250 km du nord au sud et 200 km d'est en ouest. L'approche géostatistique a été mise en œuvre.

Dans la suite du document, on parlera plus brièvement de « l'étude EMD-2003 » et de « l'étude INERIS-2005. »

4.1.1 Application des méthodes géostatistiques

4.1.1.1 Principe

La géostatistique s'applique à l'estimation de phénomènes d'aspect irrégulier et néanmoins structurés dans l'espace : la pollution de NO₂ entre dans ce cadre (Fouquet, 2003 a).

Si la physique du problème n'est pas considérée en tant que telle, par la résolution des équations de transport des polluants, la géostatistique, à la différence des techniques classiques d'interpolation, tient compte de la variabilité spatiale des données. Sa mise en œuvre repose d'une part, sur la réalisation de mesures et d'autre part, sur la construction d'un modèle probabiliste qui décrit et quantifie la corrélation spatiale du phénomène. La variable d'intérêt, ici la concentration de NO₂, est ensuite estimée par combinaison linéaire des données expérimentales.

Parmi les techniques géostatistiques, celle du **krigeage ordinaire** (géostatistique linéaire monovariable) est l'une des plus employées. L'introduction d'informations secondaires par **cokrigeage** (géostatistique multivariable), **krigeage du résidu**, ou **krigeage avec dérive externe** (géostatistique non stationnaire) a prouvé son efficacité pour accroître la précision des cartographies, notamment dans les zones où les mesures sont plus éparses (Bobbia M. et al., 2001). Elle est également passée dans la pratique (ASPA, 2005 ; ATMO Rhône Alpes, 2005).

Une présentation didactique du krigeage et de son application à l'évaluation de la pollution atmosphérique est fournie dans Fouquet (2003 a). L'annexe 2 propose quelques éléments théoriques utiles à la compréhension de la méthode. Le krigeage étant une technique complexe dans ses détails et ses possibilités, nous renvoyons également à la consultation de mémoires ou rapports (Fouquet, 2006 ; Wackernagel, 2004 ; Rouïl et Malherbe, 2002 ; Perdrix et Houdret, 2000 ; Plaisance et al., 1999) et d'ouvrages spécialisés (Rivoirard, 2003 ; Arnaud et Emery, 2000 ; Armstrong, 1998.)

4.1.1.2 Données d'entrée

Mesures de concentration

Même en présence de variables auxiliaires (cf. ci-dessous), l'application des méthodes géostatistiques requiert des données en nombre important. Par expérience, sur une zone donnée, une cinquantaine de points de mesure se révèle au minimum nécessaire pour :

- évaluer correctement les relations entre la concentration de NO₂ et les autres variables disponibles;
- mettre en évidence et modéliser la variabilité spatiale de la pollution ;
- garantir une certaine fiabilité de la cartographie, la précision de l'estimation se détériorant loin des points expérimentaux.

Ces mesures sont obtenues le plus couramment par des campagnes d'échantillonnage passif qui permettent un bon quadrillage du domaine d'étude. Pour des préconisations sur l'échantillonnage spatial, on peut se reporter à une étude LCSQA-EMD-INERIS (2005, 2006).

D'autre part, la cartographie saisonnière ou annuelle pose la question de l'**échantillonnage dans le temps et de la représentativité temporelle des périodes de mesure**. A ce sujet, des travaux sont en cours au sein du groupe technique *Plans d'échantillonnage et reconstitution des données* et devraient déboucher en 2007 sur des recommandations. Une pratique actuelle, conforme à la durée minimale réglementaire (14% de l'année), consiste à réaliser six à huit quinzaines de mesure par an, soit groupées dans chaque saison (ASPA, ATMO Champagne-Ardenne) soit réparties sur l'année (ATMO Rhône-Alpes). Des tests préalables sur des séries annuelles de stations fixes peuvent orienter le choix des périodes d'échantillonnage.

Exemple

L'étude EMD-2003 utilise les données de deux campagnes de prélèvement conduites sur la zone d'intérêt :

- Première campagne du 9 au 25 juillet 2002, comprenant 71 tubes ;
- Deuxième campagne du 3 au 18 avril 2003, comprenant 98 tubes.

Notons qu'il s'agit de périodes relativement brèves car s'ils visent bien un objectif de surveillance, ces travaux portent avant tout sur l'évaluation et la comparaison de méthodes de cartographie, qui ne nécessitent pas un échantillonnage de plusieurs mois.

Dans l'étude INERIS 2005, les données de pollution de fond sont issues d'une campagne conduite par LIG'Air sur la région Centre. Des tubes à échantillonnage passif ont été exposés par périodes consécutives d'un mois, de janvier à décembre 2005, dans 43 communes et agglomérations (dont six pourvues de deux points de mesure). Les données de 49 sites de mesure ont été ainsi exploitées.

Variables auxiliaires

Sur un grand domaine, **des données auxiliaires sont généralement nécessaires pour améliorer la qualité du krigage**. En effet, à l'échelle de plusieurs dizaines de kilomètres, les données de mesure, à moins d'être très denses, ne représentent qu'une information partielle sur la répartition des concentrations. La carte de krigage ordinaire élaborée sur la région Centre en est une illustration (LCSQA-INGERIS, 2005). Les valeurs estimées, relativement uniformes dans le domaine, augmentent ou diminuent localement autour des sites de mesure, créant des tâches de couleur (Figure 1). Ces dernières représentent les niveaux de fond urbains et leur étalement autour des villes, dans un certain rayon d'influence qui est fonction du variogramme (~ 15 km). En revanche, la cartographie ne dit rien des variations de concentration entre les communes. Or de telles variations existent, si l'on en croit l'échantillonnage plus fin effectué en 2003 au nord-est de la région : les concentrations mesurées montrent bien l'influence des routes (Figure 1).

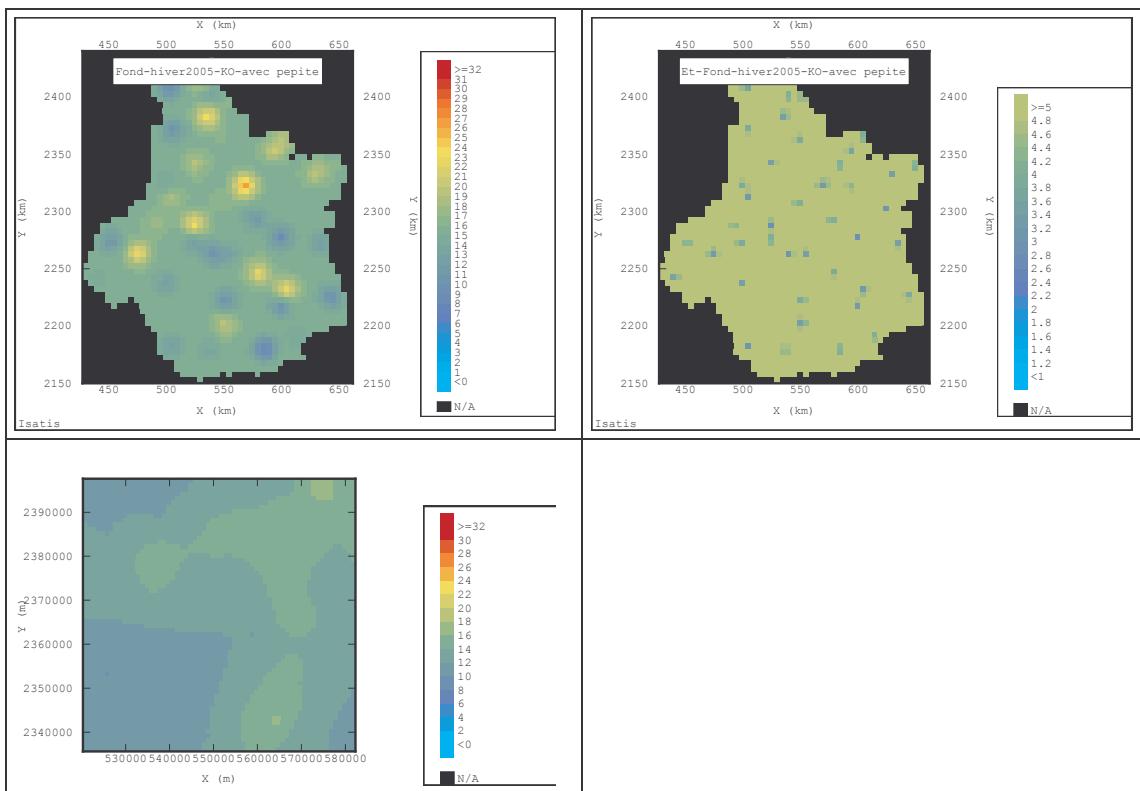


Figure 1 – Estimation de la concentration de NO₂ [µg/m³] par krigeage ordinaire.

Ligne du haut : cartographie régionale (hiver 2005) et écart-type de krigeage.

Entre les points de mesure, la précision de l'estimation se dégrade, l'écart-type de l'erreur de krigeage est sensiblement plus élevé.

Ligne du bas : cartographie sur le domaine EMD (réalisée à partir de la moyenne des campagnes de juillet et de mars).

D'autre part, lorsque l'on travaille sur des domaines de grande étendue, il n'est pas rare de rencontrer des non stationnarités.

Dans ce contexte, des variables auxiliaires soigneusement choisies peuvent pallier efficacement l'insuffisance de l'information initiale et, le cas échéant, aider à la modélisation d'une dérive. L'annexe 3 contient de plus amples renseignements sur la construction de ces variables.

Les études EMD-2003 et INERIS-2005, la lecture de rapports d'AASQA (ASPA, 2005, ATMO Rhône-Alpes, 2005) ou d'articles scientifiques (Ross et al., 2006) nous permettent de proposer, à titre indicatif, une liste de facteurs potentiellement explicatifs :

- Emissions de NOx : densité d'émissions dans un rayon de 1 à 5 km. Ce type de variable peut être obtenu à partir d'un inventaire local des émissions.
- Population : densité de population dans un rayon de 1 à 5 km. Afin que les variables calculées soient assez précises, il est conseillé de répartir préalablement les données communales en fonction de l'occupation du sol ou d'utiliser des données de plus fine échelle (IRIS et îlots).
- Occupation du sol : densité de territoires artificialisés (catégorie 1 de la nomenclature CORINE) dans un rayon de 1 à 5km ;

- Altitude ;
- Distance à la mer pour les régions côtières.

Une transformation logarithmique des ces variables (de la forme $\ln(1+X)$) a souvent pour effet de linéariser la relation avec la concentration de NO_2 .

Exemple

Sur la région Centre, les variables prises en compte sont :

- l'inventaire kilométrique des émissions de NOx élaboré en 2001 par LIG'Air pour l'année 1999 et réactualisé en 2005 (les données sont disponibles à une résolution temporelle mensuelle et/ou annuelle suivant le secteur d'activité.);
- les données de population communales (en 2005), remplacées en 2006 par les données IRIS et îlots ;
- les données d'occupation du sol de la base CORINE Land Cover.

4.1.1.3 Mise en œuvre

De façon générale, toute application de la géostatistique nécessite une analyse exploratoire des données rigoureuse et approfondie. Nous attirons plus particulièrement l'attention sur les points suivants :

Sélection des données de concentration

Toutes les données mesurent-elles le même type de pollution ? L'histogramme des concentrations, la nuée variographique, les nuages de corrélation entre les concentrations et les variables auxiliaires sont autant de moyens graphiques pour apprécier l'homogénéité de la population de données et, si besoin est, pour décider d'écartier des mesures sans rapport avec cette population.

Choix de la période de représentation

Les données collectées permettent de réaliser des cartographies moyennes pour chacune des saisons pendant lesquelles ont eu lieu des mesures.

Ce traitement différencié des données hivernales et estivales est d'autant plus recommandé que souvent, les deux saisons se distinguent non seulement par les valeurs observées mais aussi par la variabilité spatiale des concentrations.

L'obligation de satisfaire à la réglementation requiert néanmoins l'élaboration de cartes moyennes annuelles. Deux possibilités se présentent :

- l'analyse géostatistique est réalisée sur la moyenne de toutes les mesures de l'année ;
- on prend la moyenne des estimations saisonnières.

Elles n'ont de sens que si le plan d'échantillonnage temporel permet d'approcher la moyenne annuelle avec une précision satisfaisante. Sur ce point, on pourra se

rapporter aux travaux du GT *Plans d'échantillonnage et reconstitution de données* dont les résultats seront publiés au second semestre 2007. La seconde procédure ouvre la voie au **cokrigeage temporel** dont le principe et les avantages possibles sont décrits plus loin (cf. Méthodes de krigage ci-après, § 5.1.1.3, et annexe 4).

Exemple

En ce qui concerne la région Centre, la continuité des mesures sur l'année 2005 a permis de calculer des concentrations moyennes annuelles sans problème de représentativité temporelle. Chaque saison au sens strict (hiver : janvier-mars et été : juillet-septembre) a fait aussi l'objet d'une analyse séparée (cf. annexe 5).

Sélection des variables auxiliaires

L'emploi de données secondaires implique de rechercher des variables auxiliaires les plus corrélées possible aux mesures des échantillonneurs passifs. **Ces variables ont-elles été calculées selon différents supports spatiaux?**

La corrélation peut y être en effet très sensible (Figure 2), avec en outre un gain important par rapport à l'usage de la variable d'origine. Le coefficient de corrélation entre la concentration moyenne hivernale et l'inventaire kilométrique* migré sur les 49 sites de la région Centre est ainsi de 0,6. Avec la densité d'émissions calculée dans un rayon de 2km autour des points de prélèvement*, le coefficient de corrélation atteint 0,9 et le nuage de corrélation est plus resserré.

*Il s'agit plus exactement de la transformée logarithmique de ces variables.

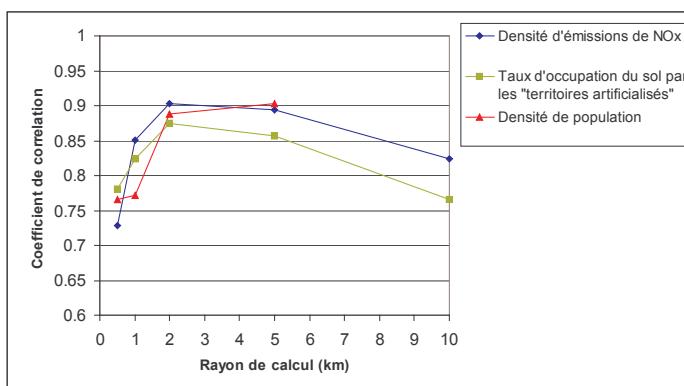


Figure 2 – Région Centre. Evolution du coefficient de corrélation entre la concentration hivernale de NO₂ (49 sites de fond, hiver 2005) et différentes variables de densité, en fonction du rayon de calcul de ces variables. Un maximum est atteint à 2km pour les émissions et l'occupation du sol et à 5km pour la population.

Sur le domaine rural exploré en 2003 une étude similaire a conduit à sélectionner la transformée logarithmique des émissions et celle de la densité d'émissions dans un rayon de 5 km. Ces variables auxiliaires présentent une corrélation de 0,77, respectivement 0,81, avec les mesures (moyennes des campagnes de juillet 2002 et mars 2003).

Si le calcul de densités permet d'améliorer la corrélation entre la concentration de NO₂ et les variables auxiliaires, il faut rester vigilant dans l'utilisation de ces variables qui peuvent influencer sensiblement la cartographie. En particulier, on

prendra soin de ne pas lisser excessivement l'information auxiliaire en considérant des rayons de calcul de très grande taille.

D'autre part, par des **calculs de corrélation et des analyses en composantes principales**¹ (cf. exemple en annexe 5), on examinera la corrélation linéaire entre les variables auxiliaires sélectionnées afin de voir si elles sont redondantes ou non.

Si l'on veut mener à bien la modélisation et le krigeage, les variables auxiliaires doivent être calculées, selon le support spatial adopté pour chacune, aux points de mesure d'une part, aux centres des mailles de la grille d'estimation d'autre part. Comme il est noté dans l'annexe 3, le pas de la grille d'estimation ne coïncide pas nécessairement avec la taille de ce support.

Modélisation - Méthode de krigeage

Dans un domaine régional englobant différents types de territoires (zones naturelles, agricoles, urbaines, routes...), l'utilisation de variables auxiliaires est un moyen efficace d'accroître la précision de l'estimation et d'obtenir une représentation plus réaliste. Au préalable, une analyse monovariable suivie d'un krigeage ordinaire se révèle cependant utile pour appréhender les caractéristiques du jeu de données et pouvoir, par la suite, apprécier l'apport des variables explicatives.

L'annexe 2 présente succinctement les diverses techniques de krigeage, leurs hypothèses et le type de modélisation requis. Dans le cas fréquent où les variables auxiliaires sont disponibles sur un maillage dense, on peut faire appel aux techniques du *krigeage du résidu* (aussi appelé *krigeage des résidus*), du *krigeage avec dérive externe* ou du *cokrigeage colocalisé*.

Les deux premières relèvent de la géostatistique non stationnaire. Elles supposent l'existence d'une **relation linéaire entre la concentration de NO₂ et la ou les variables explicatives** qui représentent la tendance du phénomène.

- Ces méthodes sont similaires dans leur principe. Le modèle de variogramme est celui des résidus de la régression multilinéaire de NO₂ en fonction des variables auxiliaires.
- On les distingue l'une de l'autre parce qu'en voisinage glissant, le krigeage avec dérive externe a la propriété d'ajuster localement les coefficients de la régression. En contrepartie, il n'est plus possible de construire directement le variogramme des résidus, ce qui complique la modélisation.
- L'hypothèse d'une relation linéaire entre les variables donne du poids à la dérive, notamment dans les zones où la densité de mesures du NO₂ est plus faible (Wackernagel, 2004). Les cartographies de fond

¹ L'ACP permet d'observer les corrélations, similarités ou oppositions entre les variables auxiliaires et de mettre en évidence des familles de variables. Afin d'analyser plus finement les corrélations, une ACP peut être ensuite effectuée au sein de chaque famille. L'ACP permet de réduire le nombre initial de variables auxiliaires en ne conservant que certaines variables représentatives de l'ensemble. Elle fournit en outre des variables synthétiques (les facteurs) qui peuvent être éventuellement employées comme variables auxiliaires.

élaborées sur la région Centre pour l'hiver 2005 l'illustrent : avec la densité d'émissions, le tracé des routes ressort nettement ; avec la densité de population, il apparaît de manière plus diffuse. Evaluées aux 49 sites d'échantillonnage, ces deux variables auxiliaires sont pourtant très corrélées entre elles mais dans les zones rurales dépourvues de données de NO_2 , là où leur influence est précisément accrue, cette corrélation se détériore (cf. Figure 38 en annexe 5). Sans échantillonnage complémentaire du NO_2 dans ces zones, le choix de l'une ou l'autre carte ne peut que reposer sur l'expertise.

- Ainsi le krigeage du résidu ou le krigeage avec dérive externe sont recommandables dans les situations où la corrélation linéaire entre la concentration de NO_2 et les variables auxiliaires est avérée et de bonne qualité (Rivoirard, 2002).

Pour des corrélations linéaires plus modérées, lorsque l'hypothèse sous-jacente aux précédentes méthodes n'est pas vérifiée, le cokrigeage colocalisé peut être préféré (ASPA et Géovariances, 2005).

- Cette méthode, qui procède de la géostatistique stationnaire, considère toutes les variables sur le même plan, supposant qu'il existe entre elles une corrélation spatiale. Elle impose d'ajuster un modèle de corégionalisation (modèles mono- et bivariables), tâche parfois malaisée du fait de l'aspect peu structuré des variogrammes (cf. Figure 30 en annexe 5). Mais à la différence du krigeage avec dérive externe, elle laisse à l'utilisateur la maîtrise des paramètres du modèle (ASPA et Géovariances, 2005).
- Le cokrigeage est dit colocalisé (ou multicolocalisé) du fait que la variable auxiliaire n'est utilisée qu'au point d'estimation et aux points de mesure de la variable principale. Cette simplification, imposée par des contraintes numériques, représente une approximation par rapport au cokrigeage ordinaire (Rivoirard, 2002, 2004).
- Lorsque la concentration de NO_2 est très corrélée aux variables auxiliaires, les résultats du cokrigeage colocalisé sont normalement proches de ceux du krigeage du résidu/avec dérive externe.
- En général, on se limite à un seul cofacteur afin de simplifier la modélisation. Quand plusieurs variables explicatives non redondantes sont disponibles, une pratique est de les combiner préalablement par régression multilinéaire (ASPA et Géovariances, 2005).

Finalement les nuances qui séparent ces trois méthodes sont assez peu importantes (cf. annexe 2). Le choix d'une méthode se fait plutôt au cas par cas, selon les corrélations et les structures spatiales observées, les résultats de la validation croisée et l'expérience de l'utilisateur.

Notons que le krigeage du résidu ou krigeage avec dérive externe est compatible avec le « cokrigeage temporel » évoqué précédemment. Dans cette technique proposée par le Centre de Géostatistique (Fouquet, 2003c), les concentrations saisonnières ne sont pas estimées de façon disjointe mais par cokrigeage ordinaire. Des variables auxiliaires propres à chaque saison peuvent être

introduites en dérive externe. La moyenne des estimations estivale et hivernale ainsi obtenues fournit l'estimation de la moyenne annuelle. Cette méthode de cokrigeage garantit la cohérence des estimations estivale, hivernale et annuelle. Si elle fait peu gagner en précision lorsque les échantillonnages de l'été et de l'hiver sont identiques et que pour chaque saison, on utilise les mêmes variables auxiliaires (cas de la région Centre), elle prend tout son intérêt lorsque les échantillonnages se distinguent (volontairement, ou à cause de données manquantes) ou que les variables explicatives sont différentes. Toute l'information disponible sur l'année est ainsi prise en compte.

Exemple

Dans l'étude EMD-2003, la prise en compte des émissions ou de la densité d'émissions par cokrigeage colocalisé a permis d'affiner la représentation cartographique sur la zone d'étude.

Dans l'étude INERIS-2005, des résultats plus satisfaisants que le krigage ordinaire ont été obtenus avec le krigage du résidu ou le krigage avec dérive externe.

Grille d'estimation

Le choix de la maille d'estimation dépend notamment :

- des dimensions du domaine ;
- de la résolution spatiale des variables d'origine (concentrations, informations secondaires) : on veillera à ne pas définir une maille d'estimation trop fine par rapport à cette résolution, afin de ne pas exagérer artificiellement le niveau de détail de la carte.

A l'échelle régionale, une résolution de l'ordre du kilomètre peut convenir. Dans l'étude INERIS-2005, une maille de 5km de côté a été adoptée sur l'ensemble de la région. Dans la partie nord-est du domaine, cette grille a été resserrée selon une maille de 1km de côté (=résolution de l'inventaire des émissions, également employée dans l'étude EMD-2003).

4.1.1.4 Résultats

Afin d'illustrer notre propos, nous présentons dans ce chapitre quelques unes des cartographies établies sur le nord-est (EMD-2003) ou la totalité (INERIS-2005) de la région Centre. On rappelle que ces estimations sont fondées :

- pour la première étude: sur la moyenne des concentrations mesurées en juillet 2002 et avril 2003 et sur l'inventaire des émissions élaboré en 2001 ;
- pour la seconde étude: sur la moyenne hivernale, estivale ou annuelle des concentrations mesurées en 2005 et soit sur l'inventaire des émissions réactualisé en 2005, soit sur les données de population IRIS et îlots.

Domaine 66 km x 66 km

1) Carte de cokrigeage colocalisé avec l'emploi de la variable $\log(\text{emission}+1)$ comme variable auxiliaire (fig 2)

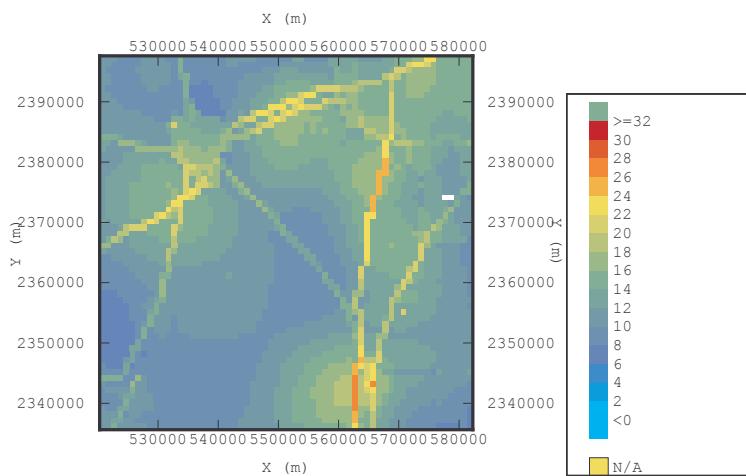


Figure 3 - Estimation des concentrations annuelles de NO_2
(variable externe : émissions)

Cette première carte fait ressortir de manière très fine les zones de fortes émissions. La contribution des axes routiers apparaît clairement.

2) Carte de cokrigeage colocalisé avec l'emploi de la variable $\log(\text{densité d'émission}+1)$. On considère la densité d'émissions dans un rayon de 5 km autour d'un point de mesure. Cette approche conduit à lisser l'influence des émissions (fig 3).

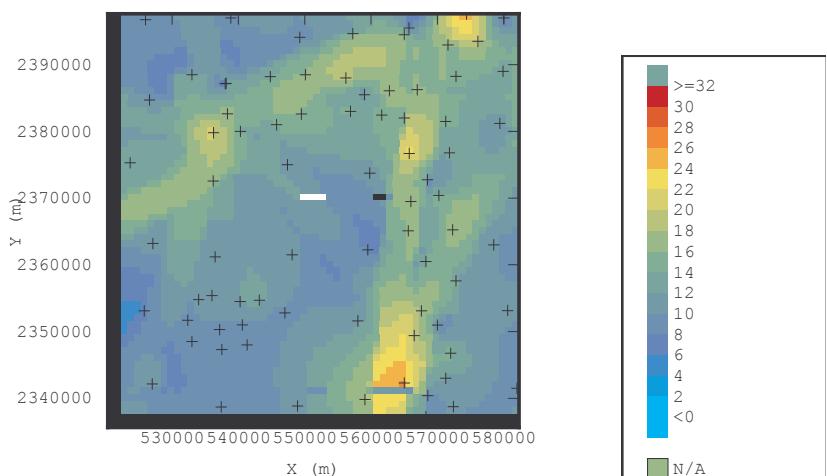


Figure 4 - Estimation des concentrations annuelles de NO_2
(variable externe densité d'émissions)

Domaine régional

Comme on l'a mentionné au paragraphe 4.1.1.3 (rubrique sur les méthodes), la variable en dérive a une influence marquée sur l'aspect de la cartographie, notamment dans les zones qui séparent les mesures. **Pour les types de territoires qui ne sont pas représentés dans l'échantillonnage** (ici les zones rurales), **et pour lesquels, de ce fait, la validité du modèle liant la concentration de NO₂ et la(les) variable(s) auxiliaire(s) n'est pas garantie, l'interprétation et l'utilisation de la cartographie nécessitent une certaine prudence.**

Par suite de l'utilisation des variables auxiliaires, l'écart-type de l'erreur de krigage est faible et relativement uniforme sur le domaine.

Les cartes ne sont ici présentées que pour l'hiver ; toutes les autres cartes sont jointes en annexe 5.

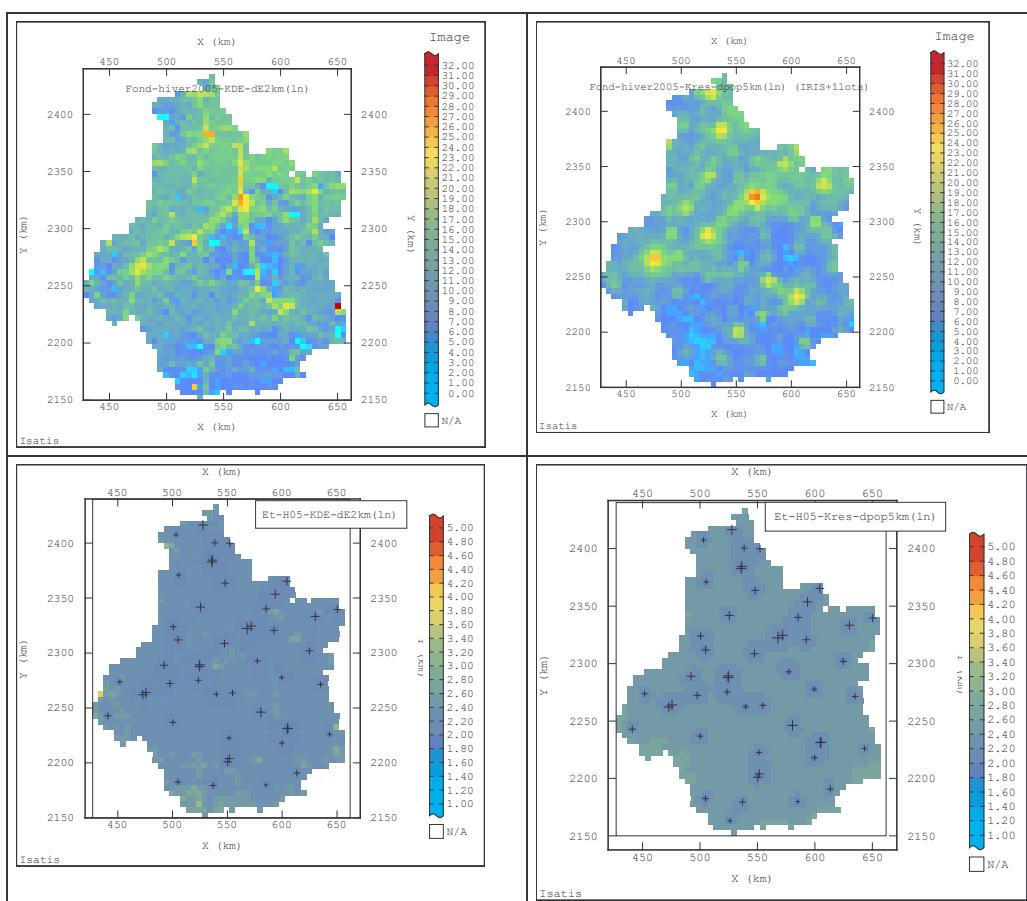


Figure 5 – Estimation des concentrations de NO₂ de l'hiver 2005 (ligne du haut) et écart-type de l'erreur de krigage (ligne du bas) [$\mu\text{g}/\text{m}^3$]. Colonne de gauche : krigage avec dérive externe en utilisant la densité d'émissions de NOx à 2km ; colonne de droite : krigage du résidu en utilisant la densité de population à 5 km.

Incertitude d'estimation

Les directives européennes incitent à la réalisation de cartographies de la pollution atmosphérique mais elles imposent également des objectifs de qualité pour les données qui sont produites, notamment en matière d'incertitude. En ce qui concerne la modélisation, cette incertitude représente l'écart relatif entre l'estimation et la mesure. Pour le NO₂, elle ne doit pas dépasser 30% en modélisation.

- L'écart-type de l'erreur d'estimation (ou « écart-type de krigage »), σ_K , ne représente pas cette incertitude, bien qu'il constitue un indicateur de précision. Fonction du modèle de corrélation spatiale et de la configuration des points de mesure, l'écart-type de krigage σ est d'autant plus élevé que le phénomène est moins structuré dans l'espace. En particulier, il est très sensible à l'effet de pépite. D'autre part, ses variations dans la zone d'étude montrent les endroits où l'erreur d'estimation est plutôt faible et ceux où, par manque de données, elle est plutôt élevée. Ainsi est-il surtout exploité qualitativement.
- Le meilleur moyen d'accéder à l'erreur relative d'estimation est de disposer d'un jeu de données indépendant, qui réponde aux caractéristiques suivantes :
 - Il n'a pas servi à l'élaboration de la carte ;
 - Les données sont de même typologie que la pollution cartographiée ;
 - Il correspond à la même période de temps que la période représentée ;
 - Il se répartit sur tout le domaine d'étude, et sur toute la gamme supposée de concentration.

Il est rare de disposer d'un tel jeu. Dans quelques situations favorables, on peut néanmoins réserver une partie des données à la validation. Il s'agit des cas où l'échantillonnage initial est très dense et où les caractéristiques spatiales de la concentration et les corrélations avec les variables auxiliaires permettent d'écartier des mesures sans dégradation notable de la cartographie (cf. étude sur la région Champagne-Ardenne, Wroblewski et al, 2006). Les erreurs relatives obtenues quantifient alors l'incertitude d'interpolation spatiale. Si les campagnes ne se font pas en continu sur l'année, elles n'incluent pas l'erreur liée à l'échantillonnage temporel.

Plus couramment, on peut calculer les erreurs relatives d'estimation aux sites de mesure fixes qui n'ont pas été pourvus de tubes. Cette évaluation est limitée par le nombre souvent réduit de stations et par le fait que les données d'échantillonneurs passifs et les données d'analyseurs automatiques sont de nature différente. Un avantage cependant est de pouvoir quantifier l'écart entre estimation et mesure non seulement sur la période d'échantillonnage mais aussi sur la période qu'on cherche à représenter (saison ou année).

- Si la validation croisée sert principalement :
 - o à s'assurer de l'adéquation entre le modèle géostatistique et les données expérimentales ;
 - o à trancher entre plusieurs modèles possibles de variogrammes et, si besoin est, entre plusieurs voisinages de krigage,

le calcul des erreurs relatives de validation croisée constitue un moyen pratique, quoique partiel, d'apprécier l'incertitude d'estimation.

Cependant il faut être attentif : les statistiques de validation croisée dépendent de l'implantation des données. Elles reflètent principalement la qualité de l'estimation dans les zones les plus riches en données, là où les points de mesure sont proportionnellement plus nombreux.

Des exemples de calcul d'erreurs relatives sont fournis en annexe 5.

Remarque : la fourniture d'une estimation sous forme d'un intervalle de confiance sort du champ d'application de la géostatistique linéaire, considérée dans ce rapport.

Comparaison avec un seuil fixé par la réglementation

En un point x du domaine d'étude, le krigage fournit une estimation de la concentration de NO_2 par combinaison linéaire des données expérimentales ainsi qu'un écart-type de l'erreur d'estimation. Mais à cause du lissage opéré par le krigage, de la généralisation à l'année de mesures souvent temporaires, et du fait que la loi de distribution de l'erreur est inconnue, cette estimation ne suffit pas pour affirmer que la concentration est supérieure ou non au seuil annuel de 40 $\mu\text{g}/\text{m}^3$. La variable qu'il convient d'estimer est une probabilité de dépassement de seuil, et il faut alors faire appel aux méthodes de la géostatistique non linéaire. L'utilisation de ces méthodes, et plus spécialement de trois d'entre elles, est décrite dans Cárdenas (2004) :

- l'espérance conditionnelle ;
- le conditionnement uniforme, variante de la précédente, qui permet d'introduire une composante non stationnaire dans le calcul ;
- les simulations conditionnelles, qui visent à reproduire la variabilité spatiale de la variable d'étude en fournissant plusieurs versions possible de la réalité.

Les deux premières ont l'avantage d'une plus grande rapidité de mise en œuvre mais toutes permettent de prendre en compte des variables auxiliaires, soit par cokrigage, soit par dérive externe. Ces méthodes reposent néanmoins sur un certain nombre d'hypothèses souvent difficiles à vérifier et des précautions d'usage s'imposent.

L'application de la géostatistique non linéaire à la pollution régionale n'est pas encore pratique courante. L'expérience du LCSQA et des AASQA dans l'utilisation de ces méthodes est limitée. Celles-ci n'ont pas été mises en œuvre dans les travaux de 2003 et 2005 sur la région Centre. Citons une étude exploratoire de l'ASPA (2005) dans laquelle des simulations conditionnelles servent à évaluer les populations d'Alsace potentiellement exposées à des concentrations supérieures à $40 \mu\text{g}/\text{m}^3$.

4.1.1.5 Représentation de la pollution de proximité

L'intégration de la pollution proche de grands émetteurs dans une cartographie de fond n'est pas une préoccupation nouvelle. Ce problème s'est posé aux AASQA dès qu'elles ont commencé à s'intéresser à la géostatistique (cf. Rapport d'assistance LCSQA, Malherbe et Rouil, 2002). Toutefois les différences d'échelle temporelle et spatiale entre les deux types de pollution et les difficultés méthodologiques qui en découlent pour la cartographie ont conduit, dans un premier temps, à favoriser la représentation du fond.

Récemment, des travaux sur la question ont été réalisés par le LCSQA dans un contexte régional (région Centre, Malherbe et Cárdenas, 2005), et au sein du groupe AIR PROCHE dans un contexte urbain (Toulon). Dans les deux études, l'échelle temporelle utilisée pour la pollution de fond est conservée (la saison ou l'année) ; l'analyse porte sur l'affinement de l'échelle spatiale et la correction de l'estimation à proximité des routes.

Quelle que soit l'approche, il est nécessaire de disposer d'un nombre suffisamment important de mesures de proximité. S'agissant de la région Centre, des données ont été recueillies pendant toute une année en dix-neuf points de mesure de proximité.

Nous rappelons ci-dessous la démarche proposée pour la région Centre :

1. analyse exploratoire des données de fond, étude géostatistique et estimation par krigage de la concentration de fond régionale, avec usage de variables auxiliaires ;
2. calcul des différences entre les concentrations mesurées aux sites de proximité et les niveaux de fond estimés en ces points. Ces écarts sont considérés comme les valeurs expérimentales d'une augmentation de concentration liée au trafic. Cet accroissement s'exprime comme une fonction linéaire (établie par régression) de la densité d'émissions de NOx (ici dans un rayon de 1km).
3. affinement du maillage dans des bandes situées de part et d'autre des principaux axes routiers et correction de l'estimation de fond par ajout de l'écart modélisé.

Cette procédure a été testée sur l'agglomération de Blois et d'Orléans. Si visuellement, et selon l'expertise de LIG'Air, les représentations obtenues sont plus conformes à la réalité du terrain, aucune validation poussée n'a pu être encore réalisée.

Moyennant des données complémentaires, en particulier des données de sites disposés perpendiculairement aux routes, un modèle de correction plus précis pourrait inclure le type de voie et la distance à la route .

Des améliorations sont donc possibles. Le principal défaut de la méthode réside actuellement dans les contraintes informatiques. En effet, malgré les capacités des SIG, le calcul de la densité d'émissions sur un maillage fin se heurte à un problème de ressources informatiques (saturation de la mémoire). Par suite, il doit être effectué en segmentant la région, ce qui représente une opération longue.

4.1.2 Corrélations

4.1.2.1 Principe

La prédiction des concentrations de NO₂ par des relations de corrélation est d'usage fréquent dans les pays anglo-saxons. Ces relations ne reposent pas sur la connaissance précise des mécanismes physiques ou chimiques sous-jacents mais sur l'étude des liens statistiques entre variables. Afin de mener à bien cette analyse et d'assurer la validité statistique des corrélations spatiales et temporelles mises en évidence, de nombreuses données d'observation sont au préalable nécessaires.

Du fait qu'elle imbrique différentes échelles d'espace, la méthode testée en 2003 s'applique spécifiquement à des domaines régionaux ou ruraux influencés. Elle a été mise au point au Royaume-Uni par John R. Stedman du bureau d'études AEAT, pour le compte de différents organismes du royaume chargés de la surveillance de la qualité de l'air. Elle se base sur des relations empiriques, permettant le calcul des moyennes annuelles en oxydes d'azote NO_x = NO + NO₂ et dioxyde d'azote NO₂ à partir des émissions d'oxydes d'azote. Cette méthode a été appliquée de 1996 à 1999 au Royaume-Uni et a permis la réalisation de cartes des teneurs moyennes annuelles en NO_x et NO₂ dans l'air ambiant, avec une résolution de 1 km² (Stedman, 1996, 2000).

La méthode repose sur l'hypothèse de base selon laquelle la concentration moyenne annuelle en oxydes d'azote dans l'air ambiant serait la somme de trois composantes :

- une composante résultant des émissions situées à longue distance ;
- une composante résultant des émissions locales ;
- une composante résultant des émissions de proximité.

$$(NO_x) = (NO_x)_{rural} + \Delta(NO_x)_{urbain} + \Delta(NO_x)_{trafic} \quad (1)$$

où :

- (NO_x) est la concentration moyenne annuelle en oxydes d'azote dans l'air ambiant, c'est-à-dire la somme de la concentration moyenne annuelle en monoxyde d'azote (NO) et de la concentration moyenne annuelle en dioxyde d'azote (NO₂) dans l'air ambiant, exprimée en microgrammes d'équivalent NO₂ par mètre cube d'air (μg/m³ équiv. NO₂) ;

- $(NOx)_{rural}$ est la concentration moyenne annuelle en NOx dans l'air ambiant d'une zone rurale (contribution des émetteurs lointains) exprimée dans la même unité ;
- $\Delta(NOx)_{urbain}$ est l'augmentation de concentration en NOx due à la contribution des émetteurs locaux, dans une zone urbaine de fond ; elle est exprimée dans la même unité ;
- $\Delta(NOx)_{trafic}$ est l'augmentation de concentration en NOx (par rapport à la concentration urbaine de fond) due à la contribution du trafic, dans une zone de proximité automobile ; elle est exprimée dans la même unité.

Les sites implantés à proximité d'émetteurs industriels importants sont exclus du champ d'application de cette relation. De même, les sites de trafic en dehors d'une zone urbanisée, tels que les abords d'une autoroute en rase campagne, sont traités différemment. Pour des sites en zone rurale, que Stedman définit comme une zone éloignée de plus de 10 km de toute source importante de NOx, la relation (1) se simplifie et équivaut à :

$$(NOx) = (NOx)_{rural} \quad (2)$$

La concentration moyenne annuelle en NOx en milieu rural est alors estimée à partir de la concentration moyenne annuelle en NO₂ en milieu rural (NO₂)_{rural}, selon la relation :

$$(NOx)_{rural} = 1,2 \times (NO_2)_{rural} \quad (3)$$

Le coefficient de proportionnalité de 1,2 utilisé dans la relation (3) a été calculé à partir des données fournies par un analyseur automatique par chimiluminescence, installé dans une zone rurale du Royaume-Uni. Précisons, qu'au Royaume-Uni, la concentration moyenne annuelle en dioxyde d'azote est estimée par interpolation spatiale sur tout le territoire, grâce au réseau secondaire de mesure du dioxyde d'azote par échantillonnage passif. Ce réseau comprend trente-deux sites de mesure répartis sur tout le territoire. Les mesures sont effectuées en continu, tous les 14 jours (durée d'exposition des échantilleurs). Les moyennes annuelles mesurées sont ensuite interpolées sur une grille de résolution de 20km x 20km (Atkins et Lee, 1995).

Pour les sites des zones urbaines, qui sont définies par Stedman comme des zones éloignées de plus de 5 m de toute voie de circulation et situées à moins de 10 km de sources de NOx, la relation (1) se simplifie et équivaut à :

$$(NOx) = (NOx)_{rural} + \Delta(NOx)_{urbain} \quad (4)$$

A la suite d'une étude de corrélation, $\Delta(NOx)_{urbain}$ est supposé proportionnel aux émissions de NOx situées dans un carré de 5 km x 5 km centré autour du site considéré, selon la relation :

$$\Delta(NOx)_{urbain} = k_{urbain} \times E_{NOx \text{ 25 km}^2} \quad (5)$$

où :

- k_{urbain} est un coefficient de proportionnalité empirique, assimilé à un coefficient de dispersion ;

- $E_{NOx\ 25\ km^2}$ est la somme des émissions annuelles de NOx dans un carré de 25 km² de surface, centré sur le site considéré, exprimée en tonnes d'équivalent-NO₂ pour 25 km² et par an (t.(25 km²)⁻¹.an⁻¹ équiv. NO₂).

La surface à considérer dans la relation (5) pour faire la somme des émissions locales de NOx a été choisie parmi plusieurs valeurs comprises entre 1 et 400 km². La valeur de 25 km² est celle pour laquelle le coefficient de corrélation linéaire entre $\Delta(NOx)_{urbain}$ et $E_{NOx\ 25\ km^2}$ est le plus grand.

Le calcul de la concentration moyenne annuelle en NO₂ est alors effectué à partir de relations empiriques reliant la concentration moyenne annuelle en NO₂ et la concentration moyenne annuelle en NOx. Deux types de relation ont tour à tour été utilisés par Stedman :

- premièrement, une relation linéaire de type :

$$(NO_2) = a \times (NOx) + b \text{ où } a \text{ et } b \text{ sont des constantes (6)}$$

- deuxièmement, une relation de type puissance :

$$(NO_2) = c \times (NOx)^d \text{ où } c \text{ et } d \text{ sont des constantes (7)}$$

Il faut noter que les coefficients constants (Kurbain, a, b, c et d) des relations (5), (6) et (7) ont des valeurs constantes pour tout le Royaume-Uni sauf pour le centre-ville de Londres. Dans cette mégapole exceptionnellement étendue, les conditions d'émission et de dispersion atmosphérique des polluants sont en effet considérées comme très différentes de celles qui sont rencontrées partout ailleurs au Royaume-Uni (Carslaw et al., 2001).

Pour les sites des zones urbaines de proximité automobile, que Stedman définit comme des zones urbanisées éloignées de moins de 5 m d'une voie de circulation importante, la relation (1) doit être considérée dans sa totalité.

$\Delta(NOx)_{trafic}$ est supposé proportionnel aux émissions linéaires de NOx dues au trafic, exprimées en moyenne annuelle par kilomètre de voie de circulation à proximité, selon la relation :

$$\Delta(NOx)_{trafic} = k_{trafic} \times E_{NOx/km} \text{ (8)}$$

où :

- k_{trafic} est un coefficient de proportionnalité empirique (il est déterminé à partir des émissions) ;
- $E_{NOx/km}$ est la somme des émissions annuelles de NOx pour 1 km de la voie de circulation située à proximité du site considéré, exprimée en kg.km⁻¹.an⁻¹ équiv. NO₂.

Pour les sites de proximité automobile situés en dehors d'une zone urbanisée (en pleine campagne), Stedman préconise d'appliquer la relation (8) mais en divisant par quatre les émissions liées au trafic. Le fait de ne considérer qu'un quart des émissions permet de rendre compte de la dispersion atmosphérique plus importante en milieu ouvert, dans une zone non construite, comparativement à un milieu plus confiné, en zone urbaine construite.

4.1.2.2 Données d'entrée

L'application de la méthode des corrélations s'appuie sur des mesures en zone rurale et sur des données d'émissions.

Dans l'étude de 2003, les résultats de deux campagnes de mesure de NO₂ effectuées par échantillonnage passif (cf 4.1.1) ont été utilisées afin de cartographier les teneurs moyennes annuelles de NO₂ en zone rurale. Les valeurs obtenues ont été multipliées par le facteur 1,2 pour en déduire les teneurs moyennes annuelles de NO_x.

D'autre part, considérant l'inventaire kilométrique des émissions de NO_x, la somme des émissions dans un carré de 25 km² autour de chaque centre de maille permet d'accéder à l'augmentation de concentration due au niveau urbain.

4.1.2.3 Mise en œuvre

Etant donné les émissions de NO_x sur le domaine de 66km² à 66km², il est possible de calculer E_{NO_x}, 25km² pour chaque maille de 1km² à 1km² incluse dans une zone de 62km² à 62km². L'augmentation de NO_x due à ces émissions se déduit alors de la relation (9) pour chaque maille J du domaine (la relation est appliquée pour toutes les mailles, que celles-ci soient réellement en zone urbaine ou non) :

$$\Delta(\text{NO}_x)_J \text{ urbain} = k_{\text{urbain}} \times E_J \text{ NO}_x 25 \text{ km}^2 \quad (9)$$

avec $\Delta(\text{NO}_x)_J \text{ urbain}$ exprimé en $\mu\text{g/m}^3$ équiv. NO₂
et $E_J \text{ NO}_x 25 \text{ km}^2$ en $\text{t.(25km}^2\text{)}^{-1}.\text{an}^{-1}$ équiv. NO₂.

Deux valeurs de k_{urbain} , déterminées par Stedman pour le Royaume-Uni, ont été testées :

- l'une relevant de la méthode 1998 ($k_{\text{urbain 98}} = 0,0311$) ;
- l'autre de la méthode 1999 ($k_{\text{urbain 99}} = 0,0394$).

En chaque maille J, la concentration de NO_x est calculée selon la relation (10), en additionnant la teneur en NO_x en zone rurale, préalablement estimée par krigeage ordinaire, et l'accroissement en NO_x dû aux émissions:

$$(\text{NO}_x) = (\text{NO}_x)_{\text{rural}} + \Delta(\text{NO}_x)_{\text{urbain}} \quad (10)$$

Remarque : k_{urbain} peut être réadapté au domaine d'étude par un calage en zone urbaine, en utilisant les mesures des stations fixes (fond urbain) et l'inventaire actualisé des émissions de NO_x.

4.1.2.4 Résultats

Nous avons obtenu deux cartes de la concentration annuelle en NO₂ sur la zone, respectivement pour les coefficients 98 et 99 (figure n°4) :

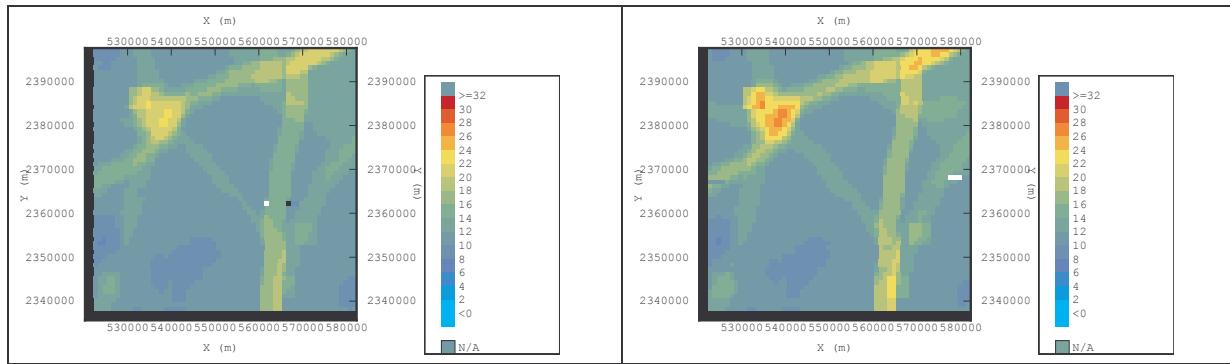


Figure 6 - Cartographies annuelles en NO₂ suivant la méthode de corrélation

On remarque l'influence importante de la valeur du coefficient empirique k : le coefficient de 1999 induit une plus grande amplitude de concentration ; l'écart-type est plus grand et les valeurs extrêmes sont mieux représentées. On peut aussi noter l'influence de la prise en compte des émissions sur une zone de 25 km² : la carte est traversée de bandes de concentrations plus élevées, qui décrivent l'impact des routes.

4.1.3 Modélisation déterministe

4.1.3.1 Principe

Sur des domaines étendus (région), il est pertinent d'utiliser des modèles déterministes dits de « chimie-transport » et notamment des modèles fondés sur une représentation eulérienne du transfert atmosphérique des polluants. Les modèles eulériens reposent sur une approximation tridimensionnelle des équations de la mécanique des fluides (équation de Navier-Stokes) pour la partie dynamique et de l'équation de conservation de la masse pour la dispersion des polluants. Le maillage tridimensionnel peut se composer de parallélépipèdes de même taille ou de tailles différentes (maillage structuré), ou d'autres éléments disposés de façon non régulière, tels que des tétraèdres (maillage non structuré). Les équations sont intégrées numériquement sur ce maillage, de façon à calculer une valeur numérique de chaque variable intégrée (vitesse, température, turbulence, concentration...) en chaque sommet ou centre de maille. Des techniques numériques souvent très élaborées permettent d'aboutir à la solution recherchée par des processus itératifs. Les modèles tridimensionnels eulériens peuvent réaliser des simulations de très grande qualité, s'ils sont utilisés convenablement : qualité du maillage, ajustement des paramètres numériques, qualité des données d'entrée.

Dans l'étude EMD-2003, le code 3D eulérien de transport-dispersion TRANSCHIM-R (Coria UMR 6614, INSA de Rouen) a été choisi parce qu'il présente les avantages suivants :

- C'est un code universitaire, en conséquence aisément modifiable dès lors que les paramétrisations ont été bien identifiées et comprises. Cette ouverture lui permet d'être ajusté au cas d'étude et transposé à d'autres sites.
- La résolution numérique des équations de transport utilise un maillage horizontal et vertical cartésien qui peut être affiné localement dans les trois dimensions.

- Les principes généraux de la méthode de résolution sont identiques à ceux qui sont utilisés dans le code TRANSCHIM-1 du modèle de transport eulérien SUBMESO.
- La chimie gazeuse peut éventuellement être activée. Elle est basée sur le système réactionnel MOCA. Ce dernier peut être utilisé dans sa version intégrale ou simplifiée en sélectionnant le nombre d'espèces et de réactions que l'on souhaite prendre en compte.

En contrepartie, il est très long à mettre en œuvre si la résolution du maillage est élevée, ce qui constraint le nombre de modélisations réalisables.

Sur de grands domaines, une autre possibilité est offerte par l'utilisation quotidienne de CHIMERE dans des systèmes opérationnels de simulation de la qualité de l'air, qu'il s'agisse de PREV'AIR à l'échelle de la France ou de plates-formes régionales ou multirégionales mises en œuvre dans des AASQA.

La simulation des concentrations horaires de NO_2 repose sur une représentation tridimensionnelle eulérienne des processus de transport, dépôt et chimie des polluants (cf. 4.1.3.1). Notons que des modèles de chimie-transport tels que CHIMERE ont été initialement conçus pour modéliser l'ozone et non le NO_2 .

Les données d'entrée requises sont :

- un inventaire des émissions ;
- des données d'occupation du sol ;
- des données météorologiques ;
- des conditions aux limites et conditions initiales.

L'inventaire des émissions constitue une donnée limitante dans la modélisation déterministe du NO_2 . En effet, comme le NO_2 a plutôt un impact local, les résultats de la modélisation sont très sensibles à l'inventaire utilisé (Rouïl, 2005). De la qualité de ce dernier dépend donc étroitement la qualité des cartographies.

L'archivage des sorties simulées de NO_2 (éventuellement analysées, *i.e.* corrigées statistiquement en fonction des mesures aux stations, si l'on dispose de données d'observation adéquates) permet, à l'issue d'une saison ou d'une année entière, d'établir des cartes moyennes. Un exemple est fourni ci-dessous pour la France (Figure 7).

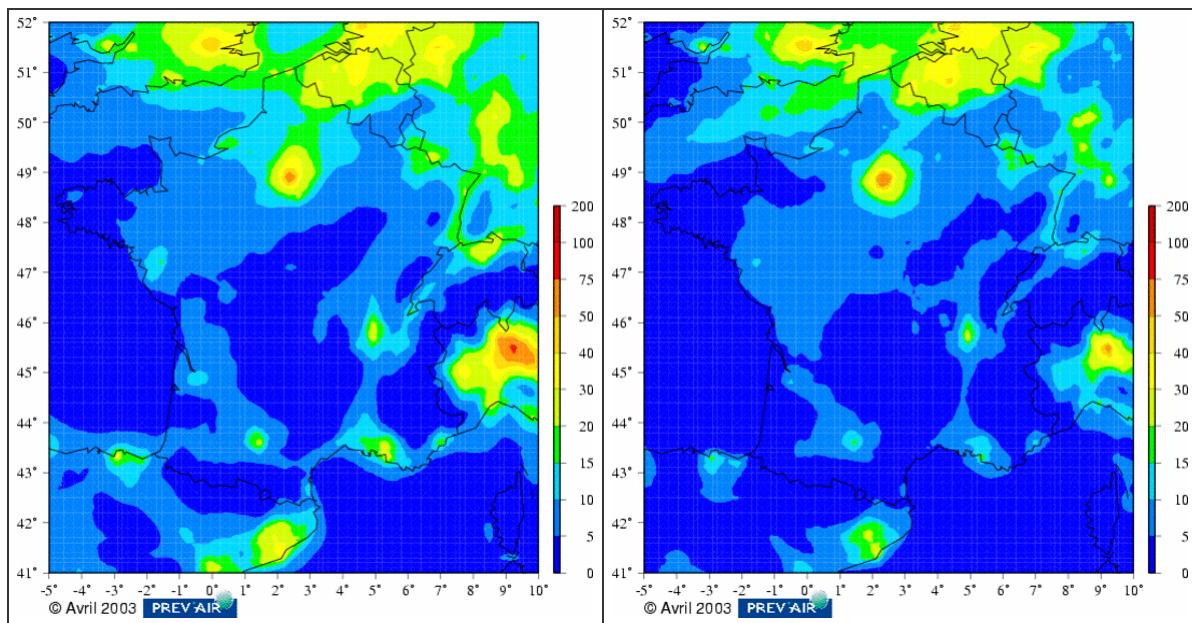


Figure 7 – Cartographies saisonnières de NO_2 sur la France issues du système PREV'AIR (CHIMERE-France) [$\mu\text{g}/\text{m}^3$]. A gauche hiver 2006 ; à droite : été 2006.

4.1.3.2 Mise en œuvre

Du fait de temps de calcul trop élevés, il n'est pas possible de simuler avec TRANSCHIM chaque heure de l'année. La modélisation eulérienne de concentrations moyennes annuelles de NO_2 a donc été faite en deux temps :

- un calage préalable ;
- l'application du modèle à différentes situations météorologiques soigneusement définies.

Le code TRANSCHIM-R a été ainsi utilisé sur le domaine de 66 km de côté (maille de 1km). Dans une première étape, il a été calé sur ce domaine grâce à l'exploitation des données des deux campagnes d'échantillonnage. Dans une seconde étape, une rose des vents représentative de la moyenne annuelle sur la zone a été prise en compte. Elle comporte 16 classes de direction de vent et 4 classes de vitesse. L'application du modèle à chaque situation composant cette météorologie a permis d'évaluer la concentration annuelle de NO_2 en tout point du maillage.

4.1.3.3 Données d'entrée

1) Emissions :

Les émissions de NO_2 proviennent de deux sources différentes, premièrement de l'inventaire réalisé par LIG'AIR sur son domaine de compétence, et de l'inventaire EMEP pour la partie du domaine AIRPARIF. L'inventaire LIG'AIR est construit suivant la méthodologie CORINAIR, sa date de réalisation est 1999 et sa résolution est le kilomètre carré. Les rejets d'oxydes d'azote sont fournis en équivalent NO_2 . L'inventaire EMEP

date de 1999 et sa maille de résolution est d'environ 1600 km². La partie nord-est du domaine a été reconstruite à partir de l'inventaire EMEP pour la plupart des sources, il a été complété par la prise en compte directe du trafic automobile sur les grands axes.

2) Conditions aux limites :

Les conditions aux limites sont fournies par le modèle de prévision CHIMERE : quatre verticales sont extraites des sorties de CHIMERE et distribuées sur les quatre faces du domaine de calcul.

3) Météorologie :

Les conditions de vent et de température proviennent de calculs statistiques effectués à partir des données de mesure de Météo France sur la zone. La rose comporte 16 classes de direction de vent et 4 classes de vitesse. Les données météorologiques prises en compte sont les valeurs d'Arpège à 10 m. La reconstruction des profils de vent se fait ensuite en utilisant un profil académique.

4.1.3.4 Résultats

Le calcul déterministe sur la zone permet de produire une cartographie annuelle de NO₂ (fig n°5).

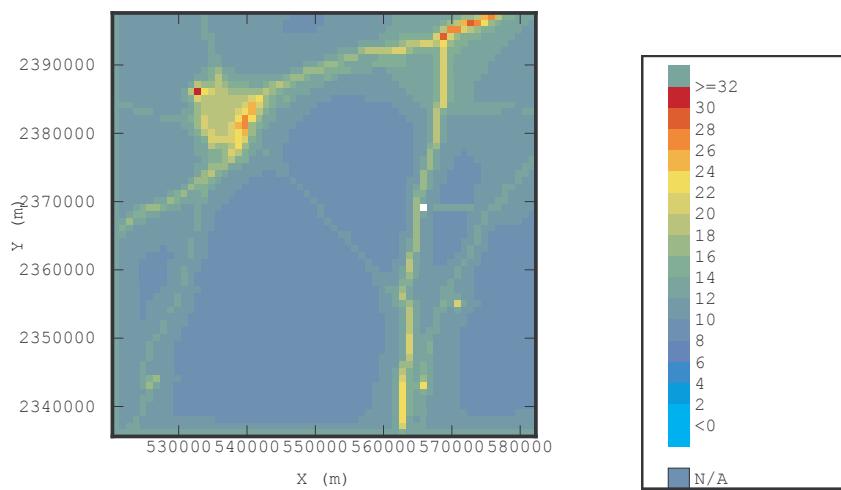


Figure 8 - Cartographie annuelle de NO₂ par modélisation déterministe

4.1.4 Comparaison des approches

Sur un domaine rural incluant des axes routiers et une agglomération, cette partie a pour objet de comparer les types d'estimation suivants :

1. estimation géostatistique à partir de campagnes de mesure locales (*i.e.* réalisées dans la zone) ;
2. estimation géostatistique à partir de mesures effectuées dans toute la région ;

3. estimation par corrélation à partir de mesures locales ;
4. modèle déterministe de haute résolution (TRANSCHIM, résolution 1km) calé sur des mesures locales ;
5. modèle déterministe de plus grande échelle (CHIMERE-France, utilisé dans PREV'AIR, résolution 10 km).

Le domaine considéré est la portion Nord-Est de la région Centre étudiée en 2003 par l'EMD.

NB : dans les travaux de 2003, il inclut également une petite partie de l'Île-de-France ; dans ceux de 2005, celle-ci n'est pas prise en compte.

Deux stations de mesure fixes sont implantées dans ce domaine et peuvent servir de sites communs de validation à l'ensemble des méthodes :

- station urbaine de fond Chartres Fulbert ;
- station urbaine de fond Lucé (commune en périphérie de Chartres).

Cette validation est partielle à cause du petit nombre de stations. D'autre part, cet exercice de comparaison est restreint par le fait que les périodes d'étude ne sont pas concomitantes: 2003 pour les travaux EMD (méthodes 1, 3 et 4), 2005 pour les travaux INERIS (méthode 2) et 2006 pour PREV'AIR (méthode 5). Si les valeurs mesurées durant ces années sont différentes, avec un rapport d'environ 0,8 entre les moyennes annuelles de 2005 (ou 2006) et 2003, en revanche, les profils annuels de concentration conservent approximativement la même allure et la corrélation temporelle entre les deux sites reste élevée. On ne détecte donc pas de modification importante dans le comportement du polluant.

4.1.4.1 Analyse ponctuelle

Pour les différentes méthodes considérées, la comparaison avec les teneurs moyennes annuelles de NO₂ calculées à partir des mesures en station fournit des résultats tout à fait compatibles avec les recommandations de la directive européenne (tableau 1). En effet les incertitudes maximales autorisées pour les zones faiblement exposées (concentration inférieure au seuil d'évaluation minimal, 26 µg/m³) sont de 75 % en analyse objective et de 30% en modélisation. Bien que la base de comparaison soit restreinte (uniquement deux stations), les écarts relatifs de chacune des techniques se trouvent en deçà de cette limite de 30%.

La modélisation de plus grande échelle conduit seule à des écarts supérieurs à 30% (entre 25% et 50% selon la station et la saison). Les sous-estimations observées se justifient par la résolution de l'inventaire des émissions utilisé. A une résolution de 10km, les contributions des sources locales de NOx sont fortement atténuées. Les techniques d'analyse, qui n'ont pas été ici mises en œuvre, mais qu'il est prévu d'appliquer en 2007 sur des domaines régionaux, peuvent être un moyen d'affiner localement l'estimation (cf. 4.2.3).

Tableau 1 - Comparaison des différentes techniques en moyenne annuelle sur deux points du domaine – année 2003

Stations	Chartres-Fulbert		Chartres-Lucé	
	Concentration en ($\mu\text{g}/\text{m}^3$)	Ecart relatif En %	Concentration en ($\mu\text{g}/\text{m}^3$)	Ecart relatif En %
Moyenne annuelle mesurée (2003)	20.90		22.09	
Calcul déterministe de haute résolution	26.05	25	25.17	14
Calcul géostatistique (à partir de mesures locales, avec la densité d'émissions)	15.23	-27	16.08	-27
Calcul Stedman 98 (à partir de mesures locales)	26.00	24	20.39	-8
Calcul Stedman 99 (à partir de mesures locales)	26.30	26	22.20	0.0

Tableau 2 – Mêmes statistiques que ci-dessus pour une **estimation géostatistique** fondée sur des données régionales (pollution de fond) – année 2005. Le krigeage est effectué aux points mêmes des stations.

	Chartres-Fulbert				Chartres-Lucé			
	Hiver 2005 (jan-fév-mars)	Eté 2005 (juil-août-sept)	Année 2005 estimée par la moyenne des estimations été et hiver	Année 2005 estimée à partir des moyennes expérimentales annuelles	Hiver 2005 (jan-fév-mars)	Eté 2005 (juil-août-sept)	Année 2005 estimée par la moyenne des estimations été et hiver	Année 2005 estimée à partir des moyennes expérimentales annuelles
Moyenne mesurée ($\mu\text{g}/\text{m}^3$)	21,3	12,3	17,3	17,3	22,3	14,3	19,1	19,1
Moyenne estimée ($\mu\text{g}/\text{m}^3$)	24.63	14.05	19.34	18.69	25.38	14.56	19.97	19.30
Ecart relatif (%)	15,5	13,9	9,5	5,8	13,7	1,6	20,4	16,4

Tableau 3 – Mêmes statistiques que ci-dessus pour une estimation géostatistique fondée sur des données régionales (pollution de fond) – année 2005. Les valeurs estimées sur la grille sont migrées aux stations.

	Chartres-Fulbert				Chartres-Lucé			
	Hiver 2005 (jan-fév-mars)	Eté 2005 (juil-août-sept)	Année 2005 estimée par la moyenne des estimations été et hiver	Année 2005 estimée à partir des moyennes expérimentales annuelles	Hiver 2005 (jan-fév-mars)	Eté 2005 (juil-août-sept)	Année 2005 estimée par la moyenne des estimations été et hiver	Année 2005 estimée à partir des moyennes expérimentales annuelles
Moyenne mesurée ($\mu\text{g}/\text{m}^3$)	21,3	12,3	17,3	17,3	22,3	14,3	19,1	19,1
Moyenne estimée ($\mu\text{g}/\text{m}^3$)	24.0	13.70	18.85	18.27	25.48	14.62	20.05	19.37
Ecart relatif (%)	12,5	11,1	6,7	3,4	14,1	2,0	20,9	16,8

Tableau 4 – Mêmes statistiques que ci-dessus pour le modèle de chimie-transport CHIMERE utilisé dans PREV'AIR – année 2006. Les valeurs estimées sur la grille sont interpolées aux stations.

	Chartres-Fulbert		Chartres-Lucé	
Période d'estimation	Hiver 2006 (15 oct. – 31 déc.)	Eté 2006 (15 juin-15 sept.)	Hiver 2006 (15 oct. – 31 déc.)	Eté 2006 (15 juin-15 sept.)
Moyenne mesurée ($\mu\text{g}/\text{m}^3$)	22,0	13,4	21,0	14,2
Moyenne estimée ($\mu\text{g}/\text{m}^3$)	11,0	10,0	11,0	9,9
Ecart relatif (%)	- 50	-25,4	-47,6	-30,3

4.1.4.2 Analyse spatiale

L'analyse précédente est complétée par une intercomparaison des estimations sur le domaine. Les résultats des différentes techniques se révèlent cohérents entre eux. De façon générale, les valeurs estimées appartiennent, en moyenne, à la même gamme de concentration (Tableau 5). De façon plus détaillée, on peut faire les remarques suivantes :

- Dans la modélisation déterministe de haute résolution, fondée sur un inventaire kilométrique des émissions de NOx, les points les plus pollués ressortent plus nettement qu'avec les autres méthodes. Les gradients de concentration entre les zones de faibles et de fortes émissions sont également plus marqués ; l'amplitude de variation des concentrations est plus grande. La cartographie décrit finement la décroissance des concentrations de part et d'autre des routes.
- La précision du krigage est augmentée par l'introduction d'une variable auxiliaire. Toutefois, avec la densité d'émissions, les cartes de cokrigage colocalisé ou de krigage avec dérive externe prennent un aspect plus diffus que la modélisation déterministe et la gamme de concentration est plus restreinte. Dans les deux études (EMD-2003 et INERIS-2005), la densité d'émissions dans un rayon de 5 km ou 2km respectivement est la plus corrélée à la concentration de NO₂, ce qui s'accorde avec la typologie des tubes, censés mesurer la pollution de fond. Par construction, elle est plus lisse que l'inventaire kilométrique initial.
- Les techniques de corrélation et de géostatistique fournissent des résultats comparables, ce qui n'est pas surprenant puisque dans les deux cas, on suppose l'existence d'une relation linéaire entre la concentration de NO₂ (ou l'incrément de cette dernière par rapport au fond rural) et les émissions de NOx calculées dans un rayon voisin de 2km. On note en particulier la ressemblance entre la cartographie obtenue par la méthode de Stedman, avec le coefficient de 1998, et l'estimation annuelle issue du krigage régional. (Rem. : Les niveaux légèrement plus bas de ce krigage s'accordent en avec la baisse de concentration constatée aux stations fixes entre les années 2003 et 2005. De même, le rapport écart-type/moyenne des données estimées est plus faible mais le domaine n'inclut pas la portion d'Île de France et le nœud routier qu'elle contient). Avec le coefficient de 99, la méthode de Stedman fait ressortir plus intensément la pollution urbaine mais la comparaison avec les données des stations de Chartres laisse penser qu'au centre-ville, les concentrations urbaines de fond sont surestimées (surestimation de 25% environ).

Tableau 5 - Statistiques des concentrations annuelles de NO₂ estimées par les différentes techniques sur l'ensemble du domaine (nord-est de la région Centre et petite partie de l'Île de France, 2003)

	Minimum (μm^3)	Maximum ($\mu\text{g/m}^3$)	Moyenne ($\mu\text{g/m}^3$)	Ecart-type ($\mu\text{g/m}^3$)
Calcul déterministe de haute résolution	14.52	49.45	17.06	2.69
Calcul géostat.(émissions)	6.49	27.84	12.86	3.30
Calcul géostat.(densité d'émissions)	9.59	17.04	12.92	1.69
Calcul Stedman 98	7.69	22.52	12.41	2.70
Calcul Stedman 99	7.81	26.81	12.98	3.41

Tableau 6 - Statistiques sur la zone nord-est de la région Centre des concentrations annuelles de NO₂ extraites de l'estimation régionale par krigage avec dérive externe (2005)

	Minimum en ($\mu\text{g}/\text{m}^3$)	Maximum en ($\mu\text{g}/\text{m}^3$)	Moyenne en ($\mu\text{g}/\text{m}^3$)	Ecart-type en ($\mu\text{g}/\text{m}^3$)
Calcul géostat.(densité d'émissions)	3.46	19.52	11.78	1.77

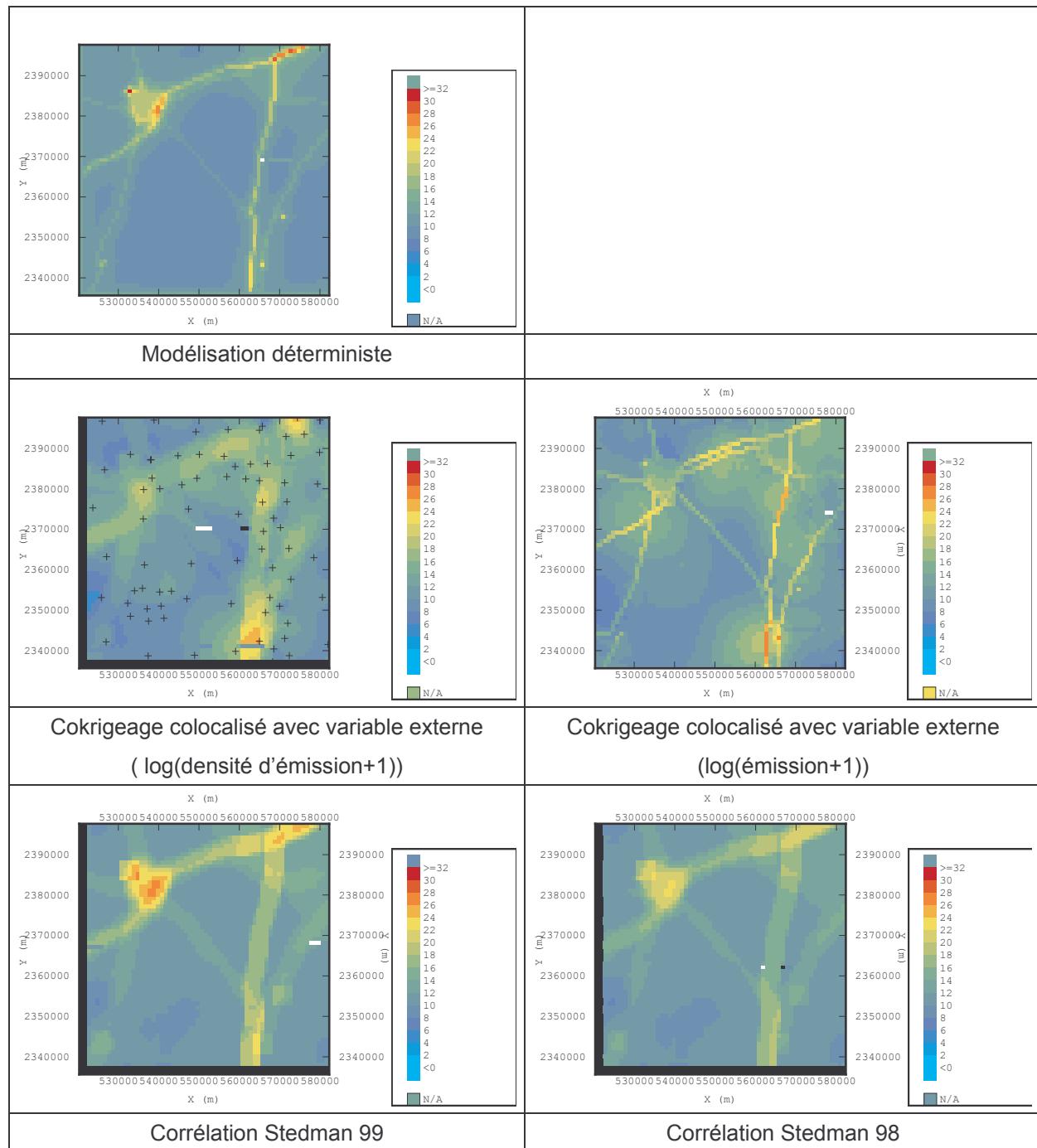
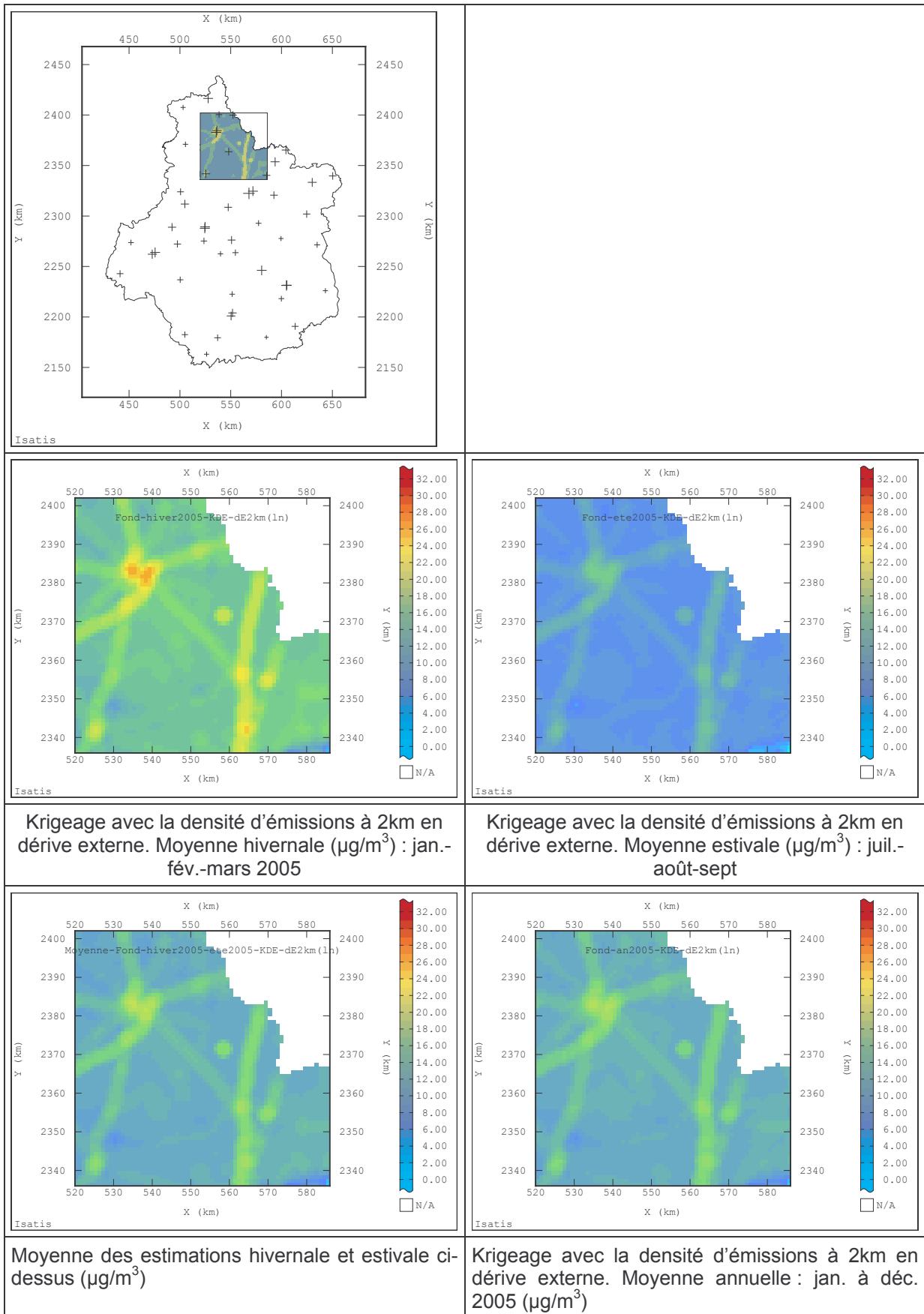


Figure 9 -: cartes annuelles de NO₂ (Etude EMD-2003) en utilisant les différentes techniques



Page précédente :

Figure 10 - : cartes annuelles de NO₂ tirées de l'estimation régionale (Etude INERIS-2005). La zone d'extraction est délimitée dans la carte du haut. Les cartes des écarts-types de l'erreur de krigage sont jointes en annexe 5.

4.1.4.3 Analyse globale

- Modélisation déterministe

La modélisation déterministe de haute résolution (1km pour TRANSCHIM) offre une représentation cohérente de la concentration annuelle en NO₂ sur le domaine étudié, avec un écart inférieur à 25% par rapport aux mesures en station fixe. Elle semble restituer le plus finement la variabilité spatiale des concentrations et décrit relativement bien les « points chauds » présents sur le domaine. Elle est cependant exigeante en données d'entrée et sa mise en œuvre réclame un personnel spécialement qualifié. Elle requiert également un temps de calcul prohibitif pour une représentation exhaustive des situations météorologiques horaires d'une année. Par suite, elle nécessite un choix statistique des conditions météorologiques à représenter pour aboutir aux concentrations saisonnières ou annuelles. L'apport de l'inventaire national spatialisé permet d'envisager un emploi opérationnel de cette technique.

Les modèles de chimie-transport comme CHIMERE, qui fonctionnent quotidiennement dans les systèmes numériques de simulation de la qualité de l'air, délivrent des représentations à une résolution de quelques kilomètres, renseignant sur les grandes orientations des panaches de pollution plus que sur les gradients de concentration locaux. Dans le domaine étudié, un écart de 30% à 50% est ainsi constaté entre les valeurs modélisées (non analysées) et les mesures saisonnières en zone urbaine. Avec des stations de mesure suffisamment nombreuses et représentatives du type de pollution simulé, il est cependant possible, grâce aux techniques d'analyse, d'accorder les cartographies avec les observations locales. Cela n'a pu être ici réalisé. La disponibilité de l'information au pas horaire permet de tenir compte de toutes les situations météorologiques rencontrées dans la saison ou l'année et de calculer des cartes moyennes.

- Géostatistique

Les méthodes géostatistiques reproduisent de façon satisfaisante la variabilité des concentrations sur le domaine, à condition d'intégrer des variables auxiliaires soigneusement choisies. Il faut en effet être attentif dans l'utilisation de ces variables qui doivent présenter une bonne corrélation avec les mesures, sans lisser exagérément l'information dans l'espace. Dans la zone considérée, les estimations s'écartent de moins de 28% des mesures en station fixe. Si l'utilisation des émissions comme variable externe ne devrait pas poser problème, grâce à l'inventaire national spatialisé et aux données des inventaires locaux, le bon usage de la technique requiert un grand nombre de mesures (au moins une cinquantaine). Bien que le coût des échantilleurs soit relativement restreint, il en résulte des coûts en ressources humaines et des coûts analytiques non

négligeables. L'acquisition des principales notions théoriques, jointe à la pratique, permet d'appréhender assez aisément les méthodes géostatistiques et de les appliquer à des situations simples. Pour un usage plus poussé de ces techniques et une interprétation approfondie des résultats, une certaine expertise en analyse des données s'impose, ainsi qu'une bonne connaissance des hypothèses sous-jacentes et des phénomènes de pollution.

- Méthode des corrélations

La méthode de corrélation utilisée (méthode de Stedman) offre également un rendu satisfaisant de la variabilité des concentrations sur la zone d'étude. On note un écart inférieur à 26% par rapport aux mesures en station fixe. Pourvu que les relations de corrélation soient déjà établies, la méthode est peu exigeante en données d'échantillonnage. En effet, la pollution rurale de fond de NO₂ demande d'être estimée sur l'ensemble du domaine. Un krigage sur une trentaine de points suffit à en donner une représentation tout à fait acceptable. Cette méthode, comme les deux précédentes, nécessite l'emploi d'un cadastre d'émissions de NOx. L'arrivée prochaine de l'inventaire national spatialisé en favorisera

l'application sur l'ensemble du territoire national, en particulier dans des zones ne disposant pas d'inventaire local. La technique des corrélations est robuste et aisée à mettre en œuvre. Il faut cependant veiller à mettre à jour annuellement les coefficients des régressions.

Le tableau suivant reprend les principaux points de comparaison entre les méthodes :

	Modélisation déterministe	Géostatistique	Méthode des corrélations
Données de - concentration (cartographie de fond)	Campagnes de mesure réalisées à des saisons différentes pour le calage du modèle (une cinquantaine de sites de mesure répartis sur le domaine et situés dans des zones de faibles, moyennes ou relativement fortes émissions).	Au minimum une cinquantaine de sites de mesure répartis sur le domaine et situés dans des zones de faibles, moyennes ou relativement fortes émissions. Plusieurs semaines de mesure pour assurer la représentativité temporelle des données ² . Ces données sont utilisées à la fois pour l'ajustement du modèle géostatistique et en entrée du calcul de krigage.	Une trentaine de sites pour cartographier le fond rural. Plusieurs semaines de mesure en hiver et en été pour assurer la représentativité temporelle des données ² . Quelques stations urbaines pour l'ajustement des coefficients k.
Autres données d'entrée	Inventaire des émissions de NOx (de	Utilisation de variables auxiliaires fortement	Inventaire des émissions de NOx.

² Préconisations en cours d'élaboration au sein du GT Plans d'échantillonnage et reconstitution de données.

	Modélisation déterministe	Géostatistique	Méthode des corrélations
	<p>l'ordre du km pour la modélisation haute</p> <p>Données météorologiques : sorties de modèles météorologiques</p>	<p>recommandée : par exemple un inventaire des émissions de NOx et/ou des données de population, si possible à fine échelle (données IRIS+îlots de l'INSEE).</p>	
Analyse de données à réaliser	Analyse statistique des données météorologiques pour définir des situations représentatives des conditions annuelles.	Analyse des corrélations entre variables de pollution et variables auxiliaires.	Mise à jour des coefficients k (adaptation au domaine d'étude, mise à jour après éventuelle réactualisation de l'inventaire des émissions).
Mode d'évaluation de l'incertitude	Comparaison aux stations fixes.	<p>Incertitude le plus souvent approchée par validation croisée et par comparaison aux stations fixes.</p> <p>Données d'observation généralement insuffisantes pour évaluer l'incertitude d'estimation spatiale/temporelle sur tout le domaine.</p> <p>Attention aux types de territoires pour lesquels le lien supposé entre concentration et variable auxiliaire n'a pu être vérifié.</p>	<p>Comparaison aux stations fixes.</p> <p>Pas d'évaluation possible dans le reste du domaine à moins de disposer de données complémentaires mesure en milieu urbain.</p>

	Modélisation déterministe	Géostatistique	Méthode des corrélations
Précision attendue	<p>Restitution fine des gradients de concentration.</p> <p>Dans cette étude, écart relatif entre estimations et mesures annuelles compris entre 14% et 25% aux stations urbaines considérées, pour une modélisation de haute résolution.</p>	<p>Précision accrue par l'introduction de variables auxiliaires ; représentation réaliste de la répartition spatiale de la pollution et des accroissements de concentration au voisinage des villes et des sources (axes routiers).</p> <p>D'après l'expérience, écart relatif entre estimation et mesure inférieur à 30% sur la majeure partie du domaine, en hiver et sur l'année. Estimation souvent moins précise en été (concentrations plus basses et phénomène moins structuré).</p> <p>Dans cette étude, écart relatif entre estimations et mesures annuelles compris entre 5% et 21% aux stations urbaines considérées.</p>	<p>Rendu satisfaisant de la variabilité spatiale des concentrations, à condition que les coefficients k utilisés soient adaptés au domaine.</p> <p>Dans cette étude, écart relatif entre estimations et mesures annuelles compris entre 0% et 30% aux stations urbaines considérées (plus faible pour une petite commune).</p>
Prise en compte de la pollution de proximité	<p>Directement représentée si l'inventaire des émissions et le modèle utilisés sont de très haute résolution.</p> <p>Imbrication possible d'un modèle de rue dans les autres cas.</p>	<p>Incrémation possible de la concentration, liée statistiquement à la densité d'émissions de NOx et éventuellement à d'autres paramètres : distance à la route, type de voie, milieu... (Malherbe et Cárdenas, 2005)</p> <p><u>Données complémentaires</u> requises : mesures de proximité en nombre suffisant pour établir une relation statistique, données fines d'émissions, caractéristiques des routes, ...</p>	<p>Incrémation possible de la concentration, proportionnelle aux émissions de NOx sur la route la plus proche ; le coefficient est différent selon l'environnement (urbain/rural).</p>
Durée de mise en œuvre	<p>Il faut considérer un mois de calage initial.</p> <p>Une semaine pour la mise en forme des</p>	<p>Une semaine pour la mise en forme des données et le calcul des variables auxiliaires aux points</p>	<p>(Une semaine pour la détermination de nouveaux coefficients à chaque mise à jour de l'inventaire.)</p>

	Modélisation déterministe	Géostatistique	Méthode des corrélations
	<p>données d'entrée du modèle</p> <p>Une semaine pour faire tourner le modèle dans les situations météorologiques nécessaires à l'évaluation d'une moyenne annuelle.</p>	<p>de mesure.</p> <p>Une à deux semaines pour analyser en profondeur les données, construire et évaluer le modèle géostatistique.</p> <p>Quelques jours de plus si la pollution de proximité est prise en compte.</p> <p>Temps supplémentaire requis par le calcul des variables auxiliaires sur la grille d'estimation.</p> <p>Les temps de calcul induits par à l'affinement du maillage le long des routes constituent une limitation à la prise en compte de la pollution de proximité sur une région entière.</p> <p>Possibilité de travailler sur des portions de domaine.</p>	<p>Une semaine pour la mise en forme des données et le calcul des émissions sur le domaine.</p> <p>Une semaine pour analyser les données, construire et évaluer le modèle</p>
Compétences requises	Compétences pointues en modélisation déterministe.	<p>Connaissances fondamentales de statistique et géostatistique linéaire mono- et multivariables.</p> <p>Expérience dans l'ajustement des paramètres du variogramme et du krigeage.</p>	Connaissances fondamentales de statistique.
Logiciels	Code eulérien tridimensionnel.	<p>Système d'information géographique pour le traitement des données auxiliaires.</p> <p>Logiciel de géostatistique tel qu'Isatis (Géovariances) afin de pouvoir conduire une analyse approfondie des données et explorer, si cela nécessaire, différents types de modélisation.</p>	Outil statistique.

4.2 CARTOGRAPHIE JOURNALIERE OU HORAIRE

4.2.1 Application des méthodes géostatistiques ou statistiques

Aucune expérience de cartographie journalière du NO₂ par géostatistique ou statistique n'a été relevée à l'échelle régionale. Théoriquement, rien n'empêche d'utiliser ces techniques. En pratique, la principale entrave réside dans les données d'entrée puisque les échantillonneurs passifs cumulent généralement quatorze jours de données. L'idéal serait que le domaine compte un nombre suffisant de stations automatiques de mesure du NO₂ et que celles-ci se répartissent régulièrement dans l'espace. Or ces stations se concentrent généralement dans les zones urbaines et périurbaines des principales agglomérations de la région, des campagnes d'échantillonnage par moyen mobile pouvant ponctuellement (dans les communes rurales par exemple) et temporairement compléter ce réseau.

Face à ce problème, qui se pose également à l'échelle urbaine (cf. §5.2.1), une méthode de cartographie journalière de NO₂ a été proposée (Cárdenas et Perdrix, 2005, 2006). Elle exploite :

- les données de campagnes d'échantillonnage réalisées antérieurement sur l'agglomération, et les corrélations entre ces données et les mesures de stations fixes ;
- les informations apportées par des variables auxiliaires telles que les émissions ou la population ;
- la capacité des logiciels de géostatistique, ici Isatis, d'automatiser le krigeage.

Transposer cette méthodologie à une région est tout à fait envisageable. A cette échelle, la définition de stations virtuelles pourrait se révéler plus délicate car l'existence de corrélations élevées entre des sites d'échantillonnage ruraux et des stations majoritairement urbaines ou périurbaines paraît moins assurée. Ce type d'approche sera considéré en 2007, non pas cependant dans un contexte purement géostatistique, mais dans le système PREV'AIR, en combinaison avec CHIMERE (voir les perspectives développées ci-après, §4.2.3).

4.2.2 Modélisation déterministe

Actuellement, plusieurs systèmes opérationnels de modélisation déterministe délivrent des simulations de NO₂ à un pas de temps horaire :

- à l'échelle nationale, et selon une résolution d'environ 10km, le système PREV'AIR ;
- plus localement, et à plus haute résolution (quelques kilomètres), des plates-formes régionales ou interrégionales de modélisation (AIRES, ESMERALDA).

Les principes de la modélisation et les données d'entrée sont décrits au paragraphe 4.1.3.1. Comme il y est mentionné, la qualité des cartes obtenues est d'autant meilleure que l'inventaire des émissions utilisé est plus précis.

Exemple

Pour reprendre l'exemple de la région Centre, on peut évoquer le système OCARINA qui exploite quotidiennement les simulations de PREV'AIR et d'ESMERALDA. Selon l'expertise de LIG'Air, les cartes de modélisation obtenues donnent une idée de la tendance générale des concentrations de NO₂ sur la région. Si elles montrent l'influence du panache d'Île-de-France, en revanche, elles ne permettent pas de voir les phénomènes de production locaux ; l'information produite n'est pas assez fine pour cela.

4.2.3 Perspective : couplage modélisation déterministe/géostatistique

En tenant compte des limites précédemment évoquées, une étude LCSQA de 2007 sera consacrée au couplage entre les sorties de PREV'Air et le krigage, afin de cartographier localement, sur un département ou une région, les concentrations journalières, et si possible horaires, de NO₂.

Une telle étude avait été entreprise en 2004 à l'échelle de la France, selon la méthodologie d'analyse adoptée pour l'ozone (Honoré et Malherbe, 2004). Aucun obstacle théorique ne s'oppose à la production de cartes analysées du NO₂, pourvu que les sites d'observation soient représentatifs de la maille du modèle et bien répartis dans le domaine. A cause du manque de stations rurales, les résultats obtenus en 2004 n'avaient pas donné satisfaction.

Le travail de 2007 cherchera à pallier ces faiblesses grâce à l'expérience acquise sur la cartographie automatique du NO₂ en agglomération (cf. §5.2.1.1). En particulier, on évaluera la possibilité de créer des stations virtuelles dans un domaine régional, par corrélation entre des sites temporaires et des stations fixes ; l'apport des variables auxiliaires sera aussi attentivement examiné.

5. CARTOGRAPHIE EN ZONE URBAINE

5.1 CARTES MOYENNES SAISONNIERES OU ANNUELLES

5.1.1 Application des méthodes géostatistiques

Régulièrement, en complément de la surveillance assurée par le réseau de stations fixes, les AASQA réalisent des campagnes de mesure dans les agglomérations et établissent des cartographies par géostatistique. Elles ont ainsi gagné une bonne maîtrise du krigage et de son application à la cartographie urbaine. C'est pourquoi nous nous limitons ici à quelques points importants et aux avancées méthodologiques qui méritent d'être signalées.

5.1.1.1 Principe

Le principe de l'estimation par krigage est le même qu'en cartographie régionale (cf. §4.1.1.1).

5.1.1.2 Données d'entrée

Mesures de concentration

Selon l'expérience acquise et les tests réalisés dans une étude sur l'échantillonnage (Wroblewski et Malherbe, 2005), un minimum d'environ 45 à 60 points de fond et une maille d'échantillonnage de 1 à 1,5 km de côté sont nécessaires pour conduire une analyse exploratoire approfondie et représenter convenablement la variabilité des concentrations urbaines. Ces exigences sont généralement respectées pour les agglomérations importantes ; le nombre de sites urbains ou périurbains s'élevait ainsi à 75 pour Mulhouse (Fouquet, 2003a, 2003b), 50 à 100 pour Montpellier selon la saison (Le Loch, 2003), 79 pour Rouen (Cori, 2005) ; 83 pour Toulon (AIRMARAIX et al., 2006), 77 pour Bordeaux (AIRAQ, 2005).

D'un point de vue temporel, ces études montrent **l'intérêt d'effectuer des mesures pendant plusieurs quinzaines afin de couvrir diverses situations de pollution**. En effet, il arrive souvent que dans une même saison (été ou hiver), les quinzaines de mesure présentent des caractéristiques différentes les unes des autres.

Variables auxiliaires

Même si l'échelle de travail est réduite par rapport à une cartographie régionale, et la densité de points de mesure (nb de sites/km²) plus élevée, **l'introduction d'informations secondaires peut améliorer significativement l'estimation** (Fouquet, 2006). Les variables auxiliaires les plus pertinentes peuvent dépendre de la saison, du contexte géographique, des caractéristiques de la ville... Cependant, de l'expérience française (AIRMARAIX et al., 2006 ; Cárdenas et Perdrix, 2005 ; Cori, 2005 ; Fouquet, 2003b, 2003c ; Le Loch, 2003) et étrangère (Hochadel et al., 2006 ; Ross et al., 2006 ; Gilbert et al., 2005 ; Sahsuvaroglu et al., 2006 ; Smith et al., 2006 ; Carslaw et al., 2001) on peut tirer une liste de possibles facteurs explicatifs :

- Densité d'émissions de NOx dans un cercle ou une maille de quelques centaines de mètres à quelques kilomètres de rayon ou de côté. Des corrélations relativement élevées sont souvent observées pour une maille de 1 x 1 km² à 1,5 x 1,5 km²; 5 x 5 km² pour la ville de Londres (Carslaw, 2001). Ce type de variable peut être obtenu à partir d'un inventaire local des émissions.
- Densité de population dans un cercle ou une maille de quelques centaines de mètres à quelques kilomètres de rayon ou de côté. Selon les études, des corrélations relativement élevées sont observées pour des rayons de 750 m à 2500 m. L'utilisation de données de fine résolution (îlots) est recommandée.
- Densité d'occupation du sol (bâti, zones découvertes, zones industrielles...) dans un rayon de quelques centaines de mètres (entre 100 et 1000m selon les études) ;
- Longueur de route (en distinguant éventuellement différents types de routes) dans un rayon de quelques centaines de mètres (entre 100 et 1000m selon les études) ;

- Densité de trafic dans un rayon de quelques centaines de mètres ;
- Trafic sur l'axe le plus proche ;
- Distance à la route la plus proche ou à une zone particulière, supposée influencer les concentrations ;
- Altitude.

Remarques :

Afin de fournir une image détaillée de la variabilité intra-urbaine, des variables auxiliaires telles que la densité par maille d'émissions ou de population font parfois l'objet d'une descente d'échelle (interpolation sur une grille plus fine). On veillera à ne pas créer artificiellement une variabilité invérifiable par la mesure.

Une transformation logarithmique des ces variables (de la forme $\ln(1+X)$) a souvent pour effet de linéariser la relation avec la concentration de NO_2 .

5.1.1.3 Mise en œuvre

Les recommandations de mise en œuvre restent les mêmes qu'en cartographie régionale. Les études réalisées par le Centre de Géostatistique sur Mulhouse et Montpellier montrent toute l'importance de l'analyse exploratoire des données de concentration et de variables auxiliaires (Fouquet, 2003b, 2003c Le Loch, 2003, Rouïl et Malherbe, 2003).

Elles mettent en exergue plusieurs problèmes déjà mentionnés pour la cartographie régionale :

- L'estimation des concentrations sur l'ensemble de l'agglomération pose la question la représentativité spatiale des points de mesure et de l'extrapolation dans l'espace du modèle multivariable calé sur ces points ;

⇒ Les relations entre variables de concentration et variables auxiliaires sont-elles représentatives de la zone à estimer ? La question se pose particulièrement lorsque les échantillons sont implantés de façon préférentielle. L'étude de Montpellier montre clairement que le type d'environnement dans lequel ont été exposés les tubes influence significativement les relations entre concentrations et variables explicatives.

- le pouvoir explicatif des variables auxiliaires dépend :

- de leur **résolution spatiale**.

Ce paramètre peut avoir un effet déterminant dans les relations avec les concentrations. Dans l'étude sur Montpellier par exemple, les émissions de NO_x fournies suivant une maille kilométrique sont plus corrélées aux concentrations de NO_2 que les émissions de NO_x fournies avec un pas de 200 m.

- de la **période de l'année**.

Ainsi à Mulhouse, la corrélation entre densité de population et concentration est plus forte en hiver qu'en été. **La plus grande variabilité spatiale de la concentration estivale de NO₂ et la corrélation moins prononcée entre cette concentration estivale et les variables auxiliaires sont des constats fréquents.**

- Les zones de fortes et de faibles concentrations de l'hiver ne coïncident pas nécessairement avec les zones de fortes et de faibles concentrations de l'été et comme il est noté ci-dessus, des différences de structure spatiale apparaissent d'une saison à l'autre. Le modèle d'une saison ne peut donc pas être déduit du modèle de l'autre saison par un simple recalage.
- Comment choisir un modèle adéquat et quelle précision attendre des modèles ? La validation croisée et le calcul des statistiques d'erreur et d'erreur relative apportent une réponse mais il faut garder à l'esprit que ces dernières **réflètent principalement la qualité du modèle dans les zones les plus riches en données**, là où les points de mesure sont proportionnellement plus nombreux. Avant de conclure sur la qualité du modèle, il convient de s'interroger sur les zones dans lesquelles on souhaite que ce modèle soit le plus fiable : centre de l'agglomération ? périphérie de l'agglomération ? Si la ou les zones d'intérêt ne sont pourvues que d'un petit nombre de points de mesure, la validation croisée est insuffisante pour y évaluer le modèle géostatistique.

Application du cokrigeage temporel

Bien que les caractéristiques spatiales de la pollution se distinguent généralement selon la saison (phénomène plus structuré en hiver), plusieurs exemples associés à des contextes géographiques différents, telles les agglomérations de Mulhouse (Fouquet, 2003b), Montpellier (Le Loch, 2003), Bourg-en-Bresse (Cárdenas et Malherbe, 2004) ou Lille (Cárdenas et Perdrix, 2005), font apparaître une corrélation linéaire entre les concentrations d'été et d'hiver. Le caractère récurrent de ce phénomène conduit à préconiser l'application du « cokrigeage temporel » (ou « cokrigeage saisonnier »). Cette méthode développée par le Centre de Géostatistique (Fouquet, 2003c) consiste à estimer la concentration d'une saison par cokrigeage avec l'autre saison, la moyenne des deux estimations saisonnières fournissant ensuite l'estimation annuelle. S'il impose que les échantillonnages d'été et d'hiver aient suffisamment de points de mesure communs, le cokrigeage temporel se révèle d'autant plus opportun que ces échantillonnages se distinguent en partie, soit que des mesures aient été perdues soit qu'ils ne couvrent pas exactement les mêmes zones. Ainsi, même les sites pour lesquels on ne dispose que d'une donnée saisonnière, estivale ou hivernale, sont pris en compte dans l'estimation annuelle. Enfin, l'intérêt de la méthode est aussi de pouvoir tenir compte, dans l'estimation de la moyenne annuelle, des particularités éventuelles de chaque saison, notamment de variables explicatives différentes, et, lorsque le gain lié aux variables auxiliaires est plus important en hiver, d'en faire profiter les estimations de l'été et de l'année.

Un exemple d'application sur Lille, tiré de Cárdenas et Perdrix (2005, 2006) est présenté en annexe 5.

5.1.1.4 Résultats

Les rapports d'étude (AASQA, LCSQA, travaux AIR PROCHE, cf. AIRMARAIX et al., 2006) offrent de nombreux exemples de cartographies réalisées sur des agglomérations de tailles diverses.

Les figures ci-dessous sont extraites d'une étude conduite sur l'agglomération de Bourg-en-Bresse (Cárdenas et Malherbe, 2004). Pendant l'été 2001 (4 quinzaines) et l'hiver 2001-2002 (4 quinzaines), 46 sites de mesure urbains ou périurbains ont fait l'objet d'un échantillonnage. A la suite de l'analyse exploratoire des données, plusieurs modélisations ont été testées et comparées par validation croisée : ajustement par différentes structures de variogramme, prise ou non en compte de la densité de population comme dérive externe, modélisation séparée de chaque saison ou modélisation bivariable. Dans cet exemple où les prélèvements ont eu lieu aux mêmes points en été et en hiver, et où la variable auxiliaire est la même pour les deux saisons, le cokrigeage avec l'hiver n'agit pas sensiblement sur la cartographie. En revanche l'introduction de la densité de population a un impact marqué sur l'estimation. Si l'effet favorable des variables auxiliaires a été constaté dans plusieurs situations, cet exemple rappelle cependant qu'en l'absence de données de validation, lorsque les résultats de validation croisée ne permettent pas de départager les modèles, le choix d'une cartographie conserve une part de subjectivité.

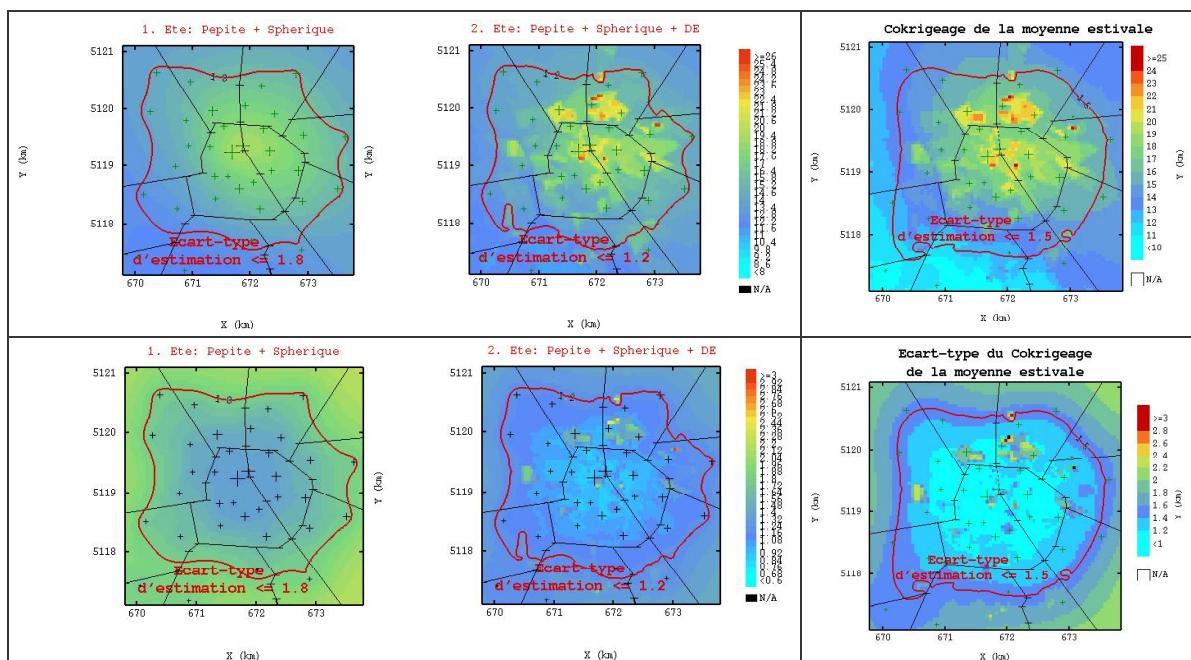


Figure 11 : Cartes d'estimation de la moyenne estivale (ligne du haut) et écart-type de krigeage (ligne du bas) [$\mu\text{g}/\text{m}^3$]. De gauche à droite : krigeage ordinaire, krigeage avec la densité de population en dérive externe, cokrigeage ordinaire avec la moyenne hivernale et la densité de population en dérive externe (attention : échelles de concentration et d'écart-type légèrement différentes pour cette dernière estimation)

S'agissant de la comparaison des estimations avec une valeur seuil, les remarques faites pour la cartographie régionale restent valables. A l'échelle urbaone, on ne relève, là encore, qu'un petit nombre d'applications de la

géostatistique non linéaire (Cárdenas, 2004, Deraisme et Bobbia, 2003, Bobbia, 2001).

5.1.1.5 Prise en compte de la proximité

Dans l'étude INERIS-2005 relative à la région Centre, la démarche proposée pour intégrer la pollution de proximité a été appliquée, pour illustration, à des parties du domaine centrées sur des agglomérations (Blois, Orléans) (cf. 4.1.1.5). Ces calculs, qui exploitent les données de fond et de proximité issues de toute la région et les modèles fondés sur ces données, relèvent ici de l'échelle régionale. Aucune validation poussée dans les villes considérées n'a pu être effectuée, faute de mesures en nombre suffisant sur la même période.

Cette méthodologie reste applicable à l'échelle urbaine proprement dite, c'est-à-dire dans un domaine limité à l'agglomération, et avec des données contenues dans ce domaine. Une variante moins coûteuse en temps de calcul est proposée par Cárdenas (2006) dans l'étude sur la cartographie journalière automatique du NO₂. Elle consiste, dans une bande de 150 m de part et d'autre des principaux axes, à ajouter un terme à la dérive externe. Celui-ci est fonction de la distance à l'axe et du nombre de voies de la chaussée. Cette procédure a été introduite dans l'estimation journalière. Les concentrations estimées aux sites fixes de proximité ont été comparées aux observations, en moyenne sur deux années. L'amélioration est significative. Cette approche pourrait donc être évaluée et, si besoin est, ajustée pour des cartographies moyennes saisonnières ou annuelles. La contrainte est de posséder une base détaillée des routes présentes dans l'agglomération (base GEOROUTE, IGN).

Une autre démarche a été récemment développée au sein du groupe AIR PROCHE (AIRMARAIX, 2006 ; Jeannée et al., 2006). Comme dans la méthodologie présentée en 4.1.1.5, elle consiste à incrémenter la concentration de fond d'un terme de proximité (appelé « surconcentration »). La différence réside dans l'estimation de cette surconcentration :

- L'incrémentation est calculée non plus par différence avec des données de proximité mais par différence avec des résultats de STREET préalablement calés sur des mesures ;
- L'incrémentation est spatialisée non plus par modélisation statistique, en fonction des émissions de NO_x, mais par krigage simple, selon une décroissance exponentielle.

Cette démarche a permis d'obtenir une représentation cohérente de la pollution de fond et de proximité de NO₂ sur l'agglomération toulonnaise et elle semble pouvoir se généraliser. Néanmoins, elle exige un nombre important de mesures de fond et de proximité, aussi bien pour l'élaboration de la cartographie que pour la validation du modèle.

5.1.1.6 Remarque : cartes bilans

Les paragraphes précédents concernent principalement les traitements géostatistiques effectués sur des données d'échantillonnage passif, à l'issue de quelques quinzaines de mesure en été et en hiver. Bien que certains points de méthode appellent des approfondissements, la cartographie journalière automatique (cf. description de la méthode en 5.2.1) est un autre moyen d'élaborer des cartes moyennes saisonnières ou annuelles : il suffit de faire la

moyenne des estimations obtenues pour tous les jours de la saison ou de l'année considérée.

Si dans l'espace, l'information est généralement réduite par rapport à l'utilisation directe de données de campagnes (en chaque ancien point d'échantillonnage, on ne peut pas nécessairement créer de station virtuelle), dans le temps, elle est en revanche plus riche, intégrant les mesures quotidiennes de stations fixes effectuées sur toute l'année.

5.1.2 Modélisation déterministe ou semi-empirique

La modélisation déterministe à l'échelle de la ville connaît un essor récent dans les AASQA et fait actuellement l'objet d'études. Des travaux ont été notamment conduits dans le cadre du groupe AIR PROCHE (AIRMARAIX et al., 2006, INERIS, 2006).

Les outils numériques utilisés, fondés sur une approche gaussienne de la dispersion, sont plus légers que les modèles 3D eulériens et reposent sur des formulations simplifiées des processus qui gouvernent la dispersion et la transformation chimique du NO_2 dans l'atmosphère. Ils permettent également de traiter les phénomènes de rue-canyon.

Parmi les modèles disponibles, citons ADMS-Urban, développé par la société Cambridge Environmental Research Consultants (CERC), pour la cartographie de la pollution à l'échelle de la ville. Le modèle SIRANE, développé au Laboratoire de Mécanique des Fluides et d'Acoustique de l'Ecole Centrale de Lyon, relève plus spécialement de l'approche analytique « modèle de rue » mais son usage peut être étendu à tout un quartier (Soulhac et al., 2004). Il inclut de plus une modélisation gaussienne des panaches, qui permet de prendre en compte la dispersion des polluants émis par les cheminées des toits.

Ces deux outils ont été appliqués par l'INERIS à l'agglomération lyonnaise (Tognet, 2006a), et pour ADMS-Urban, à l'agglomération Nantaise (Tognet, 2006, b). Pour une description détaillée des données d'entrée, de la mise en œuvre et des résultats obtenus, nous renvoyons le lecteur aux rapports correspondants.

En bref, ces modèles nécessitent les informations suivantes :

- un niveau de pollution de fond issu de stations de mesure soigneusement choisies ou de sorties de modèles à plus grande échelle (CHIMERE France ou local par exemple) ; pour ADMS-Urban, un cadastre d'émissions surfaciques de NO_x ;
- des séries de données météorologiques sur la période considérée ;
- une description de la configuration des rues ;
- les émissions des rues.

Ils ne peuvent fournir de résultats satisfaisants **sans un ajustement correct du niveau de fond urbain**. Si celui-ci est estimé par modélisation (CHIMERE pour SIRANE, dispersion gaussienne dans ADMS-Urban), un inventaire d'émissions précis se révèle indispensable, associé, pour ADMS-Urban, à la définition d'un domaine de grande taille.

Moyennant le respect de ces contraintes, ces modèles conviennent bien à l'évaluation de concentrations moyennes de NO₂ sur le long terme puisqu'ils demandent des temps de calcul raisonnables. Ils sont aussi adaptés à la simulation de concentrations horaires de NO₂, pourvu que le modèle de chimie intégré soit suffisamment développé.

5.1.3 Comparaison des approches

Si la géostatistique a toute sa place dans la cartographie du fond urbain, la nouveauté de certaines approches et l'expérience encore récente dans l'usage de certains modèles ne permet pas de comparer les méthodes avec tout le recul nécessaire.

	Modélisation déterministe gaussienne ou semi-empirique	Géostatistique
Données de concentration	Plusieurs sites de mesure de la pollution de fond urbaine ou périurbaine, afin de fournir, si nécessaire, une donnée de fond au modèle (SIRANE), et caler le modèle.	45 à 60 sites de mesure, de type urbain ou périurbain Plusieurs semaines de mesure en été et en hiver pour assurer la représentativité temporelle des données ³ .
Autres données d'entrée	Inventaire des émissions surfaciques de NOx sur l'agglomération (pour ADMS-Urban) Emissions des rues Configuration des rues Données météorologiques : T, DV, VV, nébulosité totale, précipitations, pression (pour SIRANE)	Utilisation de variables auxiliaires recommandée : par exemple un inventaire des émissions de NOx et/ou des données de population à fine échelle (si possible données îlots).
Mode d'évaluation de l'incertitude	Comparaison aux stations fixes, éventuellement à des données d'échantillonnage passif.	Validation croisée et comparaison aux stations fixes. Incertitude temporelle approchée sur des séries annuelles de stations fixes.
Précision attendue	Pour le fond urbain, les études réalisées pour AIR PROCHE (Tognet, 2006 ; AIRMARAIX, 2006) indiquent un écart inférieur à 30% entre les moyennes annuelles de NO ₂ modélisées et estimées (Toulon, Villeurbanne).	Pour le fond urbain : Précision généralement accrue par l'introduction de variables auxiliaires, plus spécialement en hiver ; représentation réaliste de la répartition spatiale de la pollution et des accroissements de concentration au centre-ville.

³ Préconisations en cours d'élaboration par le GT Plans d'échantillonnage et reconstitution de données.

	Modélisation déterministe gaussienne ou semi-empirique	Géostatistique
	<p>Pas d'incertitude liée à une couverture temporelle incomplète ; l'année entière est prise en compte.</p>	<p>D'après l'expérience, écart relatif entre estimation et mesure inférieur à 30% en hiver sur la majeure partie du domaine (en validation croisée). Estimation moins précise en été (concentrations plus basses et phénomène moins structuré).</p> <p>Attention aux types de zones pour lesquels la validité du lien concentration-variable auxiliaire n'est pas assurée (manque de mesures).</p> <p>Incultitude supplémentaire liée à la couverture temporelle incomplète⁴.</p>
Prise en compte de la proximité	<p>Prise en compte des phénomènes de rue-canyon (exemple : OSPM dans ADMS-Urban, SIRANE) et des échanges de polluants au niveau des carrefours (SIRANE).</p> <p>Incultitude de modélisation difficilement évaluable sans un échantillonnage adapté⁵.</p>	<p>Incrémantation possible de la concentration (ajout d'une « sur-concentration »), par spatialisation de la différence entre un modèle de rue et le fond urbain estimé (méthode AIRMARAIX / AIR PROCHE, 2006).</p> <p>Données complémentaires requises : sorties d'un modèle de rue (Street par exemple) pour les principaux axes de la ville, mesures de proximité en nombre suffisant pour réajuster le modèle de rue et contraindre par des observations l'estimation de la sur-concentration.</p> <p>Incultitude d'estimation difficilement évaluable sans un échantillonnage adapté⁵.</p>
Durée de mise en œuvre	<p>Pour une ville entière : un mois de travail incluant la mise en forme des données, la paramétrisation du modèle, les calculs (plusieurs dizaines d'heures pour des simulations horaires sur une année), l'analyse des résultats.</p> <p>Temps supplémentaire pour le calcul des émissions des rues.</p>	<p>Durée variable selon les particularités du jeu de données. De façon générale une quinzaine de jours incluant la mise en forme des données, le calcul des variables auxiliaires, l'analyse exploratoire, la modélisation géostatistique, l'évaluation du modèle, la cartographie.</p>

⁴ Pour huit quinzaines de mesure réparties entre l'été et l'hiver, l'écart entre la moyenne des données et la vraie moyenne annuelle est généralement compris entre 5% et 15% ; il peut atteindre 25% à 30% dans certains cas moins favorables. Ces informations sont issues de tests préliminaires réalisés au sein du GT Plans d'échantillonnage et reconstitution de données.

⁵ Ce point fait l'objet d'une étude LCSQA de 2007.

	Modélisation déterministe gaussienne ou semi-empirique	Géostatistique
Compétences requises	Compétences en modélisation déterministe, surtout pour l'interprétation, les outils étant assez simples à utiliser.	Connaissances fondamentales de statistique et géostatistique linéaire mono- et multivariable. Expérience dans l'ajustement des paramètres du variogramme et du krigeage.
Logiciels	Outils propres à la simulation de la pollution urbaine comme ADMS-Urban (CERC) à l'échelle d'une agglomération, SIRANE (Ecole Centrale de Lyon) à l'échelle d'un quartier	Système d'information géographique pour le traitement des données auxiliaires. Logiciel de géostatistique tel qu'Isatis (Géovariances) afin de pouvoir conduire une analyse approfondie des données et explorer, si cela nécessaire, différents types de modélisation.

5.2 CARTES JOURNALIERES OU HORAIRES

5.2.1 Application des méthodes géostatistiques (cartographie automatique)

5.2.1.1 Principe

L'application de la géostatistique à la cartographie horaire urbaine du NO₂ est limitée par le nombre de stations fixes, trop faible pour élaborer des cartes dans des conditions satisfaisantes.

Afin de combler ce manque, le LCSQA a eu l'idée de tirer profit de campagnes par échantillonnage passif (Perdrix et al., 2004), et d'explorer une voie, pour l'instant restreinte à la **cartographie journalière**, qui consiste à :

- compléter le réseau fixe par des stations « virtuelles », créées aux points où les données de tubes sont bien corrélées dans le temps avec une station existante. Chaque jour, les concentrations moyennes journalières en ces points sont estimées à partir des relations de corrélation mises en évidence. Une incertitude d'estimation leur est affectée sous la forme d'une variance de l'erreur de mesure.
- définir sur l'historique des campagnes le modèle géostatistique qui s'ajuste le mieux aux données (recherche et sélection de variables auxiliaires, structure du variogramme, prise en compte de la variance de l'erreur de mesure, distinction été/hiver....) ;
- construire automatiquement des cartographies en utilisant comme données d'entrée les données journalières des stations réelles et virtuelles ainsi que les variables auxiliaires sélectionnées, et comme modèle, le modèle établi et évalué sur l'historique des campagnes. L'estimation est contrôlée par validation croisée aux points des stations fixes.

5.2.1.2 Données d'entrée

La mise au point de la méthode requiert :

- un historique de campagnes par échantillonnage passif : **plusieurs quinzaines de mesure sur plusieurs saisons et si possible sur plusieurs années** ;
- les données des stations fixes pour les années correspondantes ;
- des informations secondaires : population, émissions, altitude....

Son application au jour le jour ne requiert que les nouvelles données enregistrées aux stations fixes et les variables auxiliaires définies dans l'étape préparatoire.

5.2.1.3 Mise en œuvre

La phase préparatoire, relativement longue et approfondie, a pour objet :

- la recherche des tubes corrélés aux stations fixes, sous certaines contraintes (imposées de façon à assurer la fiabilité des corrélations trouvées), et la création de stations virtuelles. Cette étape est facilitée par une macro Excel conçue notamment à cet effet ;
- l'analyse exploratoire des données de campagnes, la recherche de variables explicatives et la construction d'un modèle géostatistique robuste, capable de représenter la variabilité spatiale des concentrations de NO₂ dans une majorité de situations. L'expertise de l'utilisateur reste indispensable pour cette étape.

La phase opérationnelle, plus rapide, consiste chaque jour :

- à estimer les concentrations aux stations virtuelles. Cette étape peut être effectuée automatiquement par la macro Excel ;
- à fournir à Isatis le fichier actualisé des données de concentration (stations fixes et stations virtuelles) ;
- à lancer l'estimation par krigeage, grâce à une interface Windows spécialement développée pour cette action. Des cartographies mais aussi des résultats de validation croisée sont fournis.

5.2.1.4 Résultats

La méthode de cartographie automatique a été appliquée aux villes de Lille et de Reims, en utilisant les inventaires locaux d'émissions et les données de population comme informations auxiliaires. La technique d'estimation sélectionnée à l'issue de la phase exploratoire, et programmée dans Isatis, est celle du cokrigeage colocalisé. L'incertitude sur les données des stations virtuelles est prise en compte par une variance additionnelle qui est affectée à ces dernières dans la matrice de krigeage (« variance de l'erreur de mesure »).

La méthode a été évaluée jour par jour sur plusieurs années par validation croisée aux points des stations fixes. De multiples statistiques d'erreurs ont pu être ainsi

calculées. Les résultats sont tout à fait satisfaisants au regard de l'objectif réglementaire de qualité, qui exige une incertitude sur la moyenne journalière modélisée inférieure à 50%. Les journées pour lesquelles les estimations ne satisfont pas à ce critère se révèlent être des journées de plus faibles concentrations de NO₂, caractérisées par des conditions météorologiques anticycloniques.

La méthode peut être transposée à d'autres villes, voire à de plus larges domaines. Des améliorations et développements peuvent être envisagés dans la définition des stations virtuelles et le choix des techniques d'estimation.

5.2.1.5 Prise en compte de la proximité

Dans un souci d'automatisation de la procédure, une méthode plus rapide d'exécution que l'approche développée à l'échelle régionale (étude INERIS-2005) ou que la méthode conçue pour la cartographie urbaine annuelle (AIRMARAIX, 2006) a été proposée.

Elle consiste à ajouter à la variable auxiliaire définie pour le fond (variable normalisée, fonction des émissions et de la population) un terme normalisé complémentaire, nul partout sauf dans une bande de ± 150 m autour des grands axes de circulation. La valeur de ce terme dépend du nombre de voies et de la distance par rapport à l'axe.

Cette méthode a été appliquée à l'agglomération de Lille. Visuellement, les cartographies obtenues sont réalistes, montrant une augmentation des concentrations à proximité des grands axes. En revanche, le manque de données en bordure de ces axes ne permet pas d'apprécier l'incertitude.

5.2.2 Modélisation déterministe / semi-empirique

Les modèles régionaux de chimie-transport peuvent simuler finement l'évolution des concentrations urbaines de NO₂ auxquelles la production locale de l'agglomération contribue grandement, pour peu que des inventaires d'émissions adaptés soient disponibles et que les modèles soient appliqués avec une résolution suffisante (cf. §4.2.2). Les modèles de qualité de l'air fondés sur des méthodes gaussiennes ou semi-empiriques (cf. §5.1.2), et consacrés spécifiquement à la représentation de la pollution urbaine (fond et proximité), suscitent aujourd'hui un intérêt. Ils sont capables de fournir une information au pas horaire. Bien que des précautions d'usage s'imposent, un calibrage correct du niveau de fond urbain et un modèle de chimie suffisamment développé, leur permet de reproduire de manière satisfaisante les cycles diurnes.

Des expériences voient le jour, comme celle de l'ASPA qui met en ligne quotidiennement des cartographies sur Strasbourg (système Urban air, fondé sur ADMS-Urban, www.atmo-alsace.net), ou celle d'ATMO Rhône-Alpes qui diffuse des cartes horaires de qualité de l'air dans les rues du centre de Lyon (www.atmo-rhonealpes.org/site/lyon_direct/lyon_direct.php). Pour ce second exemple, le modèle SIRANE, alimenté par des mesures en temps réel de la pollution de fond, est employé. La contribution du trafic est évaluée pour une journée type, ce qui, comme le note ATMO Rhône-Alpes, exclut les situations exceptionnelles.

6. CARTOGRAPHIE AU VOISINAGE DES ROUTES

Il n'existe pas encore de méthodologie aboutie pour la cartographie au voisinage des routes. Si des outils existent, aucun résultat sûr n'a été encore obtenu ni aucune méthode d'évaluation certaine encore définie.

6.1 CARTES MOYENNES SAISONNIERES OU ANNUELLES

6.1.1 Application des méthodes géostatistiques

Les cartographies au voisinage de routes sont moins courantes que les cartographies urbaines. Les pratiques en matière d'échantillonnage, qu'il s'agisse du nombre et de la disposition des tubes ou de la durée des campagnes, sont aussi plus disparates. Souvent les AASQA se limitent à présenter les valeurs mesurées aux points d'échantillonnage ou à évaluer, en une dimension, la décroissance des concentrations perpendiculairement à la route (AIR Languedoc-Roussillon, 2004). Ces études ont généralement pour fin d'évaluer l'impact d'une route particulière, par exemple avant ou après un aménagement urbain.

La difficulté de la cartographie réside dans le fait de **modéliser, à petite échelle (quelques centaines de mètres de part et d'autre de la route), un phénomène non stationnaire** (les concentrations diminuent rapidement à mesure qu'on s'éloigne de la route).

Des exemples de cartographies existent cependant. Ils reposent sur un échantillonnage constitué de sites de fond (maillage dense) et de points de mesure disposés perpendiculairement à la route. Un krigage ordinaire a été mis en œuvre (ORAMIP, 2004 ; AIRMARAIX, 2003). Pour l'une des campagnes (AIRPARIF, 2005), des données de modèles de rue ont été intégrées dans l'algorithme de krigage.

Dans tous les cas, conformément aux mises en garde des auteurs, l'objet de la cartographie est de mettre en évidence les **grandes tendances de la pollution autour de l'axe étudié et non pas d'estimer précisément les niveaux de concentration**, qui dépendent grandement de l'environnement local. Cette précision n'est d'ailleurs pas examinée.

En outre, la plupart des variables auxiliaires (notamment l'inventaire des émissions) n'ont pas la résolution suffisante pour rendre compte des gradients de concentration sur quelques centaines de mètres. Deux variables cependant peuvent jouer le rôle facteurs explicatifs, comme l'indique une étude sur la vallée de la Thur (Séguret, 2003) : la **distance par rapport à la route** et, si la géographie du domaine le demande, l'altitude.

L'usage des méthodes géostatistiques dans la cartographie de proximité reste un sujet d'étude.

6.1.2 Modélisation empirique ou semi-empirique

La modélisation de la pollution de proximité sert plutôt à évaluer de façon systématique les concentrations le long des principaux axes d'une zone (par exemple, afin de tester des scénarios de réduction des émissions routières). Dans ce contexte simple et opérationnel qui exclut les modèles tridimensionnels eulériens ou lagrangiens, trois types d'outils se présentent (Rouïl, 2004) :

- les modèles analytiques, qui reposent sur une formulation mathématique des concentrations attendues dans la rue canyon (ex : SIRANE, OSPM) ;
- les modèles paramétriques, qui font appel à une bibliothèque de résultats issus d'un grand nombre d'expérimentations et de modélisations fines (ex : STREET, AEOLIUS) ;
- les modèles semi-paramétriques, fondés sur une formulation analytique de la concentration dans la rue canyon, et dont les paramètres clefs sont définis à partir de valeurs connues pour des cas de référence (ex : CAR).

Alors qu'ils sont souvent utilisés en milieu bâti (AIRMARAIX, 2006, AIRPARIF, 2007), leur pertinence dans des environnements ouverts appelle plus de réserves. Pour ces environnements, on sait en effet qu'aucun modèle n'est satisfaisant dès que le vent n'est plus perpendiculaire à la route. De plus, reposant sur une représentation très simplifiée de la situation, les modèles de rue sont plus adaptés à une **analyse relative** qu'à une cartographie des valeurs absolues de concentration. Si en moyenne annuelle, les concentrations calculées constituent une donnée exploitable, au pas horaire ou journalier, l'incertitude sur les valeurs simulées est plus grande à cause de l'incertitude sur les émissions.

Des données d'observation restent indispensables pour caler le modèle et évaluer les résultats obtenus.

7. CONCLUSION

Afin de fournir une information sur l'ensemble de leur territoire, comme le réclame la réglementation, offrir une image détaillée de la qualité de l'air urbaine et répondre aux préoccupations touchant à la pollution de proximité, les AASQA ont de plus en plus recours à la cartographie, à des échelles d'espace et de temps très variées. Les outils sont multiples, différents dans leurs principes, dans la nature des données d'entrée, la facilité de mise en œuvre, la résolution temporelle et spatiale des estimations obtenues. Tous ne présentent pas le même degré d'aboutissement, ni la prise en compte et la connaissance des incertitudes le même degré d'avancement. Leurs capacités peuvent aussi dépendre du type de territoire et du support de temps considéré (heure, jour, saison...).

Face à ce grand choix d'outils disponibles, se pose la question de leur pertinence pour le type de cartographie que l'on cherche à produire. Partant des études conduites par le LCSQA, des travaux d'AASQA et d'une analyse bibliographique, ce rapport dresse un état des lieux des méthodes disponibles selon l'échelle spatiale et temporelle du phénomène à représenter. La pollution par le dioxyde d'azote est ici exclusivement considérée. Pour la cartographie en zone rurale influencée, les études réalisées par le LCSQA sur la région Centre entre 2003 et 2006 permettent de compléter plus concrètement ce bilan et de comparer sur un même domaine différentes méthodes.

D'une façon générale, on distingue :

- les méthodes paramétriques fondées sur des corrélations ;
- l'interpolation géostatistique ;
- la modélisation déterministe, qui recouvre à son tour des techniques plus ou moins complexes, et la modélisation empirique.

La cartographie moyenne saisonnière ou annuelle, sur une région ou une agglomération, est aujourd'hui relativement bien maîtrisée, pourvu que les données d'entrée soient suffisamment nombreuses et précises : au moins une cinquantaine de points de mesure en géostatistique, données de calage, inventaire des émissions de fine résolution spatiale pour les trois méthodes et/ou d'autres variables auxiliaires (population, ..) pour la géostatistique; pourvu aussi que ces données représentent correctement la période que l'on souhaite représenter (choix des périodes d'échantillonnage pour la méthode des corrélations et la géostatistique, définition des situations météorologiques en modélisation). Moyennant ces contraintes, et sous réserve d'une utilisation rigoureuse des méthodes (analyse exploratoire approfondie en géostatistique, calage du modèle, ...), on constate un bon accord entre l'estimation et la mesure. On peut espérer un écart inférieur à 30% par rapport aux mesures des stations fixes. En outre, les représentations obtenues sont généralement cohérentes entre elles, ce qui n'est pas surprenant puisque toutes s'appuient sur l'inventaire des émissions et/ou des données représentatives de l'activité humaine.

Plus délicate est l'estimation à fine résolution temporelle. A l'échelle régionale, les modèles de chimie-transport sont capables de fournir une information horaire sur les grandes tendances de la pollution. Cette information est d'autant plus précise que l'est aussi l'inventaire des émissions utilisé. Cette condition vaut également

pour les modèles urbains, fondés sur une formulation gaussienne de la dispersion. Mises le plus souvent en œuvre pour exploiter des données d'échantillonnage passif intégrées sur plusieurs semaines, la géostatistique a été principalement réservée à la cartographie moyenne saisonnière ou annuelle. Des développements ont été récemment réalisés afin de permettre la construction automatique de cartes moyennes journalières. Ils ont été évalués sur deux villes, Lille et Reims. Les résultats sont prometteurs. Si elle ne dispense pas d'une analyse préliminaire approfondie des données, l'approche proposée constitue un outil efficace de cartographie dans un contexte opérationnel.

La cartographie à proximité d'axes de circulation pose des problèmes particuliers en raison de la variabilité spatiale et temporelle inhérente à la petite échelle et de la sensibilité des résultats au type d'environnement (bâti, ouvert). Ces questions font l'objet d'investigations, aussi bien dans le domaine de la modélisation (utilisation de modèles de rue) que dans celui de la géostatistique. Ces travaux sont notamment guidés par deux objectifs : d'une part approcher aussi précisément que possible les concentrations autour de l'axe, d'autre part, coupler ces représentations avec des cartographies de plus grande échelle (ville, région). L'incertitude sur les résultats reste mal connue. Pour des moyennes annuelles, des campagnes de mesure appropriées permettent une comparaison entre estimation et mesure ; à un pas de temps plus fin, cette évaluation se limite aux stations trafic, s'il s'en trouve près des axes considérés.

8. REFERENCES

AIRAQ , 2005. Campagne de mesure par échantillonnage passif. Cartographie de l'agglomération de Bordeaux 2004/2005, Tome 1. Rapport n°ET/TP/05/05 disponible sur www.airaq.asso.fr

AIR Languedoc-Roussillon, 2005. Etat initial de la qualité de l'air avant le dédoublement de l'autoroute A9, de la déviation Est et du contournement Ouest de Montpellier. Principaux résultats disponibles sur le site www.air-lr.org.

AIRMARAIX, avec la participation de Géovariances et de Numtech, 2006. Cartographie de la pollution à Toulon : approches géostatistique et déterministe. Rapport d'études AIR PROCHE (AFSSET), disponible sur le site www.aimaraix.org.

AIRMARAIX, 2003. Etat initial de la qualité de l'air. Axe littoral sens nord-sud. Campagne de mesure du 14 novembre 2002 au 2 mars 2003. Réf. PGR/LP02.

AIRPARIF, 2007. L'évolution de la qualité de l'air à proximité des rues parisiennes entre 2002 et 2007. Rapport disponible sur le site www.airparif.asso.fr.

AIRPARIF, 2006. Campagnes de mesure de la qualité de l'air en sites trafic parisiens à forte fréquentation piétonne. Rapport disponible sur le site www.airparif.asso.fr.

AIRPARIF, 2004. Caractérisation de la qualité de l'air au voisinage d'un échangeur autoroutier urbain. L'échangeur entre le boulevard périphérique et l'autoroute A3 au niveau de la Porte de Bagnolet. Rapport disponible sur le site www.airparif.asso.fr.

ARMSTRONG M., 1998. Basic linear geostatistics. Springer-Verlag, Berlin Heidelberg.

ARNAUD M. et EMERY X. , 2000. Estimation et interpolation spatiale - *Méthodes déterministes et méthodes géostatistiques*. Hermès Science Publication.

ASPA, 2003. Caractérisation de la qualité de l'air sur l'aire d'étude du grand Ouest de Strasbourg. Rapport relatif à la campagne de mesure qui s'est déroulée du 19 juin au 3 juillet 2002 et du 3 au 17 décembre 2002. Source d'information ASPA 03032101-ID. Rapport disponible sur www.atmo-alsace.net.

ASPA, Géovariances, 2005. Cartographie régionale NO₂, C₆H₆, O₃ en Alsace. Source d'information ASPA 05020802-ID. Rapport disponible sur www.atmo-alsace.net.

ATMO RHONE-ALPES, 2005. Dioxyde d'azote et benzène dans l'air en 2005. Carte des concentrations annuelles en Rhône-Alpes. Note de synthèse disponible sur le site www.atmo-rhonealpes.org.

ATKINS D.H.F., LEE D.S., 1995. Spatial and temporal variation of rural nitrogen dioxide concentrations across the United Kingdom. *Atmospheric Environment* 29 (2),223-239.

BOBBIA M., 2001. Investigations en vie d'une approche géostatistique de la qualité de l'air. Analyse du risque : les simulations conditionnelles. Rapport final, AIR NORMAND.

BOBBIA M., PERNELET V. et ROTH C., 2001. L'intégration des informations indirectes à la cartographie géostatistique des polluants. *Pollution atmosphérique* n°170. Avril-juin 2001.

CARDENAS G. et MALHERBE L., 2004. Géostatistique et prise en compte de l'aspect temporel. Rapport LCSQA-INERIS, convention n°04000087, disponible sur le site www.lcsqa.org.

CARDENAS G. et MALHERBE L., 2003. Evaluation des incertitudes associées aux méthodes géostatistiques. Rapport LCSQA-INERIS, convention n°115/03, disponible sur le site www.lcsqa.org.

CARDENAS G. et MALHERBE L., 2002. Méthodes de représentation de la qualité de l'air dans les zones peu/pas couvertes par les stations de mesures fixes : Partie II, application à la problématique d'une association. Rapport LCSQA- INERIS, convention n°31/2001, disponible sur le site www.lcsqa.org.

CARDENAS G., 2004. Géostatistique et estimation d'une probabilité de dépassement de seuil. Rapport LCSQA-INERIS, convention n°04000087, disponible sur le site www.lcsqa.org.

CARDENAS G., PERDRIX E., 2005. Adaptation de méthodes géostatistiques à la cartographie automatique de NO₂. Rapport LCSQA-INERIS, convention n°05000051, disponible sur le site www.lcsqa.org.

CARDENAS G., PERDRIX E., WROBLEWSKI A, RIFFAULT V., 2006. Méthode de surveillance des concentrations de NO₂ : cartographie à partir de stations fixes et prise en compte de la proximité. Rapport LCSQA-EMD- INERIS, convention n°000070.

CARSLAW D.C., BEEVERS S.D., FULLER G., 2001. An empirical approach for the prediction of annual mean nitrogen dioxide concentrations in London. *Atmospheric Environment*, 35, 1505-1515.

CORI A., 2005. Utilisation de données environnementales pour la cartographie du NO₂ sur la région rouennaise. Première partie. AIR NORMAND, Ecole des Mines de Paris.

DERAISME J., BOBBIA M., 2003. L'apport de la géostatistique à l'étude des risques liés à la pollution atmosphérique. *Environnement, Risques & Santé*, Vol. 2, N°3, 168-175.

Directive n° 1999/30/CE du 22 avril 1999 relative à la fixation de valeurs limites pour l'anhydride sulfureux, le dioxyde d'azote et les oxydes d'azote, les particules et le plomb dans l'air ambiant.

FOUQUET C.(de), 2006. La modélisation géostatistique des milieux anthropisés. Habilitation à diriger des recherches. Université Pierre et Marie Curie, Mémoires des sciences de la Terre, N°2006-13. Document disponible sur le site <http://cq.ensmp.fr/bibliotheque/index.html>.

FOUQUET C. (de), GALLOIS D., MALHERBE L., CARDENAS G., 2004. Mapping annual nitrogen dioxide concentrations above Mulhouse urban area. *GeoEnv 2004*, Neuchâtel (Suisse), 13-15 octobre.

FOUQUET C.(de), 2003a. Des statistiques contre la pollution. *La Recherche* , N°368.

FOUQUET C.(de), 2003b. Etude sur la réalisation de cartographies de la qualité de l'air dans les zones peu/pas couvertes par les réseaux de stations fixes à l'aide de méthodes géostatistiques (complément d'étude et synthèse). Ecole des Mines de Paris, Rapport N-9/03/G, juin 2003.

FOUQUET C.(de), 2003c. Méthodologie de cartographie de la concentration annuelle de NO₂ sur l'agglomération de Mulhouse. Ecole des Mines de Paris, Rapport d'avancement N-6/03/G, avril 2003.

GILBERT N.L., GOLDBERG M.S., BECKERMAN B., BROOK J.R., JERRETT M., 2005. Assessing spatial variability of ambient nitrogen dioxide in Montréal, Canada, with a land-use regression model. *Journal of the Air & Waste Management Association*, 55, 1059-1063.

HOCHADEL M., HEINRICH J., GEHRING U., MORGESTERN V., KUHLBUSCH T., LINK E., WICHHMANN H.E., KRÄMER U., 2006. Predicting long-term average concentrations of traffic-related air pollutants using GIS-based information. *Atmospheric Environment*, 40, 542-553.

JEANNÉE N., FAYET S., MARY L., FROMAGE-MARIETTE A., CABERO C., PERRON G., ARMENGAUD A., 2006. Cartographie de la qualité de l'air en agglomération : comment intégrer pollution de fond et pollution de proximité. 2^e conférence Environnement et Transports, Reims, France, actes, n°107, Vol.1, Inrets ed., Arcueil, France, 2006, p.303-310.

LE LOC'H G., 2003. Etude exploratoire du dioxyde d'azote sur l'agglomération de Montpellier. Ecole des Mines de Paris, Rapport N-8/03/G, juin 2003.

LIMAIR, 2006. Etude ozone – été 2005. Région Limousin. Rapport référencé ETD/2005/13, disponible sur le site www.limair.asso.fr.

MALHERBE L., CARDENAS G., 2005. Application des méthodes géostatistiques pour l'exploitation conjointe des mesures de fond et de proximité. Rapport de l'INERIS dans le cadre du LCSQA, convention 05000051, disponible sur le site www.lcsqa.org.

MALHERBE L. et ROUÏL L., 2002. Assistance en modélisation de la qualité de l'air. Rapport LCSQA-INDERIS, convention n°31/2001, disponible sur le site www.lcsqa.org.

ORAMIP, 2005. Campagne de mesure de la qualité de l'air aux abords de la RN 126 Castres-Soual.

ORAMIP, 2005. Campagne de mesure de la qualité de l'air aux abords de la RN 112 Labruguière-Saint-Alby.

PERDRIX E. et HOUDRET J-L, 2000. Faisabilité d'interpolation spatiale des données de la pollution de l'air à l'échelle régionale. Etude n°2 du Rapport d'activité de l'Ecole des Mines de Douai. Laboratoire Central de Surveillance de la Qualité de l'Air. Convention n° 18/98.

PERDRIX E., WROBLEWSKI A., FOURCHÉ B., PLAISANCE H., 2004. Cartographie des concentrations en dioxyde d'azote à partir des mesures en stations fixes. Rapport LCSQA-EMD, convention n°04000088, disponible sur le site www.lcsqa.org.

PLAISANCE H., DEPELCHIN L., BOUCHAHDANE F. et SAUVAGE S., 1999. Etablissement d'un protocole d'utilisation des tubes à diffusion : intercomparaison de tubes à diffusion NO₂ et tests de méthodes géostatistiques d'interpolation. Etude n°7 du Rapport d'activité de l'Ecole des Mines de Douai. Laboratoire Central de Surveillance de la Qualité de l'Air. Convention n° 18/98.

RIVOIRARD J., 2002. On the structural link between variables in kriging with external drift. Mathematical Geology, Vol. 34, N°7, 797-808.

RIVOIRARD J., 2003. Cours de géostatistique multivariable. Cours C-17 2. Ecole des Mines de Paris, <http://cg.ensmp.fr>.

RIVOIRARD J., 2004. On some simplifications of cokriging neighborhood. Mathematical Geology, Vol. 36, N°8, 899-915.

ROSS Z., ENGLISH P.B., SCALF R., GUNIER R., SMORODINSKY B S., WALL S., JERRETT M., 2005. Nitrogen dioxide prediction in Southern California using land use regression modeling: potential for environmental health analyses. Journal of Exposure Science and Environmental Epidemiology, 16, 106–114.

ROUÏL L. et MALHERBE L., 2002. Méthodes de représentation de la qualité de l'air dans les zones peu/pas couvertes. Partie I : stratégies et préconisations. Rapport LCSQA-INERIS, convention n°31/2001, disponible sur le site www.lcsqa.org.

ROUÏL L. et MALHERBE L., 2003. Assistance à l'utilisation des outils de modélisation. *Annexe 2: Cartographie du dioxyde d'azote dans une agglomération ou le long d'un axe routier à l'aide des méthodes géostatistiques. Synthèse des travaux réalisés par le Centre de Géostatistique de l'Ecole des Mines de Paris.* Rapport LCSQA-INERIS, convention n°115/03, disponible sur le site www.lcsqa.org.

ROUÏL L., 2005. Relevé de conclusions suite au séminaire LCSQA du 17 juin 2005. Méthodes de cartographie de la qualité de l'air. Ref : DRC-05-45997-MECO-n°31/LRo, disponible sur le site www.lcsqa.org (rubrique Actualités).

SAHSUVAROGLU T., ARAIN A., KANAROGLU P., FINKELSTEIN N., NEWBOLD B., JERRETT M., BECKERMAN B., BROOK J.R., FINKELSTEIN M., GILBERT N.L., 2006. Assessing spatial variability of ambient nitrogen dioxide in Montréal, Canada, with a land-use regression model. Journal of the Air & Waste Management Association, 56, 1059-1069.

SÉGURET S., 2003. Estimation du dioxyde d'azote dans la vallée de la Thur. Ecole des Mines de Paris, Rapport N-1/03/G, février 2003.

SMITH L., MUKERJEE S., GONZALES M., STALLINGS C., NEAS L., NORRIS G., ÖZKAYNAK H., 2006. Use of GIS and ancillary variables to predict volatile organic compounds and nitrogen dioxide levels at unmonitored locations. *Atmospheric Environment*, 40, 3773-3787.

SOULHAC L., PRADELLE F., PERKINS F.J., 2004. An evaluation of the urban dispersion models SIRANE and ADMS Urban using field data from Lyon. 9th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Garmisch-Partenkirchen, Allemagne, 1^{er}-4 juin 2004.

STEDMAN J.R., 1996. Revised high resolution maps of background air pollutant concentrations in the UK: 1996. AEAT – 3133, Issue 2, March 1998. Document disponible à l'adresse <http://www.airquality.co.uk/archive/reports/list.php>.

STEDMAN J.R., BUSH T., 2000. Mapping of nitrogen dioxide and PM10 in the UK for Article 5 Assessment. AEAT - env r0707, Issue 1, 15 p. Document disponible à l'adresse <http://www.airquality.co.uk/archive/reports/list.php>.

TOGNET F., 2006a. Projet AIR PROCHE : simulation de la qualité de l'air à l'échelle urbaine. Application à Villeurbanne. Rapport d'études AIR PROCHE (AFSSET)

TOGNET F., 2006b. Imbrication de modèles de rue dans les modèles régionaux de qualité de l'air. Rapport LCSQA à paraître.

WACKERNAGEL H., 2004. Géostatistique et assimilation séquentielle des données. Habilitation à diriger des recherches. Université Pierre et Marie Curie, Mémoires des sciences de la Terre, N°2004-02. Document disponible à l'adresse <http://cg.ensmp.fr/bibliotheque/index.html>.

WROBLEWSKI A, PERDRIX E., MALHERBE L., 2005. Adaptation des plans d'échantillonnage spatial aux objectifs des campagnes. Rapport LCSQA-EMD-INERIS, convention n°05000051, disponible sur le site www.lcsqa.org.

WROBLEWSKI A, RIFFAULT V., PERDRIX E., MALHERBE L., 2006. Adaptation des plans d'échantillonnage spatial aux objectifs des campagnes, partie II, annexes. Rapport LCSQA-EMD-INERIS, convention n°000070.

WROBLEWSKI A., PERDRIX E., OLIVIER P., 2003. Cartographie de la pollution atmosphérique en zone rurale influencée au moyen d'un modèle déterministe de dispersion des polluants. Rapport de l'EMD dans le cadre du LCSQA, convention n°115/03, disponible sur le site www.lcsqa.org.

9. LISTE DES ANNEXES

Repère	Désignation	Nombre de pages
Annexe 1	Document de référence de l'étude	3
Annexe 2	Krigage	5
Annexe 3	Variables auxiliaires	4
Annexe 4	Elaboration de cartes saisonnières : application du cokrigage temporel - cartographie automatique journalière et calcul cartes bilans	13
Annexe 5	Récapitulatif des résultats obtenus sur la région Centre	21

10. ANNEXE 1

Etude 24 :Synthèse des techniques de cartographie de NO₂ : Valorisation des travaux 2003 – 2005

INERIS/EMD

CONTEXTE ET OBJECTIF

Depuis quelques années, souhaitant d'une part, répondre aux exigences réglementaires et d'autre part, informer de façon claire le public, les AASQA se sont investies dans la réalisation de cartographies. Cette activité les a conduites à développer des compétences dans la mise en œuvre de campagnes d'échantillonnage et l'usage d'outils numériques plus ou moins sophistiqués. trois approches se distinguent :

- soit l'utilisation de modèles déterministes, lorsqu'un inventaire des émissions est disponible sur la zone d'étude,
- soit la réalisation de campagnes de mesure et l'emploi de techniques d'estimation géostatistiques.
- soit le développement de relations empiriques entre les concentrations et des variables externes telles que les émissions, l'occupation du sol, la densité de population (méthodes de Stedman)

Dans ce contexte, le LCSQA a pris une part active au développement et à l'application de ces méthodes afin, notamment, de cartographier le dioxyde d'azote.

La diversité des méthodes de cartographie et la variété de situations dans lesquelles elles sont employées (échelle régionale ou urbaine, cartographie annuelle ou journalière) conduisent néanmoins à s'interroger sur le choix de la meilleure approche, en fonction des objectifs fixés.

En s'appuyant sur l'expérience gagnée ces dernières années, cette étude a pour but d'évaluer les avantages et limites des différentes méthodes de cartographie du NO₂ et, selon les données disponibles, l'environnement géographique et l'objectif de cartographie, d'émettre des recommandations sur les choix techniques les plus pertinents. Ces préconisations feront l'objet d'un guide d'utilisation à destination des AASQA.

TRAVAUX ANTERIEURS ET EN COURS

Depuis 2003 l'INERIS et l'EMD ont réalisé plusieurs travaux consacrés aux méthodes de cartographie de la pollution par les oxydes d'azotes (cf. références), en particulier sur les zones faiblement couvertes par les réseaux de mesure. Ces travaux ont montré la possibilité d'appliquer différentes techniques de reconstruction cartographique (modélisation déterministe, interpolation, statistique et géostatistique).

TRAVAUX PROPOSES POUR 2006

A partir des données récoltées lors des études précédentes, notamment les études sur la région Centre, nous proposons de définir un jeu de données de référence le plus exhaustif possible.

Ce jeu de référence, qui pourra être subdivisé en fonction de l'échelle de travail (région, zone rurale non influencée, zone urbaine), permettra de tester les différentes méthodes de cartographie (déterministe, statistique et géostatistique, empirique) étudiées par l'EMD et l'INERIS entre 2003 et 2005.

Pour chaque objectif de cartographie considéré, il s'agira de comparer les résultats obtenus et de déterminer les avantages et limites des diverses techniques utilisées. Cette évaluation sera réalisée en fonction de différents critères qui restent à définir. On considèrera a minima la précision des cartes par comparaison entre valeurs calculées et mesurées mais on s'attachera également à d'autres critères tels que l'exigence des méthodes en terme de données d'entrée, la facilité et la durée de mise en œuvre, la résolution spatiale ou temporelle, l'aspect des cartes obtenues, les contraintes financières et les coûts de mise en œuvre...

Ces exercices de comparaison donneront lieu à l'élaboration d'un guide de recommandations pour la cartographie du NO₂. Ils s'accompagneront d'une réflexion sur l'incertitude inhérente à l'élaboration de cartographies (en complément de l'incertitude évaluée aux points de mesure).

COLLABORATIONS

AASQA

DUREE DES TRAVAUX

1 an

UNITES D'ŒUVRE IMPLIQUEES

- EMD : 330 heures ingénieur : André Wroblewski , Esperanza Perdrix
- INERIS : 600 heures ingénieur : Laure Malherbe, Giovanni Cárdenas

REFERENCES :

- Cartographie en zone rurale par modélisation déterministe (LCSQA-EMD, 2003. *Cartographie de la pollution atmosphérique en zone rurale influencée au moyen d'un modèle déterministe de dispersion des polluants*) ;
- Cartographie en zone rurale par modélisation statistique (méthode de régression de Stedman, LCSQA-EMD, *Cartographie de la moyenne annuelle de dioxyde d'azote à l'échelle régionale par la méthode des régressions*)
- Cartographie régionale saisonnière ou annuelle par exploitation de données d'échantillonnage passif et estimation géostatistique (LCSQA-INGERIS, 2005. *Application de méthodes géostatistiques pour l'exploitation conjointe de mesures de fond et de proximité*) ;
- Cartographie urbaine saisonnière ou annuelle par exploitation de données d'échantillonnage passif et estimation géostatistique (LCSQA-ENSMP/INGERIS, 2003. 1. *Etude exploratoire du dioxyde d'azote sur l'agglomération de Montpellier* ; 2. *Méthodologie de cartographie de la concentration annuelle de NO₂ sur l'agglomération de Mulhouse* ; 3. *Etude sur la réalisation de cartographies de la qualité de l'air dans les zones peu/pas couvertes par les réseaux de stations fixes à l'aide de méthodes géostatistiques (synthèse)* ; LCSQA- INERIS, 2004. *Géostatistique et prise en compte du problème temporel*) ;
- Cartographie urbaine journalière par exploitation de données d'échantillonnage passif et de mesures continues de stations fixes et estimation géostatistique (LCSQA- INERIS, *DRC-07-74777-03941A*

EMD-INNERIS, 2005. *Adaptation des méthodes géostatistiques à la cartographie automatique des concentrations en NO₂* ;

- Cartographie saisonnière ou annuelle au voisinage d'un axe routier par exploitation de données d'échantillonnage passif et estimation géostatistique (LCSQA-ENSMP/INNERIS, 2002. *Estimation du dioxyde d'azote routier dans la vallée de la Thur*)

11. ANNEXE 2 : LE KRIGEAGE

La méthode de krigage est fondée sur le calcul du variogramme expérimental et de sa modélisation. Le variogramme expérimental $\gamma(h)$ permet d'apprécier la ressemblance (ou non) entre deux concentrations éloignées d'une distance h (relation 1) :

$$\gamma(h) = \frac{1}{2 \times N(h)} \sum_{i=1}^{Nh} [C(x_i + h) - C(x_i)]^2 \quad (1)$$

avec $C(x_i)$: concentration de polluant au point d'abscisse x ,

$C(x_i + h)$: concentration de polluant au point d'abscisse $(x+h)$,

$C(x_i)$ et $C(x_i + h)$ sont vues comme la réalisation de deux variables aléatoires.

$N(h)$: nombre de couples de points distants de h dans l'ensemble des données.

En pratique, on établit à partir des données un nuage de points, appelé nuée variographique, qui représente la valeur de $\frac{1}{2} [C(x_i) - C(x_j)]^2$ en fonction de $h = |x_i - x_j|$,

(x_i, x_j) parcourant l'ensemble des points de mesure (Figure 12). Puis on définit un réseau régulier de classes de distance dans chacune desquelles la moyenne de ces écarts quadratiques est calculée. A chaque classe de distances correspond donc une valeur unique de γ . Cet ensemble discret de valeurs constitue le variogramme expérimental. Ce variogramme n'est pas défini partout, notamment aux distances pour lesquelles il n'existe aucune paire de points de mesure. Aussi lui est-il ajusté un modèle mathématique, composé d'une ou plusieurs fonctions que proposent habituellement les logiciels de géostatistique.

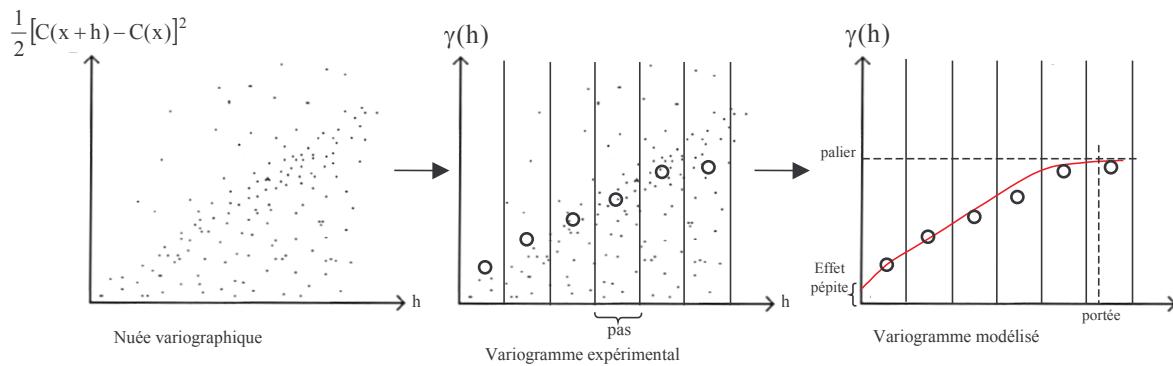


Figure 12 - Etapes de la modélisation du variogramme (d'après Plaisance et al., 1999)

Généralement, $\gamma(h)$ augmente avec la distance h jusqu'à atteindre un *palier*. Aux faibles valeurs de h , les concentrations sont plus fortement corrélées. Au-delà de la *portée*, distance à laquelle le variogramme devient horizontal, on peut considérer qu'elles sont indépendantes les unes des autres. Le variogramme, qui vaut zéro pour $h=0$, peut présenter une discontinuité à l'origine appelée *effet pépite*. Ce dernier traduit une variation de la concentration à très petite échelle et/ou l'incertitude liée aux mesures.

La « **validation croisée** » est un moyen d'apprécier la qualité de l'ajustement du modèle : ce test consiste à retirer un point de mesure du jeu de données initial et à estimer la concentration en ce point à l'aide des données restantes. Cette procédure est réitérée pour chaque point de mesure. De cette façon on obtient un ensemble de valeurs mesurées $Z(x_i)$ et un ensemble de valeurs estimées $Z^*(x_i)$. Des erreurs d'estimation peuvent être calculées, ainsi que des statistiques sur ces erreurs (Cárdenas et Malherbe, 2003).

11.1 KRIGEAGE ORDINAIRE

En un point x_0 , l'estimateur par krigeage ordinaire de la concentration s'écrit :

$$Z^*(x_0) = \sum_{x_i \in V} \lambda_i Z(x_i)$$

x_i : points d'observation de Z

V : voisinage d'estimation, unique ou glissant.

Les poids de krigeage λ_i sont tels que l'estimateur soit sans biais et de variance minimale. Ils sont solutions d'un système linéaire d'équations faisant intervenir le modèle de variogramme.

La variance de l'erreur d'estimation s'écrit :

$$\sigma_K^2(x_0) = E[Z^*(x_0) - Z(x_0)]^2 = \sum_i \lambda_i \gamma(\|x_i - x_0\|) + \mu$$

(μ : multiplicateur de Lagrange, solution avec les x_i du système de krigeage).

Cette variance, aussi appelée « variance de krigage », ou sa racine carrée, l'écart-type de krigage, constitue un indicateur de la précision de l'estimation. Elle décrit la dispersion possible de la valeur vraie, mais inconnue, autour de la valeur estimée. Elle est uniquement fonction du modèle de variogramme et de la configuration des données. Elle traduit la plus ou moins grande capacité du krigage de fournir une estimation précise, étant donné le caractère plus ou moins structuré du phénomène et la plus ou moins grande densité de mesures au voisinage du point cible x_0 . Elle ne dépend pas de la variabilité locale des données.

Afin de prendre en compte des variables auxiliaires, la géostatistique propose plusieurs méthodes que nous rappelons succinctement ci-dessous.

11.2 VARIABLES AUXILIAIRES DISPONIBLES SUR UN ENSEMBLE IRREGULIER DE POINTS OU SUR UN MAILLAGE LACHE

Méthode : cokrigage ordinaire

Hypothèse : corrélation spatiale entre la variable principale et la ou les variables auxiliaires.

Estimateur : $Z^*(x_0) = \sum_{x_i \in V} \lambda_i Z(x_i) + \sum_k \sum_{x_{kj} \in V} \mu_{kj} Y_k(x_{kj})$

x_i : points de mesure de la variable Z

x_{kj} : points de mesure de la variable auxiliaire Y_k

Le cokrigage se limite souvent à deux variables ($k=1$, une seule variable auxiliaire).

Modèle : variogrammes simples et croisés de Z et des variables auxiliaires.

11.3 VARIABLES AUXILIAIRES CONNUES PARTOUT (I.E. SUR UNE GRILLE RELATIVEMENT FINE)

Méthode : Cokrigage colocalisé

Hypothèse : corrélation spatiale entre la variable principale et la ou les variables auxiliaires.

Estimateur : $Z^*(x_0) = \sum_{x_i \in V} \lambda_i Z(x_i) + \sum_k \mu_{k0} Y_k(x_0) + \sum_k \sum_{x_i \in V} \mu_{ki} Y_k(x_i)$

x_i : points de mesure de la variable Z

Le cokrigage se limite souvent à deux variables ($k=1$, une seule variable auxiliaire) mais on peut prendre pour Y_1 une combinaison linéaire de plusieurs cofacteurs, dont les coefficients sont obtenus par régression multilinéaire :

$$Y_1(x) = a_1 \cdot F_1(x) + a_2 \cdot F_2(x) + \dots + b$$

$Y_1(x)$ est considérée comme une fonction aléatoire.

Modèle : variogrammes simples et croisé de Z et de Y_1 .

Remarque : Le cokrigeage colocalisé n'utilise les valeurs de Y_1 qu'au point cible x_0 et aux points x_i de mesure de Z. Il constitue de ce fait une approximation du cokrigeage. Il y a exacte équivalence lorsque le variogramme croisé de Z et de Y_1 est proportionnel au variogramme de Y_1 .

Méthode : krigeage du résidu

Hypothèse : $Z(x) = a_1 \cdot F_1(x) + a_2 \cdot F_2(x) + \dots + b + R(x)$

F_k : variables auxiliaires

$R(x)$: résidu aléatoire, non spatialement corrélé avec les F_k

Estimateur : $Z^*(x_0) = \hat{a}_1 \cdot F_1(x_0) + \hat{a}_2 \cdot F_2(x_0) + \dots + \hat{b} + R^*(x_0)$

\hat{a}_k, \hat{b} : coefficients estimés par régression (multi)linéaire

$R^*(x_0) = \sum_{x_i \in V} \lambda_i R(x_i)$: krigeage du résidu.

Les $R(x_i)$ sont les résidus expérimentaux de la régression (multi)linéaire :
 $R(x_i) = Z(x_i) - [\hat{a}_1 \cdot F_1(x_i) + \hat{a}_2 \cdot F_2(x_i) + \dots + \hat{b}]$

x_i : points d'observation de Z

Modèle : modèle ajusté sur le variogramme des résidus expérimentaux.

Méthode : krigeage avec dérive externe

Hypothèse : $Z(x) = a_1 \cdot f_1(x) + a_2 \cdot f_2(x) + \dots + b + R(x)$

$a_1 \cdot f_1(x) + a_2 \cdot f_2(x) + \dots + b$: partie considérée comme déterministe, fonction des variables auxiliaires. Les coefficients a_k et b ne sont pas fixés. Dans une modélisation en voisinage glissant, ils sont ajustés localement.

$R(x)$: résidu aléatoire, sans corrélation spatiale avec les f_k

Estimateur : $Z^*(x_0) = \sum_{x_i \in V} \lambda_i Z(x_i)$

x_i : points d'observation de Z

avec la condition supplémentaire sur les poids de krigeage :

$f_k(x_0) = \sum_{x_i \in V} \lambda_i f(x_i) \quad \forall k$

Modèle : la détermination directe de la structure variographique pose problème car les coefficients a_k et b , qui sont filtrés par le système de krigeage, sont considérés comme inconnus. La pratique est d'utiliser le modèle ajusté sur le variogramme des résidus expérimentaux de la régression $R(x_i) = Z(x_i) - [\hat{a}_1 \cdot f_1(x_i) + \hat{a}_2 \cdot f_2(x_i) + \dots + \hat{b}]$ ou d'effectuer une modélisation indirecte par validation croisée. Dans ce second

cas, on utilise la fenêtre de modélisation non stationnaire d'Isatis ; le choix des structures de base peut s'inspirer du variogramme des résidus.

Le tableau suivant fournit la correspondance, sous certaines hypothèses, entre ces méthodes. Pour de plus amples détails, on peut notamment se référer à Rivoirard (2002, 2003).

Méthode →	Cokrigeage colocalisé	Krigeage avec dérive externe
Conditions d'équivalence avec le krigeage du résidu	<p>Modèle à résidu, i.e. le variogramme croisé de Z et Y_1 est proportionnel au variogramme de Y_1.</p> $\gamma_{ZY_1} = a\gamma_{Y_1} \Rightarrow$ $Z^{CK}(x_0) = a \cdot Y_1(x_0) + b + R^K(x_0)$	<p>Coefficients de la dérive supposés constants sur le domaine ; modèle ajusté sur le variogramme des résidus expérimentaux de la régression.</p>

12. ANNEXE 3 : VARIABLES AUXILIAIRES

Dans l'élaboration de cartographies du NO₂, il est fréquent de recourir à des variables auxiliaires, complémentaires des données de mesure et potentiellement liées aux concentrations du ou des polluants considérés. Moyennant ou non un traitement préliminaire, selon leur nature et leur format initial, ces variables peuvent être utilisées :

- comme facteurs explicatifs dans un modèle de régression ;
- comme dérives dans un krigeage avec dérive externe ;
- dans un modèle de corégionalisation, pour une estimation par cokrigeage.

Dans ces deux derniers cas, elles ont pour fonction d'améliorer la précision de l'estimation relativement à un krigeage ordinaire.

Tableau 7 – Méthode d'estimation applicable selon les données auxiliaires disponibles

Variable auxiliaire connue sur		
	un ensemble limité de points (coïncidant ou non avec les sites de mesure) ou un maillage lâche (de densité spatiale inférieure à celle des mesures)	toute la grille d'estimation
Régression		x
Cokrigeage ordinaire	x	
Cokrigeage colocalisé		x
Krigeage avec dérive externe		x

Les variables auxiliaires pour la cartographie du NO₂ peuvent se classer en trois catégories :

1. Elles décrivent directement ou indirectement la localisation des sources de NOx et quantifient éventuellement les émissions :
 - a. inventaire spatialisé des émissions de NOx
 - b. émissions linéiques par le trafic routier
 - c. comptage automobile sur les axes routiers
 - d. données géographiques de population
 - e. base de données d'occupation du sol

f. tracé des routes

2. Elles constituent des facteurs d'influence dans la transformation et la dispersion du polluant :

- altitude

3. Comme la variable principale, elles représentent les concentrations de NO₂ mais elles proviennent :

- de campagnes de mesure d'années antérieures (données expérimentales ou estimées sur une grille) ;
- de simulations numériques (grille plus ou moins dense de valeurs).

Les variables auxiliaires brutes nécessitent généralement un prétraitement :

- transformation en variables numériques du type $f(x_i, y_j)$ des données de population, d'occupation du sol, de réseau routier..., (x_i, y_j) décrivant un maillage ;
- modification du support spatial des variables, c'est à dire de la géométrie et/ou de la taille du voisinage sur lequel sont définies les données (par exemple, à partir d'un inventaire d'émissions kilométrique, qui représente des densités d'émissions par maille de 1km x 1km, on peut calculer des densités d'émissions dans des rayons de 1km, 2km, 3km, etc., autour de chaque site de mesure et de chaque point de grille). La corrélation entre la concentration de NO₂ et la variable auxiliaire se révèle en effet sensible au support spatial de cette dernière

Ces transformations s'appliquent quelle que soit la méthode d'estimation choisie. Le tableau ci-après, issu de l'expérience et d'une analyse bibliographique, propose une liste (non exhaustive) de transformations possibles :

Exemple de variable auxiliaire brute	Variable auxiliaire traitée (calculée en chaque site de mesure / en chaque point de grille)
Inventaire spatialisé des émissions (quantité d'émissions /maille)	Densité d'émissions [kg/km ²] dans un voisinage défini par l'utilisateur*
Émissions linéiques	Densité d'émissions [kg/km ²] dans un certain voisinage*
Comptage automobile	<ul style="list-style-type: none"> - Nombre de véhicules dans un certain voisinage - Nombre de véhicules sur la route la plus proche

Population INSEE (données de communes, d'IRIS, d'îlots)	Densité de population [nb d'hab./km ²] dans un certain voisinage*
Base d'occupation du sol CORINE Land Cover	Pourcentage d'occupation du sol dans un certain voisinage*, pour chacun des postes
Base des routes	<ul style="list-style-type: none"> - Distance à la route la plus proche - Longueur totale de route par maille ou dans un certain voisinage*

*généralement circulaire. Il est conseillé de tester différents rayons de calcul.

Pour mener à bien la cartographie, les variables auxiliaires doivent être connues:

- aux points de mesure, afin d'évaluer les corrélations avec le NO₂ ;
 - pour les densités d'émissions, de population,...et les distances, la procédure la plus rigoureuse consiste à calculer directement ces variables aux points de mesure.
 - pour les autres variables (émissions brutes, concentrations, altitudes....), on effectue selon les cas une migration (attribution à un point de mesure de la valeur de variable auxiliaire la plus proche) ou une interpolation.
- en tout point de la grille d'estimation, lorsqu'une estimation par régression, krigage des résidus, krigage avec dérive externe ou cokrigage colocalisé est envisagée. Notons que la résolution de cette grille ne coïncide pas nécessairement avec le support spatial des données (Figure 13):

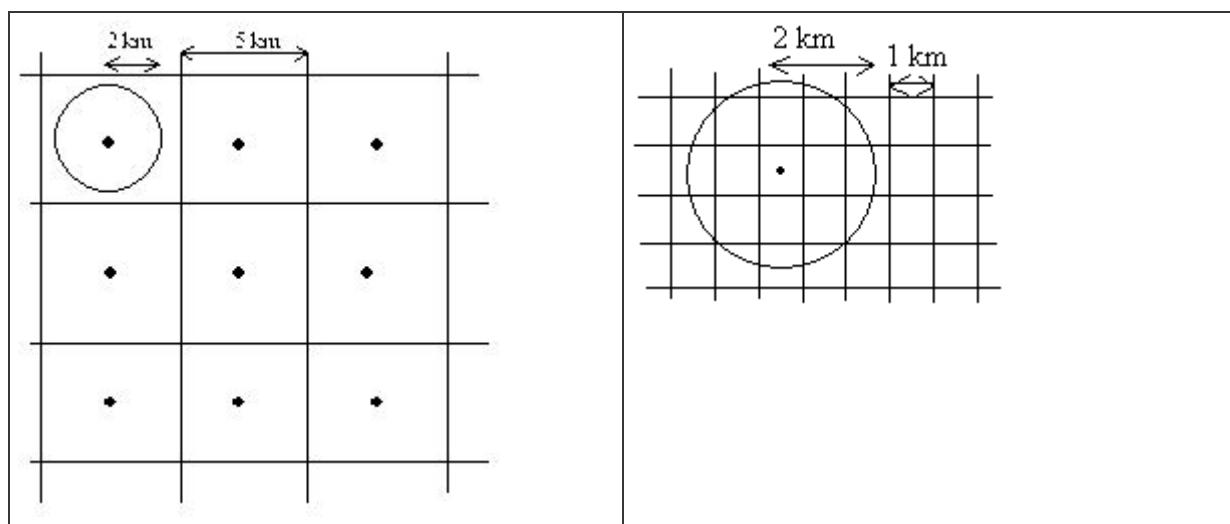


Figure 13 – Etude INERIS-2005. Calcul des densités d'émissions et de territoires artificialisés dans un rayon de 2km autour des points d'estimation (i.e. des

centres de maille). A gauche, maille d'estimation de 5km de côté ; à droite : maille d'estimation de 1km de côté.

Les calculs de densité dans des zones circulaires, les opérations d'intersection que cela implique (entre cercles et mailles, entre cercles et polygones) et les calculs de distance rendent indispensable l'usage d'un **SIG**. Plusieurs études confirment l'utilité de ce type d'outil (Hochadel et al, 2006).

Il faut noter que la recherche de variables explicatives corrélées à la concentration de NO₂ exige souvent une seconde étape de transformation mathématique des variables (exemple : passage au logarithme, racine carrée...). Cette étape peut être accomplie hors du SIG, à l'aide des logiciels de statistique ou de géostatistique utilisés (exemple : option *Calculator* d'Isatis).

13. ANNEXE 4 : ELABORATION DE CARTES SAISONNIERES – CAS D'UNE AGGLOMERATION

Deux méthodes sont appliquées pour établir des cartes moyennes saisonnières : la méthode du **cokrigage temporel** présentée dans l'étude du LCSQA 2004 : « Géostatistique et prise en compte de l'aspect temporel » (Cárdenas et Malherbe, 2004) et l'algorithme de **cartographie journalière** présenté dans l'étude LCSQA 2005 « Adaptation de méthodes géostatistiques à la cartographie automatique de NO₂ » (Cárdenas et Perdrix, 2005).

Un même jeu de données, issu de campagnes par échantillonnage passif conduites à Lille, est exploité.

13.1 COKRIGEAGE TEMPOREL

13.1.1 La méthode

Le « cokrigage temporel » consiste à estimer les concentrations saisonnières de NO₂ (moyenne estivale et moyenne hivernale) par cokrigage entre les deux saisons. Afin d'affiner ces estimations, il est possible d'utiliser une ou plusieurs variables auxiliaires comme fonctions de dérive. Celles-ci peuvent différer selon la saison.

La méthodologie pour implémenter le cokrigage temporel peut se résumer ainsi :

1. Il faut disposer de plusieurs campagnes de mesure par tubes passifs sur des périodes représentatives de chaque saison. Par exemple :

Pour l'hiver : au moins trois campagnes de mesure en janvier et février.

Pour l'été : au moins trois campagnes de mesure en juin et juillet.

2. On calcule la moyenne des campagnes de chaque saison afin d'obtenir la moyenne saisonnière, puis on vérifie la corrélation entre les mesures des deux saisons (vérifier par exemple que le coefficient de corrélation est d'au moins 0.6).
3. Comme la corrélation ne peut se calculer que sur les échantillons représentés dans les deux saisons, il faut avoir un nombre assez important d'échantillons communs aux deux saisons. Si la corrélation est forte et le nombre d'échantillons communs suffisant, on peut alors effectuer un cokrigage entre les deux moyennes saisonnières.
4. Si des données secondaires sont disponibles, on peut chercher à construire une variable auxiliaire bien corrélée à chacune des moyennes saisonnières. Les informations auxiliaires sont le plus souvent : la densité de population, le cadastre d'émissions de NO_x, la distance à la route, la densité du bâti, l'altitude, etc. Une variable peut être bien corrélée à une moyenne saisonnière mais mal corrélée à la moyenne de la saison opposée, c'est pourquoi il est conseillé de vérifier les corrélations avec les deux saisons, et s'il est besoin, de construire une variable auxiliaire spécifique à chaque saison.

5. On effectue le cokrigeage de chaque moyenne saisonnière à l'aide de la moyenne de la saison opposée et le cas échant, de la variable auxiliaire ajustée à la saison considérée. Un avantage de la méthode est le fait de pouvoir conserver les points pour lesquels les mesures d'une seule saison existent. On peut aussi calculer la variance d'estimation de chaque cokrigeage, cette variable est une mesure qualitative de la qualité de l'estimation de chaque moyenne saisonnière. Le cokrigeage suppose l'ajustement de deux variogrammes simples (un pour chaque saison) et du variogramme croisée.
6. La moyenne arithmétique des deux cokrigeages saisonniers fournit l'estimation de la moyenne annuelle du NO₂. La variance d'estimation ne peut pas être calculée de cette façon car il ne s'agit pas d'une variable additive.
7. Pour déterminer la qualité des résultats, on effectue une validation croisée (en appliquant le cokrigeage temporel) des deux moyennes saisonnières sur les échantillons, puis on calcule la moyenne de ces deux estimations et on trace le nuage de corrélation avec la moyenne annuelle des mesures.

13.1.2 Application à des données de LILLE

Il faut d'abord noter qu'en matière de pollution atmosphérique on ne considère généralement que deux saisons : l'hiver, du 1^{er} Octobre au 31 Mars, et l'été, du 1^{er} Avril au 30 Septembre. Le tableau ci-dessous affiche les statistiques des campagnes par tubes passifs organisées à Lille durant la saison estivale de 2003 (campagnes 12 à 14) et hivernale 2003-2004 (campagnes 16 à 18). Chaque saison comporte trois campagnes d'une durée d'une quinzaine de jours sur des périodes supposées représentatives des niveaux de pollution de dioxyde d'azote dans chaque saison.

Nb Q	Date Début	Date Fin	Nb Tubes	Min	Max	Moy	σ	σ^2 ($\mu\text{g}/\text{m}^3$) ²	CV
12	22-mai-03	5-juin-03	131	8	63.4	27.66	8.31	69.06	0.3
13	5-juin-03	19-juin-03	126	15.1	53.9	25.1	6.33	40.05	0.25
14	19-juin-03	3-juil-03	133	14.4	60.7	25.25	7.04	49.62	0.28
<i>Moyenne été 2003</i>	22-mai-03	3-juil-03	137	15.2	59.07	25.91	6.89	47.47	0.266

Nb Q	Date Début	Date Fin	Nb Tubes	Min	Max	Moy	σ	σ^2 ($\mu\text{g}/\text{m}^3$) ²	CV
16	8-janv-04	22-janv-04	131	17.46	52.33	31.18	5.87	34.41	0.19
17	22-janv-04	5-févr-04	131	19.04	47.73	33.15	5.9	34.86	0.18
18	5-févr-04	18-févr-04	134	22.89	52.04	37.87	6.39	40.85	0.17
<i>Moyenne hiver 2003-2004</i>	8-janv-04	18-févr-04	136	21.12	50.43	34.08	5.36	28.7	0.157

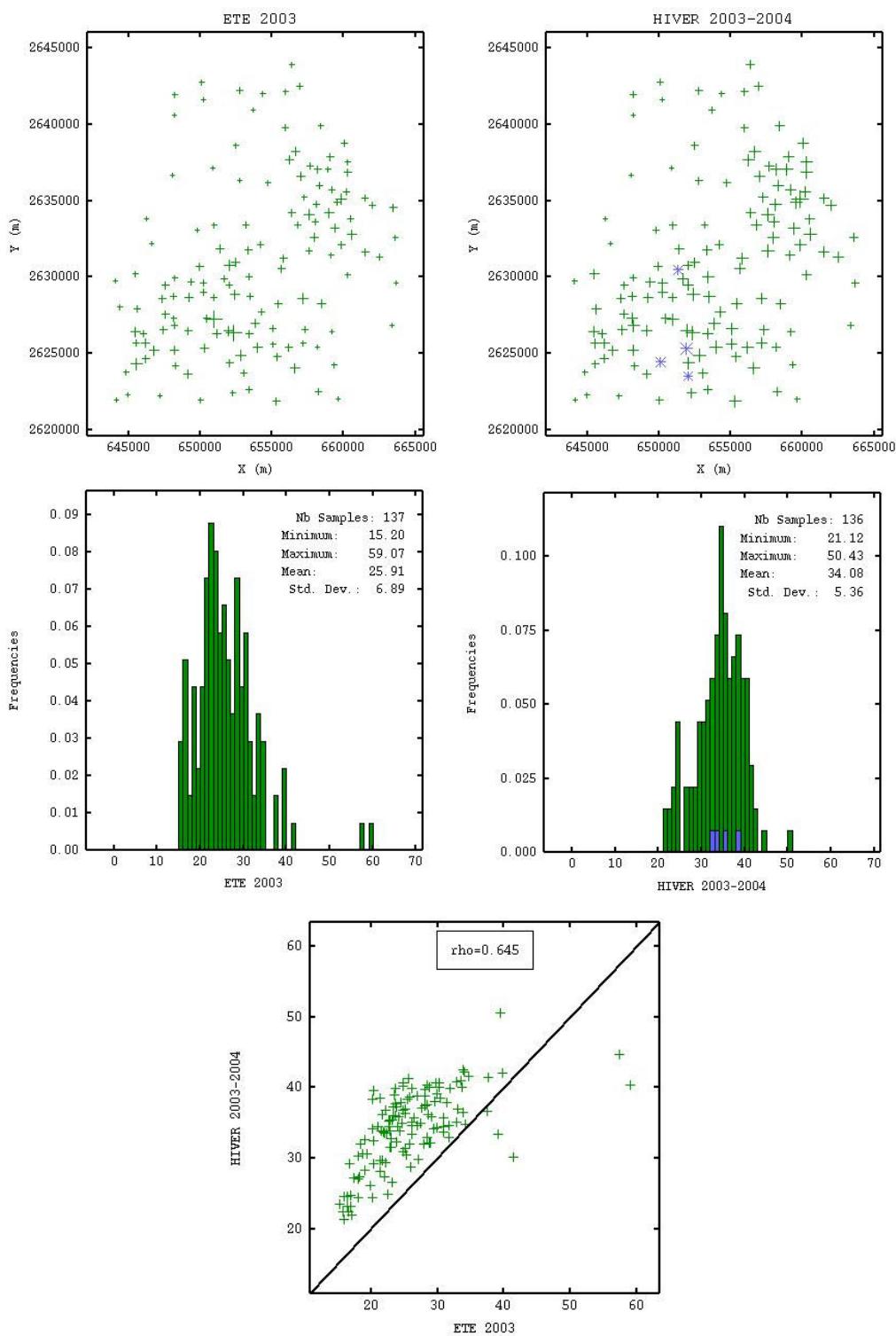
2003-2004	Date Début	Date Fin	Nb Tubes	Min	Max	Moy	σ	σ^2	CV

								($\mu\text{g}/\text{m}^3$) ²	
<i>Moyenne Annuelle</i>	22-mai-03	18-févr-04	132	18.46	50.96	29.93	5.62	31.58	0.187

Chaque moyenne saisonnière est calculée aux points où au moins deux mesures sont disponibles sur la saison. La moyenne annuelle est calculée comme la somme des deux moyennes saisonnières, divisée par deux ; ce calcul est donc réalisé sur les échantillons pour lesquels les deux moyennes saisonnières sont disponibles, c'est-à-dire sur 132 points de mesure.

Les figures suivantes illustrent les différentes étapes de la méthode :

- étude des corrélations entre saisons (Figure 14) ;
- recherche d'une ou plusieurs variables auxiliaires corrélées aux saisons (Figure 15 ; Figure 16) ;
- modélisation simple et croisée de l'été et de l'hiver (Figure 17) ;
- validation croisée (Figure 18 ; Figure 19) ;
- estimation (Figure 20 ; Figure 21 ; Figure 22).



Deux figures en haut : localisation géographique des échantillons

Deux figures au milieu : histogrammes avec statistiques des moyennes saisonnières

A gauche : Moyenne estivale 2003

A droite : Moyenne hivernale 2003-2004

Figure en bas: nuage de corrélation entre les deux moyennes saisonnières

Figure 14 : Cartes et statistiques des deux moyennes saisonnières

INERIS, DRC-07-74777-03941A

Ecole des Mines de Douai - Département Chimie et Environnement

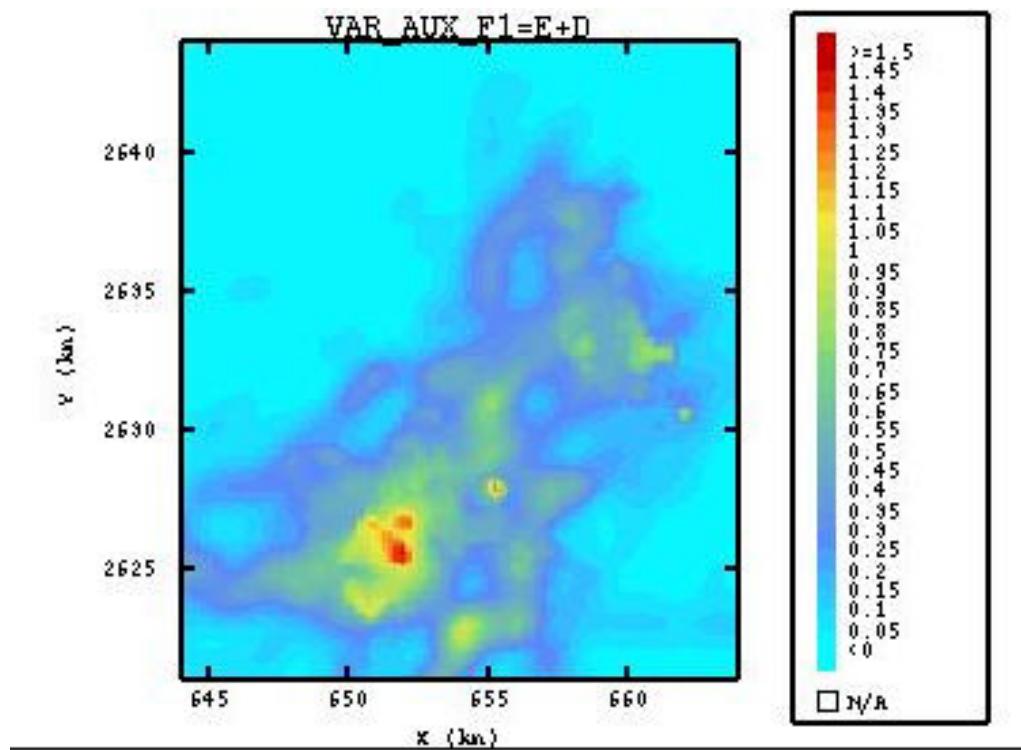


Figure 15 : Carte de la variable auxiliaire

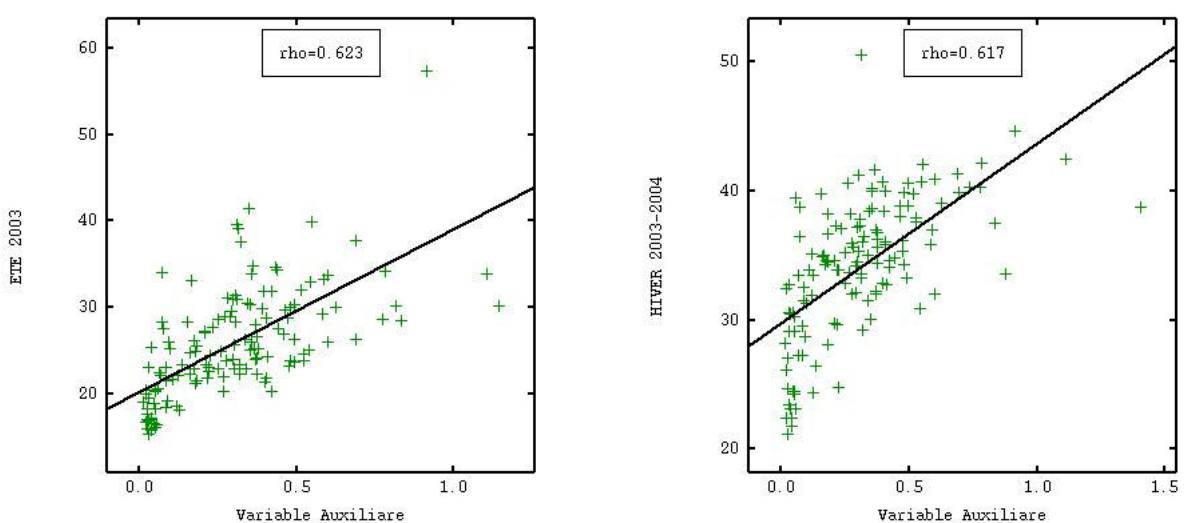
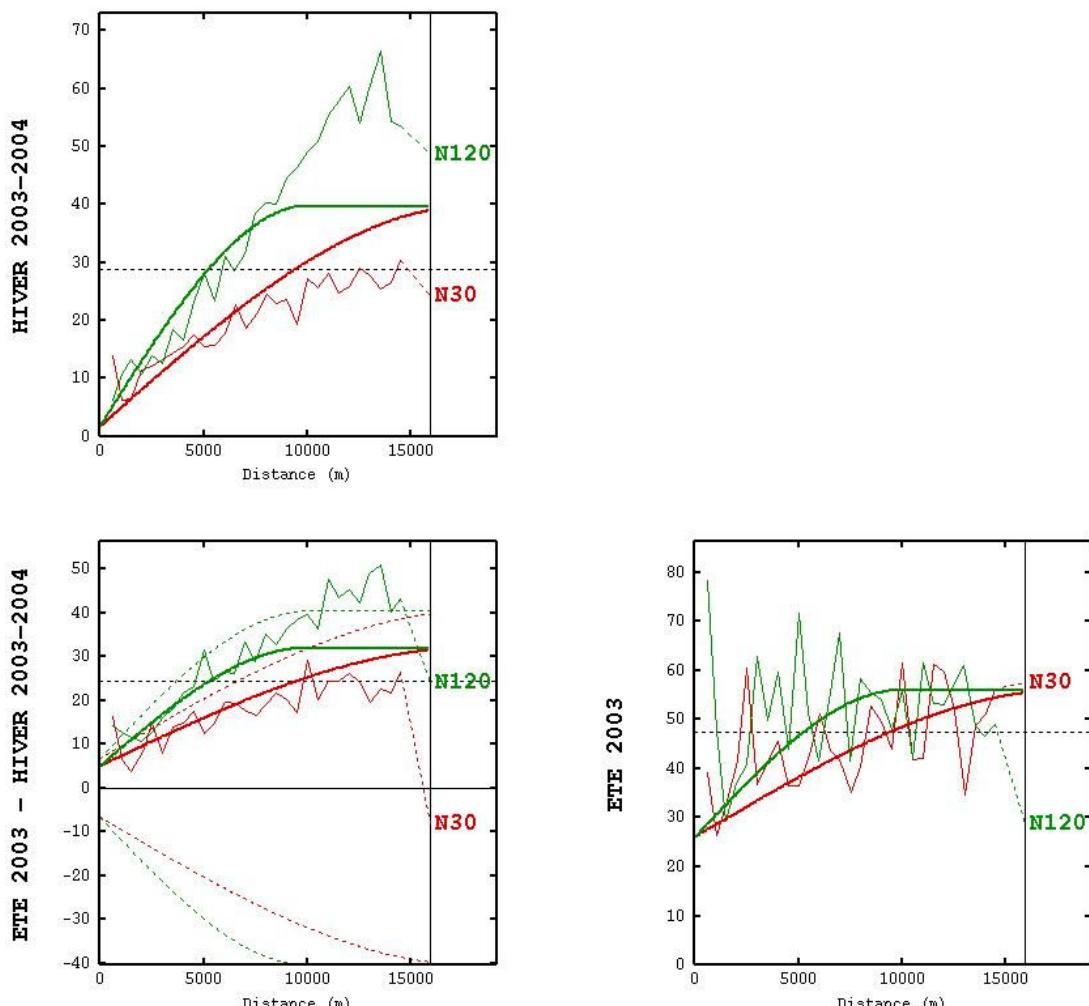
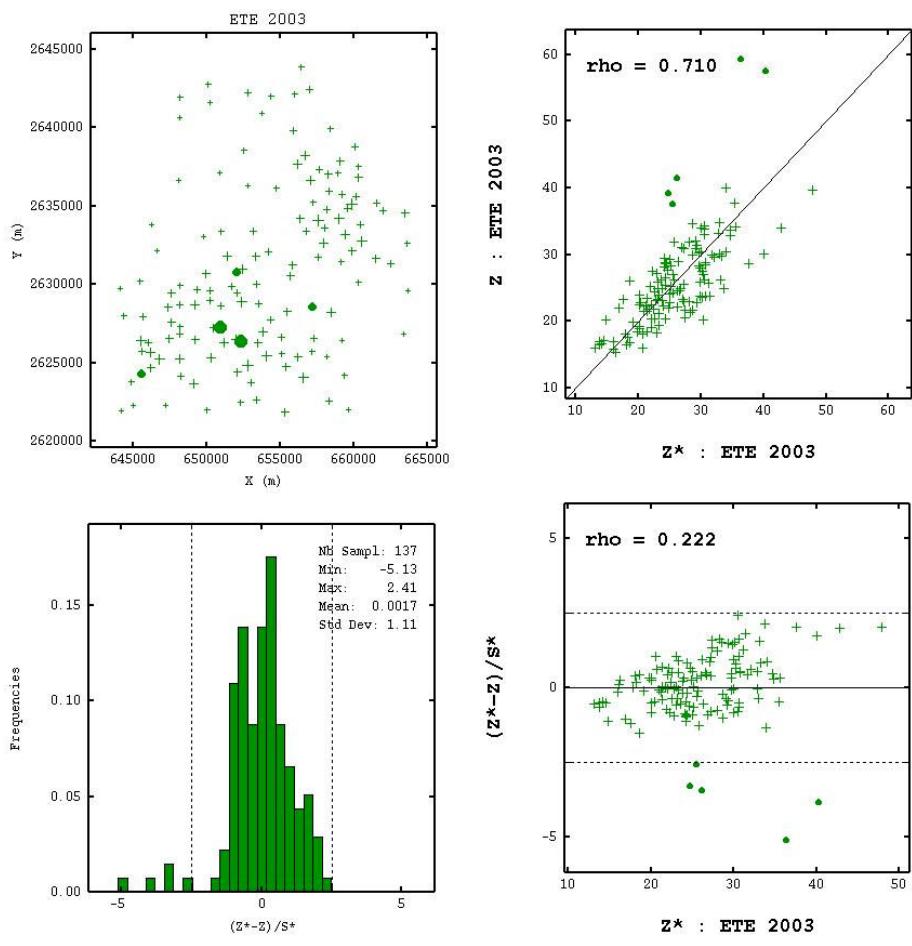


Figure 16 : Nuage de corrélation entre les deux moyennes saisonnières et la variable auxiliaire



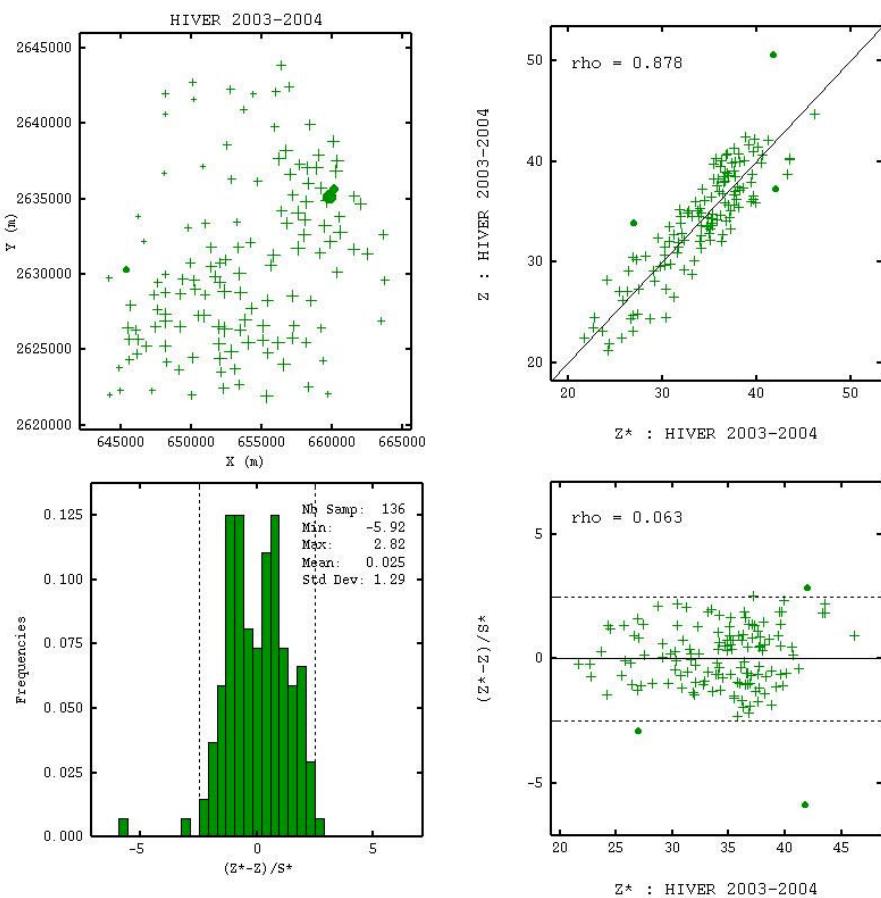
Variable	Direction	Structure	Palier	Portée (Km)
Eté 2003	N30°E (Courbe Rouge)	Pépite	26	---
		Sphérique	30	18
	N120 = N60°O (Courbe verte)	Pépite	26	---
		Sphérique	30	10
Hiver 2003-2004	N30°E (Courbe Rouge)	Pépite	1.76	---
		Sphérique	38	18
	N120 = N60°O (Courbe verte)	Pépite	1.76	---
		Sphérique	38	10
Croisée	N30°E (Courbe Rouge)	Pépite	5	---
		Sphérique	27	18
	N120 = N60°O (Courbe verte)	Pépite	5	---
		Sphérique	27	10

Figure 17 : Modèle des variogrammes simples et croisé des moyennes saisonnières



Variable (Nb=137)	Moyenne	Variance
Erreur	0.017	24.88
Erreur Réduite	0.0017	1.23

Figure 18 : Validation croisée pour la moyenne estivale



Variable (Nb=136)	Moyenne	Variance
Erreurs	0.1	6.62
Erreurs Réduites	0.025	1.67

Figure 19 : Validation croisée pour la moyenne hivernale

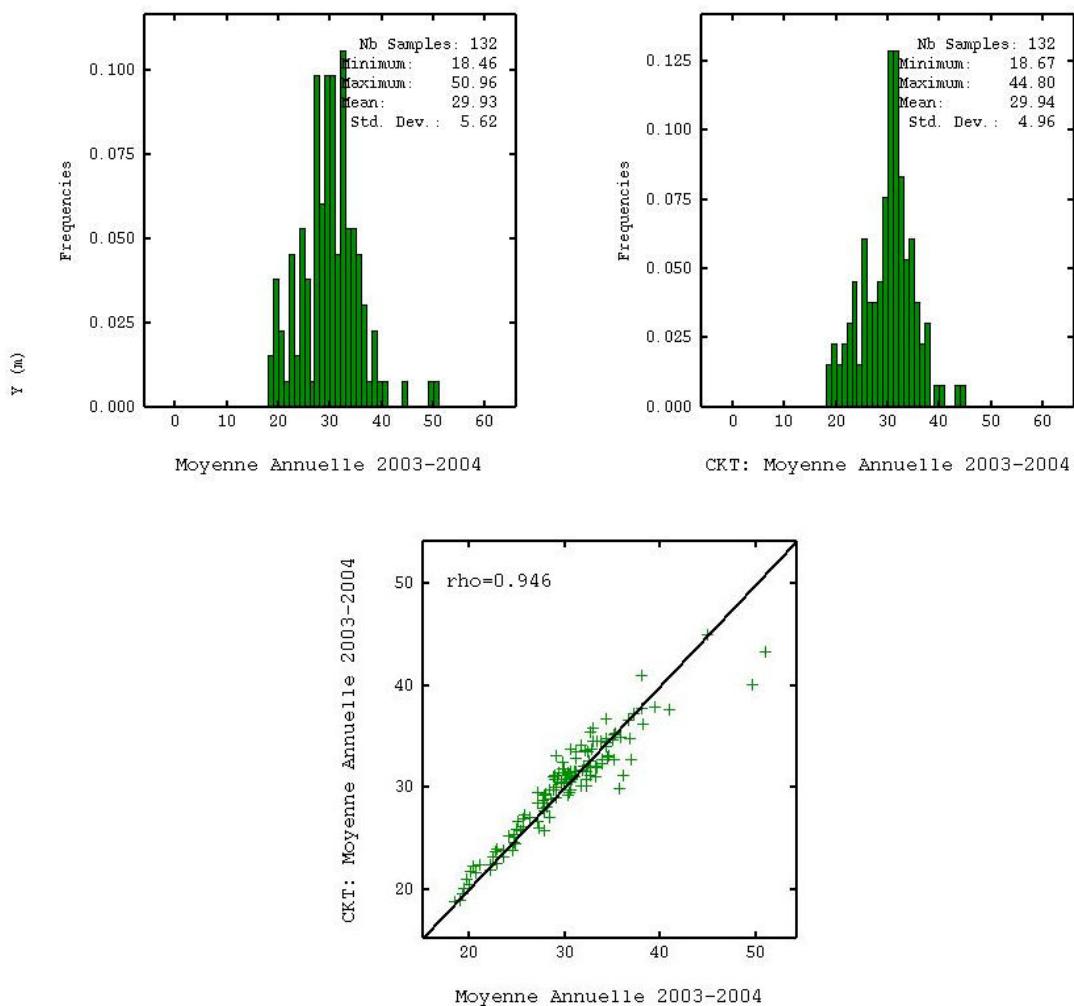


Figure 20 : Comparaison des résultats de l'estimation de la moyenne annuelle de NO₂ par cokrigeage temporel avec la moyenne annuelle statistique, aux tubes (année 2003-2004)

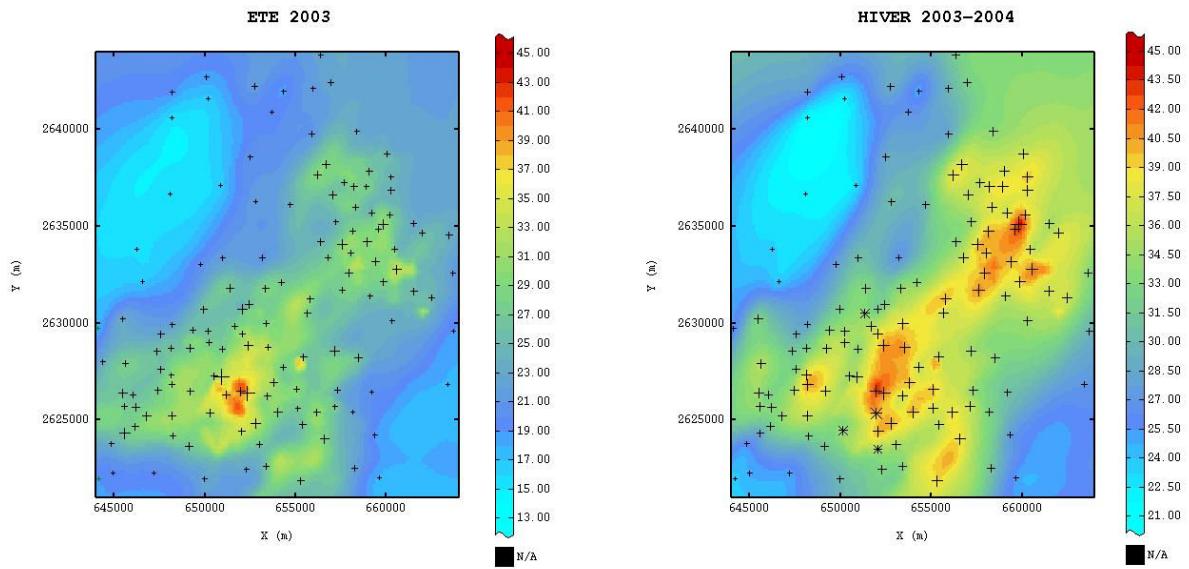


Figure 21 : Estimations des deux moyennes saisonnières par cokrigeage temporel

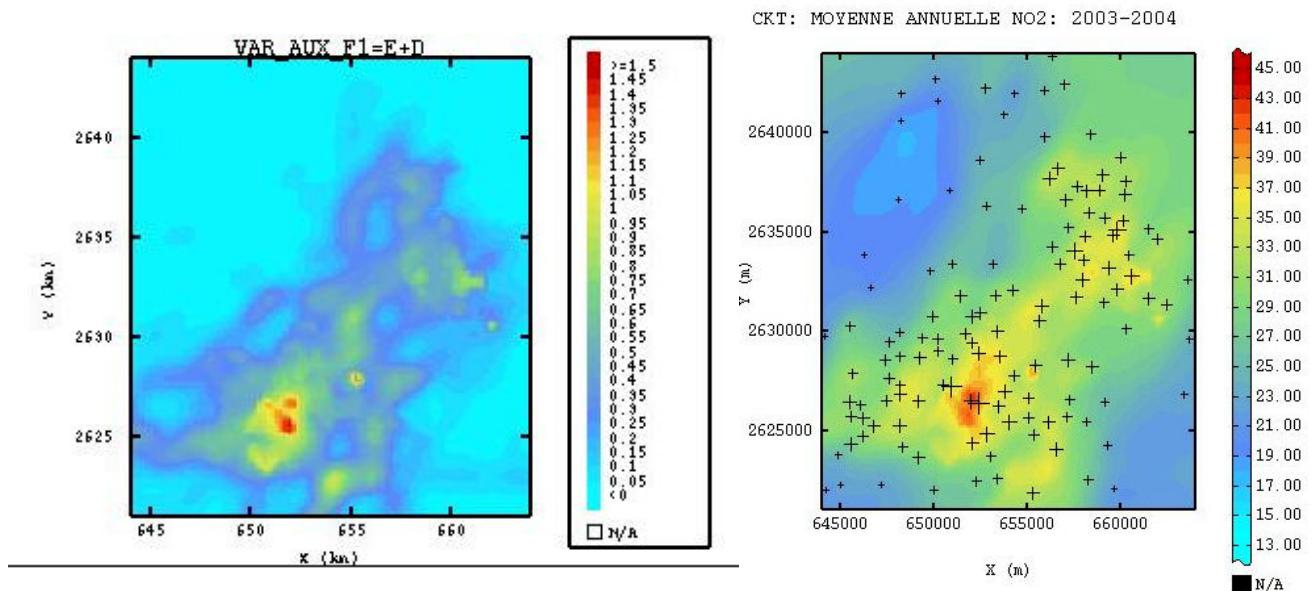


Figure 22 : Comparaison de l'estimation de la moyenne annuelle avec la variable auxiliaire (NO2, année 2003-2004)

CKT	Nb	Min	Max	Moy	σ	σ^2 ($\mu\text{g}/\text{m}^3$) ²	CV
Moyenne été 2003	7360	14.76	41.93	22.9	4.76	22.68	0.21
Moyenne hiver 2003-2004	7360	20.94	44.78	31.42	5.24	27.49	0.17
Moyenne Annuelle	7360	17.88	42.71	27.16	4.87	23.69	0.18

13.2 CARTOGRAPHIE JOURNALIERE AUTOMATIQUE

Sur le même domaine lillois, on applique la méthode de cartographie journalière automatique. Chaque jour les concentrations de NO_2 sont estimées par cokrigeage colocalisé en utilisant :

- comme données d'entrée, les mesures journalières des stations fixes et les données reconstituées aux stations virtuelles ;
- comme covariable, la variable auxiliaire construite précédemment ;
- comme modèle, un modèle bivariable construit à partir des données d'échantillonnage (variable 1 = NO_2 ; variable 2 = variable auxiliaire).

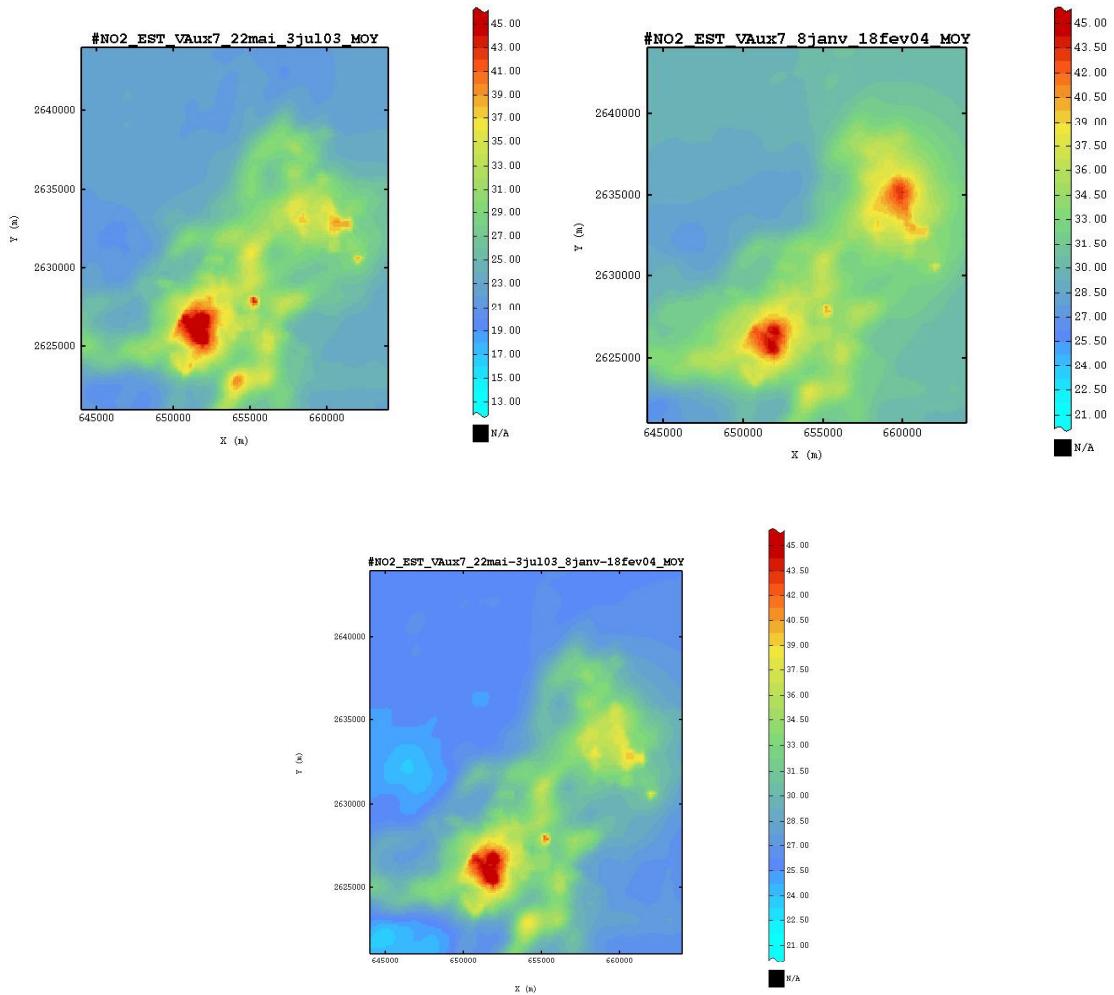
Pour de plus amples détails on peut se rapporter à Cárdenas et Perdrix (2005, 2006).

Les cartes bilans ci-dessous représentent la moyenne des estimations journalières :

- sur les périodes des campagnes exploitées précédemment (Figure 23),
- sur chaque saison (au sens large) et sur l'année (Figure 24).

Si elles restituent moins finement la variabilité des concentrations en centre-ville, à cause du manque de stations dans cette zone, les cartes de la Figure 23 sont cohérentes avec celles de la Figure 21 et de la Figure 22.

Celles de la Figure 24, qui intègrent l'information des stations fixes sur les saisons complètes, présentent des niveaux de concentration globalement plus élevés.

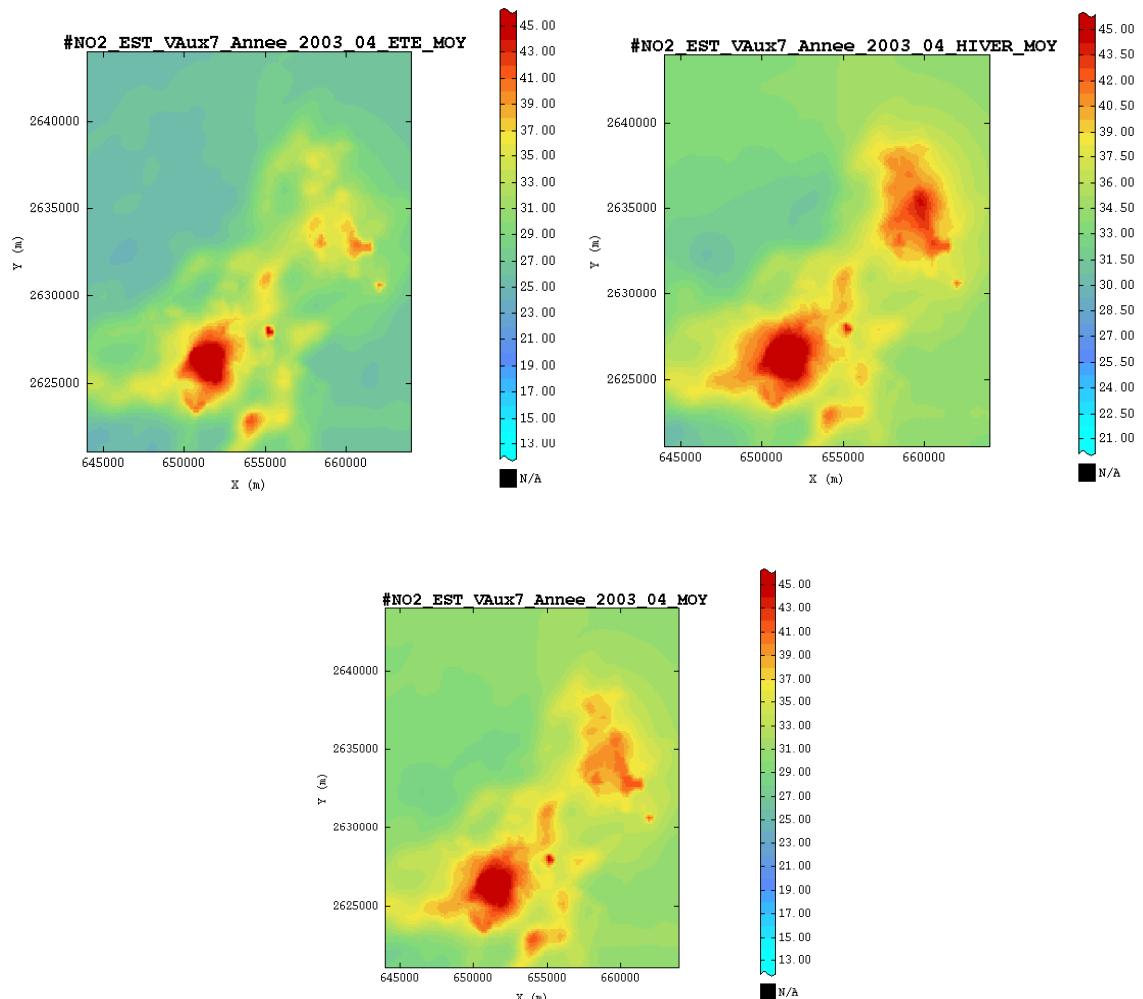


En haut à gauche : Du 23 mai au 3 juillet 2003 (période estivale 2003)

En haut à droite : Du 8 janvier au 18 février 2004 (période hivernale 2003-2004)

En bas : Du 23 mai au 3 juillet 2003 et du 8 janvier au 18 février 2004

Figure 23 : Moyennes sur les périodes des campagnes des estimations journalières obtenues par la méthode du cokrigeage journalier automatique



En haut à gauche : Du 1^{er} avril au 30 septembre 2003

En haut à droite : Du 1^{er} octobre 2003 au 8 mars 2004

En bas : Du 1^{er} avril 2003 au 8 mars 2004

Figure 24 : Moyennes saisonnières et annuelles des estimations journalières obtenues par la méthode du cokrigeage journalier automatique

14. ANNEXE 5 : APPLICATION DE LA GEOSTATISTIQUE A L'EXPLOITATION DE MESURES CONJOINTES DE FOND ET DE PROXIMITE – RAPPEL DES RESULTATS

Le déroulement de l'étude est décrit dans Malherbe et Cárdenas (2005). Nous en rappelons les étapes.

Ce travail repose sur l'exploitation de deux campagnes de mesure conduites par LIG'Air sur toute la région Centre :

- la première a eu lieu d'avril 2003 à mars 2004 ; elle comprenait 19 sites de proximité routière.
- la seconde s'est déroulée de janvier à décembre 2005 ; 49 sites de fond répartis entre 42 communes ont fait l'objet de mesures.

Dans les deux cas, des tubes à échantillonnage passif renouvelés mensuellement ont été utilisés en continu pendant douze mois.

Des informations secondaires ont été de plus considérées :

- inventaire kilométrique des émissions de NOx (élaboré par LIG'Air) ;
- occupation du sol (CORINE Land Cover, IFEN) ;
- population (données IRIS et îlots, INSEE).

La méthodologie proposée pour la réalisation d'une cartographie comprend deux grandes parties : estimation de la pollution de fond ; prise en compte d'une pollution additionnelle liée au trafic routier.

14.1 CARTOGRAPHIE DE LA POLLUTION DE FOND

Remarque : La démarche a été mise au point en 2005 sur les données d'hiver de cette même année (janvier-mars). Les résultats obtenus pour cette période sont donc plus particulièrement détaillés. En 2006, des calculs complémentaires ont été effectués pour l'été et l'année 2005 complète ; nous en présentons brièvement les résultats.

L'histogramme hivernal (estival) présente deux modalités qui ne correspondent pas à un découpage de l'espace mais sont liées à l'importance des villes et des routes présentes dans l'environnement des mesures. Il montre ainsi la contribution des émissions locales à la pollution de fond. Il ne justifie pas un traitement séparé des données de chaque modalité, puisque celles-ci caractérisent un même type de pollution, mais il suggère l'utilité de recourir à des variables descriptives des sources.

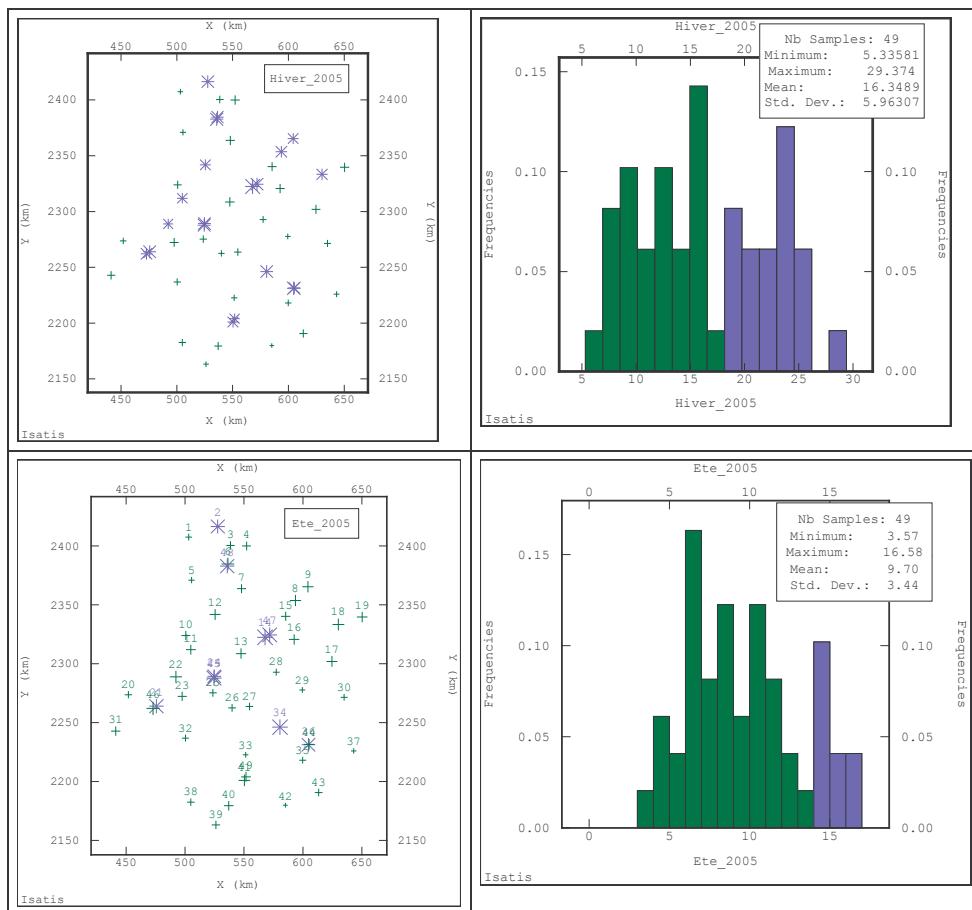


Figure 25 – Histogramme des concentrations hivernales (ligne du haut) et estivales (ligne du bas) ($\mu\text{g}/\text{m}^3$)

14.1.1 Relations avec les variables auxiliaires

A partir des données secondaires disponibles, des variables de densité ont été calculées dans différents rayons autour des sites de mesure : densité d'émissions, pourcentage d'occupation du sol par les différents postes de la nomenclature CORINE, densité de population.

La concentration de NO_2 est plus étroitement corrélée :

- avec la densité d'émissions dans un rayon de 2km ;
- avec le taux de territoires « artificialisés »⁶ dans un rayon de 2km (le nuage de corrélation n'est toutefois pas aussi linéaire) ;
- avec la densité de population dans un rayon de 5km.

Il s'agit plus précisément des transformées logarithmiques de ces variables – de la forme $\ln(1+X)$.- Dans la suite le terme de **densité** désigne plus succinctement la transformée logarithmique de la densité.

⁶ Ce terme englobe les zones de bâti continu et discontinu, les zones industrielles, les axes de circulation, cf. nomenclature CORINE.

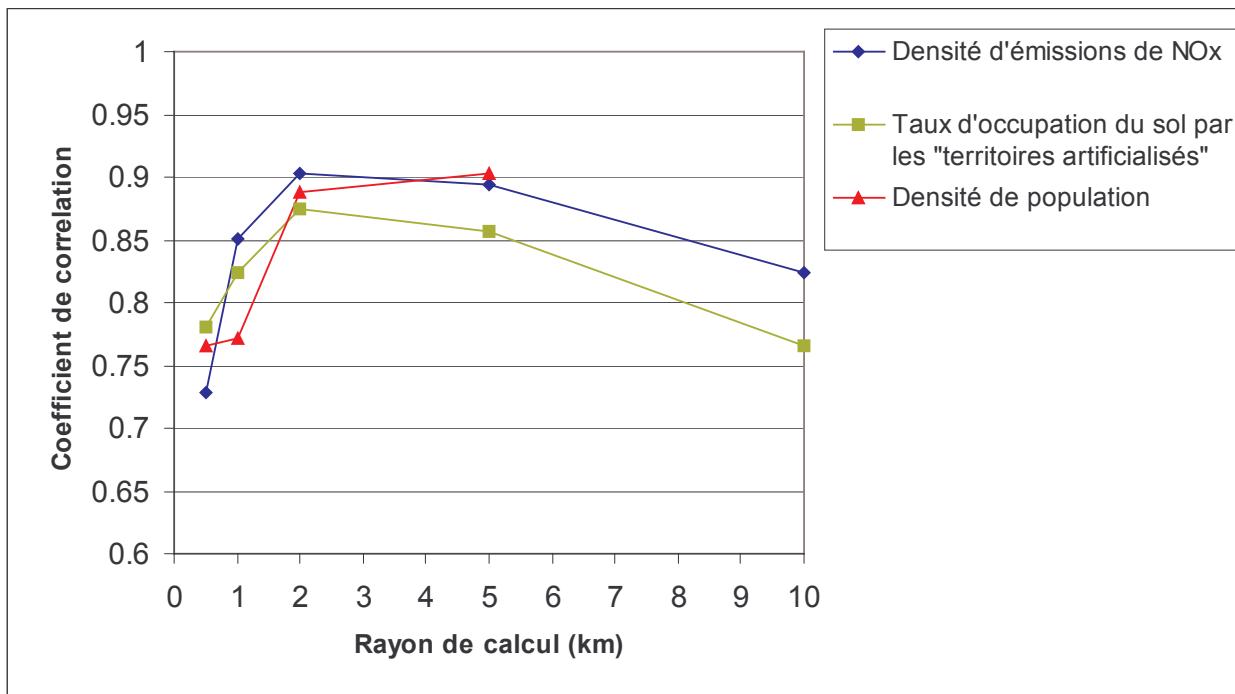


Figure 26 – Evolution du coefficient de corrélation entre la concentration hivernale de NO_2 et la variable de densité, en fonction du rayon de calcul de cette dernière

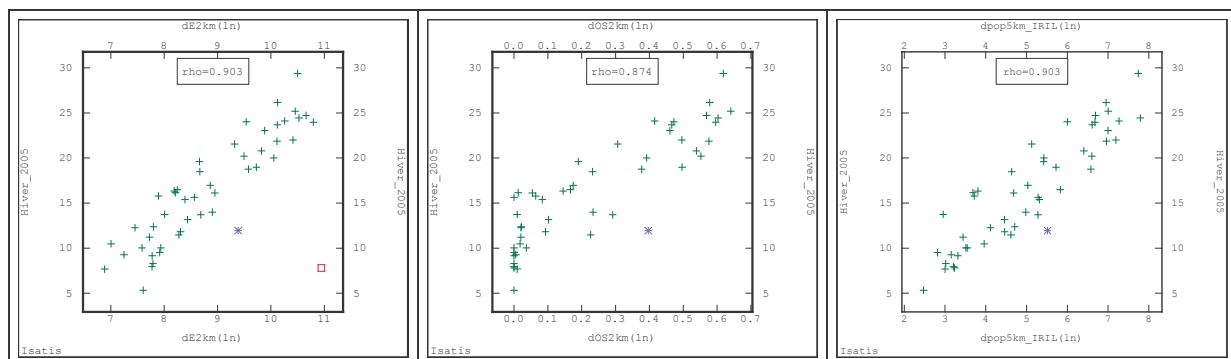


Figure 27 – Nuages de corrélation entre la concentration hivernale de NO_2 et les variables de densité sélectionnées

Bien que les nuages soient plus dispersés, les variables retenues pour l'été restent les mêmes :

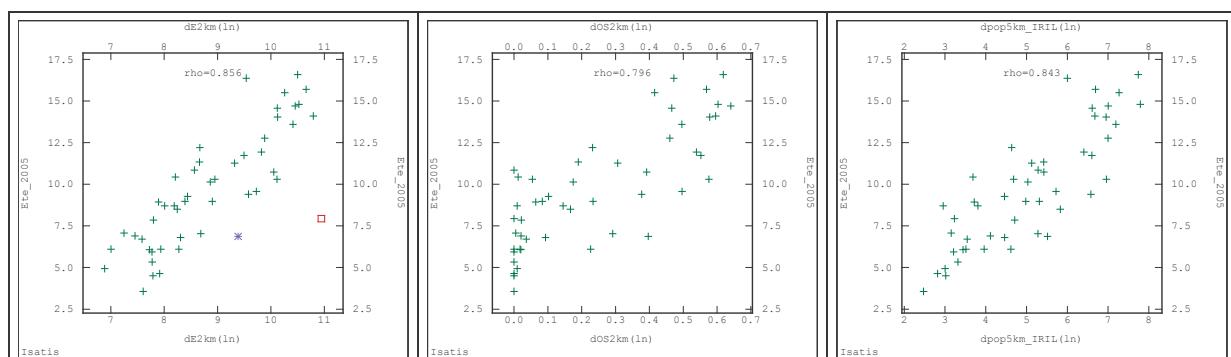


Figure 28 - Nuages de corrélation entre la concentration estivale de NO_2 et les variables de densité sélectionnées

Sur les 49 sites de fond, ces variables auxiliaires se révèlent très corrélées entre elles, comme l'indique notamment une analyse en composantes principales (Figure 29). C'est pourquoi chacune de ces variables est par la suite utilisée seule.

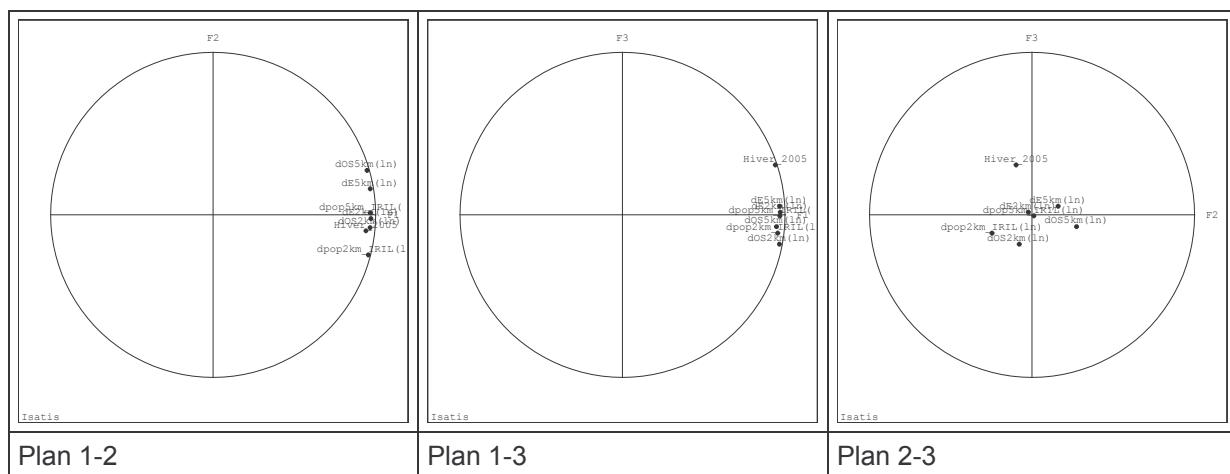


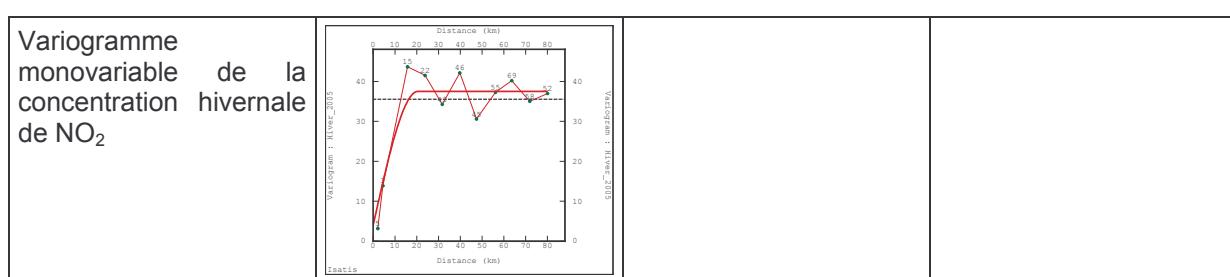
Figure 29 – Analyse en composantes principales sur les variables de densité sélectionnées

14.1.2 Modélisation du variogramme

Plusieurs modélisations sont réalisées :

- modélisation monovariable, en vue d'un krigage ordinaire ;
- pour chaque variable auxiliaire, modélisation du variogramme des résidus de la régression linéaire ($\text{NO}_2 = a.\text{var_aux} + b$), en vue d'un krigage du résidu ;
- pour chaque variable auxiliaire, modélisation non stationnaire en voisinage glissant (en s'aidant du variogramme des résidus), en vue d'un krigage avec dérive externe ;
- modélisation bivariable, en vue d'un cokrigage colocalisé.

Tous ces modèles sont testés et comparés par validation croisée. Le **krigeage du résidu** ou le **krigeage avec dérive externe** fournissent les résultats les plus satisfaisants. Des exemples de variogrammes ajustés sont présentés ci-dessous. En été, les variogrammes ont un aspect moins structuré qu'en hiver.



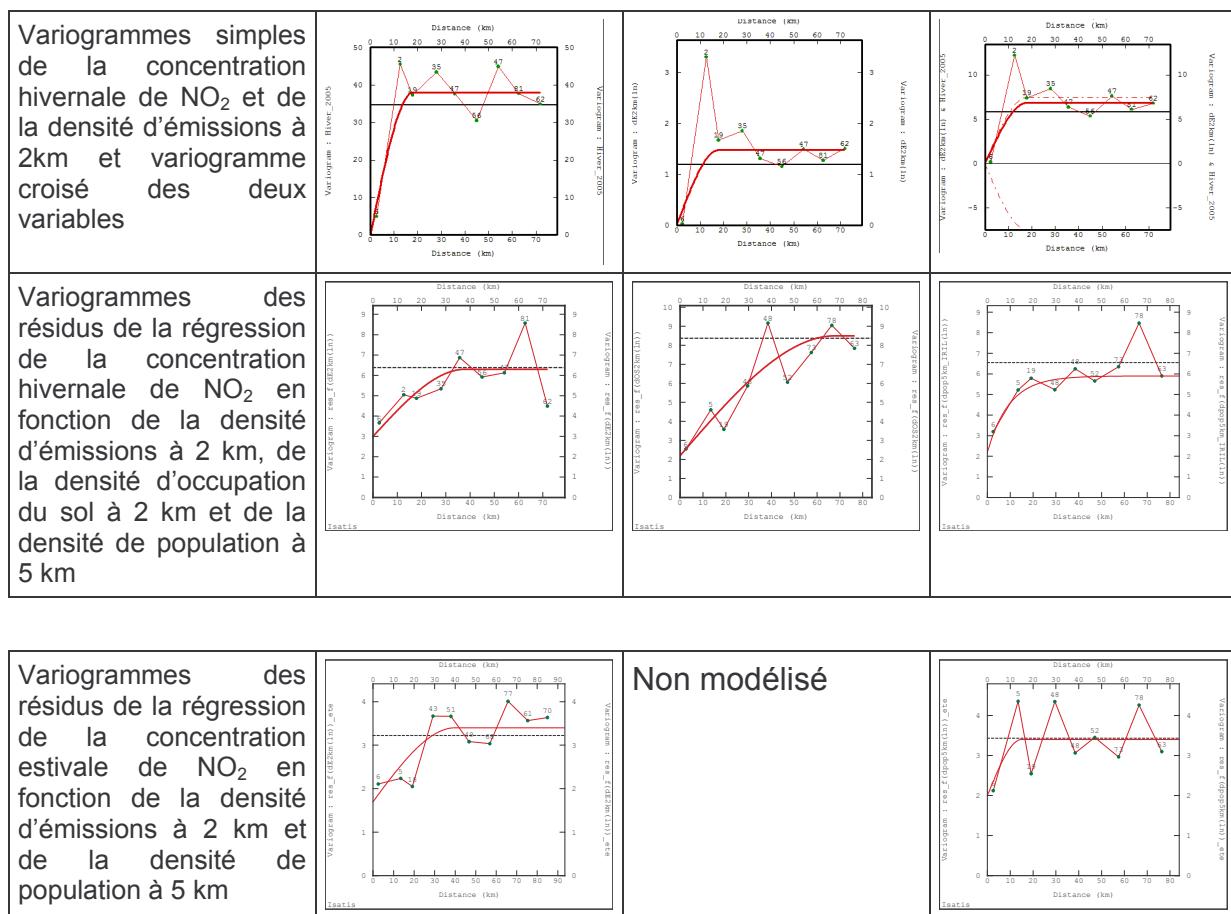


Figure 30 – Exemples de variogrammes expérimentaux et de modèles ajustés.

Dans les figures précédentes, les modèles de l'hiver et de l'été ont été construits séparément.

On a également réalisé un **modélisation bivariable de ces deux périodes** (pour la densité d'émissions seulement) en vue de tester le **cokrigeage temporel** (cf. annexe 4) :

- modélisation simple et croisée des variogrammes des résidus de l'hiver et de l'été
- estimation par cokrigeage de l'hiver et de l'été. Les variables d'entrée sont la concentration hivernale, la concentration estivale et la densité d'émissions, introduite en tant que dérive externe. Le modèle bivariable précédemment défini est utilisé, en spécifiant dans Isatis qu'il s'agit d'un modèle à dérive.

Pour l'estimation des concentrations annuelles, deux voies ont été explorées :

- On prend la **moyenne des estimations saisonnières** établies à partir des 3 mois d'hiver et des 3 mois d'été.
- On fait une **analyse géostatistique des moyennes annuelles** calculées sur les 12 mois de mesure.

14.1.3 Evaluation des modèles

A défaut de données d'échantillonnage indépendantes, les modèles variographiques sont évalués par validation croisée et comparaison aux mesures des stations fixes.

14.1.3.1 Validation croisée

Les résultats suivants concernent les modèles choisis à l'issue des tests de validation croisée. Nous présentons :

- les nuages de corrélation entre les valeurs mesurées et estimées ;
- les principales statistiques des erreurs relatives (calculées en valeur absolue) ; ces dernières correspondent à la définition de l'incertitude fixée par la réglementation. Pour le NO_2 , en modélisation, l'incertitude sur la moyenne annuelle ne doit pas dépasser 30%.

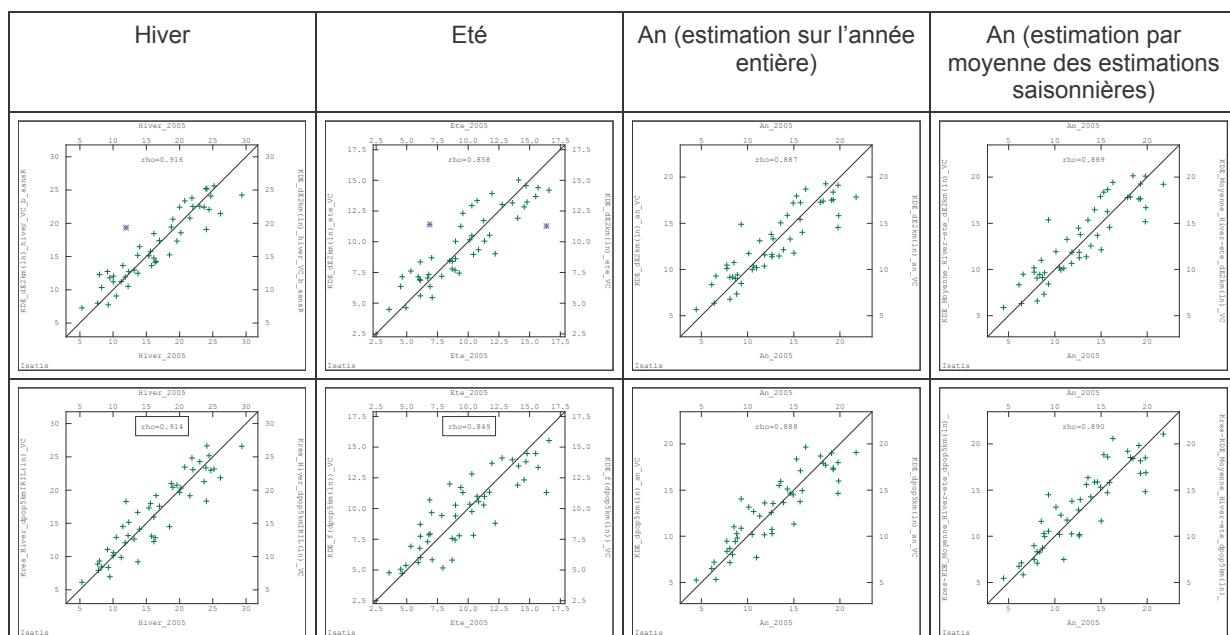
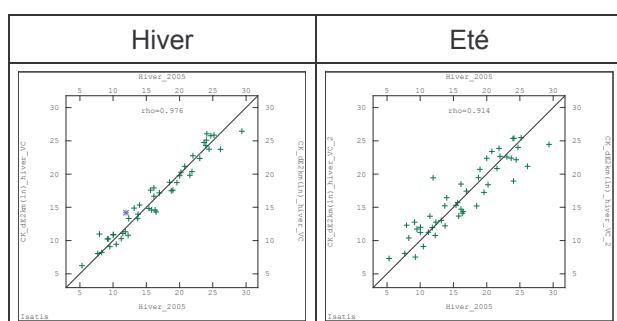


Figure 31 – Nuage de corrélation entre les concentrations de NO_2 mesurées (en abscisse) et estimées (en ordonnée) par validation croisée sur les 49 sites de fond ($\mu\text{g}/\text{m}^3$). Ligne du haut : avec la densité d'émissions à 2km. Ligne du bas : avec la densité de population à 5km.



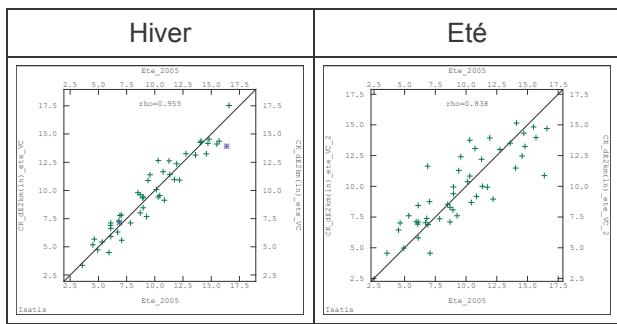


Figure 32 – Nuage de corrélation entre les concentrations saisonnières de NO_2 mesurées (en abscisse) et estimées (en ordonnée) par validation croisée ($\mu\text{g}/\text{m}^3$) en cokrigeage temporel.

Ligne du haut : le cokrigeage en un point de la concentration hivernale (resp. estivale) utilise, si elle est disponible, la valeur estivale (resp. hivernale) en ce même point ainsi que les mesures d'été et d'hiver de tous les autres sites.

Ligne du bas : le cokrigeage en un point de la concentration hivernale (resp. estivale) n'utilise que les mesures d'été et d'hiver des autres sites. Lorsqu'on estime les concentrations sur une grille, c'est dans cette configuration que l'on se trouve le plus souvent, un centre de maille coïncidant rarement avec un point de mesure.

- Tous les nuages de corrélation montrent un bon accord entre l'estimation et la mesure. En été, ils sont plus dispersés, indiquant une moindre précision d'estimation.
- D'après la Figure 32, l'estimation de la concentration hivernale (resp. estivale) en un point où cette variable est inconnue est améliorée par la connaissance de la concentration estivale (resp. hivernale). Le cokrigeage temporel peut faire ainsi gagner en précision lorsque les mesures d'été et d'hiver ne se trouvent pas toutes aux mêmes points ou qu'elles comportent des données manquantes. Il pourrait se révéler utile si les variables explicatives différaient selon la saison. Dans notre situation où toutes les données d'été et d'hiver sont disponibles exactement aux mêmes points et où la variable auxiliaire d'émissions est commune aux deux saisons, il n'apporte pas d'amélioration.
- Dans cette étude, les trois mois d'hiver et les trois mois d'été permettent une bonne représentation de la moyenne annuelle (Figure 33). La moyenne des estimations hivernale et estivale est très proche de l'estimation fondée sur les moyennes annuelles réellement mesurées.
- Selon la validation croisée, sur l'hiver comme sur l'année entière, l'objectif de qualité de la réglementation est respecté pour une majorité de sites (Tableau 8, Tableau 9). En été, le nombre de valeurs qui y satisfont est plus petit mais les niveaux de concentration sont aussi plus bas. De plus, le critère réglementaire correspond à un objectif annuel.

Remarque : ces statistiques d'erreur mesurent l'efficacité du modèle aux 49 sites de fond implantés dans des communes et agglomérations de la région. Pour ces sites, elles sont légèrement plus favorables à l'usage de la densité de population. Aucun jeu de données complémentaire ne permet de vérifier que c'est encore le cas entre les communes, dans les zones moins urbanisées.

Tableau 8 – Statistiques des erreurs relatives (en valeur absolue) – variable auxiliaire : densité d'émissions à 2km

	Hiver : de janvier à mars 2005 <i>Noir : par rapport aux 3 mois d'hiver</i> <i>Rouge : par rapport aux 6 mois de la saison hivernale</i>	Eté : de juillet à septembre 2005 <i>Noir : par rapport aux 3 mois d'été</i> <i>Rouge : par rapport aux 6 mois de la saison estivale</i>	Année : moyenne des estimations hivernales et estivales <i>Par rapport à l'année entière</i>	Année complète 2005 <i>Par rapport à l'année entière</i>
Moyenne	13.1 12.8	16.3 17.7	14.6	14.1
P90	24.2 23.7	32.9 40.6	29.7	28.4
P95	38.0 29.7	42.2 53.9	34.9	36.1
Maximum	61.8 59.7	66.1 75.5	65.1	59.9
Ecart-type	12.9 10.9	14.0 18.0	12.3	11.7
Corrélation	0.92 0.91	0.86 0.86	0.89	0.89

Tableau 9 – Statistiques des erreurs relatives (en valeur absolue) – variable auxiliaire : densité de population à 5km

	Hiver : de janvier à mars 2005 <i>Noir : par rapport aux 3 mois d'hiver</i> <i>Rouge : par rapport aux 6 mois de la saison hivernale</i>	Eté : de juillet à septembre 2005 <i>Noir : par rapport aux 3 mois d'été</i> <i>Rouge : par rapport aux 6 mois de la saison estivale</i>	Année : moyenne des estimations hivernales et estivales <i>Par rapport à l'année entière</i>	Année complète 2005 <i>Par rapport à l'année entière</i>
Moyenne	12.6 12.2	16.3 18.0	13.1	13.1
P90	23.8 22.7	33.9 37.2	25.4	24.9
P95	26.8 28.4	39.7 47.6	31.0	29.7
Maximum	53.0 50.9	55.7 64.6	55.6	50.8
Ecart-type	10.2 10.2	12.9 14.7	10.9	9.72
Corrélation	0.91 0.91	0.85 0.86	0.89	0.89

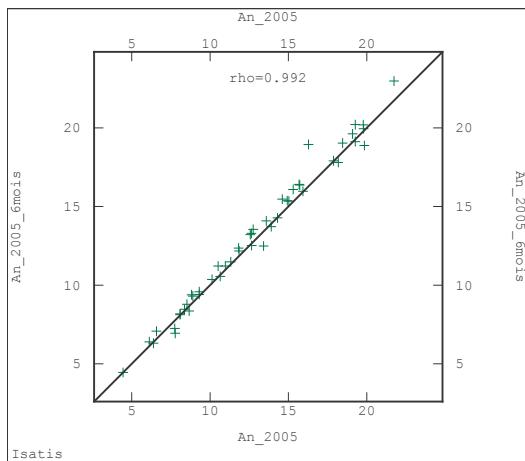


Figure 33 – Corrélation entre la moyenne des 3 mois d'hiver et des 3 mois d'été et la moyenne annuelle réellement observée.

14.1.3.2 Comparaison avec les mesures de stations fixes

Les concentrations de NO_2 sont estimées aux points des stations fixes et comparées aux données enregistrées par ces dernières. Dans les centres urbains, en hiver comme en été, l'estimation est plus précise lorsqu'on utilise la densité de population (Figure 34, Tableau 10, Tableau 11). Les concentrations observées dans les plus grandes villes de la région sont sous-estimées.

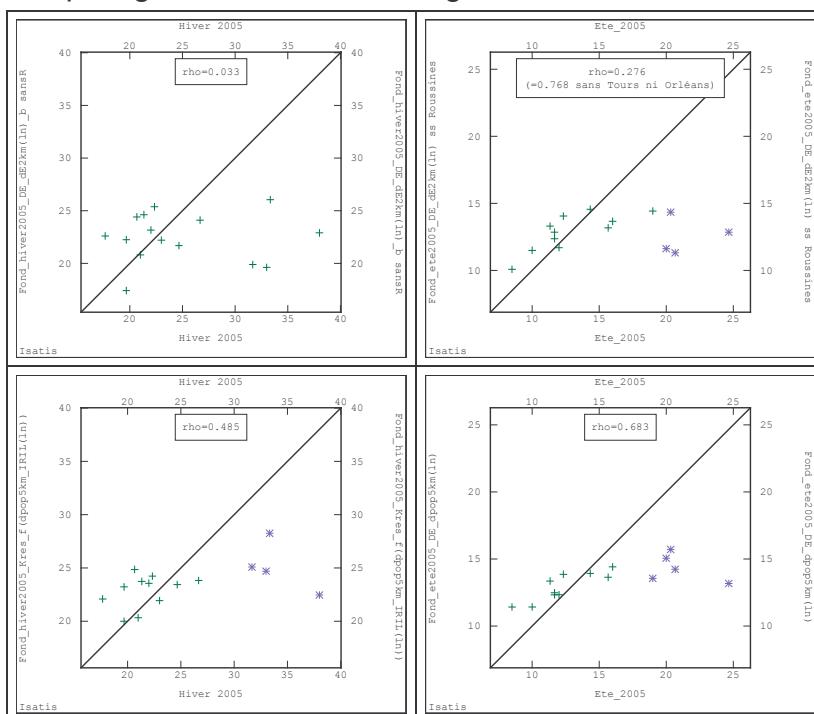


Figure 34 – Nuages de corrélation entre les mesures des stations fixes (en abscisse) et les concentrations estimées par krigeage à l'aide des émissions (à gauche), et de la population (à droite). Ligne du haut : hiver, ligne du bas : été.

En bleu : stations d'Orléans La Source, Orléans Préfecture, Tours Bruyère, Tours Jardin Botanique, Blois Centre (pour l'été).

Tableau 10 - Statistiques de l'erreur relative quand la densité d'émissions à 2km est utilisée.

	Hiver : de janvier à mars 2005 Noir : par rapport aux 3 mois d'hiver	Eté : de juillet à septembre 2005 Noir : par rapport aux 3 mois d'été
Moyenne	18.0	20.3
P90	39.7	45.3
P95	40.5	47.9
Maximum	40.5	47.9
Ecart-type	12.5	14.3
Corrélation	0.03	0.28

Tableau 11 - Statistiques de l'erreur relative quand la densité de population à 5km est utilisée.

	Hiver : de janvier à mars 2005 Noir : par rapport aux 3 mois d'hiver	Eté : de juillet à septembre 2005 Noir : par rapport aux 3 mois d'été
Moyenne	12.6	18.3
P90	25.1	34.3
P95	40.9	46.6
Maximum	40.9	46.6
Ecart-type	10.3	12.4
Corrélation	0.49	0.68

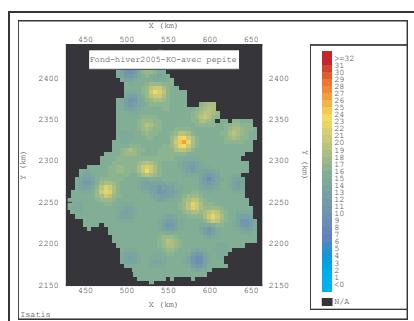
14.1.4 Estimation

Les concentrations de NO_2 sont estimées sur une grille d'estimation, selon une maille de 5km de côté, sans (krigeage ordinaire) et avec variable auxiliaire. En krigeage ordinaire, les valeurs estimées, relativement uniformes dans le domaine, augmentent ou diminuent localement autour des sites de mesure, créant des tâches de couleur (Figure 35). Ces dernières représentent les niveaux de fond urbains et leur étalement autour des villes, dans un certain rayon d'influence qui est fonction du variogramme (~ 15 km). En revanche, la cartographie ne dit rien des variations de concentration entre les communes. Hors des points d'échantillonnage, l'écart-type de l'erreur de krigeage est relativement élevé.

En krigeage du résidu ou krigeage avec dérive externe, l'hypothèse d'une relation linéaire entre la concentration de NO_2 et la variable en dérive donne du poids à celle-ci, notamment dans les zones où la densité de mesures du NO_2 est plus faible. Avec la densité d'émissions, le tracé des routes ressort nettement ; avec la densité de population, il apparaît de manière plus diffuse (Figure 36). Evaluées aux 49 sites d'échantillonnage, ces deux variables auxiliaires sont pourtant très corrélées entre elles mais dans les zones rurales dépourvues de données de NO_2 , là où leur influence est précisément accrue, cette corrélation se détériore (Figure 38). L'écart-type de l'erreur d'estimation est plus faible qu'en krigeage ordinaire.

L'estimation par krigeage de blocs ne se distingue pas sensiblement de l'estimation par krigeage ponctuel. L'écart-type de l'erreur de krigeage est logiquement plus faible puisque la concentration n'est pas estimée en un point mais en moyenne sur une maille (Figure 37).

Les cartographies estivales et annuelles ont même allure que les cartographies hivernales (Figure 39, Figure 40). Pour une variable auxiliaire donnée, la moyenne des cartes saisonnières ressemble fortement à la cartographie réalisée à partir de 12 mois complets de mesure (Figure 41).



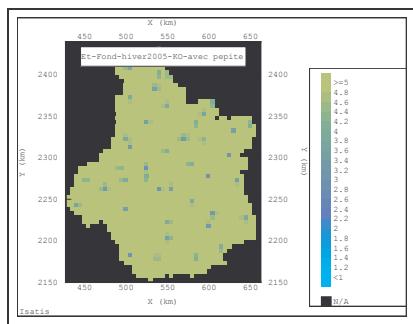


Figure 35 – Concentration hivernale de NO_2 : estimation par krigage ordinaire et écart-type de l'erreur de krigage [$\mu\text{g}/\text{m}^3$].

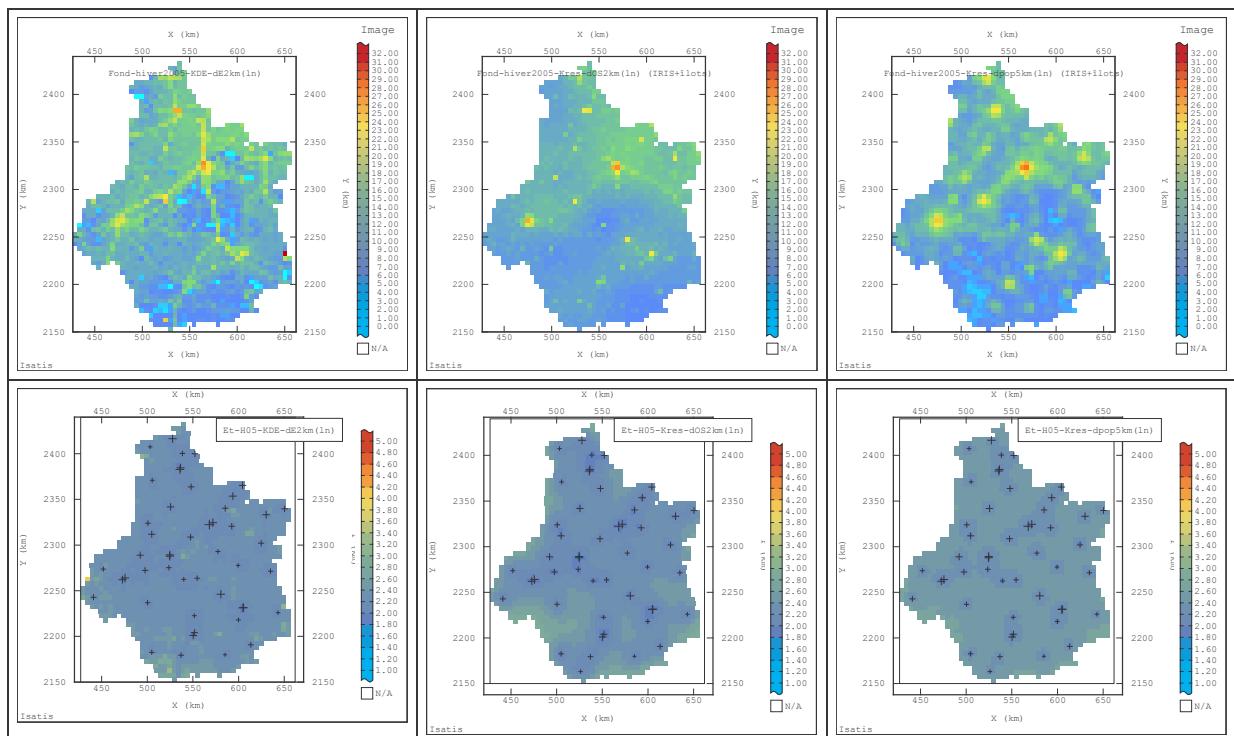


Figure 36 – Concentration hivernale de NO_2 : estimation avec variables auxiliaires [$\mu\text{g}/\text{m}^3$]. De gauche à droite : densité d'émissions dans un rayon de 2 km (krigeage avec dérive externe), densité de territoires « artificialisés » dans un rayon de 2 km (krigeage du résidu), densité de population dans un rayon de 5 km (krigeage du résidu).

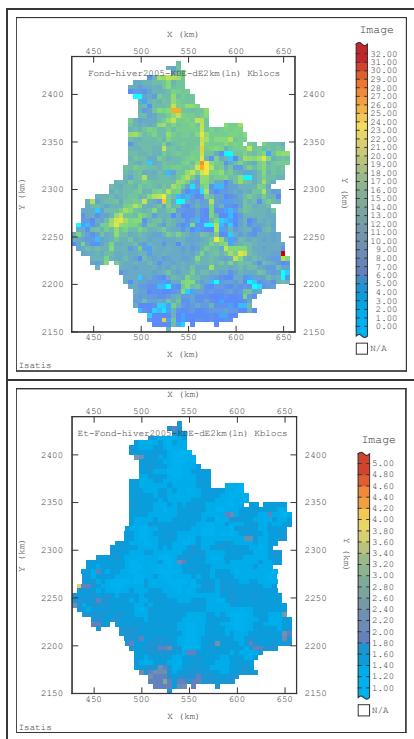


Figure 37 – Concentration hivernale de NO_2 : estimation par krigeage de blocs avec la densité d'émissions en dérive externe [$\mu\text{g}/\text{m}^3$]. Ligne du bas : écart-type de l'erreur de krigeage.

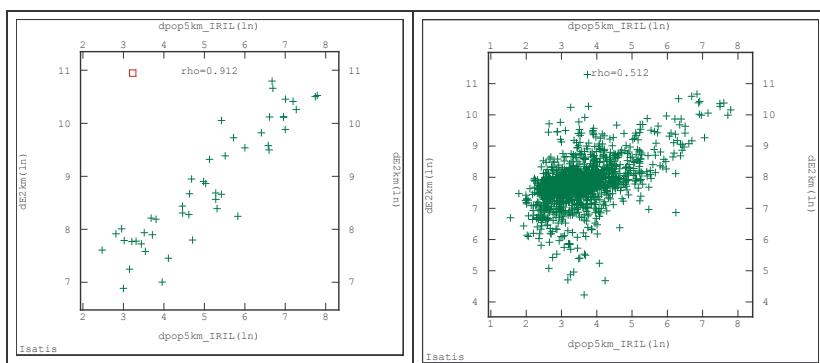


Figure 38 - Corrélation entre densité d'émissions et densité de population aux 49 sites d'échantillonnage (figure de gauche) sur la grille d'estimation.

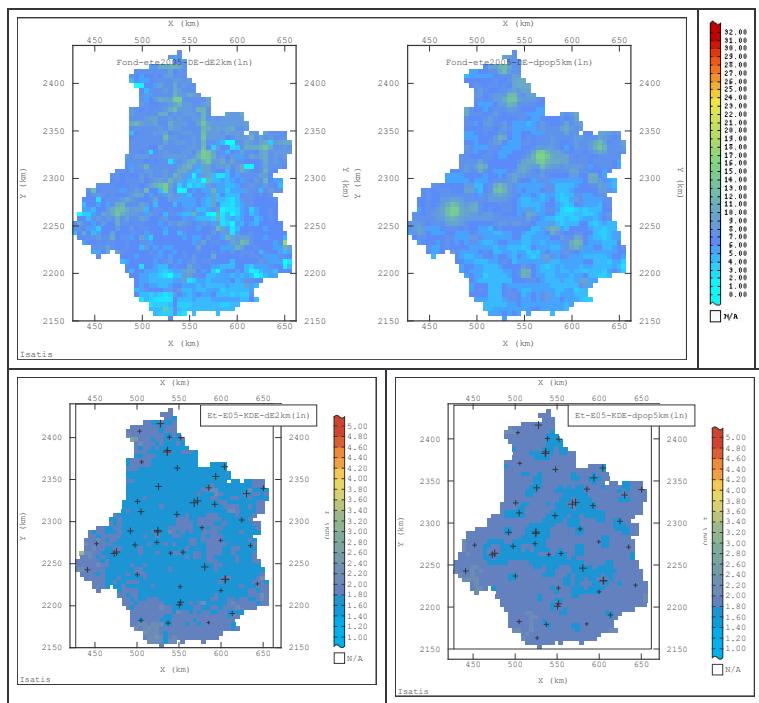


Figure 39 – Concentration estivale de NO_2 : estimation avec variables auxiliaires [$\mu\text{g}/\text{m}^3$]. De gauche à droite : densité d'émissions dans un rayon de 2 km (krigeage avec dérive externe) et densité de population dans un rayon de 5 km (krigeage avec dérive externe).

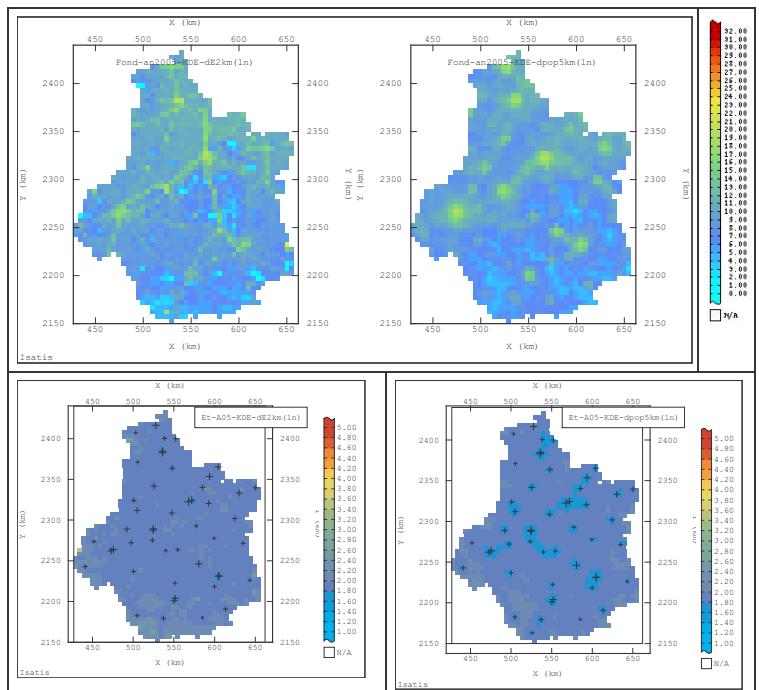


Figure 40 – Concentration annuelle de NO_2 : estimation avec variables auxiliaires [$\mu\text{g}/\text{m}^3$]. De gauche à droite : densité d'émissions dans un rayon de 2 km (krigeage avec dérive externe) et densité de population dans un rayon de 5 km (krigeage avec dérive externe).

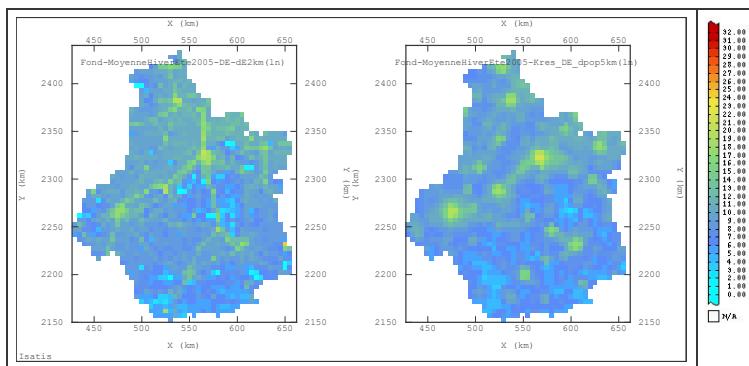


Figure 41 – Concentration annuelle de NO_2 : estimation en faisant la moyenne des estimations hivernale et estivale [$\mu\text{g}/\text{m}^3$].

14.2 PRISE EN COMPTE DE LA POLLUTION DE PROXIMITE

14.2.1 Comparaison avec les mesures de proximité

Les concentrations **de fond** hivernales de NO_2 sont estimées aux 19 points d'échantillonnage situés près des routes. Les campagnes de fond et de proximité ne s'étant pas déroulées pendant la même année, on peut seulement comparer les valeurs de fond estimées pour l'hiver 2005 avec les mesures de proximité de l'hiver 2004. Ce décalage temporel constitue une limite dans l'application de la démarche. Néanmoins, d'après les données des deux stations fixes de proximité implantées dans la région, la comparaison des hivers 2005 et 2004 reste une approximation acceptable.

Si les concentrations de fond apparaissent corrélées aux mesures de proximité, elles les sous-estiment nettement (Figure 42). Près des routes, une contribution locale du trafic routier s'ajoute donc au niveau de fond.

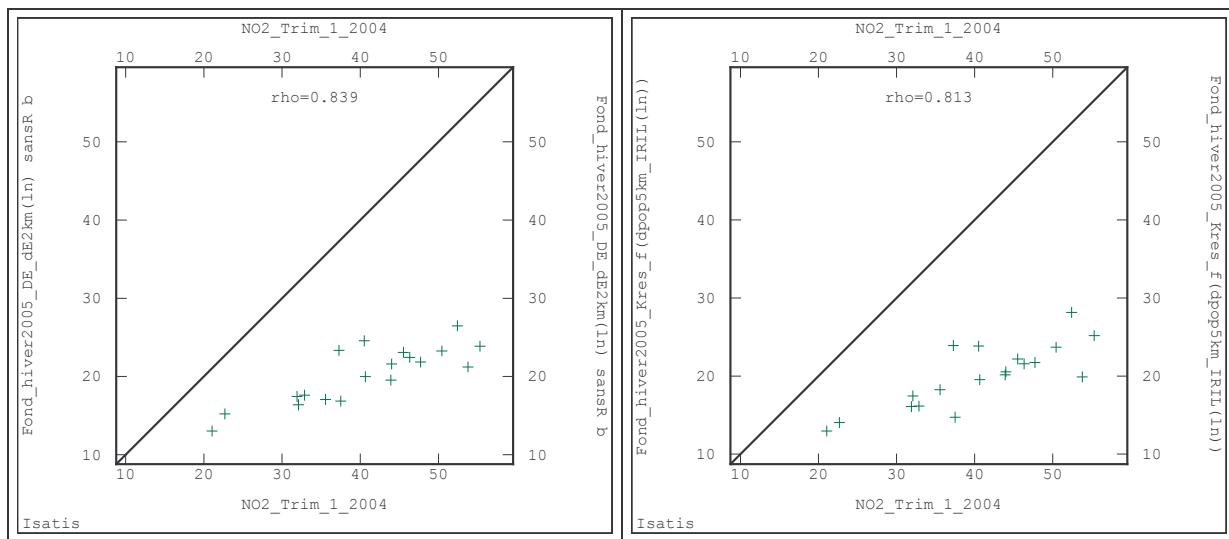


Figure 42 – Nuages de corrélation entre les concentrations mesurées aux 19 sites de proximité de la région (abscisses) et les concentrations de fond estimées par krigeage (ordonnées) avec soit la densité d'émissions (fig. de gauche) soit la densité de population (fig. de droite) comme variable auxiliaire.

14.2.2 Modélisation de l'augmentation de concentration en zone de proximité

Plusieurs variables auxiliaires (densités d'émissions, comptage routier) ont été examinées afin d'expliquer l'augmentation de concentration due au trafic (*calculée comme la différence entre mesure de proximité et estimation du fond*). La densité d'émissions de NOx dans un rayon de 1 km se révèle la plus corrélée à cette augmentation (Figure 43).

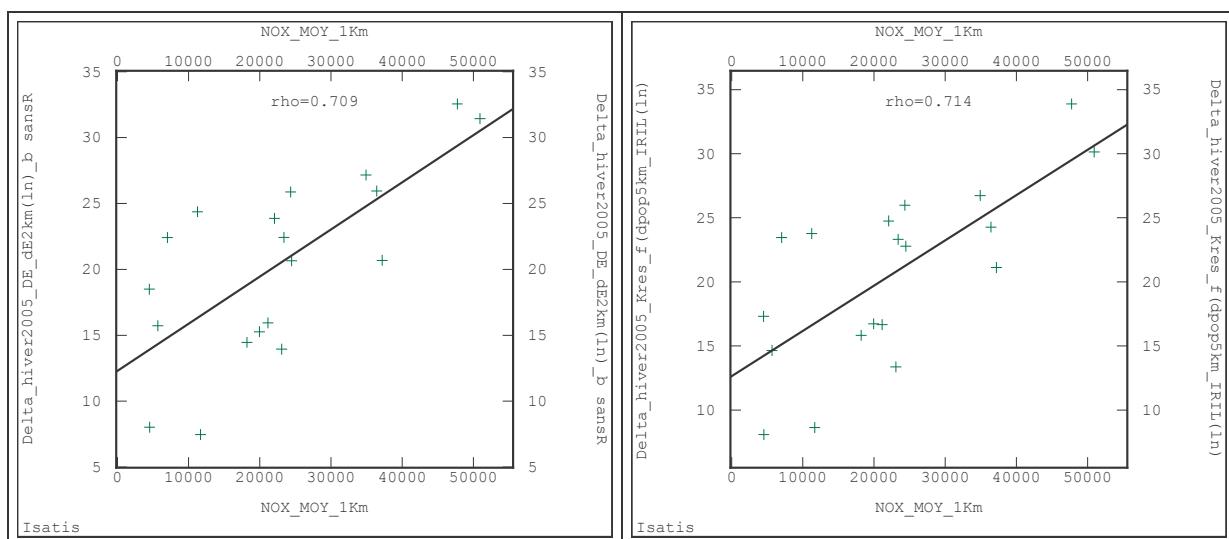


Figure 43 - Nuages de corrélation entre la densité d'émissions de NOx dans un rayon de 1 km et l'augmentation de concentration aux 19 sites de proximité. Cette augmentation est calculée par rapport à un niveau de fond estimé à l'aide de la densité d'émissions à 2 km (figure de gauche) ou de la densité de population à 5 km (figure de droite).

Ainsi, dans les zones de proximité la concentration de NO_2 peut être estimée par une relation du type : $\text{NO2}_{\text{proximité}} = \text{NO2}_{\text{fond}} + a \cdot dE_{\text{NOx}_1\text{km}} + b$.

Avec des informations supplémentaires (comme le type de route et la distance par rapport à la route), ce modèle pourrait être affiné.

De la même façon, on compare :

- les concentrations de fond de l'été 2005 estimées aux 19 points de proximité avec les mesures de proximité de l'été 2003 ;
- les concentrations de fond de l'année 2005 estimées aux 19 points de proximité avec les mesures de proximité de l'année avril. 2003-mars 2004.

Comme pour l'hiver, on constate une nette sous-estimation qui peut en partie s'expliquer grâce de la densité d'émissions à 1km (Figure 44).

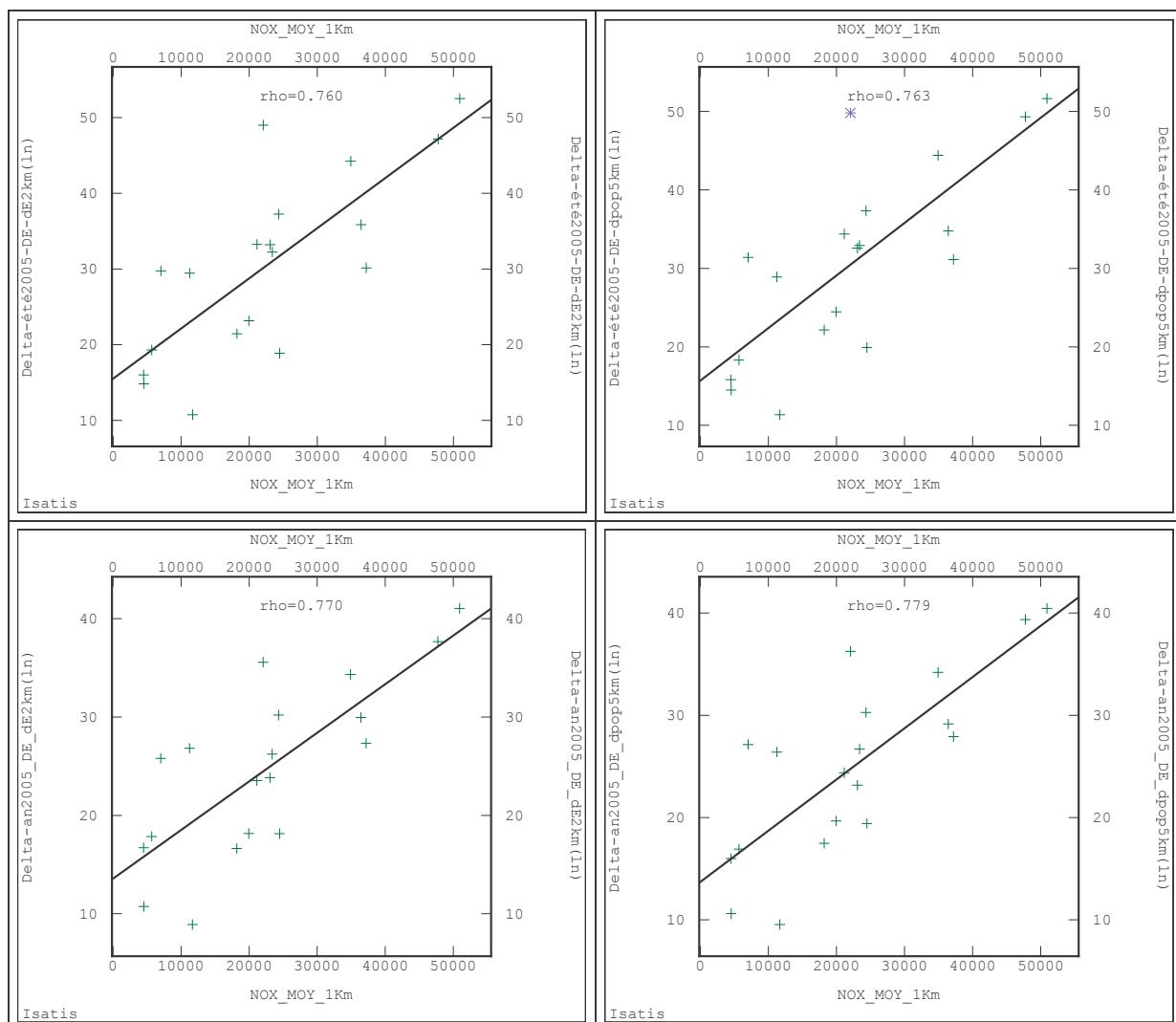
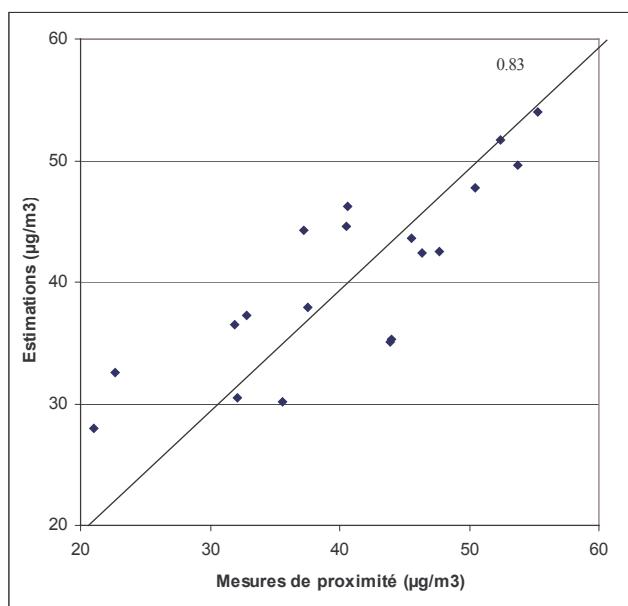


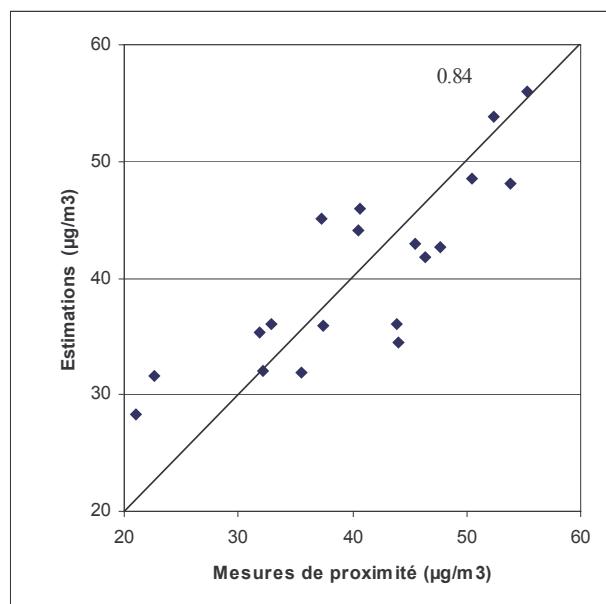
Figure 44 - Nuages de corrélation entre la densité d'émissions de NO_x dans un rayon de 1 km et l'augmentation de concentration aux 19 sites de proximité, en été (ligne du haut) et sur l'année (ligne du bas). Cette augmentation est calculée par rapport à un niveau de fond estimé à l'aide de la densité d'émissions à 2 km (figure de gauche) ou de la densité de population à 5 km (figure de droite).

14.2.3 Incertitude

Le modèle de correction proposé précédemment ne peut être pas évalué sur la région en dehors des 19 sites, faute de données d'échantillonnage complémentaires. Afin d'approcher l'incertitude d'estimation, une validation croisée est réalisée en ces 19 points. On constate une assez bonne précision d'estimation (Figure 45).



Niveau de fond estimé à l'aide de la densité d'émissions à 2 km :



Niveau de fond estimé à l'aide de la densité de population à 5 km :

Figure 45 – Validation croisée aux 19 sites de proximité. Nuages de corrélation entre valeurs estimées et valeurs mesurées. Hiver 2005-hiver2004.

14.2.4 Représentation cartographique

La procédure de correction dans les zones de proximité consiste à :

- définir des bandes (par ex de ± 200 m) autour des principales routes de la région ;
- resserrer le maillage à l'intérieur de ces bandes ;
- attribuer une valeur de fond aux petites mailles, soit en affectant à chacune la valeur de la grande maille qui l'englobe (procédure la plus rapide), soit en effectuant un nouveau krigeage (procédure plus longue puisqu'elle suppose de calculer les variables auxiliaires à plus fine résolution) ;
- calculer la densité d'émissions à 1 km au centre de chaque petite maille ;
- ajouter à l'estimation du fond le terme correctif fonction des émissions.

Cette méthodologie est coûteuse en temps de calcul. Elle n'a pas été appliquée à la région entière mais à deux parties du domaine, centrées sur Blois et Orléans. Selon l'expertise de LIG'AIR, les cartes résultantes font ressortir de façon réaliste l'augmentation de concentration près des routes (Figure 46, Figure 47).

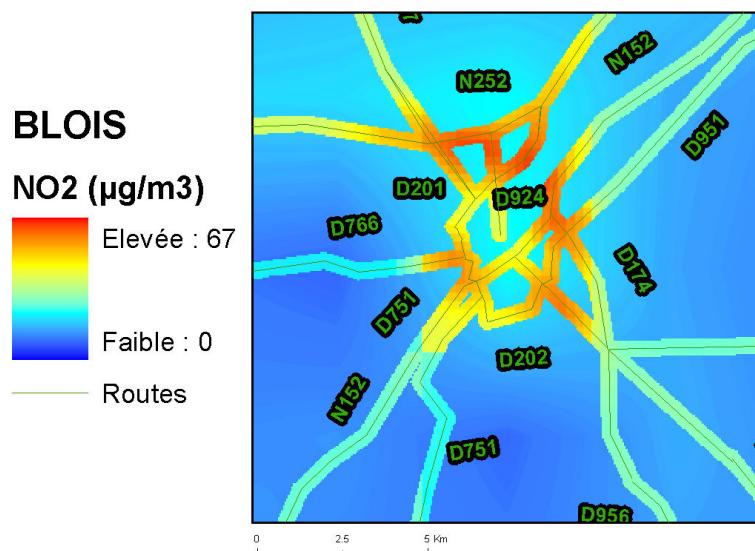


Figure 46 – Estimation de la concentration hivernale de NO2 dans l'agglomération de Blois [$\mu\text{g}/\text{m}^3$]. Autour des principaux axes, un terme correctif, fonction de la densité d'émissions à 1km, a été ajouté à l'estimation du fond.

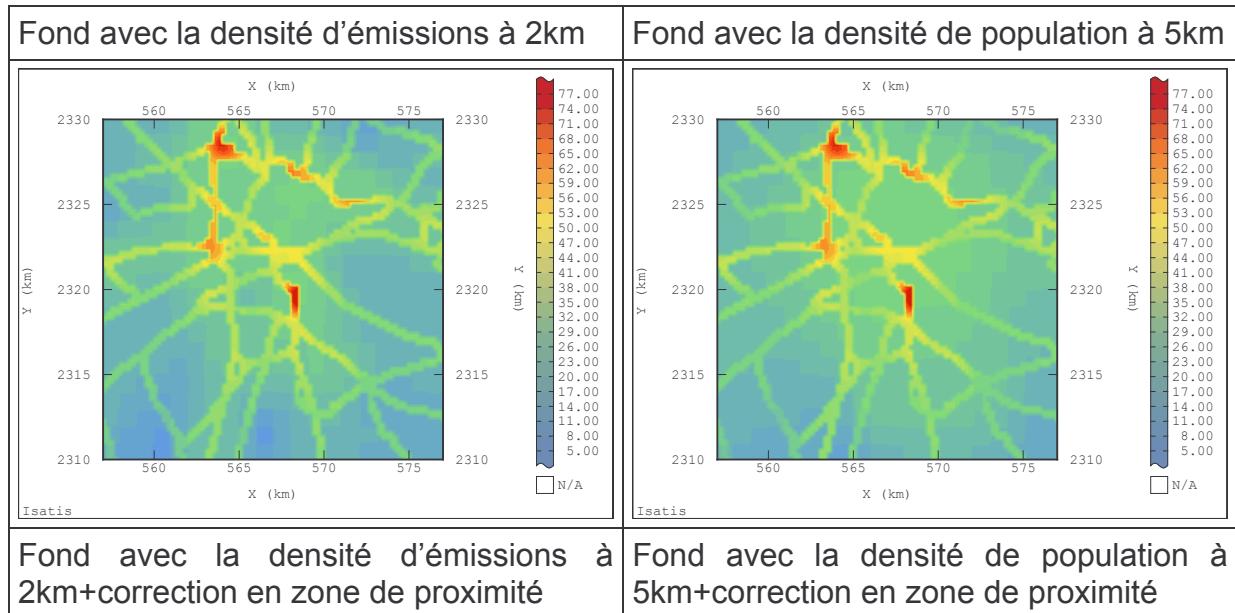


Figure 47 – Estimation de la concentration de NO₂ dans l’agglomération d’Orléans [µg/m³]. Autour des principaux axes, un terme correctif, fonction de la densité d’émissions à 1km, a été ajouté à l’estimation du fond.

En ce qui concerne la cartographie de fond, l'estimation dans le centre-ville concorde mieux avec les mesures de stations fixes lorsque la densité de population est utilisée ; en périphérie, l'accord est meilleur si l'on utilise la densité d'émissions de NO_x.

D'autre part, du 25 mars au 8 avril 2004, une campagne de mesure par échantillonnage passif a été conduite à Orléans en une quarantaine de sites de proximité. Les concentrations hivernales de NO₂ (fond + terme correctif) sont estimées en ces points et comparées aux données d'échantillonnage. La corrélation est positive (Figure 48). Une partie du nuage est centrée sur la bissectrice, tandis que le reste des points indique une surestimation d'environ 10 µg/m³. Outre la non concomitance des mesures et des estimations, des différences dans l'environnement des tubes (rue canyon, environnement ouvert, ...), facteur dont le modèle de correction ne tient actuellement pas compte, pourrait expliquer en partie ces résultats.

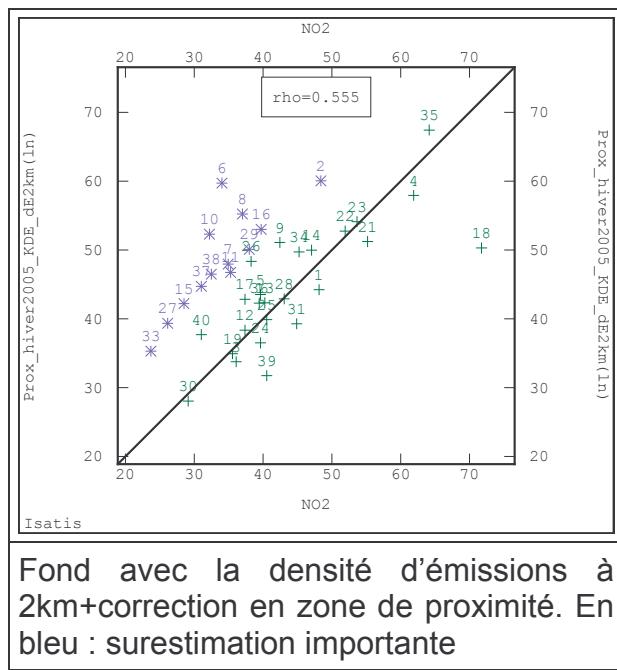


Figure 48 - Nuage de corrélation entre le fond estimé par krigage et augmenté d'un terme correctif (ordonnées) et les mesures de proximité (abscisses). Attention : les données utilisées pour l'estimation du fond et les mesures de proximité ne sont pas concomitantes.

Depuis 2006, LIG'AIR poursuit la surveillance du NO₂ à l'échelle régionale par des campagnes d'échantillonnage de proximité et de fond. Grâce aux nouvelles données collectées, il peut être envisagé d'affiner la méthodologie.