



**HAL**  
open science

## Méthode de surveillance des concentrations de NO<sub>2</sub> : cartographie automatique à partir de stations fixes et prise en compte de la proximité

Giovanni Cardenas, Laure Malherbe, Esperanza Perdrix, André Wroblewski,  
Véronique Riffault

### ► To cite this version:

Giovanni Cardenas, Laure Malherbe, Esperanza Perdrix, André Wroblewski, Véronique Riffault. Méthode de surveillance des concentrations de NO<sub>2</sub> : cartographie automatique à partir de stations fixes et prise en compte de la proximité. Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA). 2006. hal-04244603

**HAL Id: hal-04244603**

**<https://hal.science/hal-04244603>**

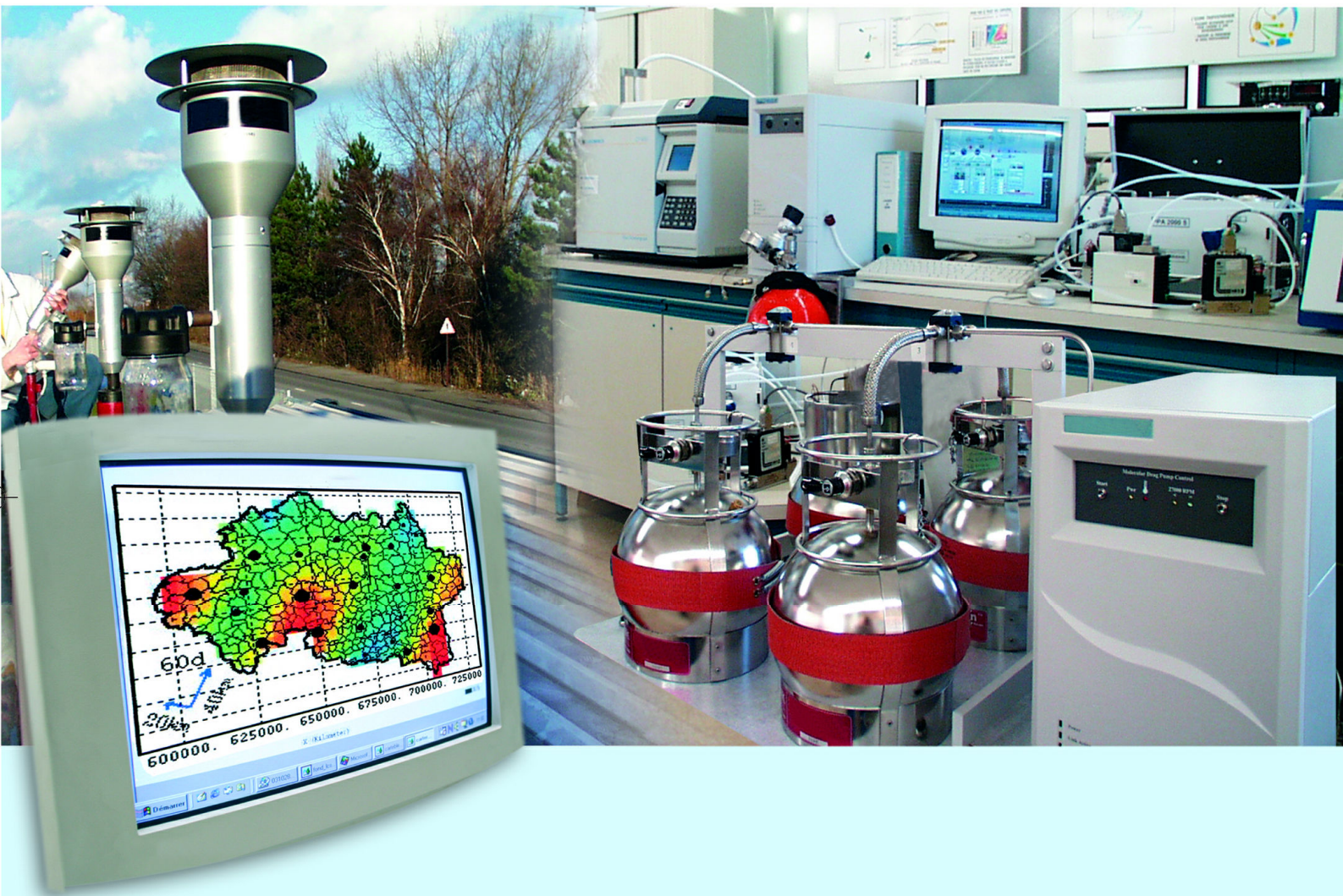
Submitted on 17 Oct 2023

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

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

# LCSQA

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



Traitements numériques

**Méthode de surveillance de concentrations de NO<sub>2</sub> :  
cartographie automatique à partir de stations fixes et prise en  
compte de la proximité**

Décembre 2006

*Version définitive*

Esperanza PERDRIX et Giovanni CARDENAS



Ministère de l'Ecologie  
et du Développement Durable

## 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.



Ministère de l'Ecologie  
et du Développement Durable

# Méthode de surveillance de concentrations de NO<sub>2</sub> : cartographie automatique à partir de stations fixes et prise en compte de la proximité

Laboratoire Central de Surveillance  
de la Qualité de l'Air

## Thème : TRAITEMENTS NUMERIQUES

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

Décembre 2006

Personne ayant participé à cette étude :  
Giovanni Cárdenas (INERIS) - Laure Malherbe (INERIS)  
Esperanza Perdrix (EMD) - André Wroblewski (EMD) - Véronique Riffault (EMD)

Ce document comporte 142 pages (hors couverture).

	Rédaction	Vérification	Approbation
NOM	G. CARDENAS	J-Y. CHATELIER L. ROUIL	M.RAMEL
Qualité	Ingénieur à l'unité 2IEN	Responsables des unités 2IEN et MECO	Responsable LCSQA/INERIS Direction des Risques Chroniques
Visa			



## TABLE DES MATIERES

<b>1. RESUME.....</b>	<b>5</b>
<b>2. INTRODUCTION .....</b>	<b>6</b>
<b>3. PRESENTATION GENERALE DE LA METHODE .....</b>	<b>8</b>
3.1 CORRELATIONS TUBES - STATIONS .....	8
3.1.1 Principe de l'étude .....	8
3.1.2 Fonctions de l'outil Excel .....	8
3.1.2.1 Données d'entrée .....	8
3.1.2.2 Corrélations .....	8
3.2 CARTOGRAPHIE PAR GEOSTATISTIQUE .....	10
3.2.1 Principales Etapes .....	10
3.2.2 Programmes de cartographie .....	13
3.2.2.1 But.....	13
3.2.2.2 Données d'entrée .....	13
3.2.2.3 Opérations Principales .....	13
3.2.2.4 Fonctionnement.....	14
3.2.2.5 Configuration requise .....	14
<b>4. VALIDATION DE LA METHODE .....</b>	<b>15</b>
4.1 AGGLOMERATION LILLOISE .....	15
4.1.1 La Méthode Du Cokrigeage Temporel.....	15
4.1.2 Comparaison entre le CKJA et le CKT Sur L'agglomeration Lilloise.....	17
4.1.2.1 Comparaison Sur Le Domaine D'etude .....	17
4.1.2.2 Comparaison Sur Les Stations Fixes .....	23
4.1.3 Analyse De L'ecart Aux Stations Fixes: « Moyenne Journalière Des Mesures » - « Estimations » .....	27
4.2 AGGLOMERATION DE REIMS .....	36
4.2.1 Les données .....	36
4.2.2 Comparaison entre le CKJA et le CKT Sur L'agglomeration Rémoise ....	40
4.2.2.1 Application du CKJA.....	40
4.2.2.1.1 Sélection des stations virtuelles et calcul de la VEM.....	40
4.2.2.1.2 Résultats des estimations journalieres .....	42
4.2.2.2 Comparaison CKT – CKJA Sur Le Domaine D'etude.....	43

4.2.2.3 Comparaison CKT – CKJA Sur Les Stations Fixes .....	47
4.2.3 Analyse De L'ecart Aux Stations Fixes: « Moyenne Journalière Des Mesures » - « Estimations » .....	49
<b>5. CONCLUSIONS SUR LA VALIDATION DE LA METHODE .....</b>	<b>54</b>
<b>6. APPLICATION DE LA METHODE D'IMBRICATION D'ECHELLE.....</b>	<b>56</b>
6.1 BREVE DESCRIPTION DE LA METHODE.....	56
6.2 ADAPTATION AU CONTEXTE DE LA CARTOGRAPHIE AUTOMATIQUE JOURNALIERE .....	57
6.2.1 Méthodologie Proposée .....	57
6.2.2 Application De La Méthodologie A L'agglomération Lilloise.....	58
6.2.2.1 Sélection de principales routes.....	58
6.2.2.2 Raffinement De La Maille Et Calcul De La Variable Auxiliaire .....	59
6.2.2.3 Analyses des resultats et conclusions .....	62
<b>7. RECOMMANDATIONS PRATIQUES POUR LA MISE EN ŒUVRE DE LA METHODE DE CARTOGRAPHIE JOURNALIERE AUTOMATIQUE DE DIOXYDE D'AZOTE .....</b>	<b>65</b>
<b>8. BIBLIOGRAPHIE.....</b>	<b>67</b>
<b>9. LISTE DES ANNEXES .....</b>	<b>68</b>

## 1. RESUME

Depuis deux ans, le LCSQA travaille sur une méthode originale de production de cartes de répartition spatiale des concentrations en NO<sub>2</sub>, à partir de mesures en stations fixes et de campagnes de mesure par échantillonnage passif.

Ces cartes doivent être réalisées à partir des données des stations fixes, tout en considérant le nombre potentiellement insuffisant de capteurs fixes pour une correcte interpolation spatiale des niveaux de pollution. La détermination de corrélations entre stations de mesure fixes et sites d'échantillonnage passif permet de générer un nombre important de « stations virtuelles ». Ces informations supplémentaires permettent de prétendre à l'utilisation d'un outil de cartographie de type géostatistique.

Ces travaux constituent le prolongement des études de faisabilité réalisées sur la région lilloise en 2004 et 2005. La méthode ainsi que l'outil informatique ont été appliqués à un autre site d'étude : la région rémoise.

Ce rapport comprend :

- Une présentation générale de la méthode.
- La validation des résultats obtenus dans deux cas d'étude : région lilloise et rémoise.
- L'application de la méthode d'imbrication d'échelle pour la prise en compte de la pollution de proximité.
- Un descriptif complet de la procédure sous EXCEL permettant d'une part, la mise en évidence des liens entre sites d'échantillonnage passif et stations fixes automatiques et d'autre part la génération automatique des cartographies des niveaux de pollution.



## 2. INTRODUCTION

Depuis 2004, le LCSQA développe une méthode originale d'estimation et de cartographie des concentrations en dioxyde d'azote, à partir d'un nombre limité de mesures effectuées en stations, de campagnes de mesures à partir desquelles sont établies des corrélations statistiques [3, 4].

Les objectifs de qualité requis par les directives européennes de qualité de l'air pour les méthodes de modélisation et d'analyse objective sont rappelés dans le Tableau 1.

Méthode	Type de moyenne estimée	Exactitude ou précision en %	Valeur limite requise en 2010, en $\mu\text{g}/\text{m}^3$
<b>Modélisation</b>	Horaire	50-60	200 : limite à ne pas dépasser plus de 18 fois par an
	Journalière	50	
	Annuelle	30	40
<b>Estimation objective</b>	non précisée	75	

*Tableau 1 : Objectifs de qualité pour l'évaluation des concentrations en dioxyde d'azote*

Il avait été montré [3, 4] qu'il existait des corrélations statistiques significatives entre les concentrations moyennes bi - hebdomadaires (moyenne sur 14 jours) en dioxyde d'azote, mesurées par une station automatique et celles, mesurées à partir de prélèvements effectués par échantillonnage passif sur la même période. L'étude d'un jeu de données, relatif à une zone présentant des caractéristiques essentiellement urbaines et périurbaines (conurbation Lille - Roubaix - Tourcoing - Villeneuve d'Ascq), a montré que plus de la moitié des sites étudiés étaient ainsi corrélés à une station de mesure automatique.

Ces travaux visaient à faciliter l'application de la méthode au sein des AASQA, notamment au travers d'une automatisation de la procédure de cartographie. Cette étude présente l'aboutissement de cette démarche d'automatisation de la cartographie.

L'étude se scinde en trois parties :

- La première se rapporte à la validation de la méthodologie sur des données provenant de la région lilloise et de l'agglomération rémoise.
- La deuxième correspond à la prise en compte de la pollution de proximité dans la cartographie journalière, il s'agit de l'adaptation de la méthodologie d'imbrication d'échelle proposée par le LCSQA en 2005 [6].

- La dernière partie correspond au guide d'application des outils informatiques développés : Le premier outil vise à l'établissement des relations liant les stations de mesures fixes aux différents sites de prélèvement à l'aide de tubes passifs. L'objectif est de mettre en œuvre une procédure automatique de détermination de ces relations entre stations et échantillonneurs passifs. Le second outil vise à faciliter la production de cartes de répartition spatiale du dioxyde d'azote, par automatisation de la procédure d'analyse géostatistique et de krigeage. Le logiciel Isatis® version 6, de Geovariances est utilisé.

### **3. PRESENTATION GENERALE DE LA METHODE**

#### **3.1 CORRELATIONS TUBES - STATIONS**

Dans les études LCSQA précédentes [3, 4] des relations mathématiques simples (régressions linéaires) ont été établies entre les concentrations bi - hebdomadaires en NO<sub>2</sub> mesurées par une station fixe automatique et celles mesurées par échantillonnage passif en d'autres points. Nous avons montré que ces mêmes relations mathématiques pouvaient servir à estimer des concentrations journalières en NO<sub>2</sub>, en divers endroits, en se basant sur la concentration journalière mesurée par la station fixe automatique.

##### **3.1.1 PRINCIPE DE L'ETUDE**

La méthodologie proposée repose sur l'exploitation des données issues de campagnes de mesures. Lorsque le nombre de campagnes de mesures par échantillonnage passif devient important (de l'ordre d'une dizaine) et que les emplacements des sites correspondants diffèrent peu d'une campagne à l'autre, il devient possible d'examiner l'existence de corrélations linéaires entre les valeurs obtenues sur un site de prélèvement et celles mesurées par un analyseur en station fixe.

Si de telles corrélations existent, on peut alors considérer les sites de prélèvement passif corrélés comme des stations fixes « virtuelles », auxquelles on peut affecter une valeur calculée, à partir de la corrélation existant avec la station automatique réelle. Cela permet de multiplier « virtuellement » le nombre de stations de mesures et donc donne accès à des méthodes d'interpolation spatiale des données (type géostatistique) pour établir des cartographies de la pollution.

##### **3.1.2 FONCTIONS DE L'OUTIL EXCEL**

L'outil informatique développé pour réaliser automatiquement la sélection des stations virtuelles a été mis au point à partir des données de l'agglomération lilloise. Nous avons testé la robustesse de l'outil au moyen des données provenant de différentes campagnes de mesures effectuées sur les agglomérations de Lille et de Reims.

###### **3.1.2.1 DONNEES D'ENTREE**

L'outil développé nécessite de disposer sur la zone étudiée et pour les mêmes périodes, d'une part d'au moins six campagnes de mesures par échantillonnage passif (trois par saison : été/hiver), et d'autre part d'au moins deux stations de mesure de typologie différente.

###### **3.1.2.2 CORRELATIONS**

A partir des différentes mesures, l'outil calcule le coefficient de détermination ( $r^2$ ) de la régression linéaire ( $y=a.x+b$ ) entre station fixe et échantillonneur passif pour

l'ensemble des couples « station/échantillonneur ». On considère la station fixe en abscisse (x) et l'échantillonneur en ordonnée (y).

*1<sup>ère</sup> étape* : L'outil trie les données en fonction de la valeur du coefficient de détermination ( $r^2$ ) et élimine les sites passifs où la valeur de  $r^2 < 0,8$ . Autrement dit, nous gardons les sites passifs pour lesquels la régression linéaire explique 80% de la variance totale des données.

*2<sup>ème</sup> étape* : L'outil recherche pour chaque couple station/échantillonneur la présence de points sur toute la gamme de concentration (cf. Figure 1). Nous définissons trois domaines de représentativité :

- [Cmin, Cmin+(plage)/3] avec plage = Cmax-Cmin
- [ Cmin+(plage)/3, Cmin+2/3(plage)]
- [ Cmin+2\*(plage)/3, Cmax]

avec Cmax et Cmin respectivement la concentration maximale et minimale de la station fixe.

Si l'un des trois domaines ne comporte pas de point, le couple (station/site) n'est pas retenu. L'objectif est de ne retenir que les corrélations représentatives de toute la gamme de concentrations.

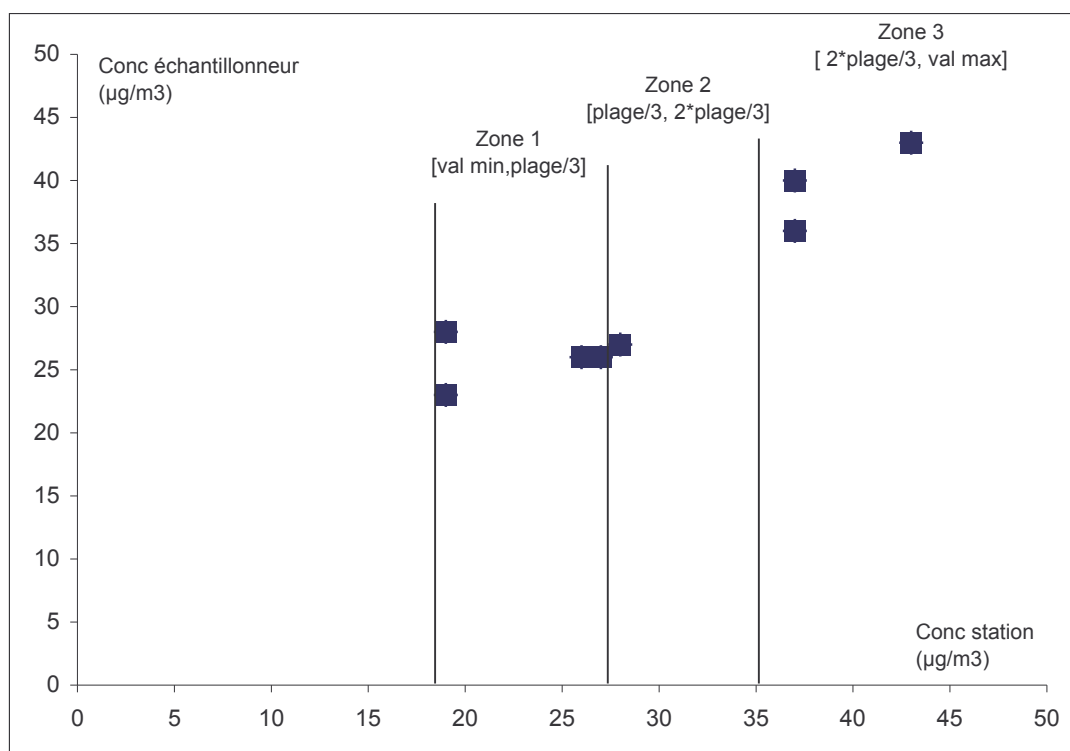


Figure 1 : Contrôle de la répartition des couples station/échantillonneur sur toute la gamme de concentration

3<sup>ème</sup> étape : Parmi les couples retenus, le logiciel sélectionne pour chaque site d'échantillonnage la station fixe la mieux corrélée. A partir du couple ainsi retenu, il va être possible de calculer les coefficients de la régression linéaire liant la station fixe et l'échantillonneur. Cet échantillonneur devient donc une station virtuelle.

A l'issue de cette étape nous disposons d'un groupe de stations virtuelles, qui sont chacune représentatives d'une corrélation station fixe - échantillonneur.

4<sup>ème</sup> étape : Calcul des coefficients de la régression linéaire ( $y=a.x+b$ ) entre station et échantillonneur passif pour l'ensemble des couples « station/échantillonneur » sélectionnés. On calcule également, pour chaque site d'échantillonnage passif, la variance de l'erreur d'estimation.

5<sup>ème</sup> étape : A partir des régressions déterminées précédemment il est possible, pour une période de mesure donnée en station fixe, de déterminer les valeurs des « stations virtuelles » associées à la même période. Ce sont ces valeurs qui permettront d'alimenter le programme de cartographie automatique journalière (cf. 3.2.2.2, page 13).

## 3.2 CARTOGRAPHIE PAR GEOSTATISTIQUE

### 3.2.1 PRINCIPALES ETAPES

La procédure expliquée précédemment permet d'estimer la concentration de dioxyde d'azote aux stations virtuelles, ce jeu des données est ensuite utilisé pour effectuer une interpolation géostatistique de concentrations de façon à obtenir une cartographie des niveaux de pollution.

Trois outils géostatistiques sont testés pour cartographier la concentration journalière de dioxyde d'azote, le premier consiste à prendre en compte la variance de l'erreur d'estimation des concentrations sur chaque station virtuelle, le second vise l'exploitation de l'information auxiliaire apportée par des variables comme le relief, la densité de population ou le cadastre d'émissions. Finalement la méthode d'interpolation sélectionnée est le « cokrigeage colocalisé ».

Le choix des paramètres d'estimation définis à l'avance a été testé dans l'étude LCSQA 2005 [4]: modèle de variogramme, variable auxiliaire et méthode d'estimation. Il s'est avéré être un bon compromis entre facilité d'application de la méthode et qualité des cartographies obtenues. Dans le chapitre 4 de la présente étude, l'analyse des résultats obtenus dans deux cas d'étude est présentée.

Le cokrigeage colocalisé est la méthode qui a obtenu les meilleurs résultats de validation croisée dans les deux cas d'étude.

Pour définir les paramètres d'estimation à l'avance on s'appuie sur l'analyse des campagnes par tubes passifs disponibles dans le domaine d'étude, de cette façon les connaissances disponibles du phénomène sont exploitées au maximum.

Le Tableau 2 affiche les principales opérations qui doivent être effectuées sur les données de campagnes par tubes passifs disponibles afin de dégager les paramètres d'estimation qui seront utilisés dans la cartographie journalière.

#	Action	Résultat
1	Dépouillement des données : Analyses de statistiques des campagnes par tubes : calcul des histogrammes, des nuages de corrélation entre campagnes, des nuages de corrélation avec les coordonnées X et Y, des variogrammes, etc.	Connaissance approfondie du phénomène : détection des données erronées, des dérives spatiales, des anisotropies, etc.
2	Calcul des moyennes saisonnières et annuelles de la concentration de NO <sub>2</sub> .	Concentrations moyennes de NO <sub>2</sub> utilisées pour l'ajustement des modèles des variogrammes et de la fonction auxiliaire.
3	Recherche des informations auxiliaires : Etude des variables disponibles : relief, émissions, densité de population, données météo, etc. Calcul des nuages de corrélation.	Construction de la ou des fonctions auxiliaires les mieux corrélées aux concentrations.
4	Calcul des variogrammes croisés entre la concentration et la fonction auxiliaire.	Analyse de la qualité des variables auxiliaires.
5	Ajustement du ou des modèles de variogrammes.	Choix des structures de base et des portées des modèles de variogramme.
6	Test des modèles par validation croisée.	Validation des modèles et choix de la méthode d'estimation : krigeage ordinaire, cokrigeage, krigeage en dérive externe ou krigeage des résidus.
7	Estimation sur une période donnée et analyses des résultats obtenus ainsi que de la variance d'estimation.	Conclusions sur la qualité des estimations et sur l'allure générale des cartes obtenues.
8	Exportation au format SIG (extension ASC d'ArcView) des grilles contenant la ou les variables auxiliaires (cf. paragraphe 1.2.2.1 de l'Annexe 2).	Fichier représentant la ou les fonctions auxiliaires qui seront utilisées pour la cartographie journalière.
9	Exportation des modèles de variogrammes (cf. paragraphe 1.2.4.2 de l'Annexe 2).	Fichier représentant le ou les modèles qui seront utilisés pour la cartographie journalière.

*Tableau 2 : Opérations à effectuer sur les données de tubes passifs*

Finalement le Tableau 3 montre les étapes qui composent la méthode de cartographie automatique journalière de NO<sub>2</sub>. Celles-ci ont été développées et automatisées au sein d'un outil informatique présenté par la suite.

#	Action
1	Application de la macro Excel « applis_lcsqa.xls » pour la création du jeu de données de concentrations journalières (stations virtuelles et stations fixes).
2	Chargement des données sur Isatis : concentrations aux stations virtuelles, grille d'estimation contenant la ou les variables auxiliaires, et modèles d'estimation.
3	Calcul et exportation du coefficient de corrélation entre la variable auxiliaire et les concentrations de NO <sub>2</sub> sur l'ensemble des stations et sur les stations fixes seulement (cf. champs 2 et 3 du tableau de l'Annexe 16).
4	Calcul et exportation des statistiques sur l'ensemble de stations (cf. champs 4 à 8 du tableau de l'Annexe 16) et sur les stations fixes seulement (cf. champs 9 à 13 sur le tableau de l'Annexe 16) ;
5	Validation croisée : <ul style="list-style-type: none"> <li>• Calcul et exportation du coefficient de corrélation entre mesures et estimations dans les stations fixes (cf. champ 1 du tableau de l'Annexe 16).</li> <li>• Calcul et exportation des statistiques des estimations obtenues sur les stations fixes (cf. champs 14 à 18 du tableau de l'Annexe 16).</li> <li>• Calcul et exportation des résultats par station fixe : concentration estimée, écart- type d'estimation, erreur, erreur carrée, racine de l'erreur carrée, erreur réduite, erreur réduite carrée, erreur relative et erreur relative carrée. (cf. Annexe 15 et champs 8 au 16 du tableau de l'Annexe 17).</li> <li>• Calcul et exportation des statistiques (moyenne et variance) de l'erreur, de l'erreur carrée, de l'erreur relative, de l'erreur relative carrée, de l'erreur réduite, de l'erreur réduite carrée et de l'erreur absolue sur l'ensemble des stations fixes (cf. l'Annexe 15 et champs 19 au 29 du tableau de l'Annexe 16).</li> </ul>
6	Par la méthode du cokrigeage colocalisé : interpolation des concentrations sur la maille d'estimation.
7	Création des cartes estimées à l'aide d'une échelle de couleur ajustée automatiquement.
8	Exportation des cartes résultantes en format d'image TIFF.
9	Exportation des fichiers résultants en format grille d'ArcView pour une exploitation ultérieure dans un SIG.

Tableau 3 : Etapes de la cartographie journalière automatique

## 3.2.2 PROGRAMMES DE CARTOGRAPHIE

### 3.2.2.1 BUT

L'objectif principal du programme est de faciliter l'application de la méthode d'estimation « Cartographie Automatique Journalière du NO<sub>2</sub> ».

Le programme lance automatiquement Isatis (mode Batch) et effectue des estimations par cokrigage colocalisé sur les données et avec les paramètres rentrés initialement, les cartes résultantes ainsi que certaines statistiques sont stockées sous forme de figures (extension TIFF et format SIG) et de tableaux (extension CSV) dans un répertoire de sortie.

### 3.2.2.2 DONNEES D'ENTREE

Les données nécessaires au fonctionnement du programme sont :

- le fichier des concentrations journalières de NO<sub>2</sub> aux stations virtuelles aux stations fixes,
- les grilles ArcView des variables auxiliaires,
- un BACKUP d'une étude isatis contenant les modèles de variogrammes choisis pour l'estimation journalière.

Parmi les paramètres d'estimation, seul le voisinage d'estimation peut être défini par l'utilisateur.

### 3.2.2.3 OPERATIONS PRINCIPALES

Afin d'être importés par le logiciel Isatis, le programme rajoute un en-tête aux fichiers contenant les données de concentrations de NO<sub>2</sub>. Ensuite, à partir des paramètres rentrés, la macro crée les fichiers de lancement d'Isatis en mode batch (extension IJNL).

Chaque fichier IJNL ainsi créé contient les instructions nécessaires au bon déroulement des interpolations et des cartographies demandées dans le domaine d'étude.

Ces fichiers sont stockés sur des répertoires et peuvent donc être consultés après la fermeture de l'application.

Si l'utilisateur confirme le lancement du programme et qu'Isatis est disponible, l'interface exécute chaque fichier IJNL et attend que le logiciel exporte les résultats des estimations<sup>1</sup>.

---

<sup>1</sup> Si les données rentrées ne correspondent pas aux formats indiqués initialement, l'utilisateur est invité à sélectionner un autre fichier pour pouvoir continuer. Dans le cas où le logiciel ISATIS n'est pas installé sur l'ordinateur et afin de donner la possibilité de quitter l'application, le lancement des interpolations se fait grâce à une boîte de dialogue qui demande une confirmation à l'utilisateur



### 3.2.2.4 FONCTIONNEMENT

Le programme a été développé en langage Visual Basic de Microsoft, le code source ainsi que le package d'installation sont disponibles sur le site web du LCSQA.

Le nom du programme d'exécution est : « CARTO\_AUTOM.EXE », l'exécution du programme produit l'affichage d'une interface composée de plusieurs boutons et cases à cocher (cf. Figure 2).

La partie supérieure de l'interface est destinée à la saisie des données. Le bouton OK permet le lancement des actions. Dans la partie inférieure de l'interface se trouvent trois boîtes de dialogue renseignant les noms et la localisation des fichiers créés et les résultats des opérations effectuées par Isatis. Une description détaillée est fournie dans la section 1.2.

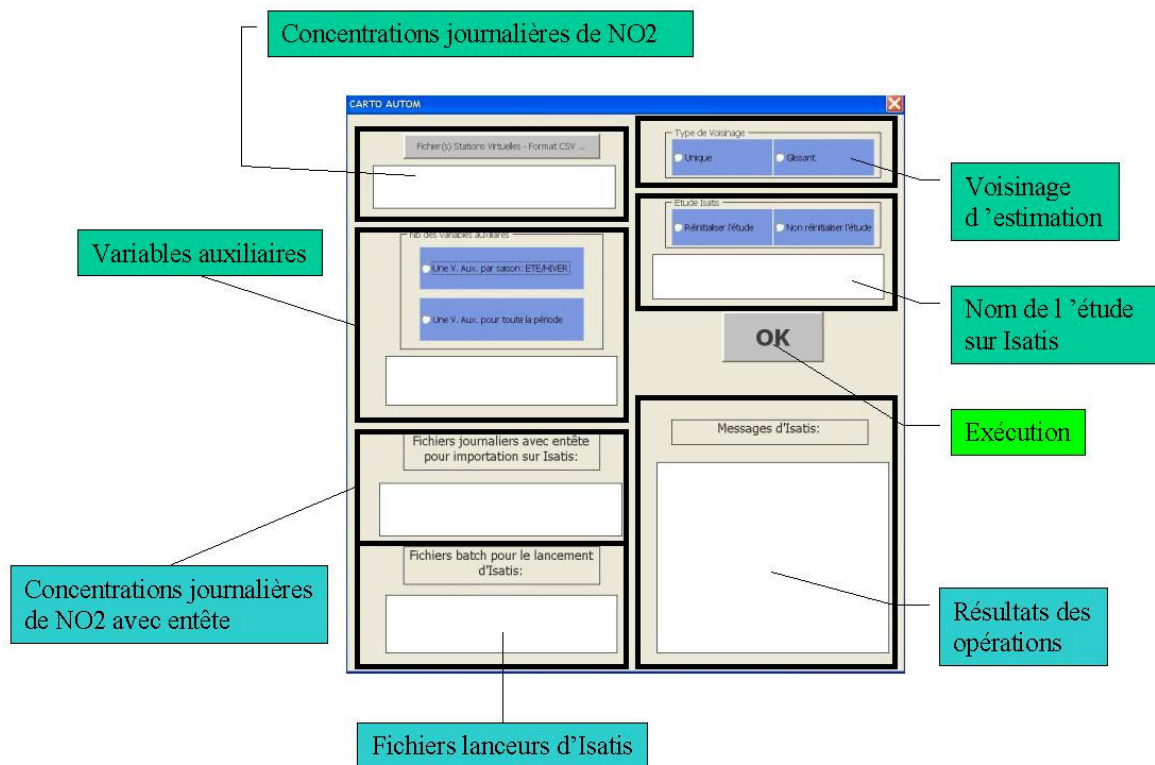


Figure 2 : Présentation de l'interface principale du programme

### 3.2.2.5 CONFIGURATION REQUISE

La configuration minimale est un ordinateur fonctionnant sous Windows (98, 2000, XP), pour effectuer les cartographies, la version 6 du logiciel Isatis sur Windows doit être installée. Cependant le programme produit les fichiers lanceurs d'Isatis (IJNL), lesquels pourront être lancés ultérieurement si le logiciel Isatis n'est pas installé.

## **4. VALIDATION DE LA METHODE**

La méthode de cartographie automatique journalière par cokrigeage colocalisé (méthode appelée CKJA par la suite) a été testée sur deux jeux de données provenant des régions lilloise et rémoise.

Deux types de validation ont été effectués, le premier a consisté à comparer les rendus cartographiques obtenus à partir de deux méthodes différentes : en l'appliquant d'une part la méthode du « cokrigeage temporel » proposée par le LCSQA en 2004 [6] (méthode appelée CKT par la suite) et en calculant d'autre part la moyenne des estimations journalières obtenues par CKJA pour la même période du temps.

La comparaison des résultats de ces deux méthodes permet de rendre compte de la différence, sur tout le domaine d'étude, entre les estimations obtenues à partir de mesures par tubes passifs et celles obtenues à partir des mesures estimées aux stations virtuelles.

Le second type de validation a consisté à analyser les statistiques par station fixe de la différence entre la « moyenne journalière des mesures » et les « estimations par CKJA ». Cette analyse permettra de comparer les résultats des estimations avec les mesures réellement effectuées pour rendre compte des différences en fonction de la typologie de la station.

Pour faciliter la compréhension, seuls les résultats de l'application des méthodes sont présentés par la suite. Une description plus détaillée des méthodologies et résultats intermédiaires est fournie dans l'Annexe 4 et l'Annexe 5.

### **4.1 AGGLOMERATION LILLOISE**

#### **4.1.1 LA METHODE DU COKRIGEAGE TEMPOREL**

Le « cokrigeage temporel » consiste à estimer la moyenne annuelle du NO<sub>2</sub> à partir des deux moyennes saisonnières (moyenne estivale et moyenne hivernale). Chaque moyenne saisonnière est estimée par cokrigeage avec la moyenne de la saison opposée.

Pour affiner chaque estimation on utilise une ou plusieurs variables auxiliaires comme fonction de dérive dans chaque cokrigeage. Ces variables auxiliaires peuvent être différentes pour chaque saison à estimer.

Une fois les deux cokrigeages des moyennes saisonnières effectués, on calcule simplement la moyenne annuelle comme la moyenne des deux cokrigeages.

Les figures 3 et 4 et le Tableau 4 montrent les résultats des estimations des moyennes saisonnières et la moyenne annuelle calculées par cokrigeage temporel (cf. Annexe 4).

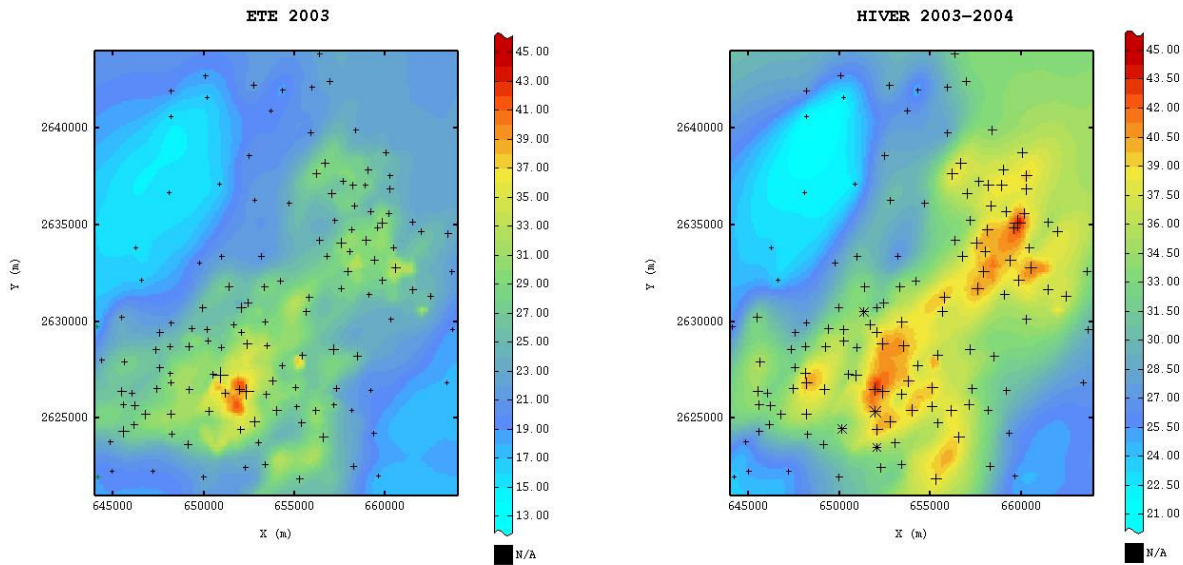


Figure 3 : Estimations des deux moyennes saisonnières par cokrigage temporel

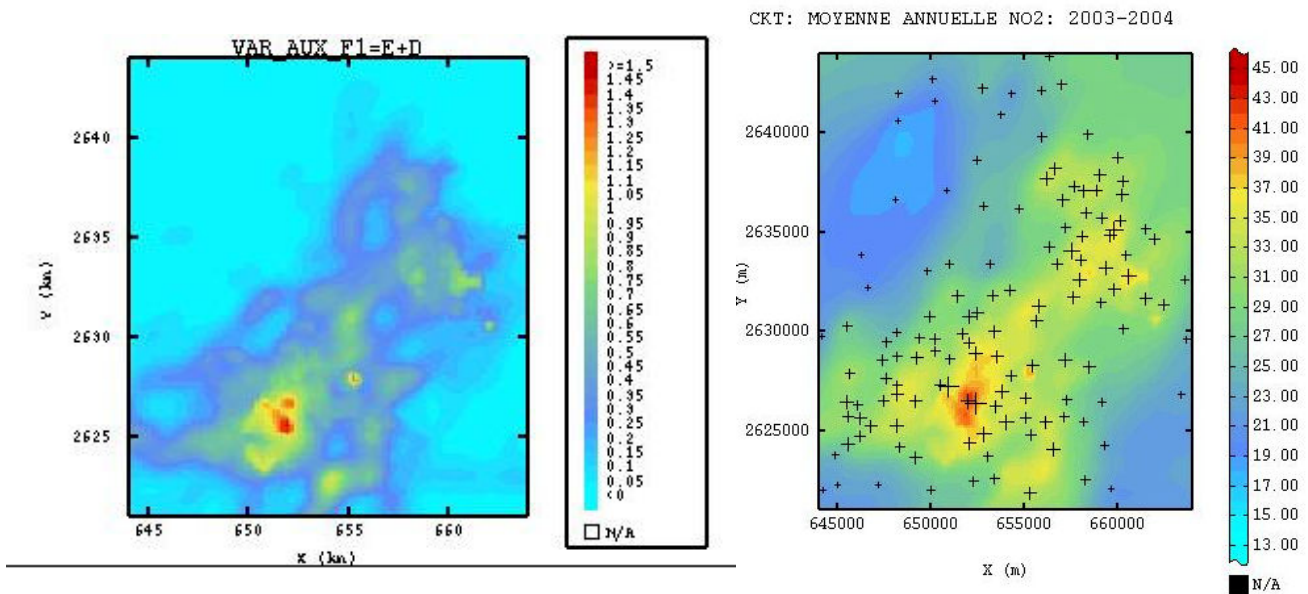


Figure 4 : Cartes de la variable auxiliaire et de la moyenne annuelle estimée de NO2 (année 2003-2004) sur la région Lilloise

Remarque : Dans toutes les cartes de cette étude la taille des croix est proportionnelle aux concentrations mesurées aux tubes passifs ou stations fixes.

CKT	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	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

$\sigma$  : écart-type

CV : coefficient de variation ( $\sigma/\text{moy}$ )

Tableau 4 : Statistiques des estimations des moyennes saisonnières et annuelles de NO2 sur la région Lilloise (année 2003-2004)

#### 4.1.2 COMPARAISON ENTRE LE CKJA ET LE CKT SUR L'AGGLOMERATION LILLOISE

Pour la cartographie journalière (CKJA) dans l'agglomération lilloise, la méthode consiste à appliquer la procédure expliquée dans la section 3.2 pour estimer les concentrations journalières de NO<sub>2</sub> entre le 1<sup>er</sup> avril 2003 et le 31 mars 2004 (voir Annexe 4).

Les paramètres définis en 2005 [4] sont conservés :

- les concentrations aux stations virtuelles sont calculées à partir de la moyenne journalière aux stations de référence Halluin - Caillou et Roubaix Château,
- la variable auxiliaire utilisée lors de l'estimation par CKT est maintenue pour l'estimation des deux saisons (somme des deux variables normées : la densité d'émissions et la densité de population, cf. carte gauche de la Figure 4),
- le fichier des concentrations journalières intègre les concentrations des stations virtuelles et les mesures des 14 stations fixes disponibles,
- dans l'Annexe 5 se trouvent les détails des modèles de variogrammes choisis,
- la méthode d'estimation est le cokrigage colocalisé avec prise en compte d'une variance de l'erreur de mesure sur les données des stations virtuelles (cf. rapport LCSQA 2005 [4]).

##### 4.1.2.1 COMPARAISON SUR LE DOMAINE D'ETUDE

Le Tableau 5, le Tableau 6 et le Tableau 7 comparent les statistiques des moyennes saisonnières obtenues par les deux méthodes : CKT et CKJA.

Les dates du début et de la fin des campagnes par tubes passifs utilisées dans le CKT sont aussi renseignées dans les tableaux. A noter que ces périodes de mesure sont censées être représentatives de la pollution de toute la saison.

La méthode du CKJA fournit en revanche l'ensemble des concentrations journalières d'une saison. Pour faciliter les comparaisons, la moyenne arithmétique des concentrations dans deux périodes du temps est considérée: d'une part l'ensemble de la saison et d'autre part la période de mesure couverte par les campagnes par tubes passifs.

Moyenne été 2003	Date Début	Date Fin	Nb jours	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
CKT	22-mai-03	3-juillet-03	43	7360	14.76	41.93	22.9	4.76	22.68	0.21
CKJA	22 mai 03	3 juillet 03	13	7360	20.22	49.53	25.88	4.48	20.1	0.17
CKJA (toute la saison)	1 <sup>er</sup> avril 03	30 sept 03	119	7360	24.59	51.65	29.1	3.92	15.4	0.14

$\sigma$  : écart-type

CV : coefficient de variation ( $\sigma/\text{moy}$ )

*Tableau 5 : Comparaison des statistiques de la moyenne estivale calculée par CKT et CKJA*

Moyenne hiver 2003-2004	Date Début	Date Fin	Nb jours	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
CKT	8-janv-04	18-févr-04	42	7360	20.94	44.78	31.42	5.24	27.49	0.17
CKJA	8-janv-04	18-févr-04	42	7360	26.6	45.17	31.81	2.97	8.82	0.09
CKJA (toute la saison)	1 <sup>er</sup> octobre 03	31 mars 04	146	7360	30.21	49.59	35.55	2.97	8.82	0.08

 $\sigma$  : écart-typeCV : coefficient de variation ( $\sigma/\text{moy}$ )

**Tableau 6 : Comparaison des statistiques de la moyenne hivernale calculée par CKT et CKJA**

Moyenne Annuelle	Date Début	Date Fin	Nb jours	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
CKT	22-mai-03	18-févr-04	85	7360	17.88	42.71	27.16	4.87	23.69	0.18
CKJA	22 mai - 3 juillet 03	8 janv.- 18 févr. 04	55	7360	23.43	47.31	28.84	3.63	13.2	0.13
CKJA (toute l'année)	1 <sup>er</sup> avril 03	31 mars 04	265	7360	27.69	50.48	32.65	3.35	11.2	0.10

 $\sigma$  : écart-typeCV : coefficient de variation ( $\sigma/\text{moy}$ )

**Tableau 7 : Comparaison des statistiques de la moyenne annuelle calculée par CKT et CKJA**

Remarque : pour la période entre le 22 mai et le 3 juillet 2003, on ne dispose que de 13 jours de mesures aux stations de référence (les concentrations des autres 30 jours sont donc manquantes pour le CKJA).

La méthode du CKJA donne en moyenne des concentrations légèrement supérieures, à celles obtenues avec la méthode du CKT (figures 3, 4 et 5).

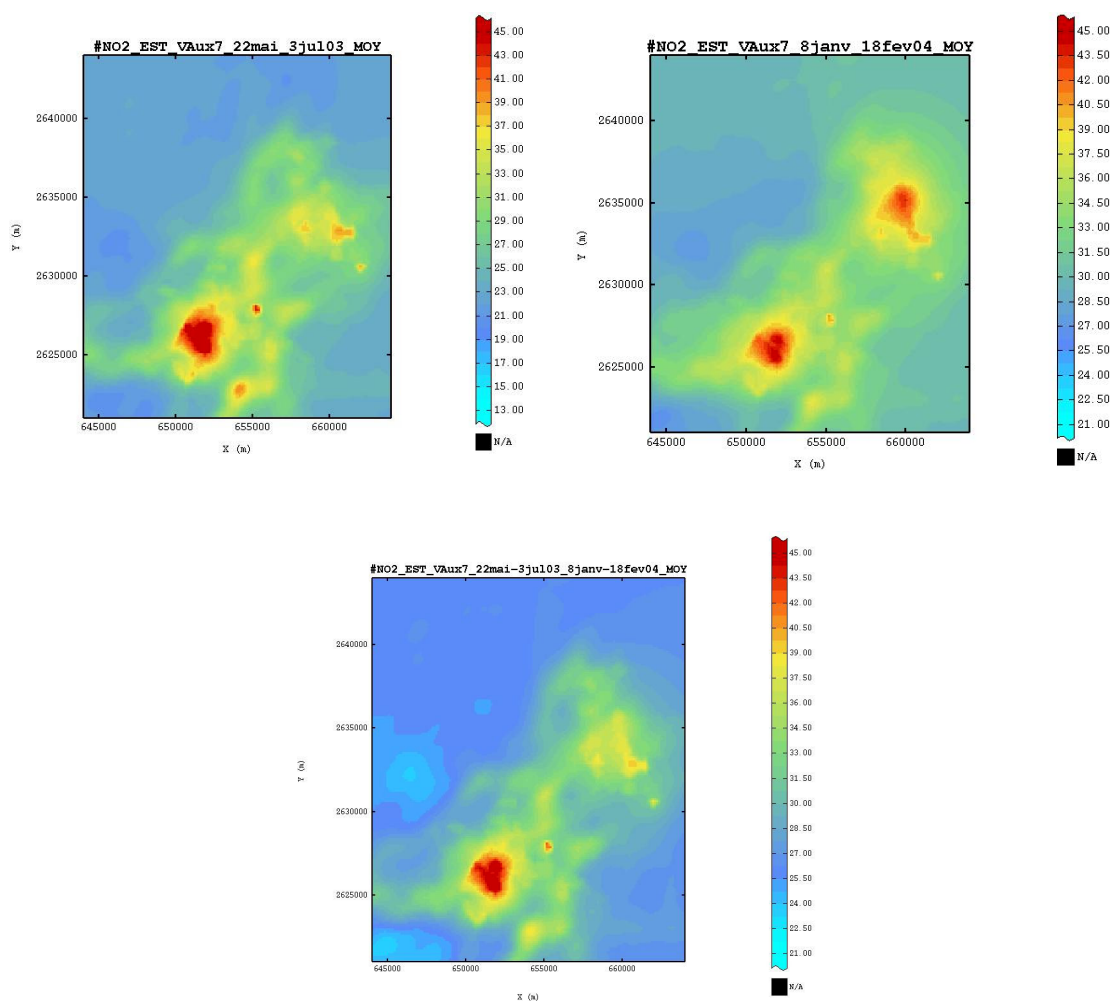
Dans le Tableau 8, les statistiques de la différence entre les deux types d'estimations sont présentées, pour la moyenne annuelle l'écart se situe à  $1.7\mu\text{g}/\text{m}^3$  en moyenne, les cartes de la Figure 6 montrent la localisation géographique des écarts.

Moyenne des écarts (CKT-CKJA)	Date Début	Date Fin	Nb jours (CKT/CKJA)	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
ETE	22 mai 03	3 juillet 03	43/13	7360	-8.2	6.2	-3.0	2.41	5.8	-0.81
HIVER	8-janv-04	18-févr-04	42/42	7360	-8.39	6.96	-0.4	3.25	10.6	-8.4
ANNEE	Eté : 22 mai - 3 juillet 03	Hiver : 8 janv.- 18 févr. 04	85/55	7360	-7.91	6.4	-1.7	2.55	6.5	-1.52

 $\sigma$  : écart-typeCV : coefficient de variation ( $\sigma/\text{moy}$ )

**Tableau 8 : Statistiques des écarts entre les estimations de NO<sub>2</sub> calculées par CKT et par CKJA (même période de temps)**

La différence entre les estimations réside d'une part dans le support de mesure des échantillons utilisés dans les deux méthodes : le CKJA prend en compte la moyenne journalière des mesures dans 14 stations fixes, celles-ci présentent une plus grande variabilité par rapport aux moyennes par quinzaine des tubes passifs prises en compte dans le CKT (le nombre de points intervenant dans le calcul des cartes est aussi différent, 132 pour le CKT et 14 stations fixes et 65 stations virtuelles dans le CKJA). D'autre part un modèle anisotrope avec structure sphérique est utilisé pour le CKT (cf. Figure 70, de l'Annexe 4) et un modèle isotrope avec structure exponentielle pour le CKJA (cf. Annexe 5).

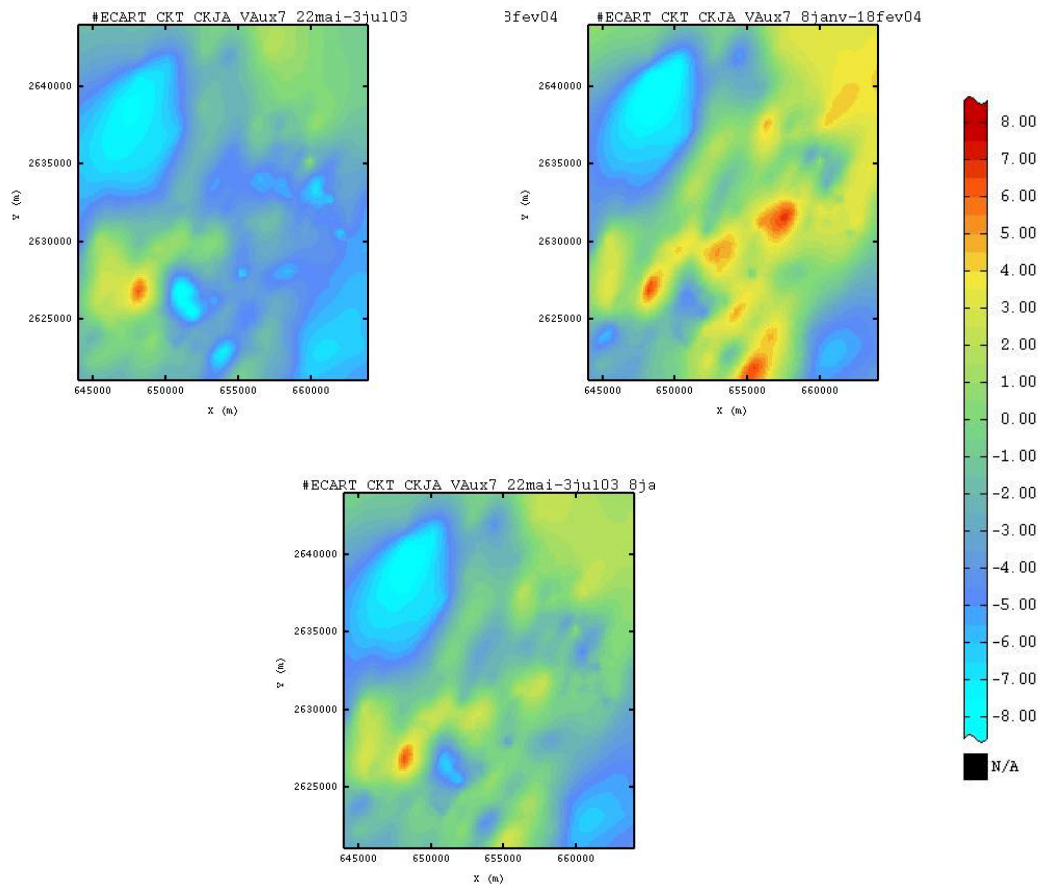


En haut à gauche : Du 23 mai au 3 juillet 2003 (période estivale 2003)

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 5 : Moyennes des estimations journalières effectuées par la méthode du cokrigage journalier automatique*



En haut à gauche : Du 22 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 22 mai au 3 juillet 2003 et du 8 janvier au 18 février 2004

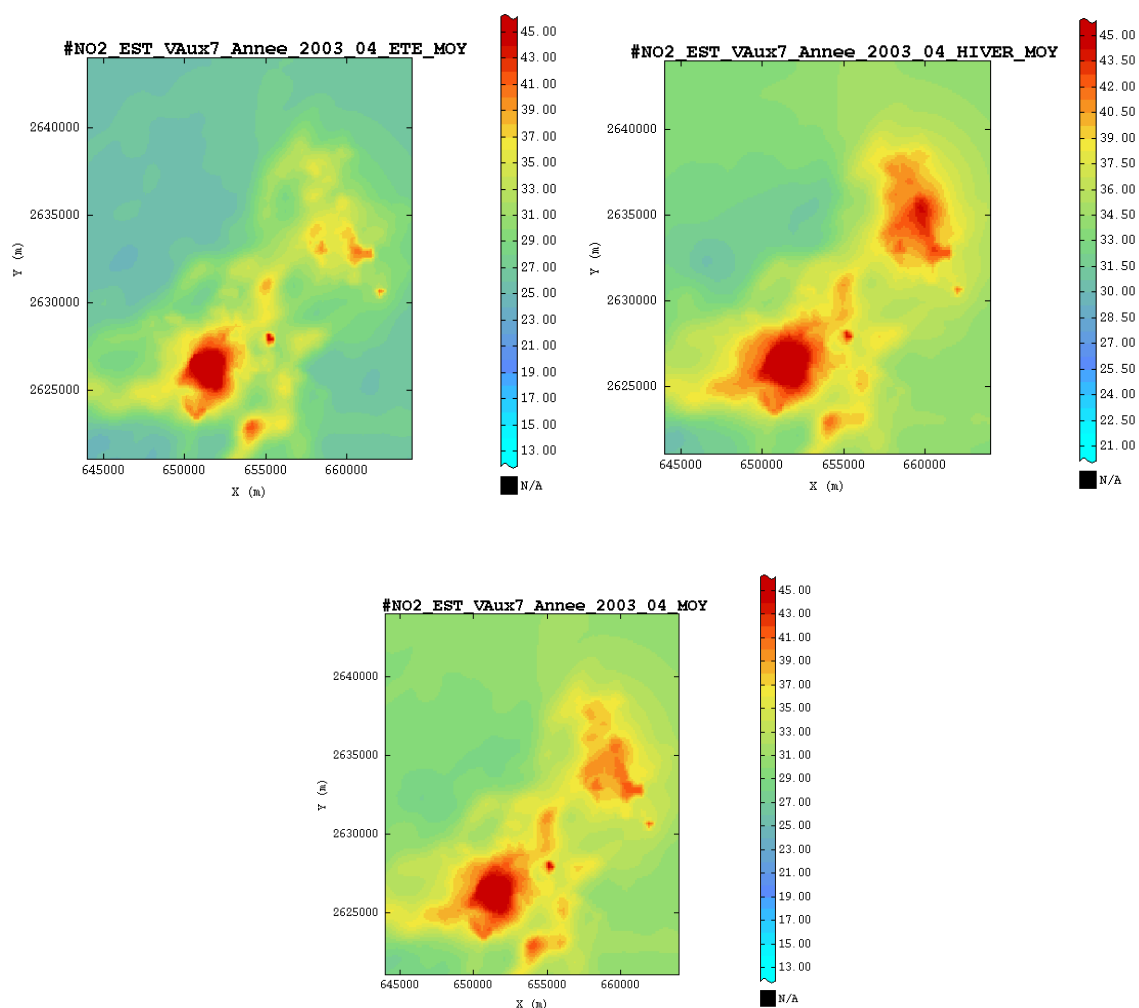
*Figure 6 : Ecart calculées entre les estimations par cokrigage temporel et par cokrigage journalier automatique (même période de temps)*

Si l'on compare l'ensemble de mesures journalières du CKJA avec le CKT (cf. Tableau 9, Figure 7 et Figure 8), les statistiques et cartes d'estimations sont en revanche très différentes. Pour la moyenne annuelle, les estimations par CKJA sont supérieures de  $5.5 \mu\text{g}/\text{m}^3$  en moyenne par rapport au CKT ; dans quelques points, les concentrations estimées par CKJA peuvent être jusqu'à  $12 \mu\text{g}/\text{m}^3$  plus élevées.

Ce résultat paraît indiquer que la prise en compte de quelques campagnes, peut conduire dans ce cas à une sous-estimation des niveaux de pollution dans le CKT (moyennes saisonnières et annuelles).

Moyenne des écarts (CKT-CKJA)	Date Début	Date Fin	Nb jours (CKT/CKJA)	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
ETE	1 <sup>er</sup> avril 03	30 sept 03	43/119	7360	-11.1	-0.2	-6.2	1.97	3.9	-0.32
HIVER	1 <sup>er</sup> oct. 03	31 mars 04	42/146	7360	-12.3	3.46	-4.1	3.32	11	-0.80
ANNEE	1 <sup>er</sup> avril 03	31 mars 04	85/265	7360	-12.0	0.33	-5.5	2.55	6.5	-0.46

Tableau 9 : Statistiques des écarts entre les estimations de NO<sub>2</sub> calculées par CKT et par l'ensemble des concentrations journalières du CKJA



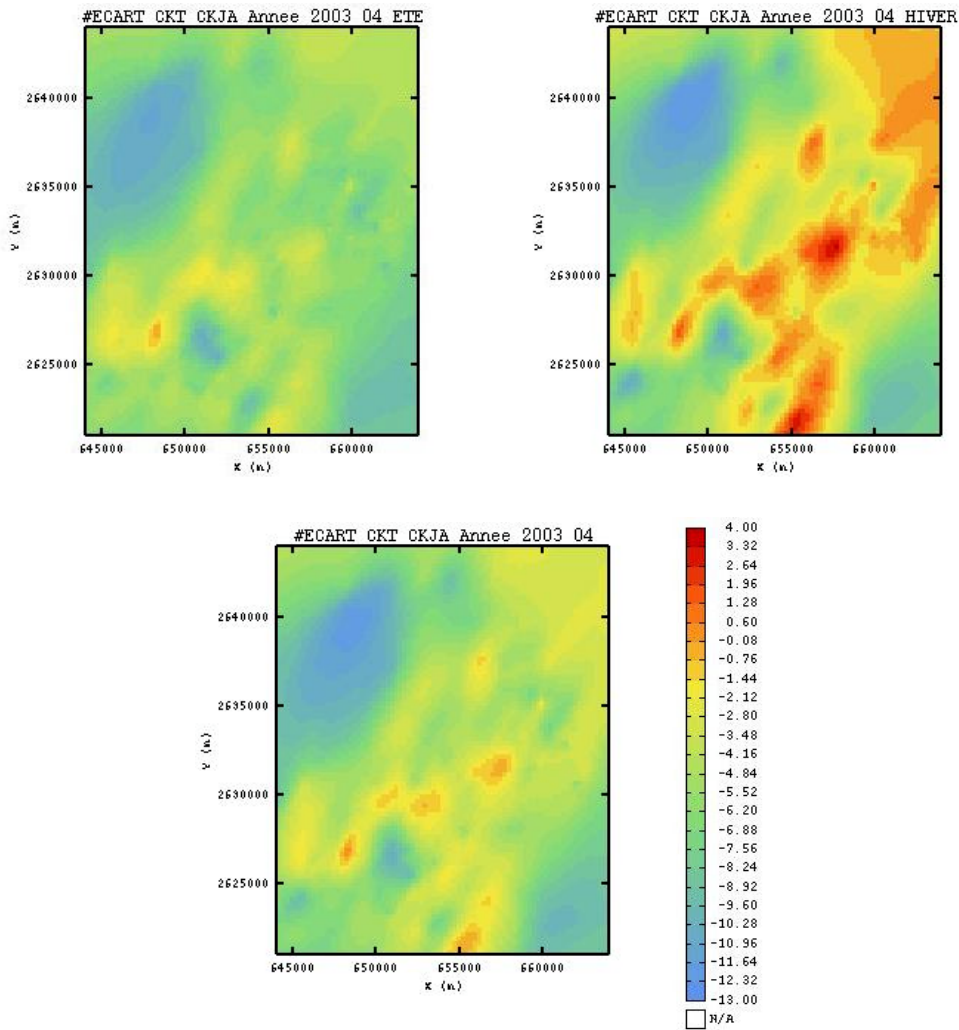
En haut à gauche : Du 1<sup>er</sup> avril au 30 septembre 2003

En haut à droite : Du 1<sup>er</sup> octobre 2003 au 8 mars 2004

En bas : Du 1<sup>er</sup> avril 2003 au 8 mars 2004

Figure 7 : Moyennes saisonnières et annuelles des estimations journalières effectuées par la méthode du cokrigeage journalier automatique





En haut à gauche : Du 1<sup>er</sup> avril au 30 septembre 2003

En haut à droite : Du 1<sup>er</sup> octobre 2003 au 8 mars 2004

En bas : Du 1<sup>er</sup> avril 2003 au 8 mars 2004

Figure 8 : Ecart calculés entre les moyennes saisonnières et annuelles estimées par cokrigage temporel et par cokrigage journalier automatique

#### 4.1.2.2 COMPARAISON SUR LES STATIONS FIXES

Pour effectuer cette comparaison, la moyenne des concentrations saisonnières et annuelles calculées par CKT et CKJA ont été affectées aux 14 stations fixes disponibles sur l'agglomération lilloise par migration à partir de la maille d'estimation. Les résultats sont présentés dans les figures 9 et 10, les périodes considérées correspondent aux mêmes périodes couvertes par les campagnes d'échantillonnage passif menées en été 2003 et hiver 2003-2004.

Les points sont plus proches de la bissectrice des nuages dans les estimations par CKJA que dans les estimations par CKT, de ce fait, l'erreur ou biais conditionnel est en moyenne plus faible pour les estimations par CKJA.

Le gain de qualité s'explique peut-être par le fait que dans le CKJA les mesures des stations fixes participent directement au calcul des estimations, tandis que dans le CKT les mesures des tubes passifs localisées aux emplacements des stations fixes sont utilisées.

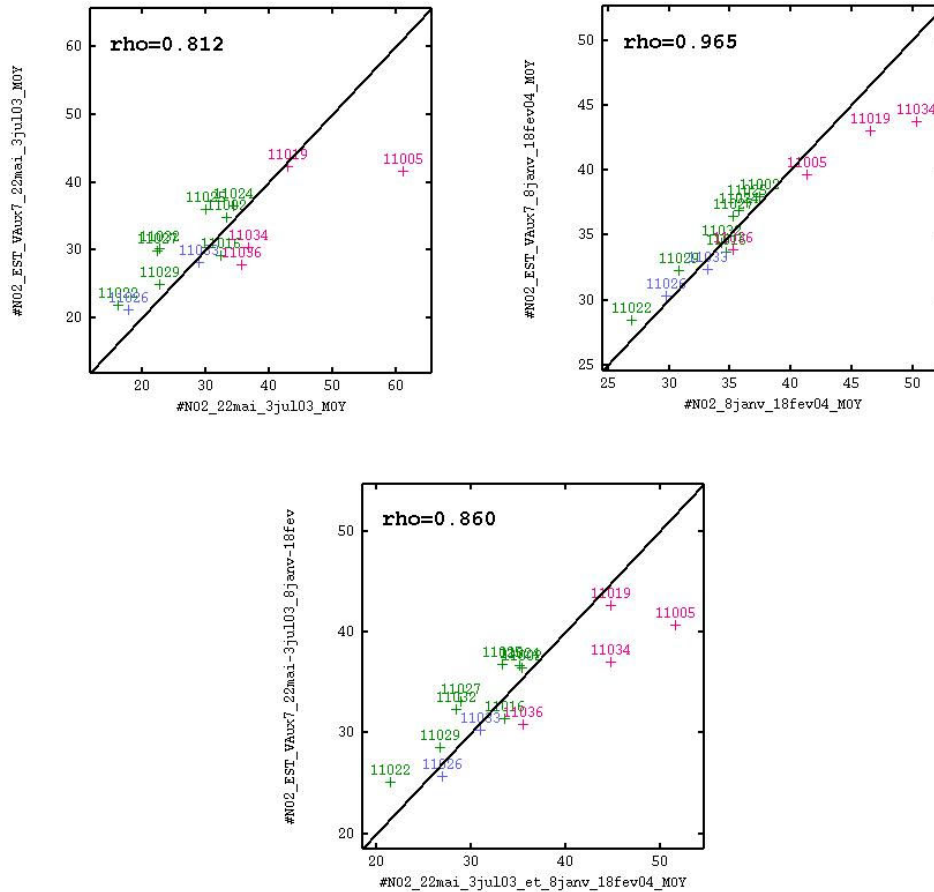


Figure 9 : Nuages de corrélation entre concentrations moyennes aux 14 stations fixes et estimations effectuées par la méthode du cokrigage journalier automatique (même période de mesure)

En haut à gauche :

Période : Du 23 mai au 3 juillet 2003 (période estivale 2003)

Axe X : Moyenne aux stations fixes de la moyenne journalière (43 jours)

Axe Y : Migration Maille 2D – points stations fixes de la moyenne des estimations par cokrigage journalier automatique (moyenne de 13 jours)

En haut à droite :

Période : Du 8 janvier au 18 février 2004 (période hivernale 2003-2004)

Axe X : Moyenne aux stations fixes de la moyenne journalière (42 jours)

Axe Y : Migration Maille 2D – points stations fixes de la moyenne des estimations par cokrigage journalier automatique (moyenne de 42 jours)

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

Axe X : Moyenne aux stations fixes de la moyenne journalière (85 jours)

Axe Y : Migration Maille 2D – points stations fixes de la moyenne des estimations par cokrigage journalier automatique (moyenne de deux périodes précédentes)

Sur toutes les figures :

Les codes AASQA des capteurs sont notés, leurs noms sont indiqués dans le Tableau 10.

Croix vertes : SF Urbaines Croix bleues : SF Périurbaines Croix rouges : SF Trafic

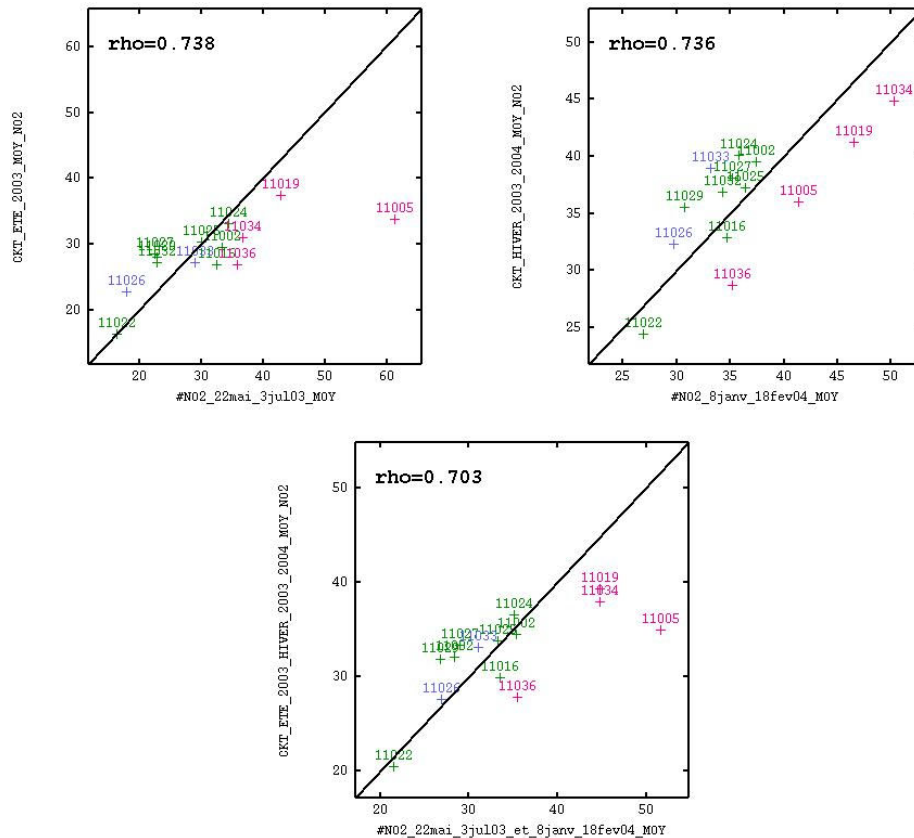


Figure 10 : Nuages de corrélation entre concentrations moyennes aux 14 stations fixes et estimations effectuées par la méthode du cokrigage temporel (même période de mesure)

En haut à gauche :

Période : Du 23 mai au 3 juillet 2003

Axe X : Moyenne aux stations fixes de la moyenne journalière (43 jours)

Axe Y : Migration Maille 2D – points stations fixes de la moyenne des estimations par cokrigage temporel (moyenne de campagnes par tubes passifs de la période estivale)

En haut à droite :

Période : Du 8 janvier au 18 février 2004

Axe X : Moyenne aux stations fixes de la moyenne journalière (42 jours)

Axe Y : Migration Maille 2D – points stations fixes de la moyenne des estimations par cokrigage temporel (moyenne de campagnes par tubes passifs de la période hivernale)

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

Axe X : Moyenne aux stations fixes de la moyenne journalière (85 jours)

Axe Y : Migration Maille 2D – points stations fixes de la moyenne des estimations par cokrigage temporel (moyenne des estimations de deux périodes précédentes)

Sur toutes les figures :

Les codes AASQA des capteurs sont notés, leurs noms sont indiqués dans le Tableau 10.

Croix vertes : SF Urbaines Croix bleues : SF Périurbaines Croix rouges : SF Trafic

### 4.1.3 ANALYSE DE L'ECART AUX STATIONS FIXES: « MOYENNE JOURNALIERE DES MESURES » - « ESTIMATIONS »

La comparaison entre les mesures et les estimations effectuées par validation croisée (CKJA) aux stations fixes est présentée par la suite.

La validation croisée consiste à estimer la concentration journalière de NO<sub>2</sub> dans chaque station fixe à partir des données des autres 13 stations fixes et des 65 stations virtuelles. Dans les estimations on prend en compte une seule fonction auxiliaire et un modèle de variogramme différent en fonction de la saison du jour à estimer (cf. paragraphe 4.1.2). La méthode de cokrigage colocalisé avec la variance de l'erreur de mesure (VEM) calculée sur les données des stations virtuelles est utilisée comme estimateur.

Dans la validation croisée la concentration de la station fixe n'est pas prise en compte dans l'estimation. Les comparaisons entre les concentrations mesurées et estimées donnent donc une indication de la qualité de la méthode d'estimation. La Figure 11 montre la localisation de 14 stations fixes disponibles dans l'agglomération lilloise. Sont aussi renseignés le code AASQA et le type de la station: Périurbaine, Trafic ou Urbaine ; les noms des stations sont indiqués dans le Tableau 10.

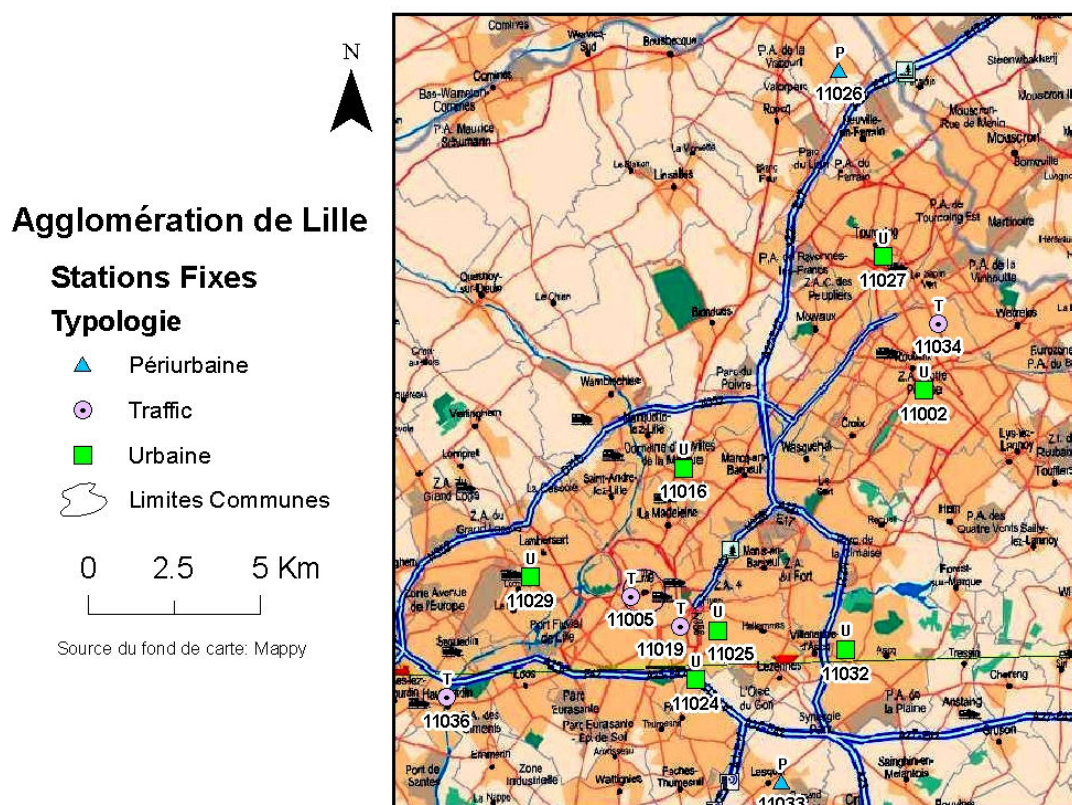


Figure 11 : Localisation et code AASQA de 14 stations fixes disponibles dans l'agglomération lilloise

La moyenne des concentrations journalières ainsi que la moyenne et l'écart - type de l'erreur et de l'erreur relative d'estimation sont affichées dans Tableau 10, le Tableau 11 et la Figure 12. Les statistiques sont calculées sur deux années différentes, du 1<sup>er</sup> avril 2003 au 31 mars 2004 et du 1<sup>er</sup> avril 2005 au 31 mars 2006, afin de considérer plusieurs épisodes de pollution.

Parmi les 729 jours composant la période de temps en étude, on dispose de 557 moyennes journalières aux deux stations fixes de référence, les concentrations des autres 172 jours ne sont donc pas disponibles pour le CKJA.

Code ST	NOM SF	Type SF	Nb jours	Min NO2	Max NO2	M NO2
11002	ROUBAIX CHATEAU	Urbaine	557	6.5	85.7	37.5
11016	MARCQ EN BAROEUL CTM		541	7.3	84.3	36.2
11022	ARMENTIERES BEAUVAIS		539	5.7	74.7	29.9
11024	LILLE FAIDHERBE		540	4.3	138.9	41.8
11025	LILLE FIVES LAKANAL		543	5.8	119.9	37.4
11027	TOURCOING BIENFAISANCE		538	4.6	91.7	34.7
11029	LOMME		548	4.2	91.1	34.2
11032	V. D'ASCQ RAMEAU		550	4.7	87.7	33.4
11026	HALLUIN CAILLOUX	Périurbaine	557	4.7	79.8	28.9
11033	LESQUIN		543	1.5	92.8	32.8
11005	LILLE LIBERTE	Trafic	524	10.2	130.8	51.3
11019	LILLE PASTEUR		548	10.3	124.5	46.8
11034	ROUBAIX SERRES		512	10.6	131.3	48.2
11036	HAUBOURDIN POMPIERS		471	8.4	90.7	39.7

M NO2 : Moyenne des moyennes journalières des mesures dans la station fixe

*Tableau 10 : Code, Nom, et Type des stations fixes et principales statistiques des moyennes des concentrations journalières*

Code ST	NOM SF	Type SF	Nb jours	Min E	Max E	M E	$\sigma$ E	Min ER	Max ER	M ER	$\sigma$ ER
11002	ROUBAIX CHATEAU	Urbaine	557	-9.5	12.3	0.75	3.1	-64.0	26.0	0.03	10.6
11016	MARCQ EN BAROEUL CTM		541	-11.4	21.2	4.23	5.0	-65.0	49.0	11.23	14.7
11022	ARMENTIERES BEAUVAIS		539	-18.8	19.5	1.93	6.1	-149.0	72.0	3.73	25.6
11024	LILLE FAIDHERBE		540	-27.1	73.0	3.49	12.4	-443.0	58.0	-8.44	51.5
11025	LILLE FIVES LAKANAL		543	-25.1	37.3	-0.96	6.5	-90.0	36.0	-5.51	20.5
11027	TOURCOING BIENFAISANCE		538	-22.7	17.5	-1.70	5.3	-201.0	29.0	-11.75	25.4
11029	LOMME		548	-19.0	27.0	1.12	5.6	-399.0	42.0	-0.93	29.1
11032	V. D'ASCQ RAMEAU		550	-18.1	21.4	-0.51	5.4	-117.0	29.0	-4.86	19.7
11026	HALLUIN CAILLOUX	Périurbaine	557	-11.6	7.5	-1.33	2.3	-171.0	17.0	-8.74	14.6
11033	LESQUIN		543	-16.1	33.0	1.84	6.7	-264.0	49.0	0.99	27.9
11005	LILLE LIBERTE	Trafic	524	-8.0	82.3	12.93	13.9	-55.0	69.0	19.91	18.9
11019	LILLE PASTEUR		548	-19.7	31.9	3.01	6.7	-96.0	41.0	5.03	15.6
11034	ROUBAIX SERRES		512	-15.5	61.2	12.17	10.8	-38.0	71.0	23.52	18.2
11036	HAUBOURDIN POMPIERS		471	-12.2	40.4	7.83	9.1	-111.0	63.0	16.26	26.0

Min E: Valeur minimale de l'erreur

Max E: Valeur Maximale de l'erreur

M E : Moyenne de l'erreur (mesure – estimation par VC)  $\mu\text{g}/\text{m}^3$

$\sigma$  E : Ecart - type de l'erreur

Min ER: Valeur minimale de l'erreur relative (%)

Max ER Valeur Maximale de l'erreur relative (%)

M ER : Moyenne de l'erreur relative (mesure – estimation par VC)\*100 / mesure

$\sigma$  ER : Ecart - type de l'erreur relative

Erreurs positives : sous - estimation, concentrations estimées plus petites que les mesures

Erreurs négatives : sur - estimation, concentrations estimées plus grandes que les mesures

**Tableau 11 : Statistiques par station des erreurs absolues et relatives d'estimation par validation croisée sur deux années de mesure (01/04/2003 - 31/03/2004 et 01/04/2005 - 31/03/2006)**

- La station Lille Pasteur (11019) est mieux estimée par rapport aux autres stations de trafic (moyenne de l'erreur plus proche de zéro).
- Les estimations des autres 3 stations de trafic ne sont pas très précises, par exemple dans la station Lille Liberté (11005) les concentrations sont sous-estimées de 20% et dans la station Roubaix Serres (11034) de 23.5%, en moyenne sur deux ans.
- La station urbaine Lille Faidherbe (11024) présente un des écart- types des erreurs les plus élevés. Ce qui signifie que l'on ne capture bien les variations.
- A l'exception de la station « Marcq En Baroeul CTM » toutes les stations urbaines et périurbaines sont surestimées ou présentent des erreurs relatives proches de zéro en moyenne sur deux ans. En revanche il est clair que les valeurs ponctuelles des maxima et minima ne sont pas très bien capturées.

La Figure 12 montre la relation entre les concentrations moyennes de NO<sub>2</sub> mesurées aux stations fixes et l'erreur relative en valeur absolue de l'estimation par validation croisée. Si la plupart des erreurs sont inférieures à 50%, on observe que les faibles concentrations peuvent présenter des erreurs relatives de plus de 100%.

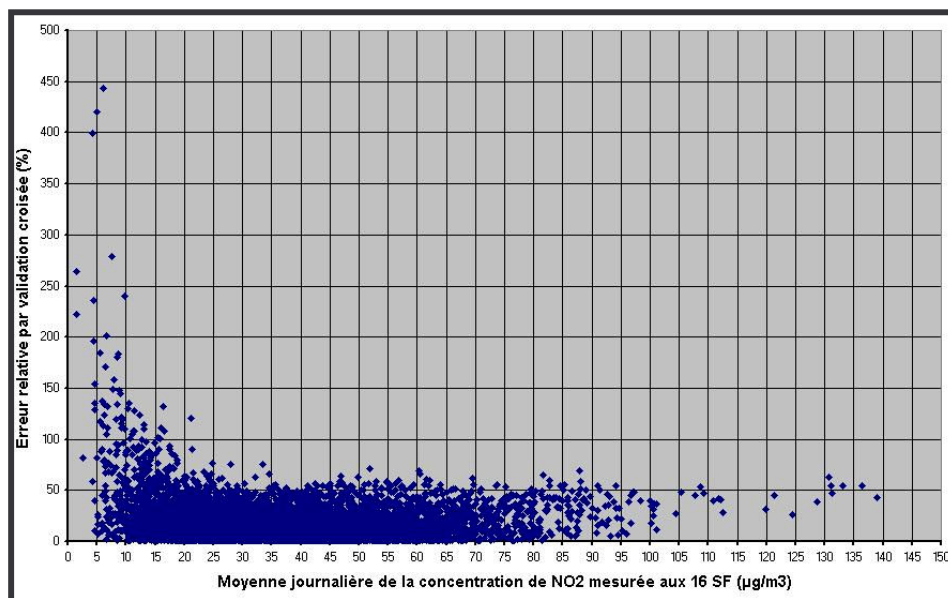


Figure 12 : Erreur relative, en valeur absolue, des concentrations de NO<sub>2</sub> pour l'ensemble des stations fixes, en fonction de la concentration mesurée

Les statistiques calculées en fonction de la typologie de la station fixe sont présentées dans le Tableau 12.

Type de SF :	Urbaine	Périurbaine	Trafic	Total
Nb de stations fixes	8	2	4	14
Nb d'estimations journalières effectuées	4356	1100	2055	7511
Minimum mesures NO <sub>2</sub>	4.2	1.5	8.4	1.5
Maximum mesures NO <sub>2</sub>	138.9	92.8	131.3	138.9
Moyenne mesures NO <sub>2</sub>	35.62	30.81	46.65	37.94
Minimum de l'erreur d'estimation par VC ( $Z - \hat{Z}$ )	-27.1	-16.1	-19.7	-27.1
Maximum de l'erreur d'estimation par VC ( $Z - \hat{Z}$ )	73	33	82.3	82.3
Moyenne de l'erreur d'estimation par VC ( $Z - \hat{Z}$ )	1.04	0.23	8.93	3.08
Ecart - type de l'erreur d'estimation par VC	6.9	5.2	11.2	8.9
Minimum de l'erreur relative d'estimation par VC (%)	-443	-264	-111	-443
Maximum de l'erreur relative d'estimation par VC (%)	72	49	71	72
Moyenne de l'erreur relative d'estimation par VC (%)	-2.05	-3.93	16.00	2.61
Ecart - type de l'erreur relative d'estimation par VC	28.0	22.7	21.1	26.8

Tableau 12 : Statistiques (par typologie des stations fixes) des résultats des estimations par validation croisée sur deux années de mesure (01/04/2003 - 31/03/2004 et 01/04/2005 - 31/03/2006)



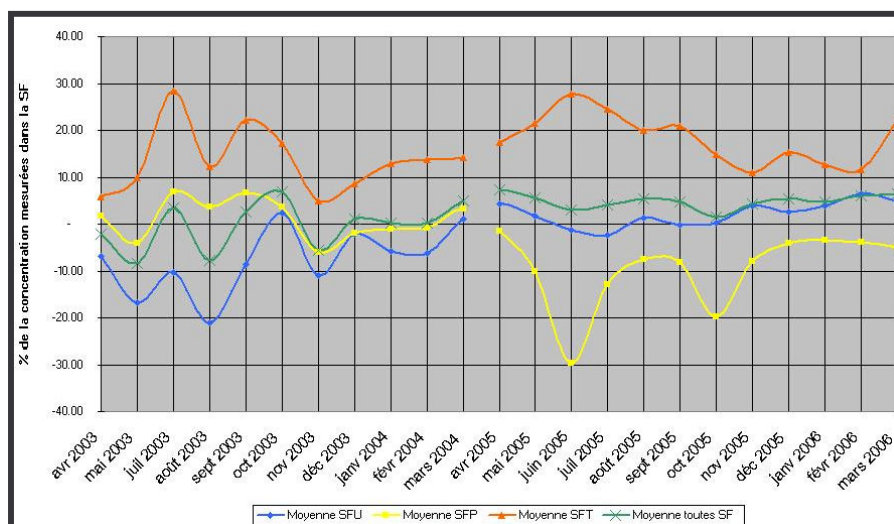
Les capteurs automatiques mesurant la pollution de proximité affichent les plus fortes erreurs absolue et relative ; en moyenne sur deux ans, ces stations sont sous-estimées de 16%.

Les stations mesurant la pollution de fond (Urbaines et Périurbaines) présentent des faibles erreurs d'estimation par validation croisée (en moyenne sur deux ans). Une surestimation de 2 à 4% est constatée.

Ces résultats sont conformes au fait que les stations de référence sélectionnées pour générer les données des stations virtuelles mesurent la pollution de fond (Halluin Caillou : station périurbaine et Roubaix Château : station urbaine).

La Figure 13 présente (en % de la concentration mesurée) la moyenne mensuelle de l'erreur relative d'estimation par validation croisée en fonction de la typologie des stations fixes :

- Pour les 4 stations de trafic la moyenne mensuelle de l'erreur relative varie entre 5 et 28%, ces capteurs sont donc fortement sous-estimés.
- Pour les deux stations périurbaines, les moyennes mensuelles de l'erreur d'estimation par validation croisée oscillent entre -6% et 7% dans la l'année 2003-2004. Dans la période 2005-2006, à l'exception de juin et octobre 2005 (dans ces deux mois les niveaux de pollution par NO<sub>2</sub> ont été très faibles dans la périphérie de l'agglomération), où l'erreur d'estimation en moyenne mensuelle est de -30% et -20% respectivement, les erreurs se trouvent dans un intervalle entre -2% et -13% ; en cette année, les concentrations sont donc la plupart du temps surestimées.
- En 2003-2004 pour les 8 stations urbaines, les concentrations sont presque toujours surestimées, jusqu'à -21% en août 2003. En 2005-2006 les erreurs sont plus stables d'un mois à un autre, elles varient entre -2% et 7%.



Erreurs positives : sous - estimation, concentrations estimées plus petites que les mesures

Erreurs négatives : sur - estimation, concentrations estimées plus grandes que les mesures

**Figure 13 : Moyenne mensuelle de l'erreur relative d'estimation par validation croisée selon le type de station**

Une autre façon d'évaluer la qualité des estimations consiste à grouper chaque estimation journalière en fonction de la valeur absolue de l'erreur relative, ainsi :

- Groupe 1 : erreur relative inférieure ou égale à 10% en valeur absolue (**G1 : ERA <= 10%**).
- Groupe 2 : erreur relative supérieure à 10 et inférieure ou égale à 20% en valeur absolue (**G2 : ERA > 10% et ERA <=20%**).
- Groupe 3 : erreur relative supérieure à 20% en valeur absolue (**G3 : ERA > 20%**).
- Groupe 4 : erreur relative inférieure ou égale à l'objectif de qualité rappelé en introduction de cette étude, soit 50% (**G4 : ERA <= 50%**).

Pour les 14 capteurs automatiques disponibles dans l'agglomération de Lille, le Tableau 13 rapporte le pourcentage des estimations journalières appartenant à chaque groupe ainsi créé.

Type SF	Code ST	NOM SF	Nb jours	G1 (%)	G2 (%)	G3 (%)	G4 (%)
Urbaine	11002	ROUBAIX CHATEAU	557	78.3	17.2	4.5	99.6
	11016	MARCQ EN BAROEUL CTM	541	35.3	37.5	27.2	99.8
	11022	ARMENTIERES BEAUVAIS	539	33.0	32.8	34.1	95.4
	11024	LILLE FAIDHERBE	540	24.3	25.2	50.6	85.4
	11025	LILLE FIVES LAKANAL	543	47.9	26.0	26.2	97.1
	11027	TOURCOING BIENFAISANCE	538	50.2	24.5	25.3	94.1
	11029	LOMME	548	44.9	31.0	24.1	95.8
	11032	V D ASCQ RAMEAU	550	47.3	29.3	23.5	97.6
Périurbaine	11026	HALLUIN CAILLOUX	557	63.2	21.4	15.4	98.9
	11033	LESQUIN	543	38.1	28.0	33.9	95.8
Trafic	11005	LILLE LIBERTE	524	31.3	19.7	49.0	94.8
	11019	LILLE PASTEUR	548	52.4	29.4	18.2	99.1
	11034	ROUBAIX SERRES	512	21.3	21.1	57.6	94.9
	11036	HAUBOURDIN POMPIERS	471	20.2	24.4	55.4	93.2

Tableau 13 : Pourcentage (par station fixe) des estimations journalières où l'erreur relative est comprise entre les intervalles indiqués

Comme l'on pouvait s'y attendre, les estimations journalières des deux stations de référence présentent le plus faible erreur relative (en valeur absolue), ce qui se traduit par de forts pourcentages des jours appartenant au premier groupe (Roubaix Château : 78.3% et Halluin Caillou : 63.2%).

En revanche, aux trois stations trafic et pour la station Lille Faidherbe, les concentrations journalières de NO<sub>2</sub> sont les moins bien estimées : forts pourcentages de jours appartenant au troisième groupe (p. ex. Roubaix Serres : 57.6% ou Haubourdin Pompiers: 55.4%).

Dans le Tableau 14 ces mêmes statistiques sont condensées en fonction de la typologie des stations fixes.

Statistiques et intervalles de l'erreur relative	Type de SF			
	Urbaine	Périurbaine	Trafic	Toutes les SF
Nb de stations fixes	8	2	4	14
Nb de jours estimés	4356	1100	2055	7511
G1 (%)	45.3	50.8	31.9	<b>42.4</b>
G2 (%)	27.9	24.6	23.7	<b>26.3</b>
G3 (%)	26.8	24.6	44.4	<b>31.3</b>
G4 (%) (Nb de jours où ERA > 50%)	95.6 (191 j)	97.4 (29 j)	95.6 (90 j)	95.9 (310 j)

Tableau 14 : Répartition par classes (%) de l'erreur relative des estimations journalières aux stations fixes

Dans 42.4% des cas, l'erreur maximale est de +/- 10%. Ce pourcentage s'élève à 68.7% pour les estimations présentant une erreur maximale de +/- 20%.

Sur l'ensemble des stations, l'erreur est supérieure à 50% dans seulement 4% des cas (310 jours).

Dans ces statistiques les estimations journalières ont été prises de façon indépendante. Par la suite, les statistiques prenant en compte de façon simultanée les estimations effectués dans une même journée sont présentées. Ces calculs permettront de choisir un critère afin de distinguer les jours pour lesquels l'estimation est satisfaisante de ceux où la méthode s'applique mal.

Le premier critère analysé consiste à choisir les jours où la valeur maximale de l'erreur relative en valeur absolue est inférieure ou égale à un seuil donné. Le Tableau 15 montre le résultat de ces calculs pour les seuils correspondant aux quatre groupes formés précédemment.

La valeur maximale journalière de l'erreur relative (en valeur absolue)...	Nb jours		Nb jours (%)		M NO2	
	03/04	05/06	03/04	05/06	03/04	05/06
est inférieure ou égale à 10%	0	0	0	0	N/A	N/A
est supérieure à 10% et inférieure ou égale à 20%	17	3	6	1	54.9	60
est supérieure à 20%	248	289	94	99	37.3	37.3
est inférieure ou égale à 50% (objectif qualité)	166	205	<b>63</b>	<b>70</b>	43.7	41.7
est supérieure à 50%	99	87	30	37	29.6	27.8
Nombre de jours de l'ensemble : 01/04/2003 - 31/03/2004 et 01/04/2005 - 31/03/2006)	265	292	100	100	38.4	37.5

M NO2 désigne la moyenne des concentrations journalières obtenues pour les classes des journées, toutes stations fixes confondues

Tableau 15 : Nombre de jours où la valeur maximale journalière de l'erreur relative de toutes les SF (en valeur absolue) est comprise entre les intervalles indiqués

Dans 63% des estimations journalières de l'année 2003-2004 et 70% des estimations de l'année 2005-2006 l'erreur relative calculée en validation croisée est toujours inférieure ou égale à 50% sur toutes les 14 stations fixes de l'agglomération lilloise. La moyenne des concentrations de NO<sub>2</sub> pour ces jours-ci est d'environ 43 µg/m<sup>3</sup>, et d'environ 29 µg/m<sup>3</sup> pour les jours ne remplissant pas le critère de qualité (erreur >= 50%).

Il est intéressant d'observer qu'il n'y a pas de cas pour lesquels l'erreur maximale parmi les 14 stations fixes est inférieure à 10%. La valeur maximale des erreurs relatives de la plupart des estimations journalières se trouve dans l'intervalle [20 - 50%].

Un autre critère moins restrictif consiste à prendre en compte, au lieu de la valeur maximale, la moyenne journalière des erreurs relatives en valeur absolue.

Le Tableau 16 montre le nombre d'estimations journalières appartenant à chaque groupe en prenant comme critère la moyenne journalière de l'erreur relative de toutes les 14 stations fixes.

La moyenne journalière de l'erreur relative (en valeur absolue)...	Nb jours		Nb jours (%)		M NO <sub>2</sub>	
	03/04	05/06	03/04	05/06	03/04	05/06
est inférieure ou égale à 10%	40	28	15	10	52.1	50.9
est supérieure à 10% et inférieure ou égale à 20%	135	184	51	63	41.1	40.9
est supérieure à 20%	90	80	34	27	28.2	25.2
est inférieure ou égale à 50% (objectif qualité)	262	290	98.9	99.3	38.7	37.7
est supérieur à 50%	<b>3</b>	<b>2</b>	1.1	0.7	16.2	10.6
Nombre de jours de l'ensemble : 01/04/2003 - 31/03/2004 et 01/04/2005 - 31/03/2006)	265	292	100	100	38.4	37.5

*Tableau 16 : Nombre de jours où la moyenne journalière de l'erreur relative de toutes les SF (en valeur absolue) est comprise entre les intervalles indiqués*

Si l'on tient compte de ce critère moins sélectif, on s'aperçoit que la moyenne des erreurs relatives des 16 stations fixes est supérieure à 50% pour seulement 5 jours, parmi les 557 estimations journalières en étude.

Il s'agit de 2 dimanches de la période estivale de l'année 2003, 2 dimanches de la même période estivale de l'année 2005 et un dimanche de novembre 2003. Ces jours se caractérisent par de faibles concentrations de NO<sub>2</sub>, la moyenne étant de 16.2µg/m<sup>3</sup> pour l'année 2003-2004 et de 10.6µg/m<sup>3</sup> pour l'année 2005-2006. Les jours pour lesquels les concentrations sont situées nettement en dessous de la moyenne sont donc assez mal cartographiés.

Les moyennes des erreurs relatives de la plupart des estimations journalières se trouvent dans l'intervalle [10 - 20%] ; 51% des estimations de l'année 2003-2004 et 63% de celles de l'année 2005-2006 sont concernées.

## 4.2 AGGLOMERATION DE REIMS

### 4.2.1 LES DONNEES

Dans l'agglomération rémoise, 8 campagnes de mesure du NO<sub>2</sub> par prélèvement passif ont été menées sur 88 sites, les quatre premières campagnes ont eu lieu durant la période hivernale 2002-2003 (du 23 décembre au 17 février), les autres campagnes correspondent à la période estivale 2003 (du 28 avril au 23 juin).

Le domaine d'étude fait environ 10 km d'Est en Ouest, et du Nord au Sud, avec un resserrement des points de mesure au niveau du centre-ville. Le réseau de surveillance dispose de 8 stations fixes de mesure du NO<sub>2</sub>, dont 4 sont définies comme urbaines et 4 comme périurbaines. La Figure 14 montre la localisation géographique des tubes passifs et des capteurs automatiques dans l'agglomération.

Pour la période hivernale la moyenne des concentrations en dioxyde d'azote mesurées par les tubes passifs sur toute l'agglomération était de 34.02 µg/m<sup>3</sup>. Dans l'été 2003, les niveaux de pollution n'ont que légèrement diminué jusqu'à 27.13 µg/m<sup>3</sup> en moyenne.

Les concentrations mesurées dans la période du 7 au 20 janvier 2003 sont les plus fortes parmi toutes les campagnes de mesure, la moyenne des concentrations pour cette campagne était de 38.34 µg/ m<sup>3</sup>. En revanche, la campagne estivale du 12 au 26 mai 2003 est celle dont les concentrations sont les plus faibles avec une moyenne de 24.75 µg/m<sup>3</sup>. Dans le rapport : « ATMO CA, étude Rs 03-03-TD-ED/EC » (cf. [7]) sont consignés les détails techniques et autres informations relatives à ces campagnes de mesure.

L'analyse des informations secondaires disponibles sur le domaine d'étude (cadastre kilométrique des émissions et la population au niveau des îlots) a permis de dégager une fonction auxiliaire corrélée linéairement avec les concentrations de dioxyde d'azote, cette fonction est la somme de deux variables normées : la densité de population et la densité d'émissions (cf. Annexe 7).

Les figures 15 et 16, représentent les modèles de variogrammes calculés pour les deux moyennes saisonnières du dioxyde d'azote et la fonction auxiliaire ainsi construite, la première montre la corrélation spatiale en période estivale et la seconde celle correspondant à la période hivernale.

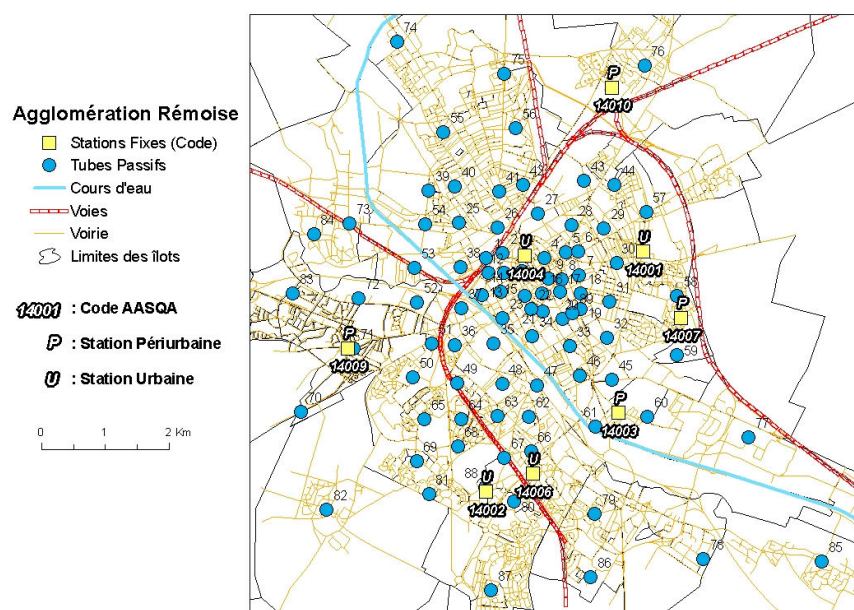


Figure 14 : Localisation géographique des tubes passifs et des 8 stations fixes dans la région rémoise

Il est nécessaire de mener au préalable une étude des données afin de définir les variogrammes qui seront implantées dans la procédure de cartographie automatique.

L'étude variographique des concentrations saisonnières montre une anisotropie géométrique dont la direction de plus grande corrélation spatiale en période estivale est N20W (N160 dans les modèles). En hiver cette direction varie légèrement de 5 degrés : l'anisotropie est N15W (N165 dans les modèles). D'autre part la corrélation spatiale des concentrations de NO<sub>2</sub> dans l'agglomération rémoise présente deux échelles spatiales (une petite et une grande échelle).

Les variogrammes expérimentaux ont été modélisés avec un effet de pépite et une structure gaussienne. Les paramètres des modèles, en fonction des deux directions d'anisotropie, des deux échelles spatiales et des deux périodes saisonnières, sont renseignés dans les tableaux ci-joints :

Paramètres	D1 : N70E		D2 : N20W	
	Petite échelle	Grande échelle	Petite échelle	Grande échelle
Pépite : 20 (32%)				
Palier	27 (44%)	15 (24%)	27 (44%)	15 (24%)
Portée	900 m	2.8 Km	2.4 Km	5 Km

Entre parenthèses : Le pourcentage de la somme de paliers, cela représente la partie de la variabilité (variance) modélisée par chaque structure (ou modèle de covariance)

Tableau 17 : Paramètres des modèles de variogrammes de la concentration de NO<sub>2</sub> en période estivale

Paramètres	D1 : N75E		D2 : N15W	
Pépite : 12 (24%)	Petite échelle	Grande échelle	Petite échelle	Grande échelle
Palier	3 (5%)	36 (71%)	3 (5%)	36 (71%)
Portée	700 m	4 Km	2.5 Km	5 Km

Entre parenthèses : Le pourcentage de la somme de paliers, cela représente la partie de la variabilité (variance) modélisée par chaque structure (ou modèle de covariance)

Tableau 18 : Paramètres des modèles de variogrammes de la concentration de  $\text{NO}_2$  en période hivernale

En période hivernale, la corrélation spatiale des concentrations est plus importante à grande échelle (portée entre 4 et 5Km) car elle représente le 71% de la variabilité spatiale (variance). En revanche, en été, le phénomène est moins bien structuré puisque l'effet de pépite représente 32% de la variabilité spatiale (variance) et la corrélation spatiale à petite échelle (portée entre 900 et 2.4Km) 44%. Ces modèles ont été conservés pour le cokrigage journalier automatique.

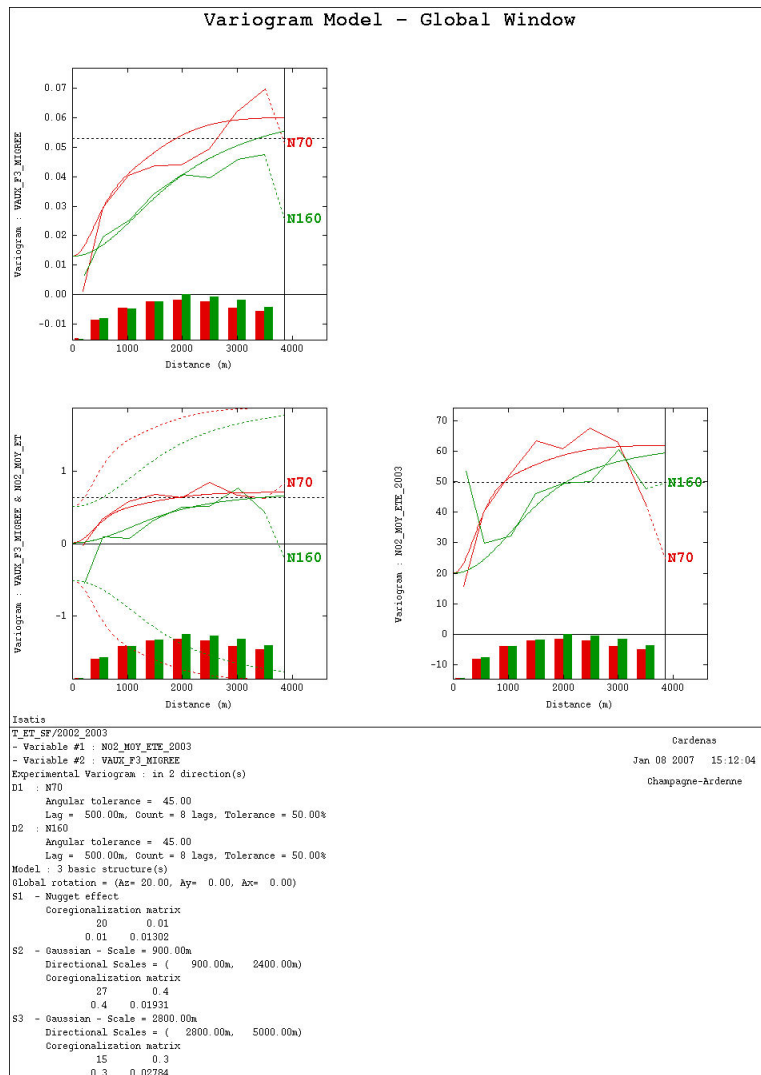


Figure 15 : Modèle utilisé pour les estimations journalières en été

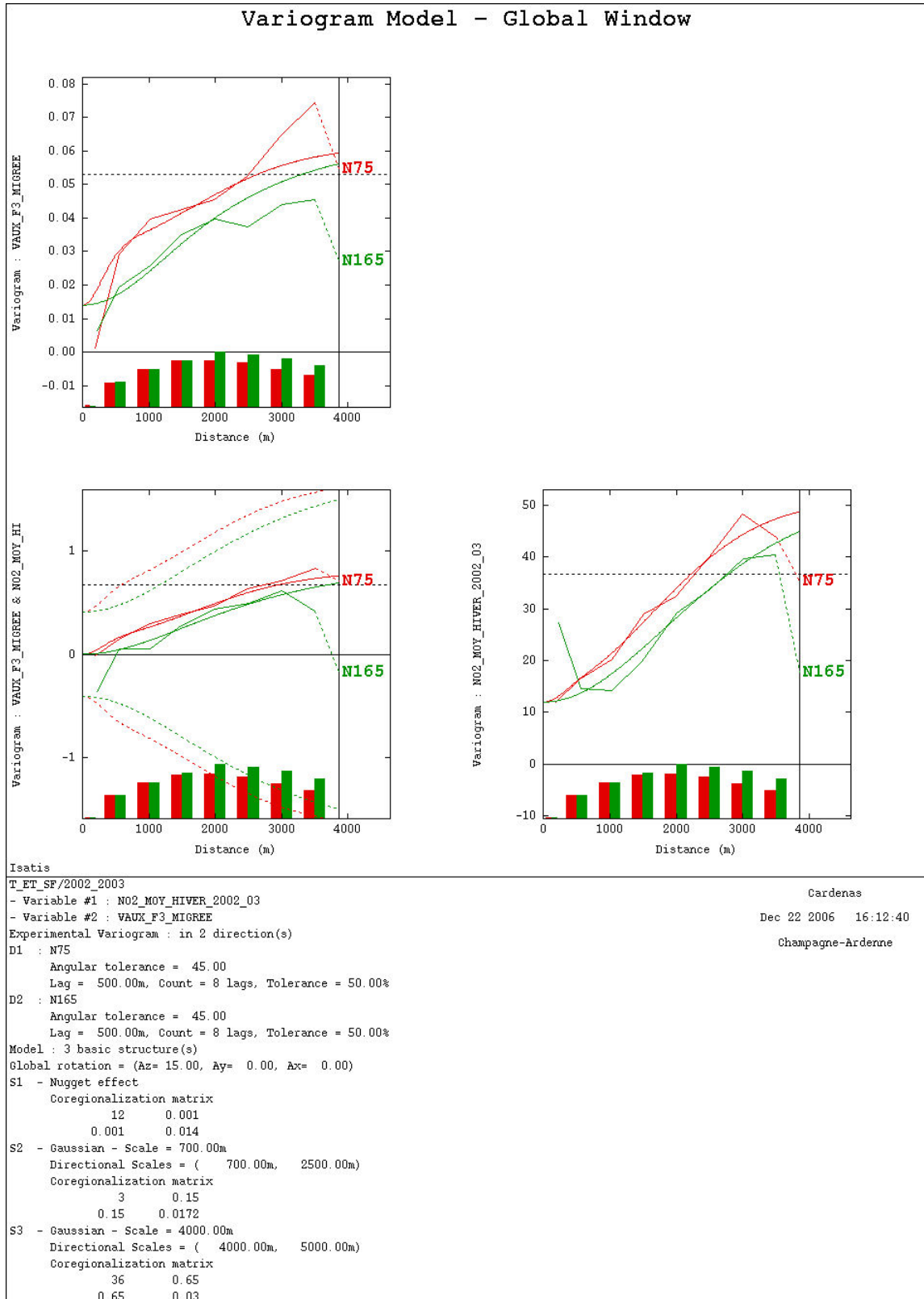


Figure 16 : Modèle utilisée pour les estimations journalières en hiver



## 4.2.2 COMPARAISON ENTRE LE CKJA ET LE CKT SUR L'AGGLOMERATION REMOISE

### 4.2.2.1 APPLICATION DU CKJA

La procédure expliquée dans la section 3.2 est appliquée pour estimer les concentrations journalières de NO<sub>2</sub>.

Le CKJA a été calculé sur deux périodes de mesure différentes, d'abord entre le 1<sup>er</sup> octobre 2002 et le 30 septembre 2003 afin de comparer avec le CKT, puis entre le 1<sup>er</sup> octobre 2005 et le 30 septembre 2006 afin de prendre en compte des épisodes de pollution variés dans la comparaison des estimations par validation croisée (cf. paragraphe 4.2.3).

#### 4.2.2.1.1 SELECTION DES STATIONS VIRTUELLES ET CALCUL DE LA VEM

La méthodologie détaillée au paragraphe 3.1 a été appliquée pour mettre en évidence des corrélations entre tubes et stations. 44 sites (stations virtuelles) remplissent les critères de sélection décrits aux étapes 1 et 2 (coefficient de détermination  $r^2 > 0,8$  et représentativité de la gamme de concentrations).

La figure ci-dessous présente par exemple les corrélations obtenues entre la station 14001 (Cernay) et 4 stations virtuelles.

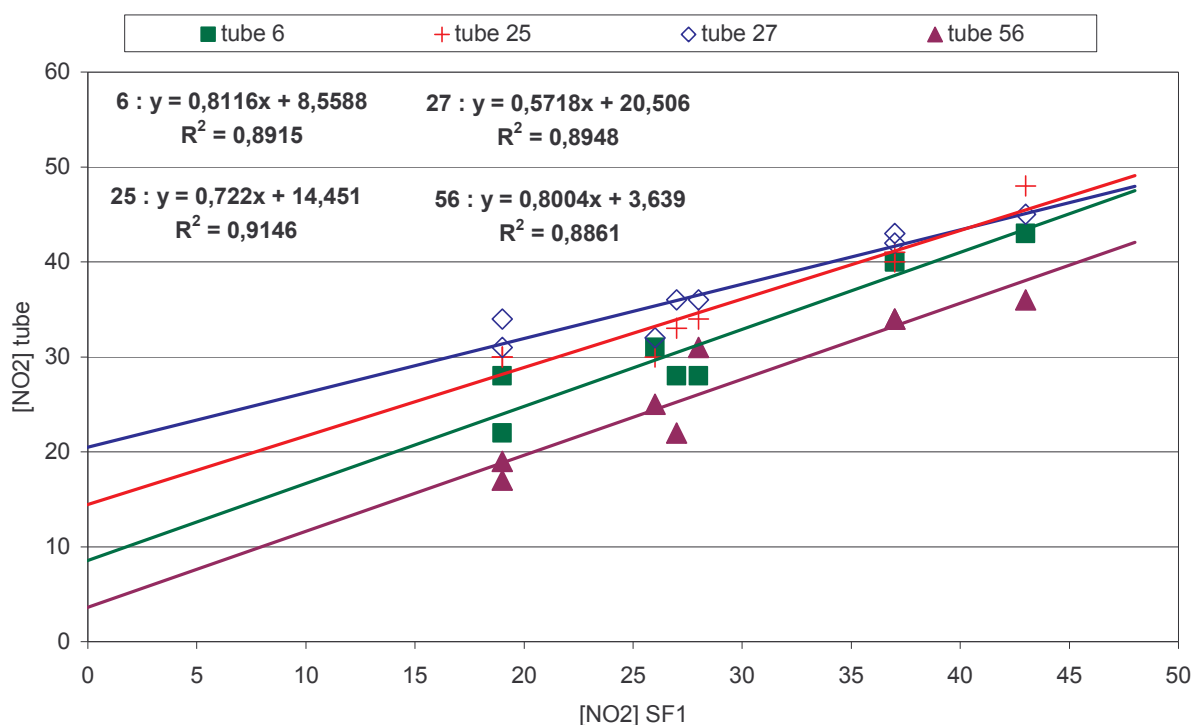


Figure 17 : Relation entre la station fixe 14001 et quatre tubes passifs corrélés

Aucun tube n'est corrélé à la station 14004 (Hôtel de Ville, localisée au centre ville et présentant de fortes concentrations). La station 14009 (Tingueux, SF périurbaine) a pour sa part été exclue de l'étude car elle ne présente pas de valeurs sur le deuxième domaine de représentativité des concentrations, comme il a été expliqué dans la deuxième étape de la méthode Corrélations tubes - stations (paragraphe 3.1.2.2). Le capteur Bétheny (code 14010, cf. Figure 18) n'a pas été pris en compte car il s'agit d'une station récemment installée (30 septembre 2005).

La répartition des corrélations pour les 5 stations fixes restantes est présentée dans la figure et le tableau ci-après.

Station fixe Type	14001 urbaine	14002 urbaine	14003 périurbaine	14006 urbaine	14007 périurbaine
Nombre de stations virtuelles corrélées	4	10	9	17	4

Tableau 19 : Nombre de corrélations observées entre stations fixes et virtuelles

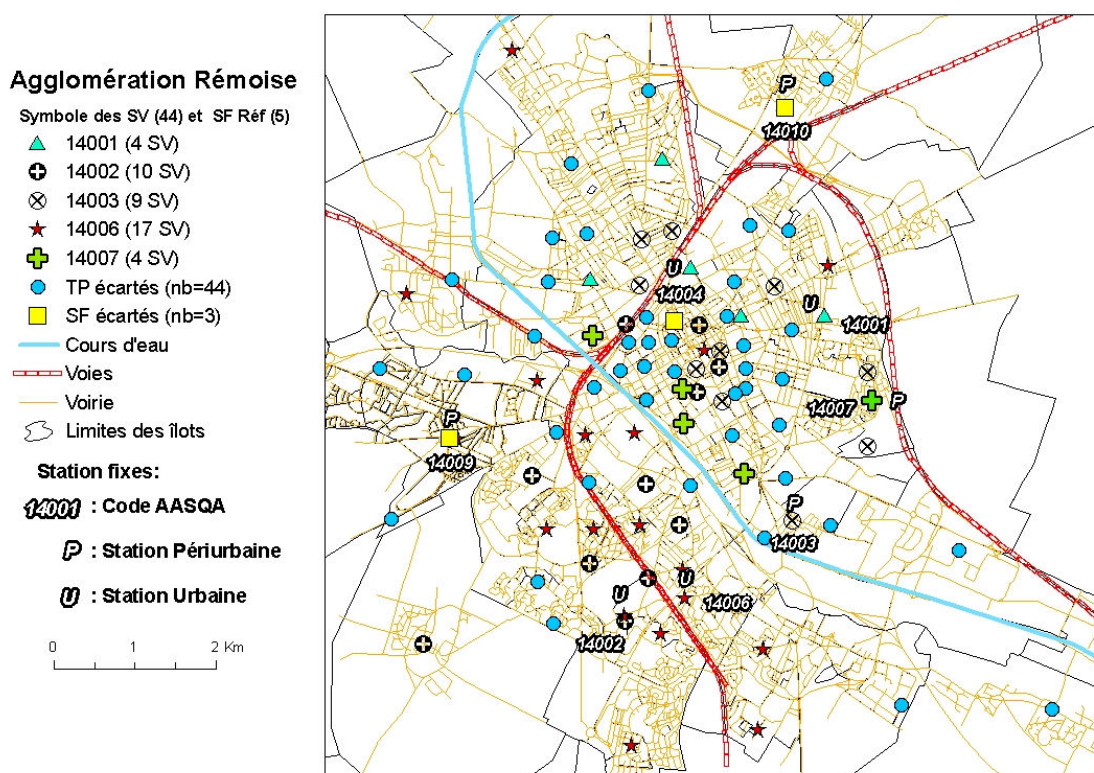
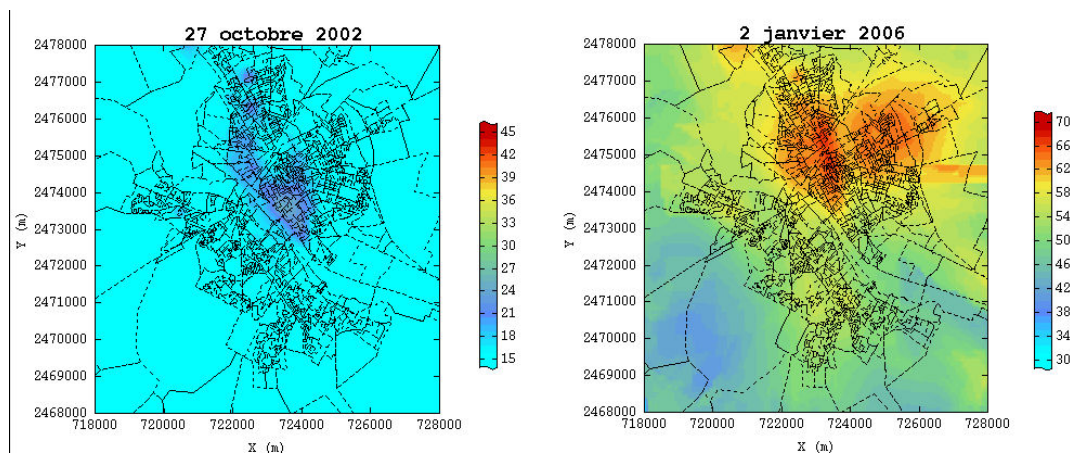


Figure 18 : Carte des stations virtuelles sélectionnées

## 4.2.2.1.2 RESULTATS DES ESTIMATIONS JOURNALIERES

La Figure 19 montre deux exemples des estimations journalières de NO<sub>2</sub> sur Reims.



Lignes pointillées : découpage îlots de l'agglomération

Concentration maximale NO<sub>2</sub>, journée du 27 octobre 2002 : 22.2 µg/m<sup>3</sup>

Concentration maximale NO<sub>2</sub>, journée du 2 janvier 2006 : 67.1 µg/m<sup>3</sup>

Figure 19 : Exemple des deux estimations des niveaux de pollution de NO<sub>2</sub> dans l'agglomération rémoise

Voici les paramètres utilisés pour le CKJA:

- Les concentrations aux stations virtuelles sont calculées à partir des stations de référence comme il a été expliqué au paragraphe précédent.
- le fichier des concentrations journalières intègre donc finalement les concentrations des stations virtuelles et les mesures des 6 stations fixes disponibles (les cinq stations du Tableau 19, et la station de l'Hôtel de Ville - code14004).
- La variable auxiliaire utilisée lors de l'estimation par CKT est maintenue pour l'estimation des deux saisons (somme des deux variables normées : la densité d'émissions et la densité de population (cf. Annexe 7)).
- Les figures 15 et 16 donnent les détails des modèles de variogrammes choisis.
- La méthode d'estimation est le cokrigage colocalisé avec prise en compte de la variance de l'erreur de mesure sur les données des stations virtuelles. La VEM est calculée comme il a été expliqué au paragraphe précédent.

#### 4.2.2.2 COMPARAISON CKT – CKJA SUR LE DOMAINE D'ETUDE

Une description détaillée de l'application de la méthode du cokrigeage temporel aux données rémoises est présentée dans l'Annexe 7.

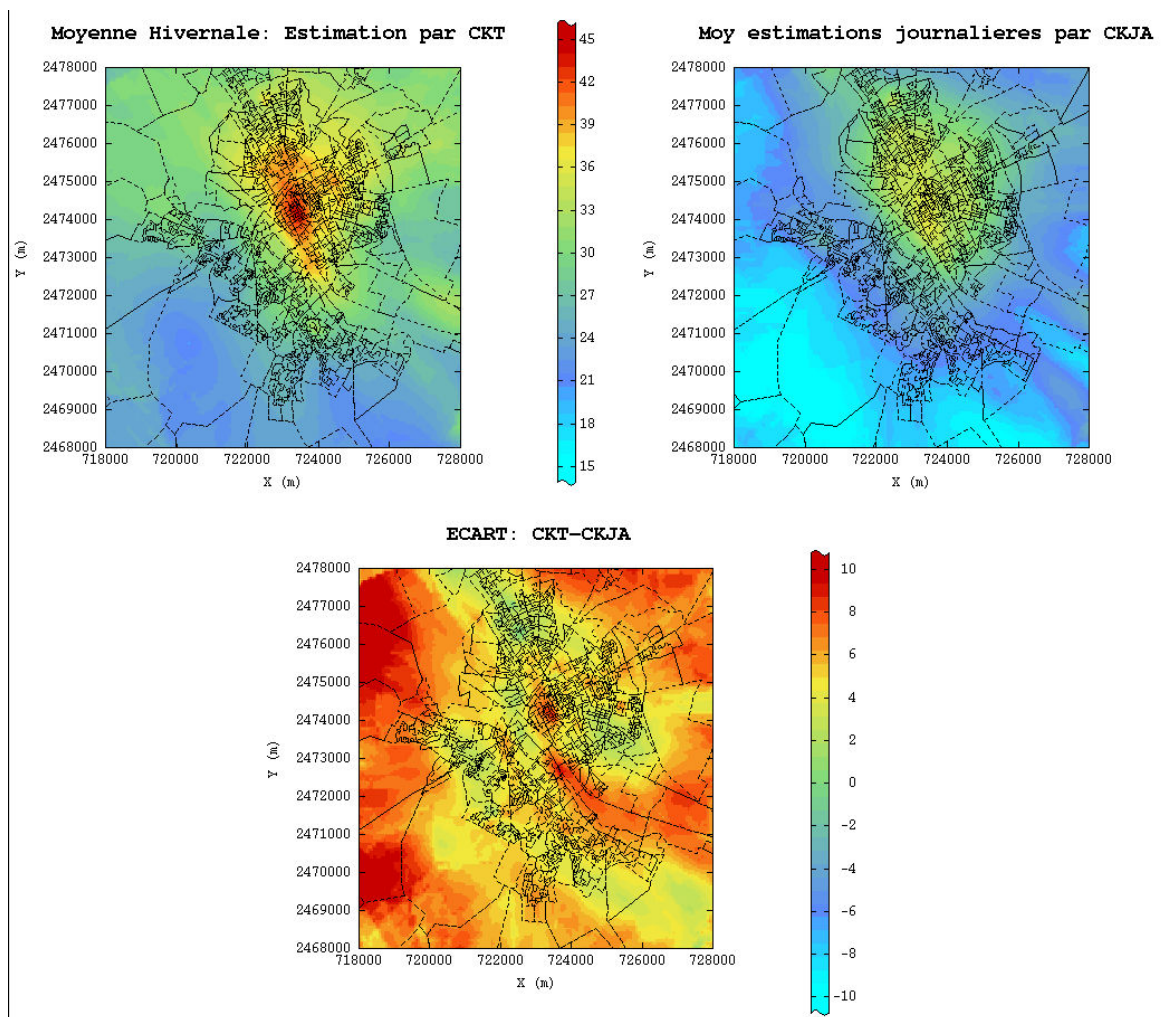
Dans les figures et tableaux ci-joints sont comparées les cartes des estimations saisonnières et annuelles effectuées avec les deux méthodes d'estimation. Pour le CKJA il s'agit de la moyenne des estimations journalières effectuées sur la même période couverte par les campagnes des tubes passifs.

Pour chaque période, les écarts absolu et relatif entre le CKT et le CKJA sont également calculés.

Contrairement aux résultats obtenus sur la région lilloise, à Reims les niveaux de pollution calculés par CKJA sont en moyenne plus faibles que ceux obtenus par CKT. Ceci est peut-être dû au manque de stations mesurant la pollution de proximité (stations de trafic) dans l'agglomération.

En effet, dans le CKJA, la seule station qui fournit de fortes concentrations est celle de typologie urbaine située à l'hôtel de ville (code 14004, cf. Figure 18), malheureusement aucune station virtuelle n'a pas été rattachée à cette station.

En revanche dans le CKT, il existe un nombre plus important de tubes passifs mesurant les fortes concentrations du centre ville. De ce fait, dans cette zone, en moyenne hivernale l'écart entre les deux estimations peut aller jusqu'à  $10 \mu\text{g}/\text{m}^3$  et jusqu'à  $5 \mu\text{g}/\text{m}^3$  en moyenne estivale.



Lignes pointillées : découpage îlots de l'agglomération

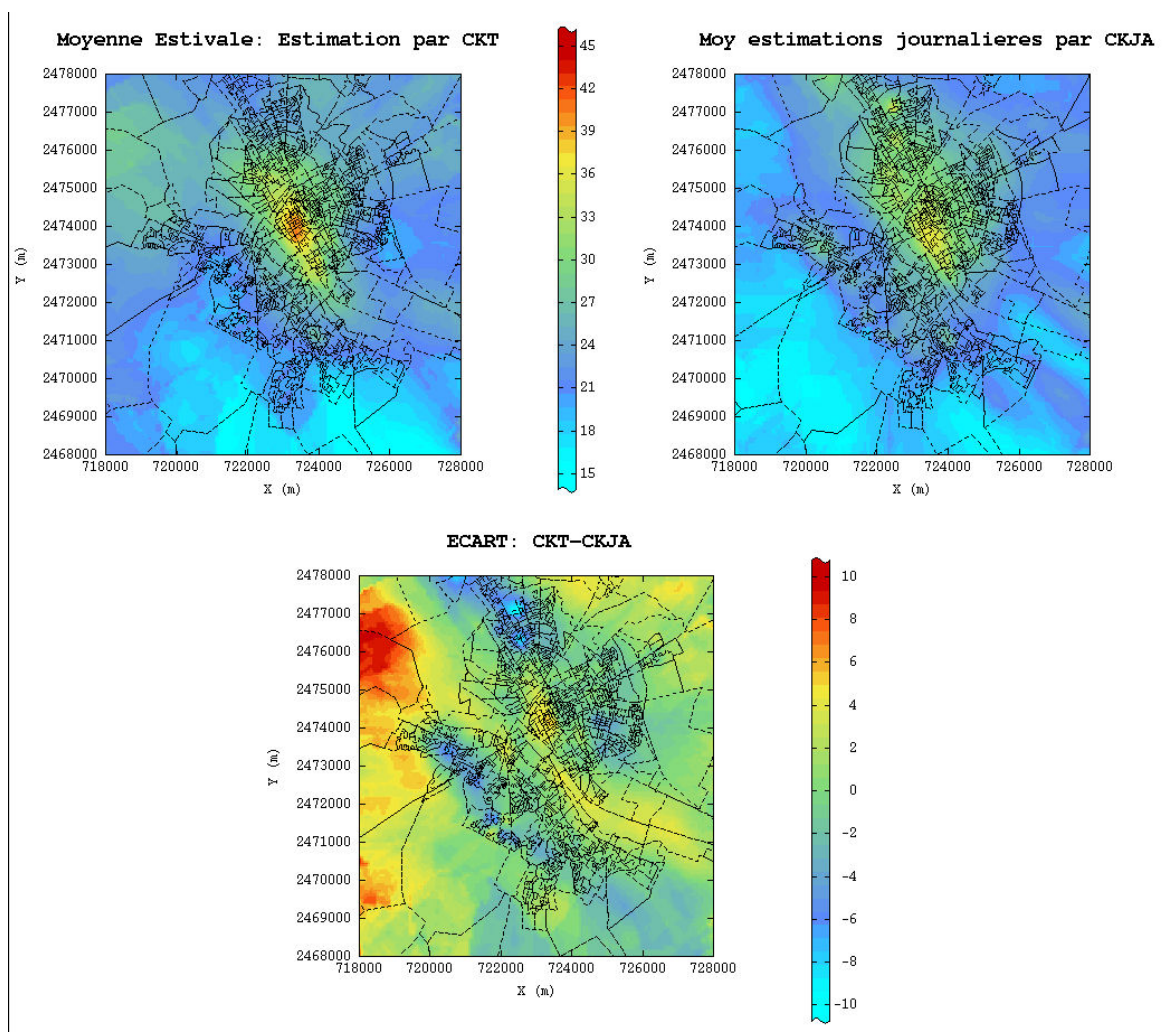
Figure 20 : Cartes comparant les estimations de la moyenne hivernale de NO<sub>2</sub> (année 2002-2003) dans la région rémoise, obtenues avec CKAJ et CKT

Moyenne hiver 2002- 2003	Date Début	Date Fin	Nb jours	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
CKT	23 déc. 2002	17 fév. 2003	57	10000	20.59	45.03	28.73	4.53	20.56	0.16
CKJA					13.55	36.05	22.81	5.05	25.47	0.22
Ecart : CKT-CKJA					0.06	12.6	5.92	1.81	3.29	0.31
Ecart Relatif (%): CKT-CKJA					19	47	21	7.3	53.38	0.35

Ecart positif : concentrations estimées par CKJA plus petites que celles estimées par CKT

Ecart négatif : concentrations estimées par CKJA plus grandes que celles estimées par CKT

Tableau 20 : Comparaison des statistiques de la moyenne hivernale calculées par CKT et CKJA



Lignes pointillées : découpage flots de l'agglomération

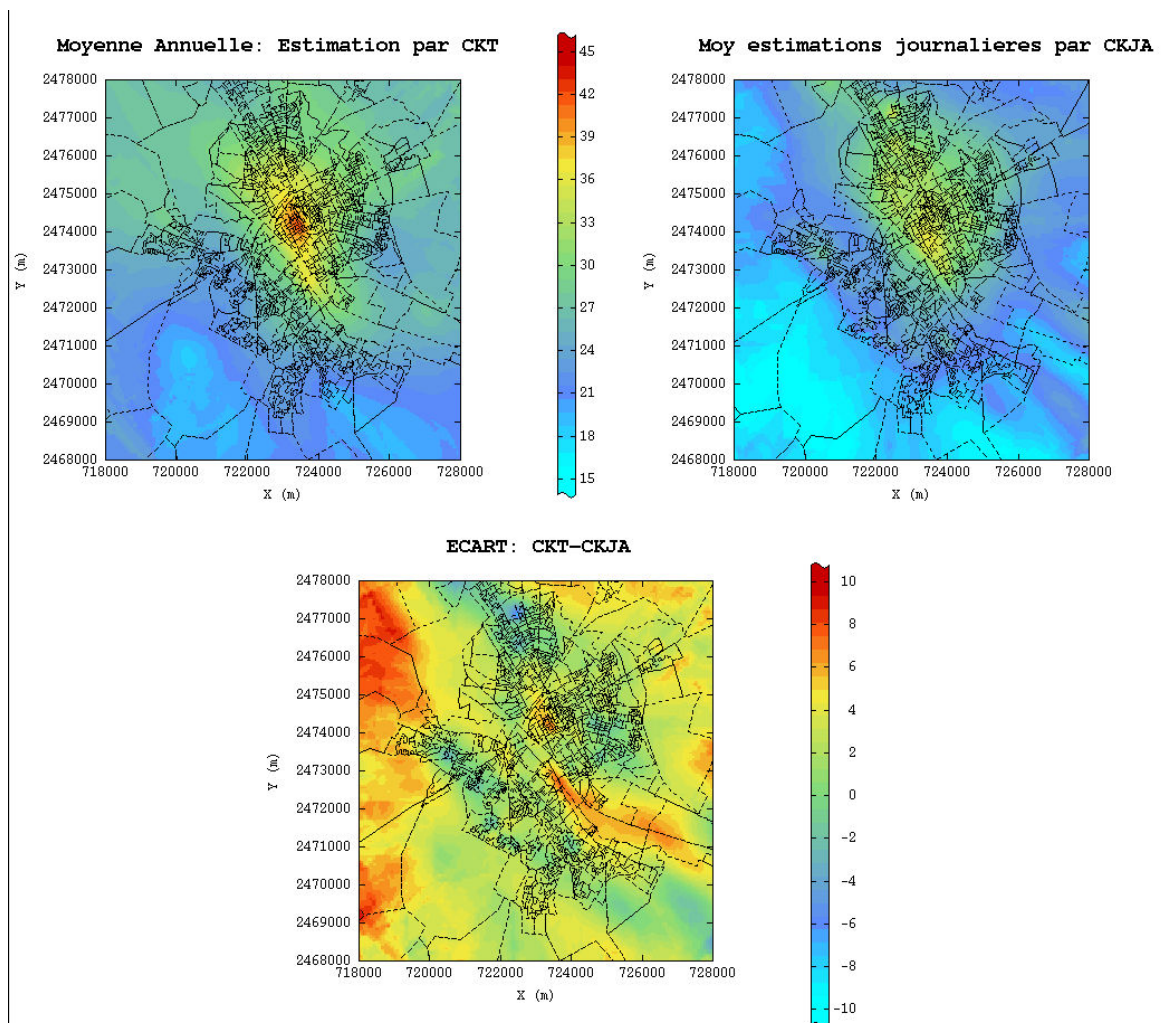
Figure 21 : Comparaison estimations de la moyenne estivale de NO<sub>2</sub> (année 2002-2003) dans l'agglomération rémoise obtenues avec les méthodes CKJA et CKT

Moyenne été 2003	Date Début	Date Fin	Nb jours	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
CKT	28 avril 2003	23 juin 2003	57	10000	14.09	40.54	22.87	3.93	15.43	0.17
CKJA					14.58	36.54	21.89	3.97	15.73	0.18
Ecart : CKT-CKJA					-12.15	10.4	0.98	2.79	7.77	2.85
Ecart relatif (%): CKT-CKJA					-57.5	36.3	3.6	12	144.48	3.3

Ecart positif : concentrations estimées par CKJA plus petites que celles estimées par CKT

Ecart négatif : concentrations estimées par CKJA plus grandes que celles estimées par CKT

Tableau 21 : Comparaison des statistiques de la moyenne estivale calculées par CKT et CKJA



Lignes pointillées : découpage îlots de l'agglomération

Figure 22 : Comparaison estimations de la moyenne annuelle de NO<sub>2</sub> (année 2002-2003) dans l'agglomération rémoise obtenues avec les méthodes CKJA et CKT

Moyenne année 2002-2003	Date Début	Date Fin	Nb jours	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
CKT	23 déc. 2002 - 28 avril 2003	23 juin 2003 - 17 fév. 2003	114	10000	18.16	42.35	25.63	4.06	16.5	0.16
CKJA					14.2	36.3	22.35	4.46	19.93	0.2
Ecart : CKT-CKJA					-7.26	9.33	3.28	2.1	4.41	0.64
Ecart relatif (%) : CKT-CKJA					-25.9	39.5	13	8.3	69.57	0.64

Ecart positif : concentrations estimées par CKJA plus petites que celles estimées par CKT

Ecart négatif : concentrations estimées par CKJA plus grandes que celles estimées par CKT

Tableau 22 : Comparaison des statistiques de la moyenne annuelle calculées par CKT et CKJA

## 4.2.2.3 COMPARAISON CKT – CKJA SUR LES STATIONS FIXES

Pour chaque station fixe, les concentrations moyennes saisonnières et annuelle du nœud le plus proche sont considérées. Ces concentrations correspondent aux estimations effectuées par la méthode du CKT et par CKJA. La période prise en compte correspond à celle couverte par les campagnes par tubes passifs.

Les nuages de corrélation et la comparaison des statistiques permet de déterminer la méthode qui donne en moyenne les valeurs les plus proches des concentrations mesurées.

Avec la station 14004 qui mesure les fortes valeurs de NO<sub>2</sub> en centre ville, dans la Figure 23 on observe d'une part un groupe de trois stations mesurant en moyenne les plus faibles concentrations, (stations 14007, 14006 et 14002), et d'autre part deux stations (14001 et 14007) présentant des concentrations intermédiaires. Ces 5 stations sont localisées en périphérie de l'agglomération (cf. Figure 18).

La méthode du CKJA paraît être plus adaptée à l'estimation de ces cinq stations, car les points mesure - estimation sont plus proches de la bissectrice dans les nuages de corrélation, surtout pour la période hivernale et la moyenne annuelle.

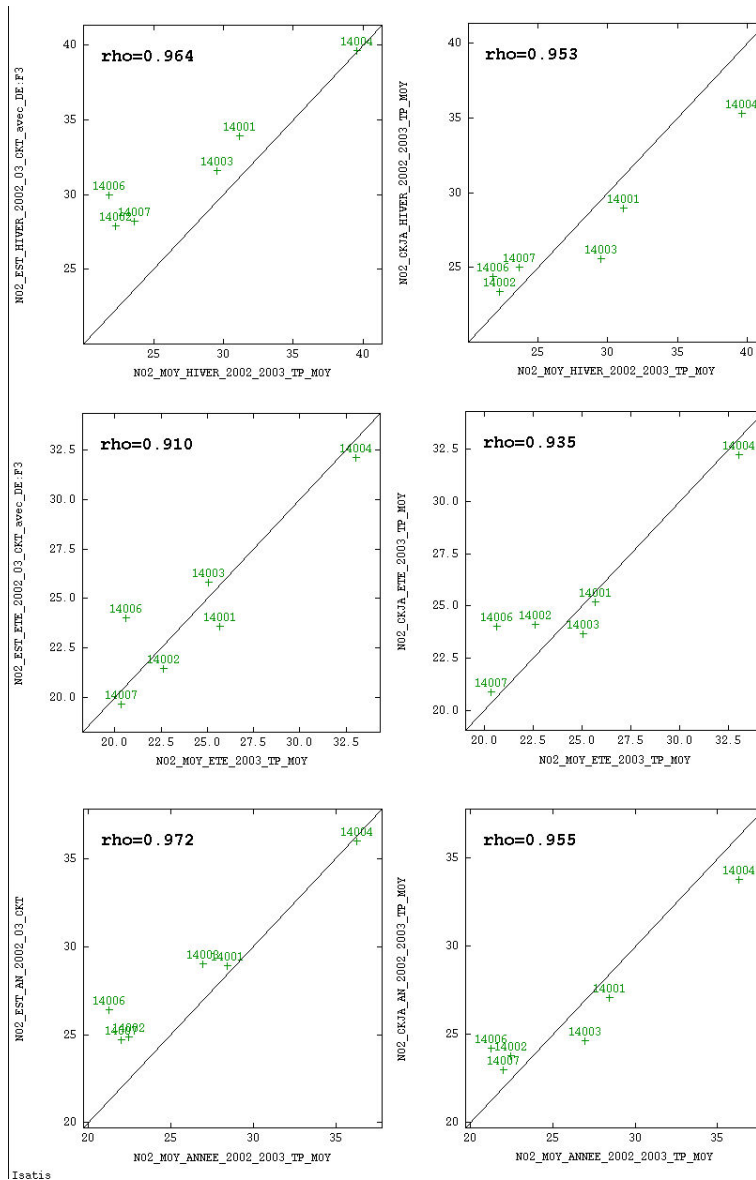
Pour le CKT le grand nombre de tubes passifs localisés en centre ville permet de mieux représenter, dans les deux saisons et en moyenne annuelle, la variabilité des concentrations de la station localisée à l'Hôtel de Ville.

En revanche, en période hivernale, le CKJA ne réussit pas à restituer les fortes concentrations de cette zone.

Statistique	HIVER		ETE		ANNEE	
	CKJA	CKT	CKJA	CKT	CKJA	CKT
Moyenne des erreurs	0.93	-3.84	-0.45	0.15	0.18	-2.06
Moyenne de l'erreur carrée	8.19	21.60	2.78	3.21	4.19	7.31
Racine de la moyenne de l'erreur carrée (RMSE)	2.86	4.65	1.67	1.79	2.05	2.70
Variance des erreurs	7.33	2.62	1.61	1.78	2.04	1.76

*Tableau 23 : Statistiques des erreurs entre concentrations moyennes aux 6 stations fixes et estimations effectuées par CKT et CKJA (même période de mesure)*





Axe X : Moyenne aux stations fixes de la moyenne journalière  
 En haut : Période hivernale: Du 23 décembre 2002 au 17 février 2003  
 Au milieu : Période estivale : Du 28 avril au 23 juin 2003  
 En bas : Moyenne des deux périodes précédents

Axe Y :

A gauche : Migration Maille 2D - points stations fixes de la moyenne des estimations par COKRIGEAGE TEMPOREL

A droite : Migration Maille 2D - points stations fixes de la moyenne des estimations par cokrigage journalier automatique

Les codes AASQA des capteurs sont notées, leurs noms sont indiquées dans le Tableau 24.

**Figure 23 : Nuages de corrélation entre concentrations moyennes aux 6 stations fixes et estimations effectuées par CKT et CKJA (même période de mesure)**

### 4.2.3 ANALYSE DE L'ECART AUX STATIONS FIXES: « MOYENNE JOURNALIERE DES MESURES » - « ESTIMATIONS »

Le Tableau 24 montre les statistiques des erreurs d'estimation calculées par validation croisée, et des moyennes journalières des mesures de NO<sub>2</sub> par station.

La station localisée en centre ville (14004), mesure les plus fortes concentrations, la valeur maximale pour la période en question est en moyenne journalière de 94 µg/m<sup>3</sup>, cette station est la moins bien estimée car elle présente le plus forte écart-type de l'erreur d'estimation.

En revanche la station 14007, localisée à l'est de l'agglomération, et qui mesure en moyenne les plus faibles concentrations, est la mieux estimée en validation croisée, la moyenne des erreurs est proche de zéro et l'écart-type de l'erreur d'estimation est le plus faible.

Code	Nom et Type SF	Nb jours	Min NO2	Max NO2	M NO2	σ NO2	Min E	Max E	M E	σ E
14001	Cernay (U)	685	4.42	68.79	28.05	11.52	-12.79	19.23	1.30	5.91
14002	Val de Murigny (U)	669	1.22	69.21	24.58	12.60	-12.55	17.58	-2.31	4.30
14003	Verrerie Couraux (P)	650	2.38	66.84	27.28	11.69	-7.93	20.54	3.87	5.51
14004	Hôtel de Ville (U)	671	8.42	94.00	35.61	14.36	-14.97	37.58	2.12	8.55
14006	Chatillons (U)	481	2.24	70.64	26.26	12.93	-11.32	16.53	-1.93	4.70
14007	Adriatique Europe (P)	295	3.88	53.07	22.91	9.52	-13.23	13.94	-0.99	4.21
L'ensemble des SF		3451	1.22	94.00	28.02	13.05	-14.97	37.58	0.60	6.35

M NO<sub>2</sub> : Moyenne des moyennes journalières des mesures dans la station fixe

Min E: Valeur minimale de l'erreur (mesure – estimation par VC)      Max E: Valeur Maximale de l'erreur

M E : Moyenne de l'erreur (µg/m<sup>3</sup>)      σ E : Ecart - type de l'erreur

Erreurs positives : sous - estimation, concentrations estimées plus petites que les mesures

Erreurs négatives : sur - estimation, concentrations estimées plus grandes que les mesures

**Tableau 24 : Statistiques par station de l'erreur d'estimation par validation croisée sur deux années de mesure (01/10/2002 - 30/09/2003 et 01/10/2005 - 30/09/2006)**

En utilisant la même méthodologie que celle appliquée dans l'agglomération de Lille, nous avons trié les résultats des estimations journalières pour les deux années 2002-2003 et 2005-2006 selon le critère de la valeur absolue de l'erreur relative.

L'erreur relative est obtenue en calculant la différence entre la moyenne journalière mesurée par la station fixe et la valeur estimée pour la validation croisée, divisée par la valeur mesurée. Si la valeur mesurée est très faible, l'erreur relative peut donc atteindre des pourcentages très forts.

Par exemple dans la station 14002, localisée au sud de l'agglomération, la concentration minimale mesurée (1.22 µg/m<sup>3</sup>) a été surestimée par validation croisée de 800%, soit une concentration estimée d'environ 9.8 µg/m<sup>3</sup> (Tableau 25).

Code	Nom et Type SF	Nb jours	Min NO2	Max NO2	M NO2	$\sigma$ NO2	Min ER	Max ER	M ER	$\sigma$ ER
14001	Cernay (U)	685	4.42	68.79	28.05	11.52	-210.18	39.14	-2.76	29.39
14002	Val de Murigny (U)	669	1.22	69.21	24.58	12.60	-800.00	27.60	-26.85	57.83
14003	Verrerie Couraux (P)	650	2.38	66.84	27.28	11.69	-252.52	49.58	7.33	23.51
14004	Hôtel de Ville (U)	671	8.42	94.00	35.61	14.36	-168.41	42.41	-3.72	30.90
14006	Chatillons (U)	482	2.24	70.64	26.26	12.93	-505.36	23.40	-22.48	45.39
14007	Adriatique Europe (P)	295	3.88	53.07	22.91	9.52	-200.38	36.71	-11.28	28.76
L'ensemble des SF		3451	1.22	94.00	28.02	13.05	-800.00	49.58	-9.19	40.23

M NO2 : Moyenne des moyennes journalières des mesures dans la station fixe

Min ER: Valeur minimale de l'erreur relative (mesure – estimation par VC)\*100 / mesure (%)

Max E:R Valeur Maximale de l'erreur (%)

M E : Moyenne de l'erreur relative (%)

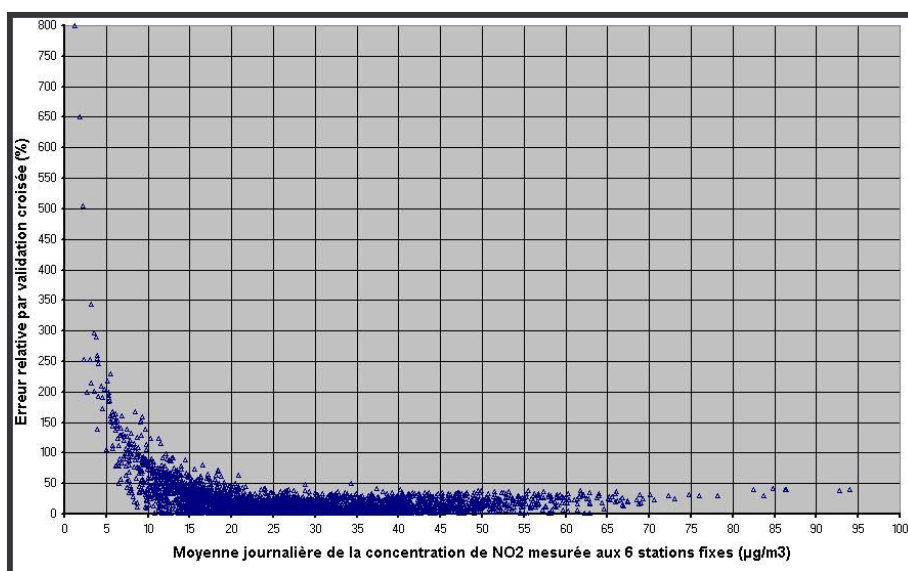
$\sigma$  ER : Ecart - type de l'erreur (%)

Erreurs positives : sous - estimation, concentrations estimées plus petites que les mesures

Erreurs négatives : sur - estimation, concentrations estimées plus grandes que les mesures

**Tableau 25 : Statistiques par station de l'erreur relative d'estimation par validation croisée sur deux années de mesure (01/10/2002 - 30/09/2003 et 01/10/2005 - 30/09/2006)**

La Figure 24 montre l'erreur relative, en valeur absolue, des concentrations de NO<sub>2</sub> pour l'ensemble des stations fixes, en fonction des concentrations moyennes mesurées de NO<sub>2</sub>. On observe qu'au-delà de 21  $\mu\text{g}/\text{m}^3$ , les estimations par validation croisée ont toujours une erreur relative inférieure à 50%.



**Figure 24 : Erreur relative, en valeur absolue, des concentrations de NO<sub>2</sub> pour l'ensemble des stations fixes, en fonction des concentrations moyennes mesurées de NO<sub>2</sub>**

Les estimations journalières ont également été groupées en fonction de la de l'erreur relative, les mêmes intervalles appliquées dans l'agglomération de Lille ont été utilisés, le Tableau 26 présente les résultats obtenues.

Code ST	NOM SF	Nb jours	G1 (%)	G2 (%)	G3 (%)	G4 (%)
14001	Cernay	685	33.7	28.2	38.1	93.87
14002	Val de Murigny	669	40.2	20.2	39.6	83.6
14003	Verrerie Couraux	650	28.8	29.4	41.8	97.7
14004	Hôtel de Ville	671	31.7	27.3	41.0	92.7
14006	Chatillons	481	38.9	23.1	38.0	85.9
14007	Adriatique Europe	295	43.1	27.1	29.8	92.9
<b>Ensemble des SF</b>		3451	35.2	25.9	38.9	91.2

G1 : ERA  $\leq$  10%

G2 : ERA  $>$  10% et ERA  $\leq$  20%

G3 : ERA  $>$  20%

G4 : ERA  $\leq$  50% (objectif qualité)

Avec ERA : Erreur relative absolue

*Tableau 26 : Pourcentage (par station fixe) des estimations journalières où l'erreur relative est comprise entre les intervalles indiqués*

On remarque que la station du centre ville 14004 et la station 14003, localisée dans la périphérie au sud-est de l'agglomération, sont les moins bien estimées, car elles comportent le plus grand nombre de jours appartenant au troisième groupe (erreurs supérieures à 20%).

Globalement, les estimations sont moins précises qu'à Lille, à cause peut-être du faible nombre de stations fixes disponibles. Dans 35.2% des estimations journalières, l'erreur maximale est de  $\pm$ 10%, ce pourcentage s'élève à 61.1% pour les estimations présentant une erreur maximale de  $\pm$ 20%. Finalement, si l'on considère le critère de qualité, dans 91.2% des cas, l'erreur est inférieure à 50% de la moyenne de la concentration mesurée.

Cependant si l'on prend en compte les estimations sur l'ensemble des stations fixes effectuées dans une même journée, de façon simultanée, on obtient les résultats exposés dans les tableaux 27 et 28. Afin de distinguer les jours pour lesquels l'estimation est satisfaisante de ceux où la méthode s'applique mal, plusieurs critères peuvent être employés.

La Tableau 27 montre les statistiques des jours où la valeur maximale de l'erreur relative est comprise entre les quatre intervalles définis précédemment.

On obtient seulement 6 jours où chaque station fixe est estimée avec une erreur inférieure ou égale à  $\pm$ 10%.

Si on prend en compte l'objectif de qualité (erreur de  $\pm$ 50%), les jours pour lesquels la méthode de cartographie automatique ne s'appliquerait pas de façon satisfaisante représenteraient 30% des données, soit l'équivalent de 206 jours (103 jours par an). Ils correspondraient à des jours où les valeurs de NO<sub>2</sub> mesurées sont faibles (17.4  $\mu\text{g}/\text{m}^3$  en moyenne).

La valeur maximale journalière de l'erreur relative (en valeur absolue)...	Nb jours	Nb jours (%)	M NO2
est inférieure ou égale à 10%	<b>6</b>	0.87	31.93
est supérieure à 10% et inférieure ou égale à 20%	106	15.34	33.27
est supérieure à 20%	579	83.79	26.82
est inférieure ou égale à 50% (objectif qualité)	485	70.19	32.3
est supérieur à 50%	<b>206</b>	<b>29.81</b>	<b>17.4</b>
Nombre de jours de l'ensemble : 01/10/2002 - 30/09/2003 et 01/10/2005 - 30/09/2006	<b>691</b>	<b>100</b>	

*Tableau 27 : Nombre de jours où la valeur maximale journalière de l'erreur relative de toutes les SF (en valeur absolue) est comprise entre les intervalles indiqués*

Un critère moins restrictif, pour déterminer la fiabilité de la méthode consiste à prendre en compte plutôt la moyenne journalière des erreurs relatives en valeur absolue. Dans ce cas, les jours pour lesquels la méthode de cartographie automatique ne s'appliquerait pas de façon satisfaisante (appelés JNS) représenteraient 6% des données, soit l'équivalent de 23 jours par an. Ils correspondraient à des jours où les valeurs de NO<sub>2</sub> mesurées sont faibles : 10,8 µg/m<sup>3</sup> en moyenne (Tableau 28).

La moyenne journalière de l'erreur relative (en valeur absolue)...	Nb jours	Nb jours (%)	M NO2
est inférieure ou égale à 10%	353	51.09	30.45
est supérieure à 10% et inférieure ou égale à 20%	194	28.08	32.66
est supérieure à 20%	144	20.84	15.02
est inférieure ou égale à 50% (objectif qualité)	649	93.92	28.96
est supérieur à 50%	<b>42</b>	<b>6.08</b>	<b>10.81</b>
Nombre de jours de l'ensemble : 01/10/2002 - 30/09/2003 et 01/10/2005 - 30/09/2006	<b>691</b>	<b>100</b>	

*Tableau 28 : Nombre de jours où la moyenne journalière de l'erreur relative de toutes les SF (en valeur absolue) est comprise entre les intervalles indiqués*

Une étude plus approfondie des caractéristiques de ces JNS démontre effectivement qu'il s'agit des jours où les concentrations de NO<sub>2</sub> sont généralement les plus faibles : les samedis et les dimanches d'une part et en printemps et en été d'autre part (Figure 25).

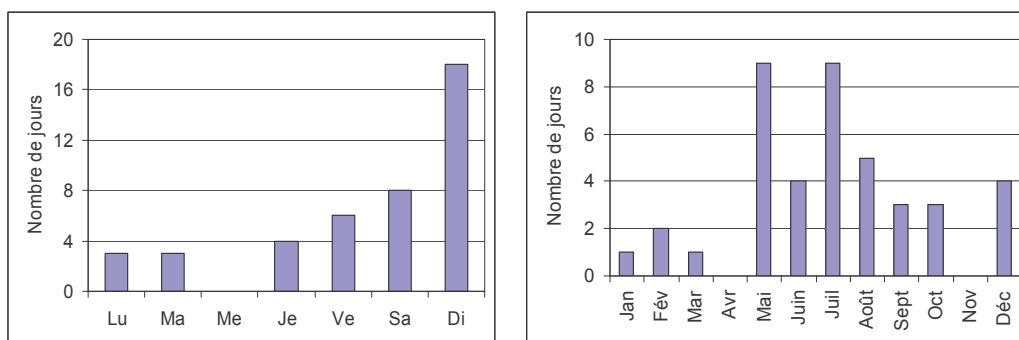


Figure 25 : Fréquence des jours pour lesquels l'estimation des concentrations de NO<sub>2</sub> est non satisfaisante en fonction du jour de la semaine et du mois, pour l'ensemble des stations fixes, calculée sur les années 2002-2003 et 2005-2006

On peut comparer les profils hebdomadaires de concentrations en NO<sub>2</sub> et O<sub>3</sub> mesurées par les stations fixes pour l'ensemble des jours et pour les JNS (figure ci-dessous). Alors que les profils moyens de NO<sub>2</sub> et d'ozone sont cohérents avec une baisse du trafic en fin de semaine et une hausse de l'ozone, le profil moyen observé pour les JNS traduit un comportement différent des deux polluants : les concentrations en NO<sub>2</sub> sont particulièrement faibles et celles d'ozone plus élevées et stables.

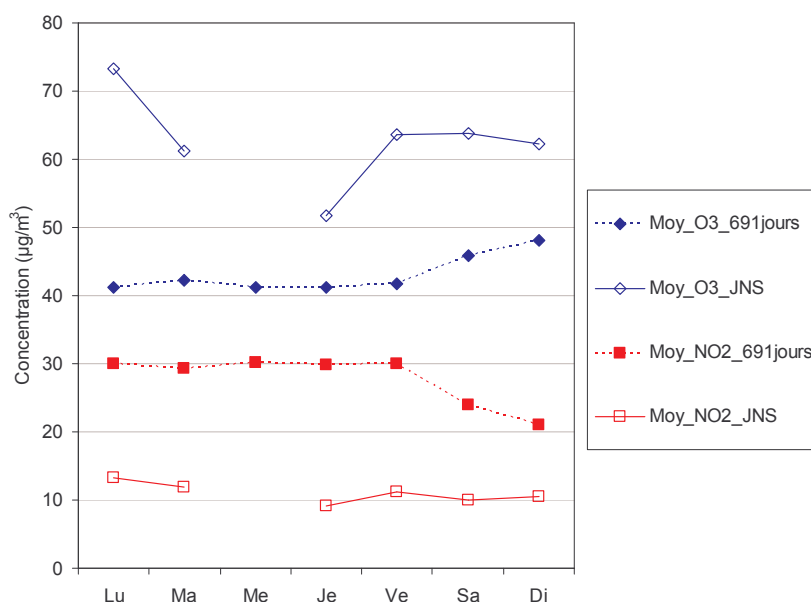


Figure 26 : Concentrations moyennes de NO<sub>2</sub> et de O<sub>3</sub> pour l'ensemble des jours et pour les JNS

## **5. CONCLUSIONS SUR LA VALIDATION DE LA METHODE**

La méthode de cartographie automatique journalière de NO<sub>2</sub> a été appliquée sur deux cas d'étude:

- la ville de Lille et sa périphérie
- et la ville de Reims.

Cette méthode a dans un premier temps été comparée à la méthode de cokrigeage temporel développée par le LCSQA en 2004. Sa performance a été montrée.

La méthode de la validation croisée a été utilisée pour estimer l'erreur sur le calcul de la concentration de NO<sub>2</sub> aux stations fixes disponibles dans les deux agglomérations. Elle a été calculée comme la différence entre la moyenne journalière des concentrations mesurées et la concentration journalière estimée par validation croisée.

La validation de la méthode a été effectuée relativement aux objectifs de qualité de la directive 1999/30/CE pour le calcul de valeurs journalières par modélisation. Celle-ci exige que l'erreur relative (définie comme l'erreur divisée par la valeur mesurée) soit au maximum égale à 50%.

Pour une estimation journalière donnée, deux critères pour décider de la validité de la méthode ont été étudiés, le premier calcule l'erreur relative sur toutes les stations fixes disponibles et considère la valeur moyenne, le seconde critère prend en compte plutôt la valeur maximale des erreurs relatives.

Le premier critère permet de sélectionner 99% d'estimations dans l'agglomération lilloise et environ 94% d'estimations dans l'agglomération rémoise, en revanche le seconde critère plus restrictif ne permet la validation que d'environ 70% des estimations journalières dans les deux agglomérations.

Pour les deux critères, les niveaux de concentration de NO<sub>2</sub> des jours où les estimations ne remplissent pas l'objectif de qualité sont plus faibles que les autres jours ; l'erreur commise est alors une surestimation de la concentration en NO<sub>2</sub>.

Ces jours particuliers sont caractérisés par :

- des concentrations en NO<sub>2</sub> relativement faibles ;
- des concentrations en O<sub>3</sub> plutôt élevées ;
- des conditions météorologiques anticycloniques, printanières ou estivales, qui correspondent à des situations dans lesquels la contribution du chauffage urbaine est faible ;
- des jours de week-ends (samedi ou dimanche), qui correspondent à des situations dans lesquels il y a moins d'émissions de NO<sub>2</sub> dues au trafic routier.

Nous en avons déduit que deux variables influaient principalement sur la qualité de la méthode :

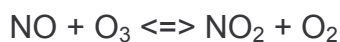
- d'une part, les teneurs en NO<sub>2</sub> ;
- d'autre part, les teneurs d'ozone.

Des teneurs en NO<sub>2</sub> trop faibles ne conviennent pas pour l'application de la méthode. En effet, les analyseurs automatiques, sur lesquels se base la méthode, ne sont pas étalonnés pour les faibles teneurs (< 5 µg/m<sup>3</sup>), mais plutôt pour les valeurs hautes, proches de la valeur-limite annuelle (40 µg/m<sup>3</sup>).

Par ailleurs, les corrélations temporelles, entre analyseurs automatiques et stations virtuelles, ont été établies pour des valeurs supérieures à 10 µg/m<sup>3</sup>.

Il apparaît donc justifié de ne pas appliquer la méthode lorsque les concentrations en NO<sub>2</sub> des stations fixes sont trop faibles, c'est-à-dire < 10 µg/m<sup>3</sup>.

Des teneurs en O<sub>3</sub> trop élevées ne conviennent pas non plus pour l'application de la méthode. On connaît en effet la réaction globale entre l'ozone et les oxydes d'azote (ci-dessous), qui montre que O<sub>3</sub> et NO<sub>2</sub> sont deux espèces antagonistes, l'une tendant à diminuer lorsque l'autre augmente :



Or l'ozone se forme en période d'ensoleillement (printemps - été), avec des teneurs plus importantes en dehors des villes (plutôt en site périurbain qu'en site urbain), notamment les jours de la semaine où le trafic est moins important (le week-end).

Ces deux variables étant liées, un critère de validité portant sur une seule d'entre elles suffit. Il est donc conseillé de ne pas considérer les estimations journalières lorsque les niveaux mesurés aux stations fixes sont inférieurs de 10 µg/m<sup>3</sup> en moyenne journalière.

En conclusion, cette étude montre que la méthode de génération automatique de cartographies journalières des concentrations en NO<sub>2</sub> proposée est performante au regard de l'objectif de qualité de la directive 1999/30/CE, à condition d'être dans le domaine de validité de la méthode.

La méthode a été appliquée avec succès à deux zones urbaines du nord et de l'est de la France : Lille et Reims. La transposition de la méthode à une autre zone de surveillance, située plus au sud, devrait être conditionnée par la fréquence des épisodes d'ozone. Dans ce cas, la variable « ozone » pourrait être prise en compte dans les corrélations temporelles entre stations fixes et stations virtuelles, et dans la procédure géostatistique de cartographie.



## **6. APPLICATION DE LA METHODE D'IMBRICATION D'ECHELLE**

### **6.1 BREVE DESCRIPTION DE LA METHODE**

L'objectif de la méthode est d'inclure dans une même cartographie de NO<sub>2</sub>, la pollution atmosphérique de fond et celle de proximité (routière, industrielle).

La méthodologie proposée par le LCSQA en 2005, et qui a été appliquée dans un contexte régional (Région Centre) repose d'une part sur l'existence des données de concentrations de fond et de proximité mesurées par tubes passifs et d'autre part sur une analyse approfondie des corrélations entre ces variables et différentes variables secondaires potentiellement liées aux concentrations, par exemple la densité de population, l'occupation de sol ou les émissions.

Ces variables sont représentées avec un SIG sur différents rayons autour des points de mesure : 500m, 1Km, 2Km, 5Km et 10Km.

Dans le cas d'étude présenté en 2005, les données de proximité provenaient des campagnes mensuelles de mesure du NO<sub>2</sub> au voisinage de principaux nœuds routiers de la région Centre.

Dans un premier temps une cartographie de la moyenne saisonnière de la pollution de fond avec introduction de variables auxiliaires est établie à l'aide des méthodes géostatistiques.

Dans la région Centre, la taille de la maille d'estimation était de 5Km et la méthode qui a présenté les meilleurs résultats en validation croisée a été le krigeage des résidus [6].

Dans un second temps, sur tous les points de proximité on calcule la différence entre la concentration de fond estimée et la mesure; puis une corrélation linéaire avec les variables calculées par SIG est recherchée.

Dans le cas de la région Centre, la densité d'émissions dans un rayon de 1Km a été retenue comme variable explicative de cette différence (écart aux sites de proximité entre les concentrations mesurées et la concentration de fond estimée par krigeage de résidus).

Il est proposé ensuite de raffiner la maille d'estimation dans les zones comprises entre 0 et 200 mètres autour des principales routes.

Sur cette nouvelle maille d'estimation, on effectue trois opérations, à savoir :

1. On calcule la pollution de fond avec les modèles et variables auxiliaires similaires à ceux employés lors de la cartographie de la pollution de fond.
2. La différence entre les deux types de pollution est ensuite spatialisée, grâce au modèle de corrélation linéaire établi.

Par exemple dans la région Centre pour calculer l'écart entre la pollution de fond et la pollution de proximité dans une maille de 250m de coté, l'expression suivante a été employée:

$$\text{« Ecart (NO}_2\text{ Fond} - \text{NO}_2\text{ Proximité) = (0.000356 X Densité d'émissions dans un rayon de 1Km) + 12.87 »}$$

A noter que les variables auxiliaires doivent être préalablement calculées par SIG dans la nouvelle maille d'estimation. Ceci peut éventuellement être un inconvénient pour la mise en place de la méthode, car à cause de l'augmentation du nombre de mailles la durée des calculs SIG peut être importante.

3. La troisième opération consiste à additionner les deux valeurs obtenues, le résultat final est la pollution de NO<sub>2</sub> dans les zones de proximité routière.

Pour la représentation finale on propose de superposer les deux grilles d'estimation dans un SIG (les grilles de 5Km et de 250m dans le cas de la région Centre).

## **6.2 ADAPTATION AU CONTEXTE DE LA CARTOGRAPHIE AUTOMATIQUE JOURNALIERE**

L'application de la méthode d'imbrication d'échelle dans le contexte du CKJA suppose l'automatisation du calcul de la différence entre concentrations mesurées et concentrations de fond estimées dans les sites de proximité et de la modélisation de cette différence à l'aide des différentes variables secondaires potentiellement liées aux concentrations.

De façon à simplifier la mise en œuvre de la méthode une autre solution a été cherchée. L'objectif est d'éviter les deux démarches citées précédemment et d'utiliser le krigeage pour calculer directement les concentrations de proximité.

L'analyse des données de l'agglomération lilloise a permis de dégager cette nouvelle méthodologie pour produire directement une cartographie automatique journalière qui prend en compte la pollution de proximité.

Pour faciliter la compréhension de la démarche, on présente d'abord les principales étapes de la nouvelle procédure. L'application effectuée sur l'agglomération lilloise illustrée des principaux résultats et graphiques est fournie par la suite.

### **6.2.1 METHODOLOGIE PROPOSEE**

La procédure proposée se décompose en quatre étapes :

- Raffinement de la maille d'estimation : la taille de la nouvelle maille dépendra de l'étendue du domaine d'étude, par exemple dans la région Centre une maille de 250m avait été définie, pour l'agglomération lilloise une maille plus fine de 50m est utilisée (cf. 6.2.2.2).
- Sélection des principales routes traversant la zone d'étude : la base GEOROUTE de l'IGN peut être utilisée pour effectuer cette sélection [8].
- Calcul d'une nouvelle variable auxiliaire qui tiendra compte de la pollution de fond et de la pollution de proximité dans les zones traversées par les principales routes: une façon de calculer cette nouvelle variable est présentée par la suite (cf. 6.2.2.2).
- Estimation par cokrigeage colocalisé des concentrations journalières : on emploie la même démarche que pour l'estimation de la pollution de fond (mêmes modèles de variogramme), la seule différence est la variable auxiliaire, laquelle représente maintenant les deux types de pollution.

## 6.2.2 APPLICATION DE LA METHODOLOGIE A L'AGGLOMERATION LILLOISE

### 6.2.2.1 SELECTION DE PRINCIPALES ROUTES

La sélection des principales routes dans l'agglomération lilloise a été effectuée à l'aide du thème « RESEAUX ROUTIERS » de la base GEOROUTE V2 Produit Adresse de l'IGN, Edition 1.6 de septembre 2001.

Ce thème est composé des « tronçons de routes » qui sont définis comme « *toute portion de l'espace routier public sur laquelle on peut circuler, un tronçon de route est la plus grande division de route entre deux nœuds figurant dans la base GEOROUTE* ».

Afin de sélectionner les routes les plus importantes le classement fonctionnel est employé (champ « CL\_FONC »), celle-ci « *définit des niveaux en fonction de la capacité des voies à acheminer les véhicules à vitesse élevée* ». La classification est composée de 6 types de caractères : *Principal, Primaire, Secondaire, Tertiaire, Desserte et Sans Objet*. Les deux premiers caractères permettent d'identifier les principales routes de l'agglomération.

La Figure 27 montre les routes ainsi sélectionnées dans la région Lilloise, l'attribut « NB\_VOIES » est également affiché, il s'agit du nombre des voies de la chaussée. Dans la figure on observe que quelques tronçons de routes peuvent présenter jusqu'à 5 voies.

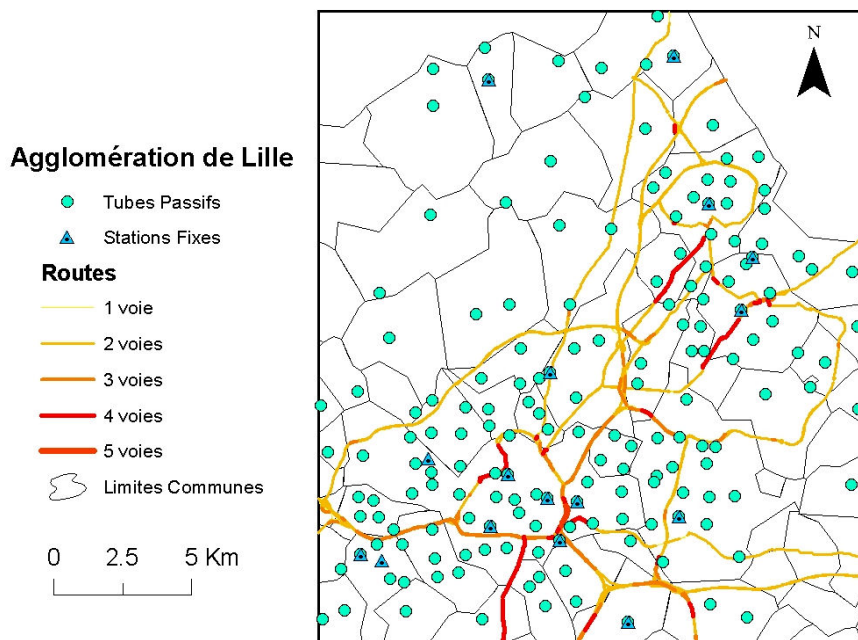


Figure 27 : Carte des principales routes de l'agglomération lilloise

## 6.2.2.2 RAFFINEMENT DE LA MAILLE ET CALCUL DE LA VARIABLE AUXILIAIRE

La taille de la maille d'estimation des concentrations de fond présentées dans le paragraphe 4.1 est de 250m. Afin de calculer la variable auxiliaire qui prend en compte la pollution de proximité, une maille plus fine de 50m a été définie (cf. Figure 28).

La variable auxiliaire représentant la pollution de fond selon la méthode établie en 2005 [4] (appelée F1), est égale à la somme de deux variables normées : la densité d'émissions et la densité de population (cf. Annexe 4 : Figure 68, page 117 et Figure 69, page 118). Cette fonction est recalculée dans la maille de 50m.

Pour la recherche de la variable auxiliaire liée aux concentrations de proximité, les mailles se trouvant à une distance de 150m autour des principales routes ont été sélectionnées. Cette fonction (appelé F2) prend en compte deux variables liées aux routes : la distance du nœud de la maille à la route la plus proche et le nombre de voies de celle-ci.

La fonction F2 a été construite de façon empirique, avec les hypothèses suivantes :

- Relation linéaire entre F2 et le nombre des voies : pour une distance à la route nulle, la fonction auxiliaire construite vaut 1 pour les routes de 5 voies, 0.9 pour les routes de 4 voies, etc. (diminution de 10% par voie).
- Relation linéaire entre F2 et la distance à la route : pour un tronçon de route de 5 voies, la fonction auxiliaire construit vaut 1 pour une distance à la route nulle, 0.9 pour une distance de 50m, 0.8 pour une distance de 100m, et 0.7 pour distance de 150m (diminution de 10% par 50m d'éloignement à la route).

Si l'on applique ces deux hypothèses sur les nœuds se trouvant à une distance maximale de 150m autour des principales routes, on obtient les valeurs de la fonction F2 (Tableau 29).

F2	Nb Voies = 1	Nb Voies = 2	Nb Voies = 3	Nb Voies = 4	Nb Voies = 5
D = 150 m	0.42	0.49	0.56	0.63	0.7
D = 100 m	0.48	0.56	0.64	0.72	0.8
D = 50 m	0.54	0.63	0.72	0.81	0.9
D = 0 m	0.6	0.7	0.8	0.9	1

Tableau 29 : Construction de la fonction F2 liée à la pollution de proximité

Voici la formule qui résume la relation entre la distance à la route, le nombre de voies de celle-ci et la fonction F2 :

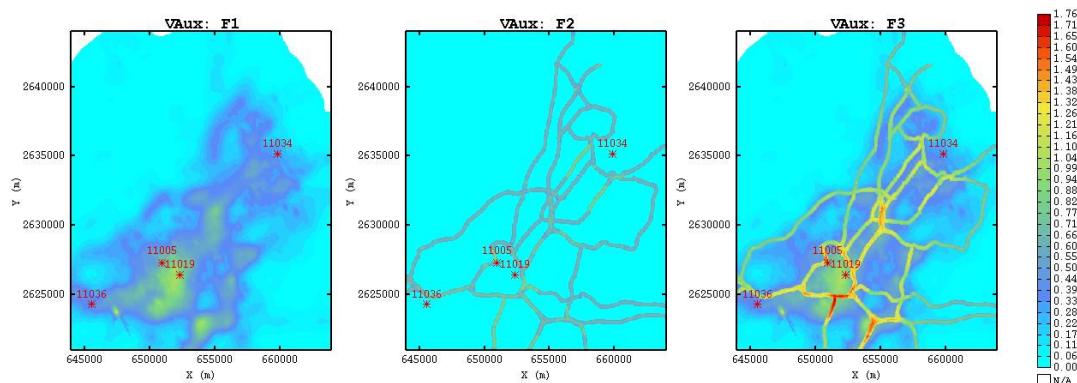
$$F2 = \left[ 0.6 + \left( \frac{NbVoies - 1}{10} \right) \right] * \left[ 0.7 + \left( \frac{150 - d}{500} \right) \right]$$

$$NbVoies = 1, 2, 3, 4 \text{ ou } 5$$

$$d = 0, 50, 100, \text{ ou } 150m$$

Pour affiner le modèle il peut être envisagé d'effectuer de campagnes par tubes passifs localisés le long de transects perpendiculaires aux principales routes de l'agglomération, une étude approfondie de ces transects permettrait de trouver un modèle plus détaillé de la décroissance des niveaux de pollution.

Finalement les deux fonctions F1 et F2 sont additionnées pour obtenir une fonction auxiliaire qui tient compte des possibles sources de la pollution de fond (émissions et densité de population) et celles de la pollution de proximité (distance à la route et nombre de voies de celle-ci). La Figure 28 présente, les trois variables auxiliaires calculées sur la maille de 50m.



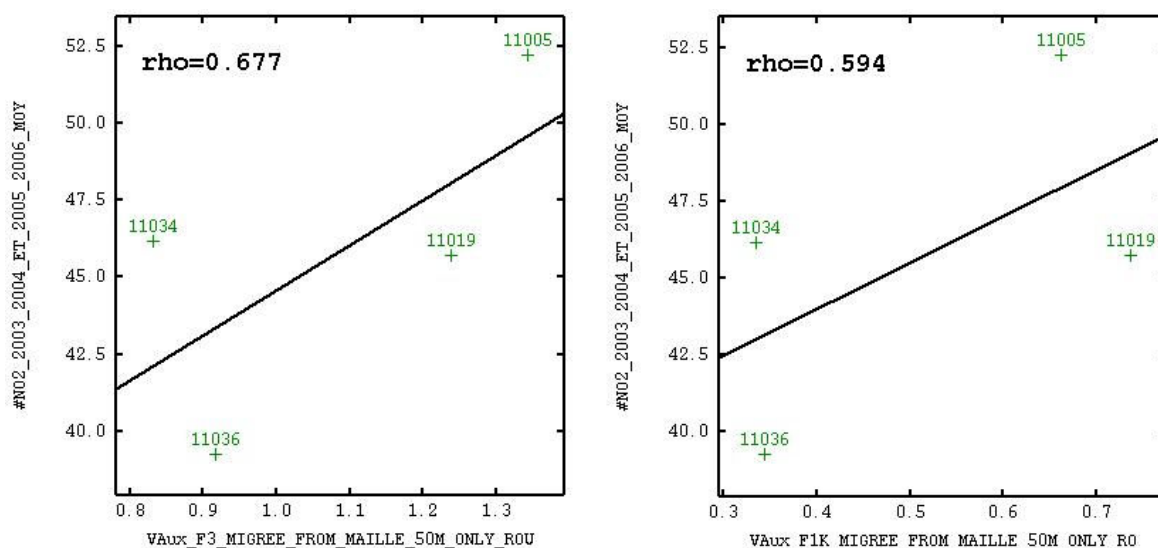
A gauche : Fonction F1 (représentant les sources de pollution de fond) recalculée dans la maille de 250m (somme des deux variables normées : densité de population et de densité d'émissions).

Au milieu : Fonction F2 (représentant les sources de pollution de proximité) calculée dans la zone de 150m autour des principales routes (elle dépend de la distance à la route la plus proche et du nombre de voies de celle-ci).

A droite : Fonction F3 (représentant les sources de pollution de fond et de proximité) somme des fonctions F1 et F2.

*Figure 28 : Carte des fonctions représentant les paramètres définissant les sources de la pollution du NO<sub>2</sub> et la localisation géographique des 4 stations fixes de proximité pour l'agglomération de Lille*

La Figure 29 montre les nuages de corrélation entre la moyenne des concentrations mesurées aux quatre stations fixes de proximité (années 2003-2004 et 2005-2006), et les valeurs des fonctions F1 et F3 du nœud le plus proche à chaque station (se trouvant à une distance maximale de 150m autour des principales routes). Pour les stations de trafic, la corrélation s'améliore lorsqu'on prend en compte la fonction F3 (elle passe de 0.594 à 0.677).



Axe Y: Moyenne des concentrations mesurées aux stations fixes du 1<sup>er</sup> avril 2003 au 30 mars 2004 et du 1<sup>er</sup> avril 2005 au 30 mars 2006

Axe X: Fonctions auxiliaires :

A gauche : migration de la valeur de la fonction auxiliaire F3 du nœud se trouvant à une distance maximale de 150m autour des principales routes (représentant les sources de pollution de proximité et de fond)

A droite : migration de la valeur de la fonction auxiliaire F1 (représentant les sources de pollution de fond)

**Figure 29 : Nuage de corrélation entre les moyennes des concentrations mesurées et estimées aux 4 stations localisées en situation de proximité**

## 6.2.2.3 ANALYSES DES RESULTATS ET CONCLUSIONS

L'analyse porte sur la comparaison entre les mesures et les estimations des nœuds les plus proches aux stations fixes. Cette comparaison permettra de mesurer l'amélioration des estimations dans les zones de proximité.

Pour cela, deux estimations ont été effectuées sur la maille de 50m, la première prend en compte la fonction F1 et la seconde la fonction F3. Les mêmes modèles que ceux utilisés dans la section 4.1.2, et présentés dans l'Annexe 5 ont été conservés pour les deux estimations.

Les estimations ont été appliquées pour estimer les niveaux de pollution journaliers sur une période de deux ans : du 1<sup>er</sup> avril 2003 au 30 mars 2004 et du 1<sup>er</sup> avril 2005 au 30 Mars 2006. L'Annexe 6 montre un exemple des estimations effectuées avec les fonctions F1 et F3 sur deux jours : le 28 mai 2003 (saison estivale) et le 11 décembre 2005 (saison hivernale).

Les moyennes (sur une période de deux années) des estimations aux quatre stations de proximité sont ensuite obtenues par migration (à partir du nœud le plus proche à chaque station de trafic).

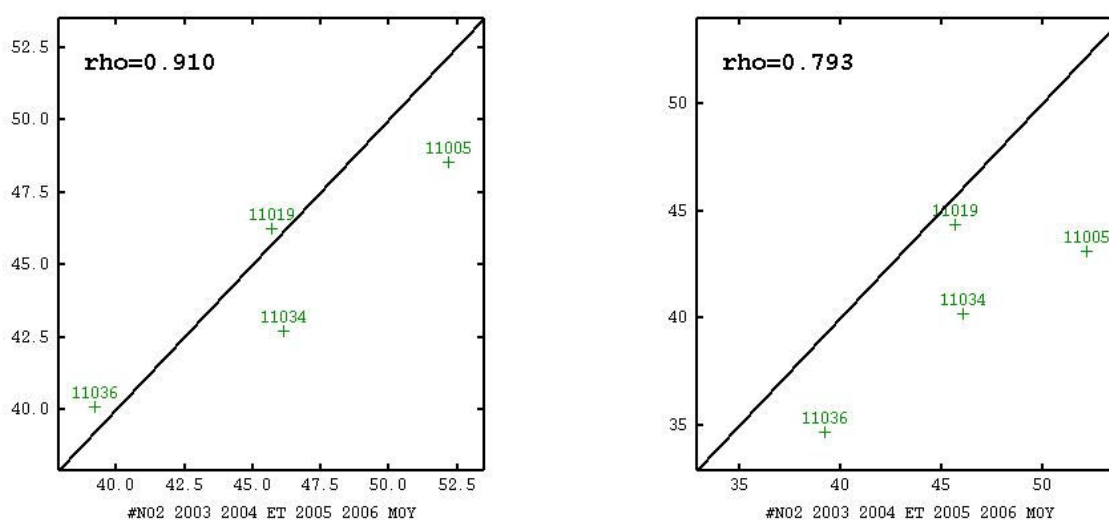
Le Tableau 30 et la Figure 30 montrent les statistiques des erreurs calculées par station entre les estimations et la moyenne des mesures sur les deux années étudiées d'une part et les nuages de corrélation mesure - estimation d'autre part. Les statistiques générales calculées sur les quatre stations sont montrées dans le Tableau 31.

STATISTIQUE	LILLE PASTEUR	HAUBOURDIN POMPIERS	ROUBAIX SERRES	LILLE LIBERTE
CODE STATION	11019	11036	11034	11005
Moyenne des mesures : 01/04/2003 - 30/03/2004 et 01/04/2005 - 30/03/2006	45.68	39.21	46.11	52.19
Moyenne des estimations avec F1 migrées	44.32	34.67	40.16	43.04
Moyennes des estimations avec F3 migrées	46.19	40.05	42.67	48.51
Erreur d'estimation avec F1	1.36	4.54	5.95	9.15
Erreur d'estimation avec F3	-0.51	-0.84	3.44	3.68
Erreur carrée d'estimation avec F1	1.85	20.61	35.4	83.72
Erreur carrée d'estimation avec F3	0.26	0.706	11.83	13.54

*Tableau 30 : Statistiques par station trafic des erreurs d'estimation de la moyenne journalière sur deux années de mesures*

Statistiques sur les 4 Stations Fixes Trafic	Estimation avec F1	Estimation avec F3
Moyenne de l'Erreur	5.25	1.44
Moyenne de l'Erreur carrée	35.40	6.59
Racine de la moyenne de l'erreur carré	5.9	2.57
Variance de l'erreur	7.83	4.50

Tableau 31 : Statistiques générales des erreurs d'estimation de la moyenne sur deux années de mesures aux stations fixes de trafic



Axe X : Moyenne des concentrations mesurées aux stations fixes du 1<sup>er</sup> avril 2003 au 30 mars 2004 et du 1<sup>er</sup> avril 2005 au 30 mars 2006

Axe Y : Pour la même période de mesure, moyennes des estimations journalières migrées:

A gauche : estimation avec la fonction auxiliaire F3 (pollution de proximité et de fond)

A droite : estimation avec la fonction auxiliaire F1 (pollution de fond)

Figure 30 : Nuage de corrélation entre les moyennes sur deux années des concentrations mesurées et estimées aux 4 stations localisées en situation de proximité

On constate la nette amélioration des estimations lorsqu'on prend la fonction F3 comme variable auxiliaire : pour les quatre stations, la moyenne des erreurs sur les deux années étudiées diminue de 5.25 à 1.44  $\mu\text{g}/\text{m}^3$ . Les stations Lille Pasteur et Haubourdin Pompiers présentent en moyenne une légère surestimation des concentrations: -0.51 et -0.84  $\mu\text{g}/\text{m}^3$  respectivement.

On peut donc conclure qu'une cartographie journalière des concentrations de dioxyde d'azote peut intégrer de façon satisfaisante la pollution de proximité en incorporant dans une variable auxiliaire des informations disponibles comme la géométrie du réseau routière, la distance aux routes et le nombre des voies.



*Néanmoins la difficulté réside dans la façon dont ces nouvelles variables doivent être incorporées à la variable auxiliaire : en l'absence de mesures permettant d'ajuster des modèles plus complexes, un modèle simple de décroissance linéaire des concentrations de NO<sub>2</sub> en fonction de la distance à la route et du nombre des voies, a été utilisé dans ce cas d'étude; cependant pour modéliser de façon détaillée le comportement du dioxyde d'azote près des sources de pollution, des campagnes d'échantillonnage par tubes passifs localisés le long de transects perpendiculaires aux principales routes de l'agglomération seraient nécessaires.*

## **7. RECOMMANDATIONS PRATIQUES POUR LA MISE EN ŒUVRE DE LA METHODE DE CARTOGRAPHIE JOURNALIERE AUTOMATIQUE DE DIOXYDE D'AZOTE**

1. L'application de la méthode de cartographie automatique de NO<sub>2</sub> repose sur l'existence d'un historique de campagnes par échantillonneurs passifs et sur la présence de plusieurs stations fixes mesurant le dioxyde d'azote sur une agglomération. Pour établir les corrélations temporelles entre tubes et capteurs, il faut disposer d'au moins six campagnes de tubes (trois campagnes par saison : hiver / été) et de deux stations fixes de typologie différente.
2. Il est nécessaire d'effectuer des analyses géostatistiques les plus détaillées possibles sur les campagnes par échantillonneurs disponibles, dans les deux cas d'étude ces analyses ont porté sur les moyennes saisonnières et annuelles ; les modèles de variogrammes, les variables auxiliaires et la méthode d'estimation ont pu être définis grâce à l'analyse de ces campagnes.
3. Comme, le plus souvent, le nombre de points se réduit lors du choix des stations virtuelles, la variable auxiliaire gagne en influence sur les résultats des estimations lors de l'application du cokrigage journalier automatique. En conséquence il est très important de bien choisir la variable auxiliaire et de ne pas oublier qu'elle peut prendre en compte des facteurs tels que la météo ou le trafic, et qu'elle peut évoluer en fonction de l'actualisation des bases des données existantes comme par exemple la mise à jour de l'inventaire d'émissions ou la prise en compte de la mobilité de la population en période de vacances, des week-ends, etc.
4. Pour sélectionner les stations virtuelles et les stations de référence, il est conseillé d'utiliser la macro EXCEL expliquée en annexe et disponible sur le site WEB du LCSQA ([www.lcsqa.org](http://www.lcsqa.org)).

Cependant ce programme prend en compte deux critères de sélection qui peuvent être modifiés, le premier concerne le coefficient de corrélation lequel doit être supérieur à 0.8 ; le second ne permet de retenir que les couples station/tube comportant des points dans les trois domaines de concentrations définies dans la section 3.1.2.2. Ce dernier critère pourrait par exemple être assoupli dans certaines zones où les stations mesurent continuellement des concentrations très faibles ou très fortes. Si l'on prend le cas de l'agglomération de Reims où aucun tube passif n'a pas été couplé à la station 14004 localisée en centre ville et mesurant les plus fortes concentrations, on s'aperçoit que la modification du critère aurait permis la sélection des couples et peut être l'amélioration des estimations dans cette zone.

Une autre modification consisterait à conserver des stations virtuelles bien corrélées aux stations fixes, mais seulement pour certaines gammes de valeurs, et à les écarter du calcul lorsque la station fixe fournit des valeurs trop faibles. Ceci afin de prendre en compte le fait que dans cette étude il a été conseillé de ne pas employer la méthode de cartographie automatique lorsque les concentrations aux stations fixes sont inférieures de 10 µg/m<sup>3</sup> en moyenne journalière.

D'autres critères peuvent aussi être imaginés, par exemple :

- Imposer que les couples stations fixes de référence - stations virtuelles correspondent à un même type d'environnement urbain.
- Limiter la séparation entre stations fixes de référence et stations virtuelles de façon à qu'elles ne soient pas trop distantes.
- S'assurer que la répartition de stations virtuelles soit uniforme au sein de l'agglomération.
- Choisir manuellement les stations de référence, dans le cas par exemple où le nombre de stations est limité, où lorsqu'une station fixe qui est censée être plus représentative d'une situation particulière n'est pas choisie par le programme ou quand on veut exploiter les mesures fournies par un moyen mobile (camion laboratoire).

Si l'on veut prendre en compte des critères supplémentaires, une adaptation de la méthode de corrélations temporelles et de la macro Excel serait nécessaire.

5. La cartographie a été facilitée par le développement du programme qui exécute le logiciel Isatis en mode automatique (batch). La réalisation automatique des interpolations n'est cependant possible que si la version Windows du logiciel Isatis (v. 6.0) est installée sur la machine.

L'interface produit les fichiers « IJNL » qui exécutent le logiciel Isatis et les fichiers sources des données avec l'entête nécessaire pour leur importation sur Isatis. Il est donc possible, pour les utilisateurs qui possèdent d'autres versions d'Isatis, d'éditer ces fichiers et d'adapter les instructions. Ces changements demandent néanmoins, un niveau avancé des connaissances sur le langage de programmation en mode batch d'Isatis.

Le package d'installation du programme et le code source en langage Visual Basic sont disponibles sur le site WEB du LCSQA ([www.lcsqa.org](http://www.lcsqa.org)).

6. Le programme ne prend en compte qu'une méthode d'estimation, il s'agit du cokrigeage colocalisé, car elle a donné les meilleurs résultats lors des analyses géostatistiques des campagnes par tubes. Cependant, il est possible qu'une autre méthode d'estimation comme le krigeage avec dérive externe ou le krigeage des résidus soit plus précise sur d'autres agglomérations. Dans ce cas il est conseillé d'utiliser directement le logiciel Isatis, ou pour les utilisateurs experts d'adapter le code source du programme.
7. Dans l'Annexe 15 et suivants, sont détaillées les principales statistiques par station fixe et de l'ensemble des stations des résultats des interpolations. Dans cette étude deux critères principalement ont été proposés pour déterminer sur la validité d'une estimation journalière, le premier établi que la valeur maximale des erreurs relatives aux stations fixes doit être inférieure à 50%, le seconde prendre en compte de préférence la moyenne des erreurs. Cependant l'utilisateur pourra choisir d'autres critères plus ou moins restrictifs, en fonction de l'utilisation finale qui sera faite de la carte obtenue.

## **8. BIBLIOGRAPHIE**

1. Directive 1999/30/CE du Conseil 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, J.O. des Communautés Européennes 163 du 29 juin 1999.
2. Arrêté du 17 mars 2003 relatif aux modalités de surveillance de la qualité de l'air et à l'information du public, J.O. 167 du 22 juillet 2003.
3. Perdrix E., Wroblewski A., Fourché B. et Plaisance H. Cartographie des concentrations en dioxyde d'azote à partir des mesures en stations fixes. LCSQA/EMD - Etude 11 - 2004,
4. Cardenas G. et Perdrix E., « Adaptation de méthodes géostatistiques à la cartographie automatique de NO<sub>2</sub> ». LCSQA/INERIS/EMD - Etude 31 - 2005.
5. Cardenas G. et Malherbe L., « Géostatistique et prise en compte de l'aspect temporel ». LCSQA/INERIS - Etude 12 - 2004.
6. Malherbe L. et Cardenas G., « Application des méthodes géostatistiques pour l'exploitation conjointe des mesures de fond et de proximité ». LCSQA/INERIS – Novembre 2005.
7. ATMO Champagne Ardenne, « Etude de la répartition spatiale du dioxyde d'azote sur l'agglomération rémoise à l'aide de tubes à diffusion passive : Rapport sur les campagnes de mesures effectuées du 23/12/02 au 17/2/03 et du 28/4/03 au 23/6/033 ». Réf : ATMO CA, étude Rs 03-03-TD-ED/EC.
8. « RESEAUX ROUTIERS» de la base GEOROUTE V2 Produit Adresse de l'IGN, Edition 1.6 de septembre 2001.

## 9. LISTE DES ANNEXES

Repère	Désignation	Nombre de pages
Annexe 1	: Document de référence de l'étude	2
Annexe 2	: Guide d'utilisation des outils informatiques	39
Annexe 3	: Format des fichiers outil Excel	4
Annexe 4	: Application de la méthode de cokrigeage temporel pour l'estimation de la moyenne annuelle de la concentration de NO2 dans l'agglomération lilloise	9
Annexe 5	: Modèles choisis pour la méthode de la cartographie automatique journalière de NO2 dans l'agglomération lilloise	2
Annexe 6	: Exemple de deux estimations avec les fonctions F1 et F3 réalisées par CKJA dans l'agglomération lilloise	1
Annexe 7	: Principales figures et statistiques de l'application de la méthode de cokrigeage temporel pour l'estimation de la moyenne annuelle de la concentration de NO2 dans l'agglomération rémoise	7
Annexe 8	: Exemple de fichier CSV avec entête (réinitialisation d'une étude isatis)	1
Annexe 9	: Exemple de fichier CSV avec entête (importation dans une étude Isatis existante)	1
Annexe 10	: Exemple des premières lignes d'un fichier BATCH_INIT.IJNL	2
Annexe 11	: Exemple des premières lignes d'un fichier BATCH_RUN.IJNL	2
Annexe 12	: Exemple des premières lignes d'un fichier ASC (« grille SIG »)	1
Annexe 13	: Exemple des premières lignes d'un fichier contenant les messages d'isatis (« _isatis.txt »)	1
Annexe 14	: Exemple des premières lignes d'un fichier contenant les temps de calcul d'Isatis (« _isatis_cpu_timer.txt »)	1
Annexe 15	: Formules pour le calcul des statistiques produites par Isatis	1
Annexe 16	: Description détaillée des champs du fichier: «..._STATS_ALL_CKCL.CSV »	1
Annexe 17	: Description détaillée des champs du fichier: «..._STATS_PAR_SF_CKCL.CSV »	1

Annexe 1 : Document de référence de l'étude

## **TRAITEMENTS NUMERIQUES**

### **Etude 25 : Méthode de surveillance des concentrations de NO<sub>2</sub> : cartographie automatique à partir de stations fixes et prise en compte de la proximité.**

**INERIS/EMD**

#### **CONTEXTE ET OBJECTIF**

Le but de cette fiche est de finaliser la mise au point d'une méthode de surveillance des concentrations de fond du dioxyde d'azote, avec une résolution spatiale fine et prise en compte de la pollution de proximité.

En effet dans le cadre de l'application de l'arrêté du 17/03/2003 Art.9-III, les AASQA doivent réaliser et publier régulièrement une information indiquant les niveaux de polluants dans chaque zone surveillée. La représentation cartographique de l'information permet une présentation claire et pertinente des situations rencontrées.

En ce qui concerne le NO<sub>2</sub>, la surveillance des zones situées hors de proximité des sources principales (industrie, voies routières) peut souvent s'exercer au moyen de méthodes alternatives, telles que la modélisation et l'estimation objective.

#### **TRAVAUX ANTERIEURS ET EN COURS**

Depuis deux ans, le LCSQA travaille sur une méthode originale de production de cartes de répartition spatiale des concentrations en NO<sub>2</sub>, à partir des mesures de stations fixes et de campagnes de mesure par échantillonnage passif.

Elle repose sur la mise en évidence de relations statistiques entre les mesures effectuées en des sites passifs (stations virtuelles) et celles effectuées par une station automatique (station réelle). A partir de ces relations statistiques simples et d'une actualisation des mesures de la station automatique, on génère des cartes de répartition du NO<sub>2</sub>, avec une fréquence qui peut être choisie (horaire, journalière, bi-hebdomadaire ou plus).

Par ailleurs, l'étude menée en 2005 sur « l'adaptation des méthodes géostatistiques à la cartographie automatique du NO<sub>2</sub> » a montré que l'utilisation judicieuse des méthodes de cartographie géostatistiques permettait de simplifier, pour l'utilisateur, la génération de cartes de pollution, tout en continuant à respecter les objectifs de qualité requis par la Directive 1999/30/CE.

Le cas concret, sur lequel a été construite la méthode, se rapportait à une zone urbaine et périurbaine du nord de la France. Nous avons montré que la méthode conduisait à la production, de manière simple, de cartes de pollution de fond et qu'elle constituait donc une donnée d'entrée correcte pour élaborer des cartographies de NO<sub>2</sub> réalistes, en constituant une base pour l'intégration d'une information propre aux situations de proximité.

Sur ce dernier point, le LCSQA a également mené une étude en 2005, dédiée à l'imbrication de cartes de pollution de fond et de pollution de proximité, en menant un exercice d'application sur la région Centre, en collaboration avec LIG'AIR.

## TRAVAUX PROPOSES POUR 2006

Le programme proposé en 2006 a pour objet de finaliser les méthodologies développées en 2005 (cartographie automatique de la pollution de fond – imbrication de cartographies de pollution de fond et de proximité) et de tester leur transposition sur de nouvelles régions d'étude :

1) la technique visant à générer automatiquement, à partir des mesures des stations automatiques et des relations statistiques établies entre les sites d'échantillonnage passif et ces stations, des cartographies de fond de NO<sub>2</sub> sera finalisée. Elle sera en particulier appliquée à un autre domaine d'étude que celui sur lequel elle a été mise au point, la région Nord-pas de Calais.

Après cette phase de validation, l'ensemble de la méthode sera décrite dans un guide d'application, accompagné d'un « package » informatique, qui comprendra :

- la procédure sous EXCEL<sup>®</sup> pour mettre en évidence les liens entre sites d'échantillonnage passif et station automatique\*;
- la procédure sous ISATIS<sup>®</sup> pour générer automatiquement les cartes, actualisées à une fréquence choisie, à partir des seules valeurs fournies par une station automatique.

2) la méthode d'imbrication de cartographies de proximité mise au point par le LCSQA en 2005 sera alors appliquée pour produire une cartographie réaliste de la situation. Une procédure de validation de l'ensemble de la démarche sera développée.

3) La possibilité de produire des cartographies imbriquées avec une fréquence temporelle importante ( journalière, horaire...) sera enfin examinée.

Ces points d'étude permettront de finaliser un guide d'application et de recommandations pour la réalisation de cartographies automatiques de NO<sub>2</sub> de fond et d'imbrication de cartographies de proximité. Cette méthode permettra de compléter l'éventail des méthodes mises à disposition des AASQA pour la surveillance du dioxyde d'azote, notamment pour les territoires non couverts par un système de cartographie déterministe, tels que les territoires des départements d'outre-mer.

- *La procédure sous EXCEL<sup>®</sup> permettra aussi de rechercher les liens et similitudes entre stations automatiques.*

## COLLABORATIONS

AASQA, en particulier LIG'AIR

## DUREE DES TRAVAUX

1 an

## UNITES D'ŒUVRE IMPLIQUEES

- EMD : 330 heures ingénieur (0.2 Ing/an)  
Esperanza Perdrix, André Wroblewski, Hervé Plaisance
- INERIS : 600 heures ingénieur (0.4 Ing/an)  
Giovanni Cardenas, Laure Malherbe

## Annexe 2 : Guide d'utilisation des outils informatiques

### 1.1 MACRO DE CORRELATIONS « TUBES – STATIONS »

L'outil informatique permettant la détermination des corrélations entre stations fixes et « stations virtuelles » a été développé en VBA sous Excel. Dans ce chapitre nous proposons un descriptif des paramètres d'entrées et de sorties du logiciel. La manipulation de l'outil est expliquée dans le paragraphe 1.1.5.

#### 1.1.1 CONFIGURATION REQUISE

L'outil informatique, est développé à partir d'une macro Excel. La configuration minimale est un ordinateur fonctionnant sous Windows (98, 2000, XP) équipé du logiciel Microsoft Excel (version minimale 2000).

#### 1.1.2 DESCRIPTIF

L'outil se compose d'un fichier Excel nommé : « applis\_lcsqa.xls »

Cette application Excel permet d'effectuer tous les calculs nécessaires à la détermination :

- des régressions linéaires entre stations fixes et échantillonneurs
- des valeurs de concentration aux stations virtuelles, pour une période donnée.

Dans la suite du rapport nous parlerons :

- de « **phase de détermination** » pour la partie de calcul conduisant à l'établissement des régressions linéaires entre stations fixes et échantillonneurs.
- de « **phase d'exploitation** » pour le calcul des valeurs de concentrations des stations virtuelles.

Cette première partie de la chaîne logicielle, permet d'alimenter le module qui établit, en prenant en compte les stations virtuelles, une cartographie automatique de la zone d'étude avec le logiciel Isatis (cf. paragraphe 1.2).

#### 1.1.3 DONNEES D'ENTREE

##### 1.1.3.1 PHASE DE DETERMINATION

Dans cette partie, le logiciel nécessite deux fichiers d'entrée :

- Un fichier de type Excel « \*.xls » contenant les concentrations des échantillonneurs passifs pour les différentes périodes de mesure (cf. fichier « tubes.xls » en Annexe 3).
- Un fichier de type Excel « \*.xls » contenant les concentrations des stations fixes pour les mêmes périodes (cf. fichier « stations.xls » en Annexe 3).

##### 1.1.3.2 PHASE D'EXPLOITATION

Dans cette partie, le logiciel nécessite deux fichiers d'entrée :



- Un fichier de type Excel « \*.xls » contenant les concentrations des stations fixes pour la période choisie (cf. fichier « NO2\_23\_12\_2001.xls » en Annexe 3).
- Un fichier de type « \*.csv » contenant les coefficients de régression des stations virtuelles (cf. « virtuelles\_param.csv » en Annexe 3).

#### 1.1.4 DONNEES DE SORTIE

##### 1.1.4.1 PHASE DE DETERMINATION

Dans cette partie, le logiciel fournit un fichier de sortie, au format « CSV » nommé « virtuelles\_param.csv ». (cf. Annexe 3)

Ce fichier comprend :

- le nom et les coordonnées des stations fixes,
- le nom et les coordonnées des sites d'échantillonnage passif,
- les coefficients de régression linéaires (pente, ordonnées à l'origine), la variance de l'erreur d'estimation, le coefficient de corrélation (r) pour chaque couple station/ échantillonneur

##### 1.1.4.2 PHASE D'EXPLOITATION

Dans cette partie, le logiciel fournit un fichier de sortie, au format \*.csv nommé en fonction de la période étudiée « NO2\_JJ\_MM\_AAAA.csv ». (Il s'agit du fichier qui sera demandé par l'interface de cartographie, cf. section 1.2.1 page 79, Tableau 32, page 81 et Tableau 33, page 82).

Ce fichier comprend :

- le nom et les coordonnées des stations fixes,
- les concentrations estimées aux stations virtuelles pour la période,
- le nom et les coordonnées des sites d'échantillonnage passif retenus comme stations virtuelles,
- les coefficients de régression linéaire (pente, ordonnées à l'origine), la variance de l'erreur d'estimation, le coefficient de corrélation ( r ) pour chaque couple station/échantillonneur

#### 1.1.5 FONCTIONNEMENT

##### 1.1.5.1 OUVERTURE DU FICHIER

Le fichier comprenant l'ensemble de la macro est le fichier : « applis\_lcsqa.xls »

Ouvrir le fichier sous Excel. Après ouverture la boîte de dialogue suivante apparaît à l'écran (Figure 31).

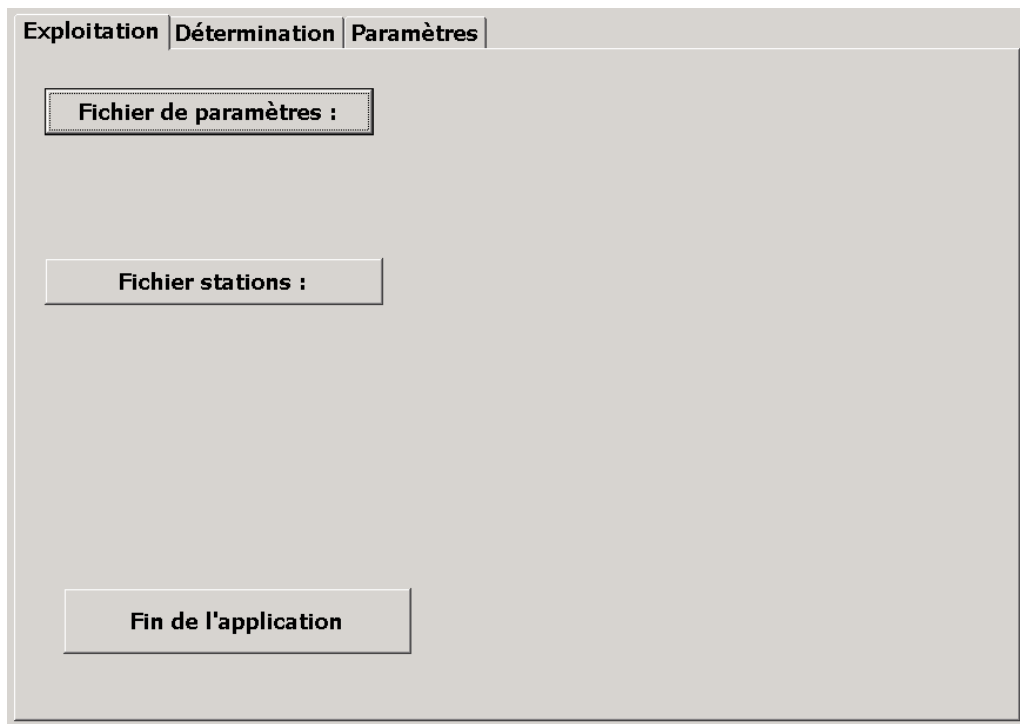


Figure 31 : Feuille de contrôle du fichier « applis\_lcsqa.xls »

La feuille de contrôle comprend trois onglets qui permettent de sélectionner les différentes fonctions du logiciel (Figure 32).



Figure 32 : Onglets de commande

Ces onglets donnent accès à trois menus qui permettent d'effectuer les opérations souhaitées.

- L'onglet « Exploitation » permet d'effectuer le calcul des concentrations sur les stations virtuelles pour une période donnée
- L'onglet « Détermination » permet d'effectuer le calcul des corrélations pour l'étude donnée
- L'onglet « Paramètres » permet d'ajuster les paramètres.

## 1.1.5.2 EXPLOITATION

Après la sélection de l'onglet « Exploitation », l'écran suivant apparaît (Figure 33).

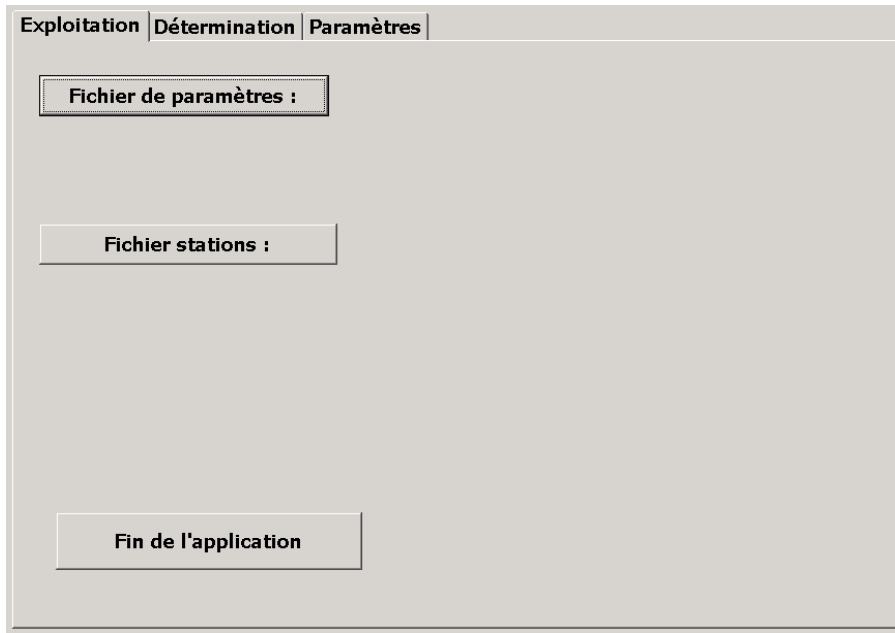


Figure 33 : Feuille de gestion de la partie exploitation

En sélectionnant le bouton « fichier de paramètres : » on accède à une fenêtre de recherche du fichier de paramètres (Figure 34). Le fichier recherché est obligatoirement « virtuelles\_param.csv ».

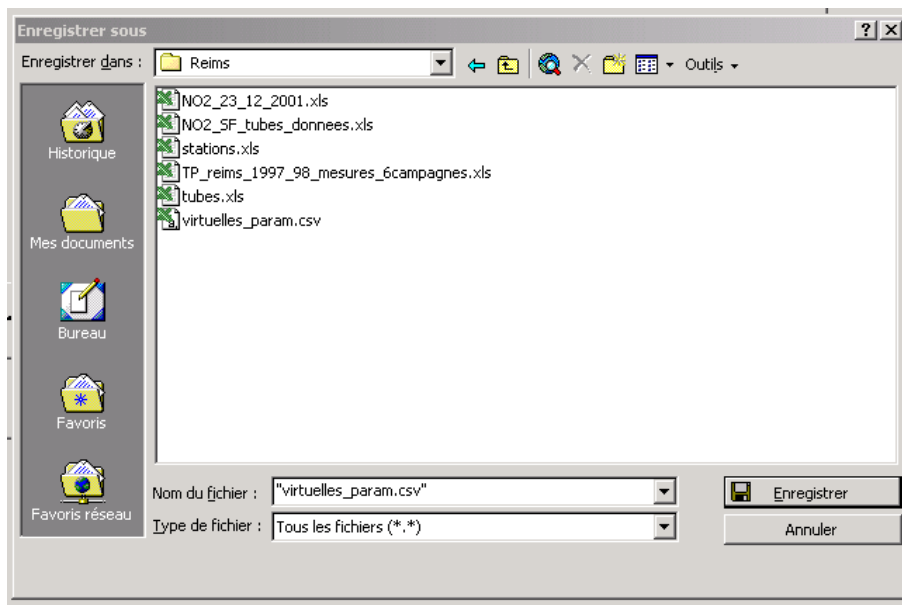


Figure 34 : Recherche du fichier « virtuelles\_param.csv »

Après sélection du fichier, la feuille fournit en retour le chemin d'accès au fichier.

En sélectionnant le bouton « fichier stations : » on accède à une fenêtre de recherche du fichier de mesure aux stations fixes pour la période considérée. Après sélection du fichier la feuille fournit en retour le chemin d'accès au fichier.

Lorsque les deux fichiers sont sélectionnés la feuille se présente ainsi (Figure 35).

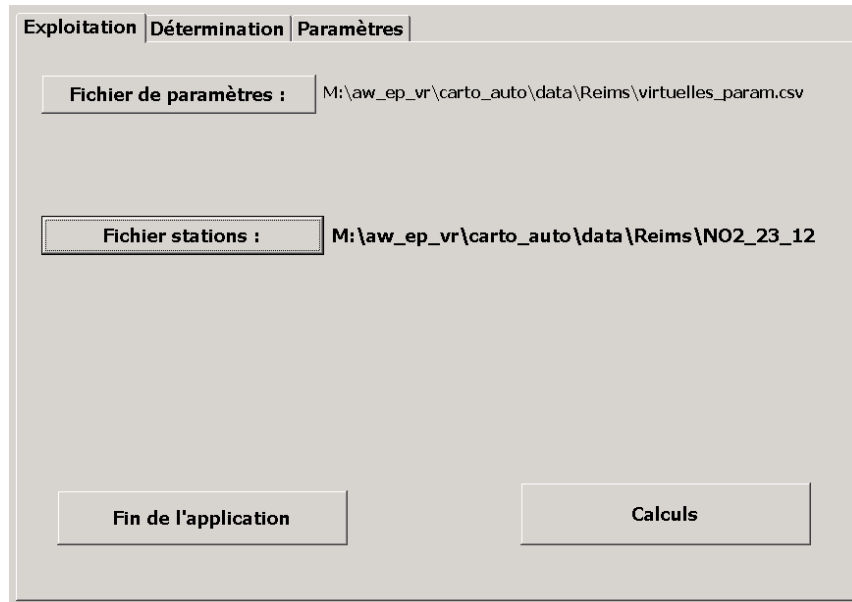


Figure 35 : Feuille prête pour le calcul « exploitation »

Le bouton « Calculs » est apparu. En sélectionnant ce bouton, le calcul du fichier de stations « virtuelles » s'effectue, le résultat est stocké dans le fichier « NO2\_JJ\_MM\_AAAA.csv » correspondant à la période choisie et dans le même répertoire que le fichier d'entrée concernant les stations fixes.

On peut sortir à tout moment de l'application par le bouton « fin de l'application ».

#### 1.1.5.3 DETERMINATION

Après la sélection de l'onglet « Détermination », l'écran suivant apparaît (Figure 36).

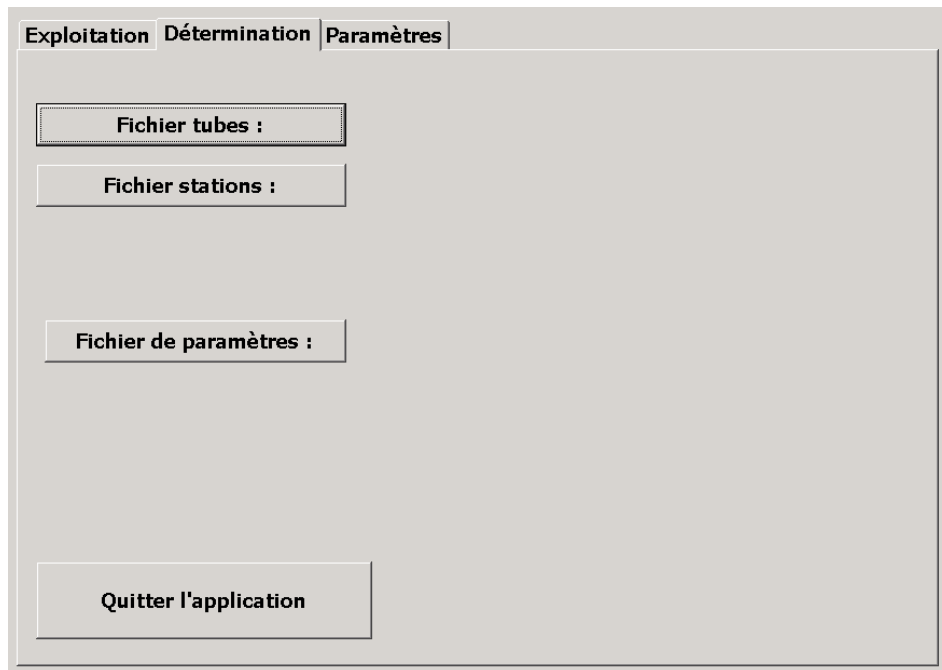


Figure 36 : Feuille de gestion de la partie « détermination »

En sélectionnant le bouton « fichier tubes : » on accède à une fenêtre de recherche du fichier de concentration des campagnes échantillonneurs passifs (Figure 37).

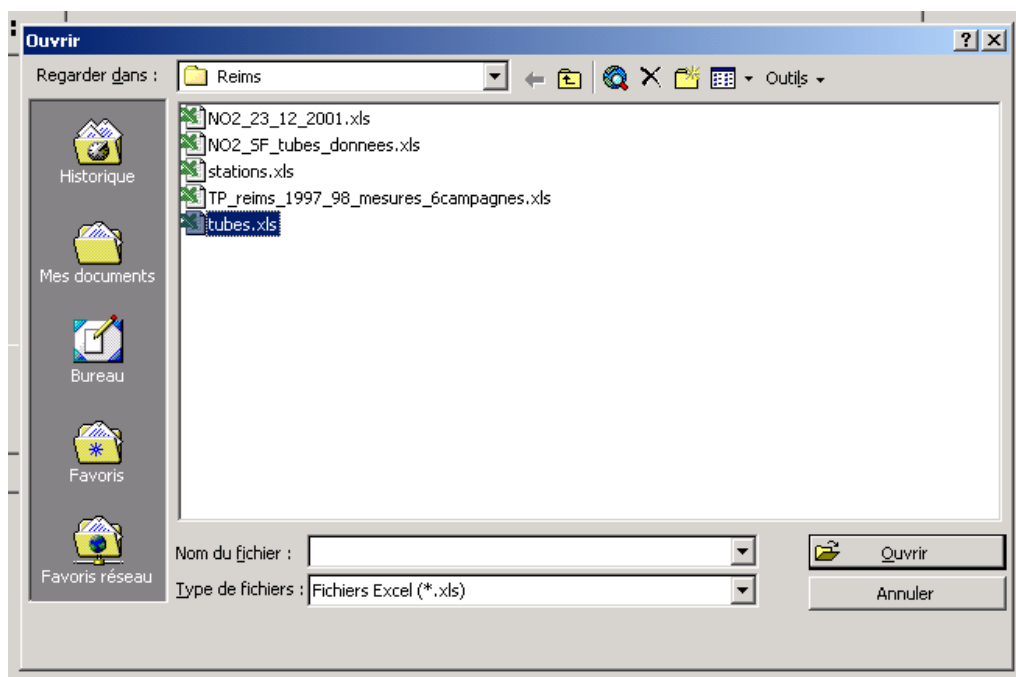


Figure 37 : Recherche du fichier « échantillonneurs »

Après sélection du fichier la feuille fournit en retour le chemin d'accès au fichier.

- En sélectionnant le bouton « fichier stations : » on accède à une fenêtre de recherche du fichier de mesure aux stations fixes pour la période considérée. Après sélection du fichier, la feuille fournit en retour le chemin d'accès au fichier.
- En sélectionnant le bouton « fichier de paramètres : » on détermine le lieu de stockage du fichier de paramètres « virtuelles\_param.csv ». Après sélection du chemin la feuille fournit en retour le chemin d'accès au fichier.

Lorsque les deux fichiers sont sélectionnés la feuille se présente ainsi (Figure 38).

The screenshot shows a software interface with three tabs: 'Exploitation', 'Détermination', and 'Paramètres'. The 'Détermination' tab is active. It contains three input fields, each with a label and a text box. The first is 'Fichier tubes :' with the path 'M:\aw\_ep\_vr\carto\_auto\data\Reims\tubes.xls'. The second is 'Fichier stations :' with the path 'M:\aw\_ep\_vr\carto\_auto\data\Reims\stations.xls'. The third is 'Fichier de paramètres :' with the path 'M:\aw\_ep\_vr\carto\_auto\data\Reims\virtuelles\_param.csv'. At the bottom of the interface, there are two buttons: 'Quitter l'application' on the left and 'Lancement des calculs' on the right.

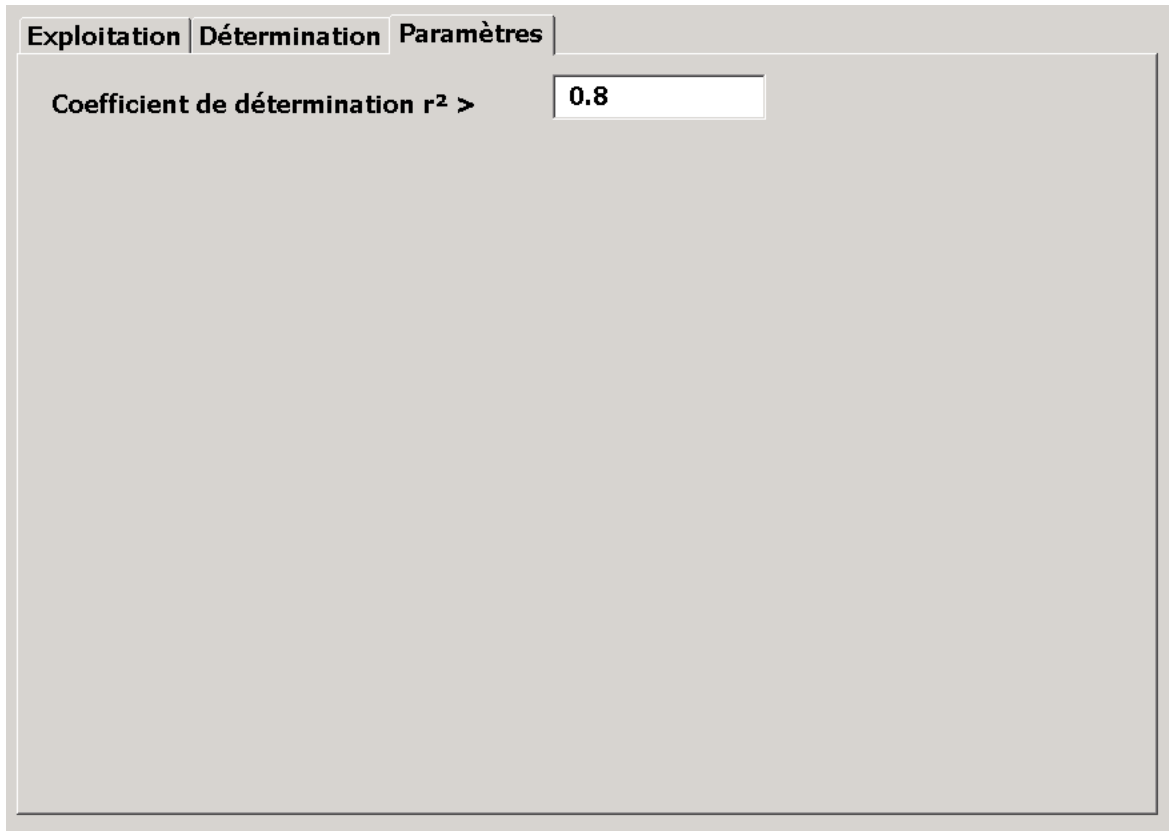
Figure 38 : Feuille prête pour le calcul « Détermination »

Le bouton « Lancement des calculs » est apparu. En sélectionnant ce bouton le calcul du fichier de paramètres de corrélation s'effectue, le résultat est stocké dans le fichier «virtuelles\_param.csv ».

On peut sortir à tout moment de l'application par le bouton « Quitter l'application ».

## 1.1.5.4 PARAMETRES

Après la sélection de l'onglet « Paramètres », l'écran suivant apparaît (Figure 39).



The screenshot shows a software interface with three tabs: 'Exploitation', 'Détermination', and 'Paramètres'. The 'Paramètres' tab is active. Below the tabs, there is a label 'Coefficient de détermination r² >' and a text input field containing the value '0.8'.

Figure 39 : Feuille de gestion de la partie Paramètres

La feuille permet de changer la valeur du seuil de sélection du coefficient de détermination (cf. 3.1.2.2).

## 1.2 MACRO DE CARTOGRAPHIE AUTOMATIQUE

La partie supérieure de l'interface du programme est destinée à la rentrée des données, elle est composée de 4 contrôles interactifs et de trois boîtes de dialogue.

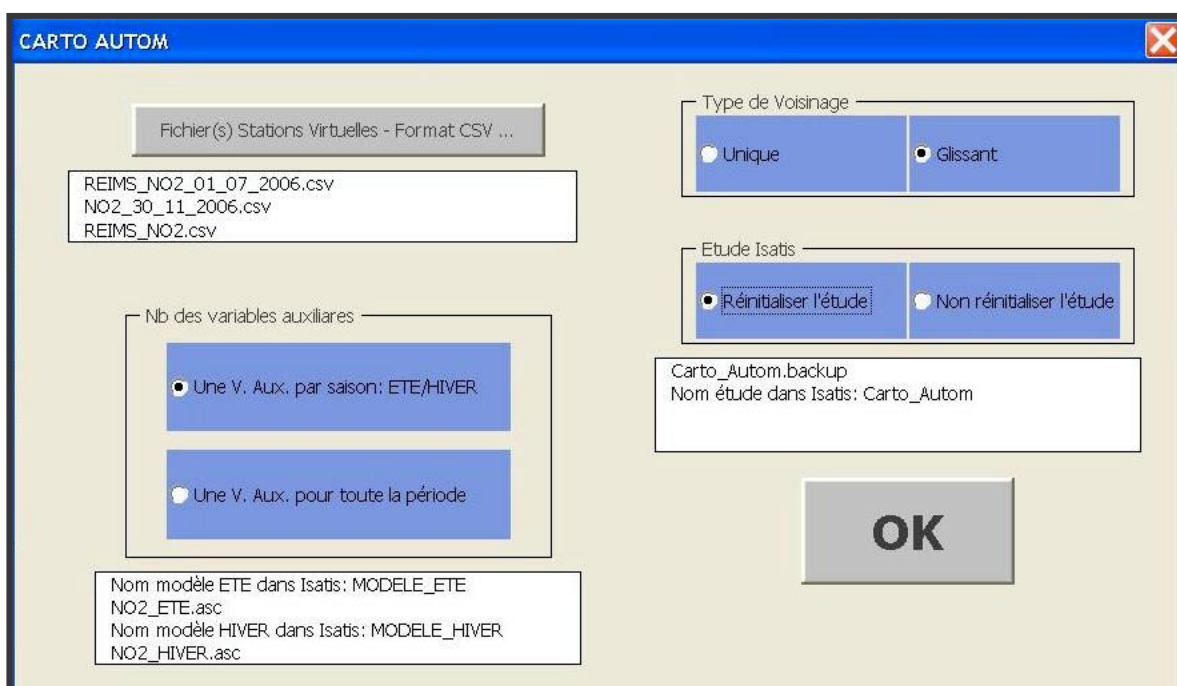


Figure 40 : Partie supérieure de l'interface destinée à la rentrée des données

### 1.2.1 FICHIER DE CONCENTRATIONS PONCTUELLES DE NO2

Le bouton « Fichier(s) Stations Virtuelles - Format CSV » permet d'intégrer les concentrations ponctuelles de NO<sub>2</sub> (stations virtuelles et stations fixes).

Il s'agit des fichiers contenant les concentrations journalières des stations virtuelles et des stations fixes. Ceux-ci ont pu être créés par la macro Excel de corrélations, comme expliqué dans la section 1.1.4.2.

La boîte de dialogue « Ouvrir un fichier » apparaît au moment de cliquer sur ce bouton ; les fichiers « CSV » peuvent alors être recherchés, sélectionnés et chargés.

Le bouton « Ouvrir » de cette fenêtre, permet de finaliser les actions et d'afficher les noms des fichiers avec le chemin d'accès dans la boîte de dialogue en bas du bouton.

Les fichiers doivent respecter les conventions suivantes :

- Extension « CSV » (comma-separated values): les données doivent être rangées en colonnes séparées par le symbole « ; ». Le point décimal des chiffres doit être le symbole « . ».



- La première ligne doit être réservée aux noms des champs.

Le programme vérifie le nom des champs de la première ligne du fichier CSV, si un des champs du Tableau 32 est manquant, le message : « Erreur dans le fichier: « *NOM\_FICHIER* » » apparaît lors de l'exécution du programme, suivi de l'indication du nom du champ manquant : « Il manque la colonne : « *NOM\_DU\_CHAMP* ». Dans ce cas, l'utilisateur doit sélectionner le bouton « OK » de la boîte de dialogue afin de continuer (cf. Figure 42).

A noter que l'ordre des colonnes n'est pas important.

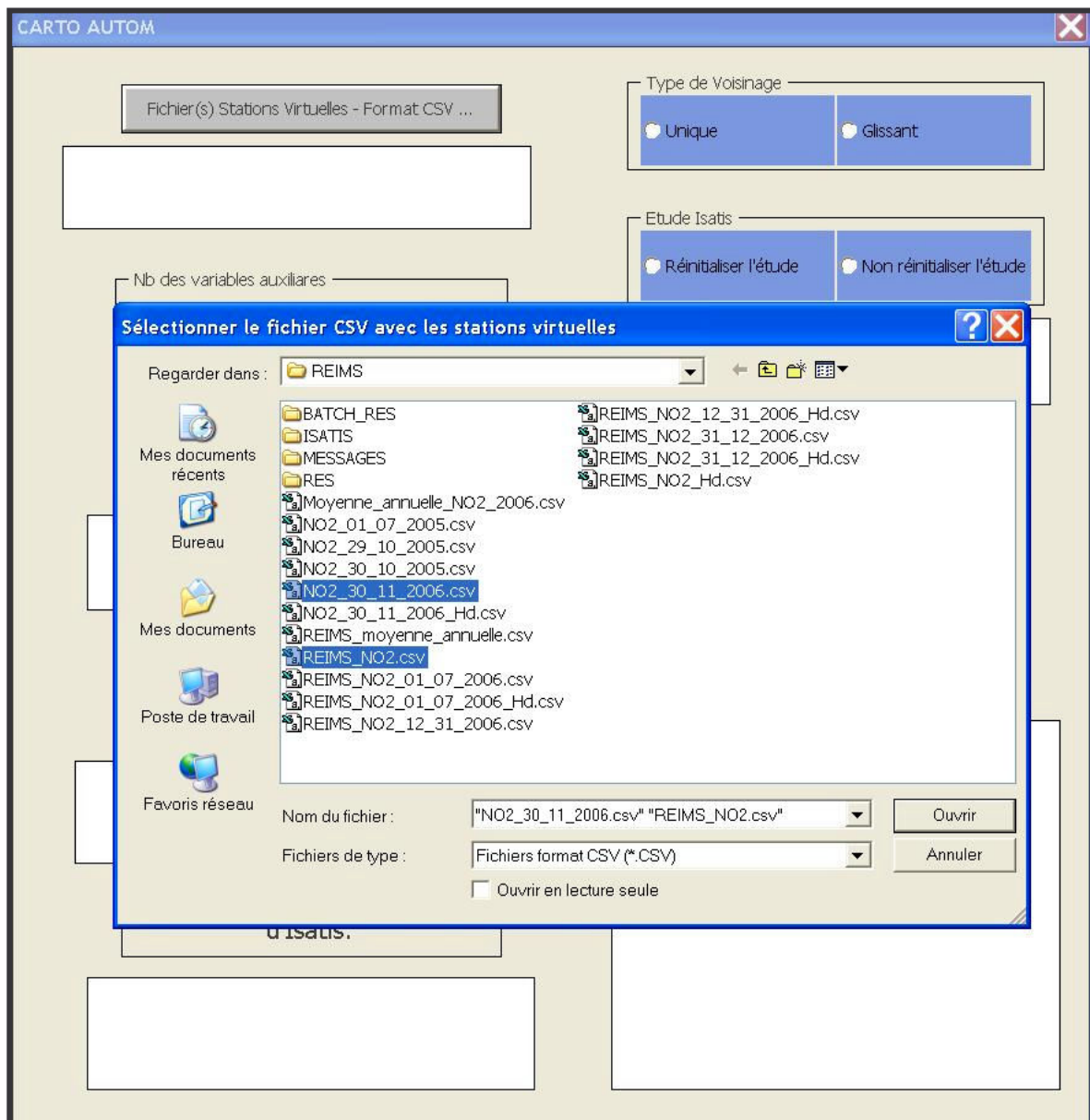


Figure 41 : Fenêtre « Sélectionner les fichiers CSV avec les stations virtuelles »

# Colonne	Nom de champ	Description
1	NOM	Nom des SV et SF
2	X	Coordonnée X (m)
3	Y	Coordonnée Y (m)
4	SFREF	Nom de la station fixe attachée à la station virtuelle
5	M	Pente de la régression linéaire (corrélations temporelles)
6	B	Intercepte de la régression linéaire (corrélations temporelles)
7	NO2	Concentrations des SV et SF ( $\mu\text{g}/\text{m}^3$ )
8	VEM	Variance de l'erreur de mesure
9	SEL_SF	Variable indiquant à Isatis quels points sont des stations fixes (Elle vaut « 0 » pour les SV et "1" pour les SF)

Tableau 32 : Nom des champs du fichier CSV

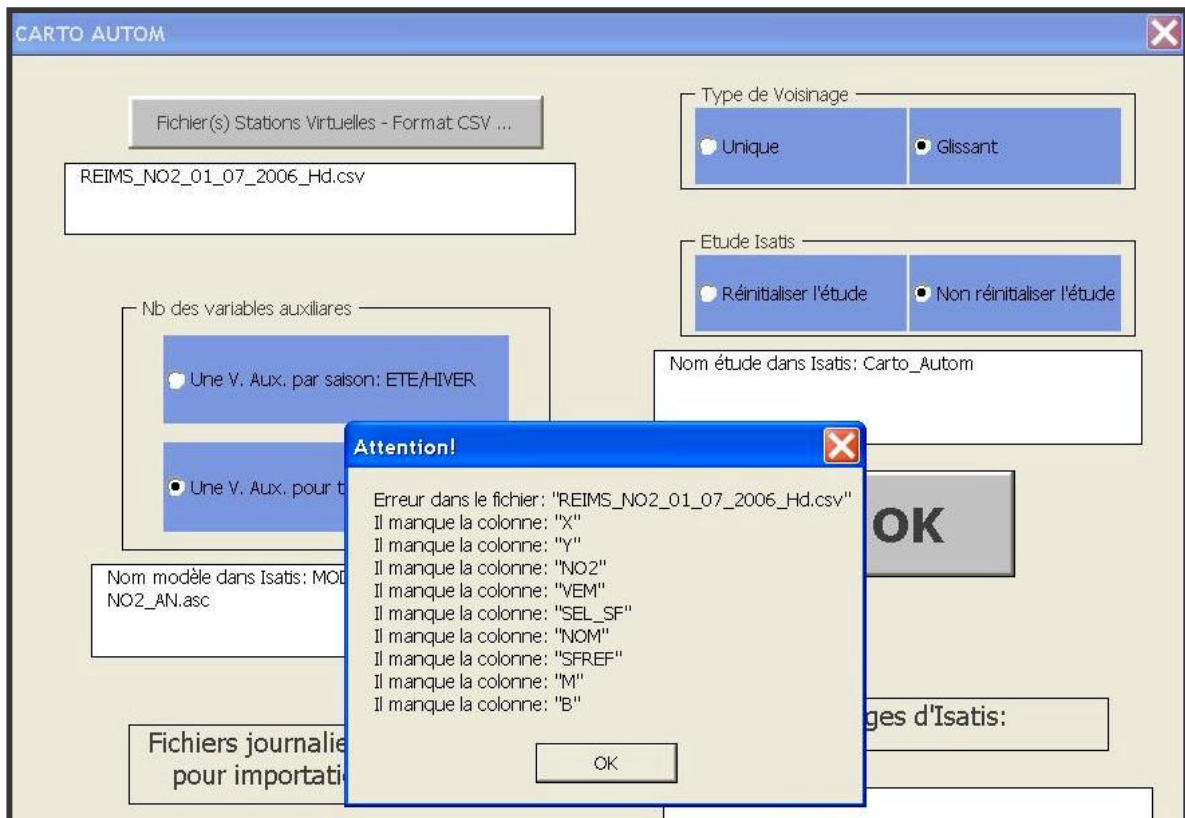


Figure 42 : Message d'erreur lorsqu'un des 9 champs renseignés dans le Tableau 32 est manquant (indication du nom du champ manquant)

Voici un exemple d'un fichier CSV valable.

NOM;X;Y;SFREF;M;B;NO2_01/07/2006;VEM;SEL_SF
1;722850;2474714;N/A;N/A;N/A;28.05;8.22;0
2;723758;2474707;N/A;N/A;N/A;32.94;9.15;0
3;724282;2474816;N/A;N/A;N/A;25.59;6.96;0
4;724024;2474384;N/A;N/A;N/A;28.01;16.68;0
5;723826;2474385;N/A;N/A;N/A;31.18;13.11;0
6;723722;2474154;N/A;N/A;N/A;31.7;25.55;0
7;724004;2474184;N/A;N/A;N/A;30.96;4.23;0
8;724045;2473764;N/A;N/A;N/A;30.13;23.1;0
43;722925;2469509;N/A;N/A;N/A;17.75;10.69;0
44;722829;2471116;N/A;N/A;N/A;21.78;11.03;0
45;725303.11;2474813.41;N/A;N/A;N/A;20.99;0;1
46;722839.57;2471046.2;N/A;N/A;N/A;20.31;0;1
47;724914.16;2472285.13;N/A;N/A;N/A;26.13;0;1

Tableau 33 : Exemple de fichier CSV

Avant d'utiliser la macro, l'utilisateur a dû choisir la variable auxiliaire ainsi que le modèle de variogramme à utiliser lors de l'estimation par cokrigage.

Le programme permet à l'utilisateur de choisir parmi deux types d'estimation :

- le premier type d'estimation s'applique lorsqu'on dispose de paramètres différents d'estimation selon la saison du jour à estimer (variable auxiliaire ou modèle de variogramme différent pour chaque saison),
- le second type d'estimation s'applique lorsqu'on ne dispose que d'une variable auxiliaire et d'un modèle de variogramme, par exemple lors de l'interpolation directe des concentrations annuelles.

Dans le premier cas, la date du jour doit être renseignée afin de choisir les paramètres d'estimation en fonction de la saison du jour (été ou hiver). Cette date doit être rajoutée à la dernière partie du nom du fichier, sous la forme : « JJ\_MM\_AAAA ».

Si la date n'est pas interprétée, le message : « Le fichier : « *NOM\_FICHER* » n'a pas de date pour calculer la saison du jour » apparaît lors de l'exécution du programme et le fichier n'est pas traité. Dans ce cas, l'utilisateur doit sélectionner le bouton « OK » (cf. Figure 43) pour continuer avec le traitement d'un autre fichier ou pour réinitialiser le programme.

Les deux tableaux ci-après donnent quelques exemples de noms de fichiers valides et non valides.

NOM FICHIER CSV	VALIDE	Commentaires
AREMA_NO2_01_12_2006.CSV	OUI	Date : 1er décembre 2006
NO2_AREMA_01_12_2006.CSV	OUI	Date : 1er décembre 2006
NO2_01_12_2006.CSV	OUI	Date : 1er décembre 2006
AREMA_01_12_2006.CSV	OUI	Date : 1er décembre 2006
AREMA_NO2_12_31_2006.CSV	OUI	Date : 31 décembre 2006 (Inversion MM/JJ interprétée)
12_31_2006_AREMA_NO2.CSV	NON	La date n'est pas dans la dernière partie du nom
REIMS01_12_2006.CSV	OUI	Date : 1er décembre 2006
AREMANO2_01_12.CSV	NON	La date est incomplète (il manque l'année)

Tableau 34 : Exemples de noms de fichiers dans le premier type d'estimation : un krigeage différent selon la saison du jour

NOM FICHIER CSV	VALIDE	NOM FICHIER CSV	VALIDE
AREMA_NO2_01_12_2006.CSV	OUI	12_31_2006_AREMA_NO2.CSV	OUI
NO2_AREMA_01_12_2006.CSV	OUI	AREMANO2_01_12_2006.CSV	OUI
NO2_01_12_2006.CSV	OUI	AREMANO2_01_12.CSV	OUI
AREMA_01_12_2006.CSV	OUI	NO2_moyenne_annuelle	OUI
AREMA_NO2_12_31_2006.CSV	OUI	NO2_ETE	OUI

Tableau 35 : Exemples de noms de fichiers dans le seconde type d'estimation : même type de krigeage pour tous les fichiers

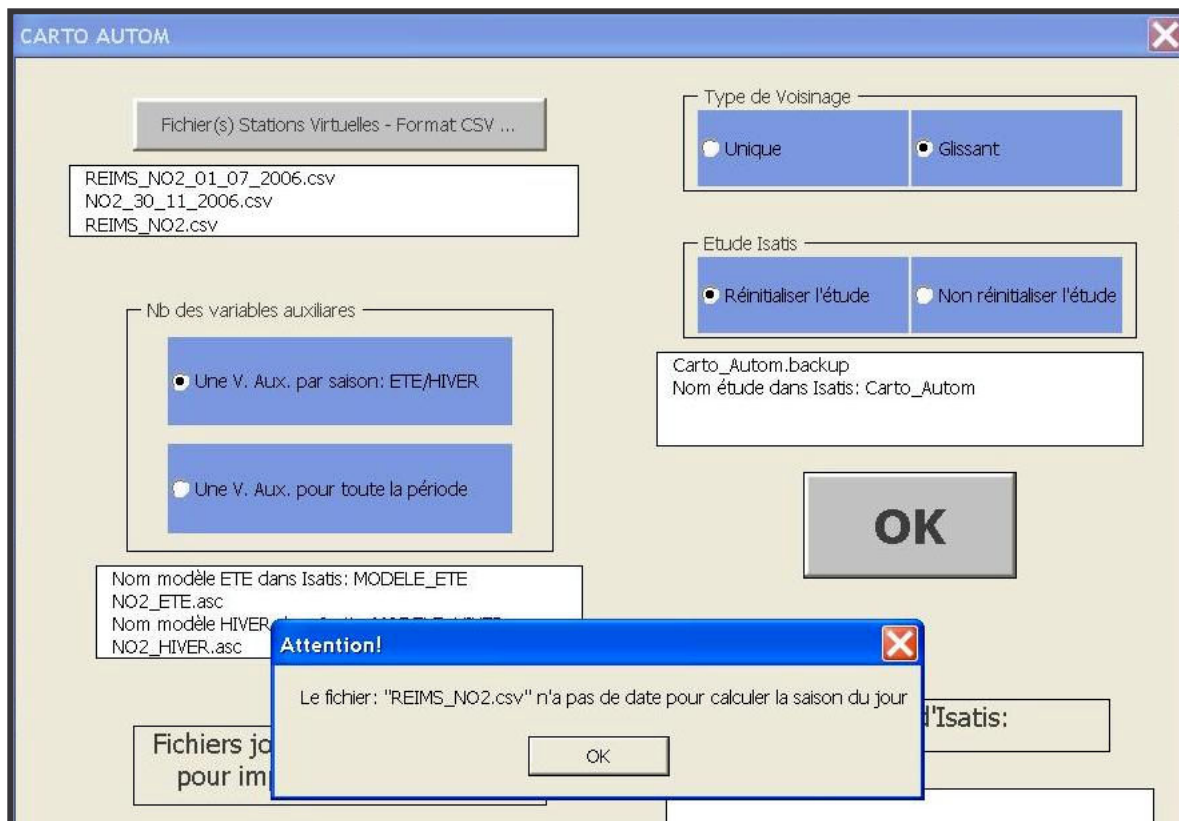


Figure 43 : Message d'erreur lors de la sélection d'un fichier « CSV » sans la date du jour, dans le premier type d'estimation : un krigeage différent selon la saison du jour

## 1.2.2 VARIABLES AUXILIAIRES

### 1.2.2.1 EXPORTATION SUR ISATIS DES VARIABLES AUXILIAIRES

L'exploitation de l'information secondaire dans la méthode de cartographie automatique, permet d'obtenir des cartographies plus réalistes de la répartition spatiale des niveaux de pollution.

Pour mener à bien une estimation par la méthode du cokrigeage colocalisé, il est nécessaire de disposer d'une seconde variable informée dans la maille d'estimation finale.

Cette variable auxiliaire a dû être définie lors de l'étude des campagnes par tubes passifs disponibles dans la zone d'étude.

Pour prendre en compte les caractéristiques saisonnières de la pollution par NO<sub>2</sub>, la méthode considère la possibilité d'utiliser une variable auxiliaire différente pour chaque saison.

Une fois que la ou les variables auxiliaires ont été calculées sur Isatis, elles peuvent être exportées à l'aide du menu de la Figure 44 comme grille SIG (extension ASC).

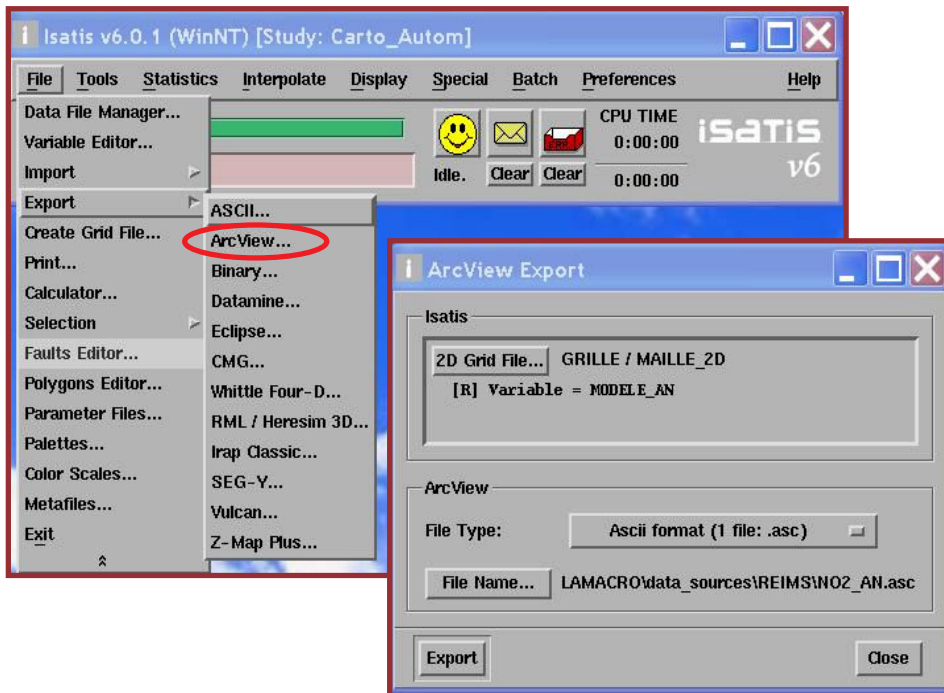


Figure 44 : Exportation de la variable auxiliaire sur Isatis

Il s'agit de la grille à deux dimensions qui sera ultérieurement demandée par l'interface, elle contient les valeurs numériques de la variable auxiliaire et détermine à la fois les dimensions de la maille d'estimation finale : les estimations par cokrigage colocalisé des concentrations de NO<sub>2</sub> s'effectueront dans cette maille. Un exemple de premières lignes d'un fichier ASC est montré dans l'Annexe 12.

#### 1.2.2.2 SELECTION DES VARIABLES AUXILIAIRES DANS L'INTERFACE

Les deux cases à cocher groupées dans le cadre dénommé « Nb des variables auxiliaires » permettent de choisir parmi les deux types d'estimation expliqués dans la section 1.2.1. Dans cette partie du programme sont rentrés le nom du modèle dans l'étude Isatis (cf. 1.2.4) et la variable auxiliaire (extension ASC).

Lorsque l'on sélectionne une case à cocher, une boîte de dialogue demande le nom du « modèle de variogramme ».

Dans la zone de texte de cette boîte apparaît par défaut le mot « MODELE » suivi du nom de la saison selon le type d'estimation choisi : « MODELE\_ETE », « MODELE\_HIVER » ou « MODELE\_ANNEE ». Ces noms peuvent être changés par l'utilisateur (cf. Figure 45).

Le nom doit être celui donné au modèle lors de modélisation du variogramme expérimental sur Isatis (cf. 1.2.4.2).

L'option « OK » de cette boîte permet de valider le nom de modèle choisi et de faire apparaître un bouton indiquant la saison et le type de format de la variable auxiliaire à rentrer (cf. Figure 46).

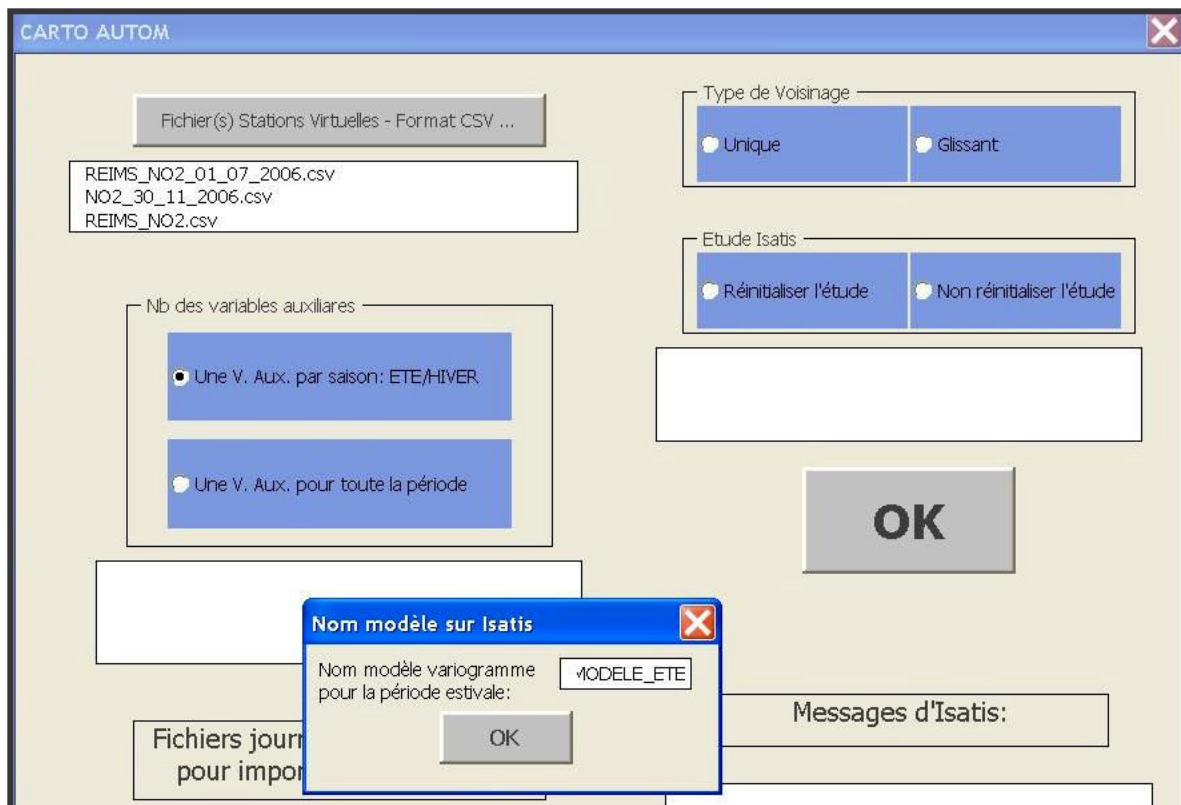


Figure 45 : Boîte de dialogue « Nom modèle sur Isatis »

Avec ce bouton, on accède à la boîte de dialogue « Ouvrir un fichier », laquelle permet de chercher et charger la variable auxiliaire au format grille d'ArcView (extension ASC, cf. Figure 47).

Si l'utilisateur a choisi l'option « Une V. Aux. par saison : ETE/HIVER », le bouton « ouvrir » de cette boîte permet de sélectionner le nom du modèle et de la maille correspondant à l'hiver.

Finalement les noms des modèles et des fichiers ArcView sélectionnés sont affichés dans la boîte de dialogue localisée dans la partie inférieure du cadre « Nb de variables auxiliaires ».

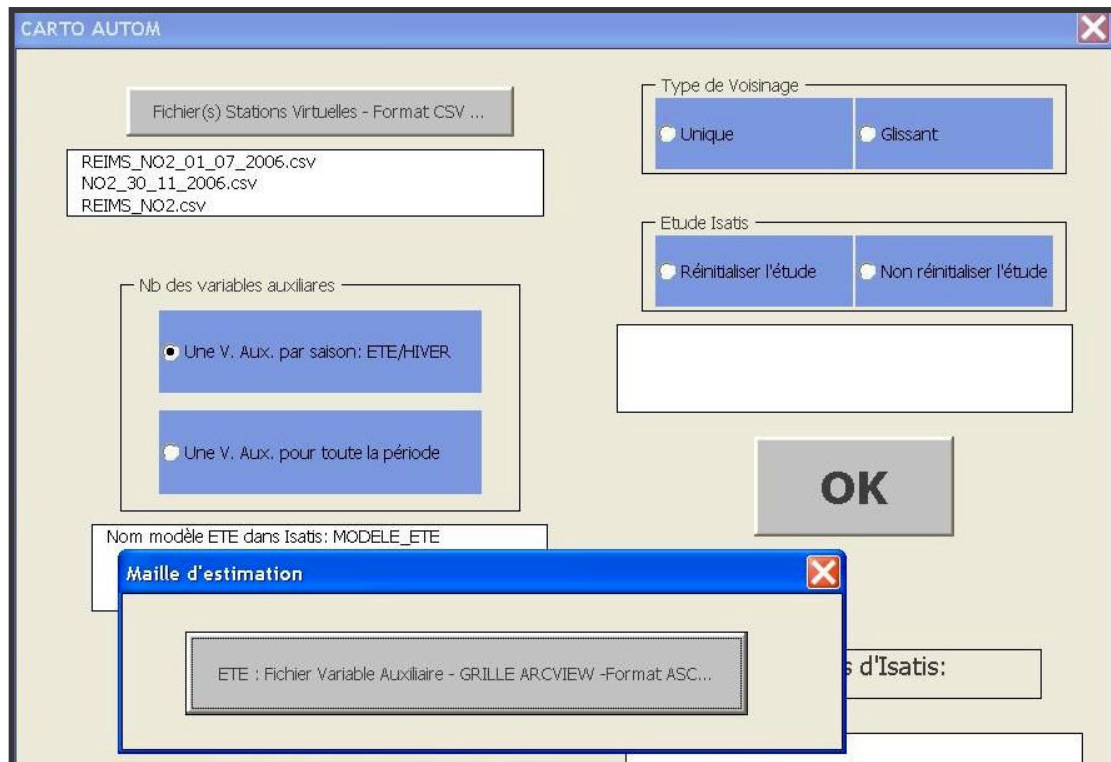


Figure 46 : Bouton permettant le chargement de la Variable Auxiliaire

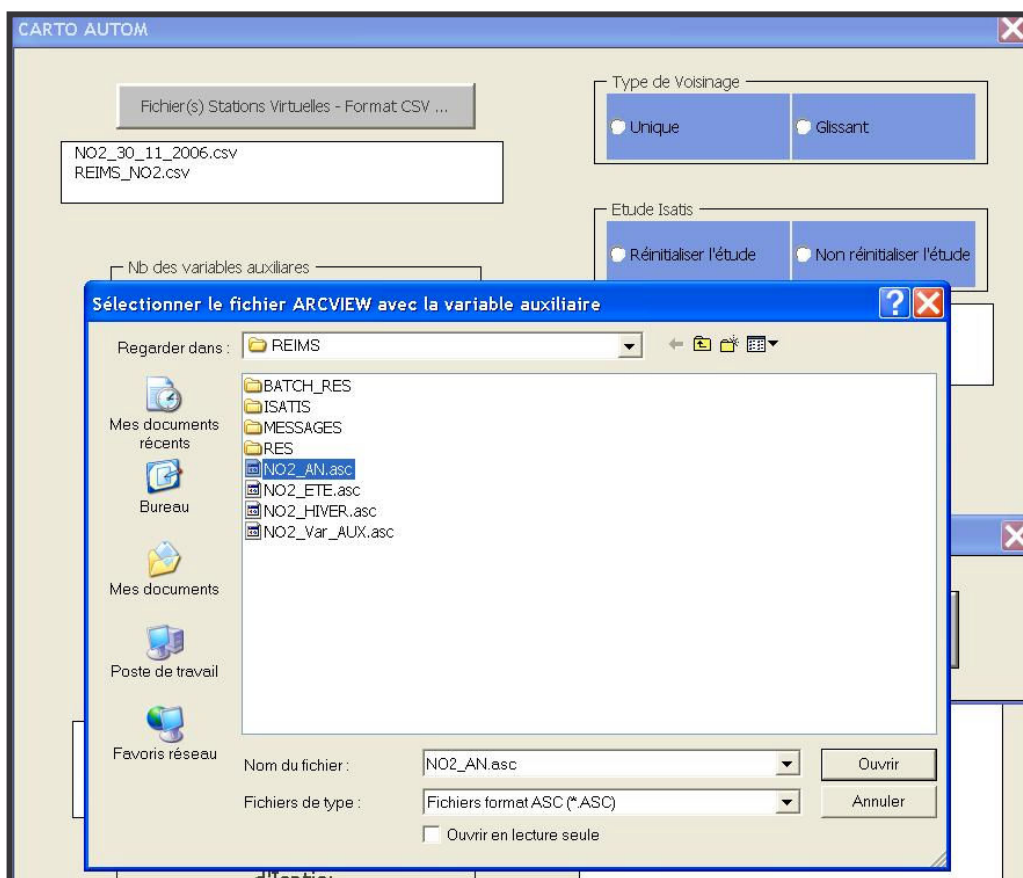


Figure 47 : Fenêtre « Sélectionner le fichier ARCVIEW ... »



### 1.2.3 TYPE DE VOISINAGE

Le cadre localisé dans la partie supérieure droite de l'interface du programme permet de choisir le type de voisinage d'estimation. L'option « Unique » correspond au cas où tous les points sont considérés lors de l'estimation de chaque nœud de la grille finale.

L'option « Glissant » permet de restreindre le voisinage aux seuls points respectant certains paramètres de voisinage : points localisés à une distance et une direction données du nœud de la grille finale.

Les conventions d'Isatis ont été conservées, l'angle doit être compris entre  $0^\circ$  et  $180^\circ$ . Un angle de  $45^\circ$  correspond par exemple à la direction E45N pour l'axe X et N45W pour l'axe Y (cf. Figure 49).

La direction  $0^\circ$  est équivalent à la direction  $180^\circ$  (direction EW pour l'axe X et NS pour l'axe Y). Si l'angle n'est pas compris entre  $0^\circ$  et  $180^\circ$ , le message d'erreur "Angle dehors de la plage valable" apparaît invitant l'utilisateur à saisir un angle correct.

La sélection de l'option « Glissant » dans l'interface fait apparaître la boîte de dialogue « Paramètres Voisinage Glissant » où doivent être renseignées la direction et la distance à respecter lors du cokrigeage (cf. Figure 48).

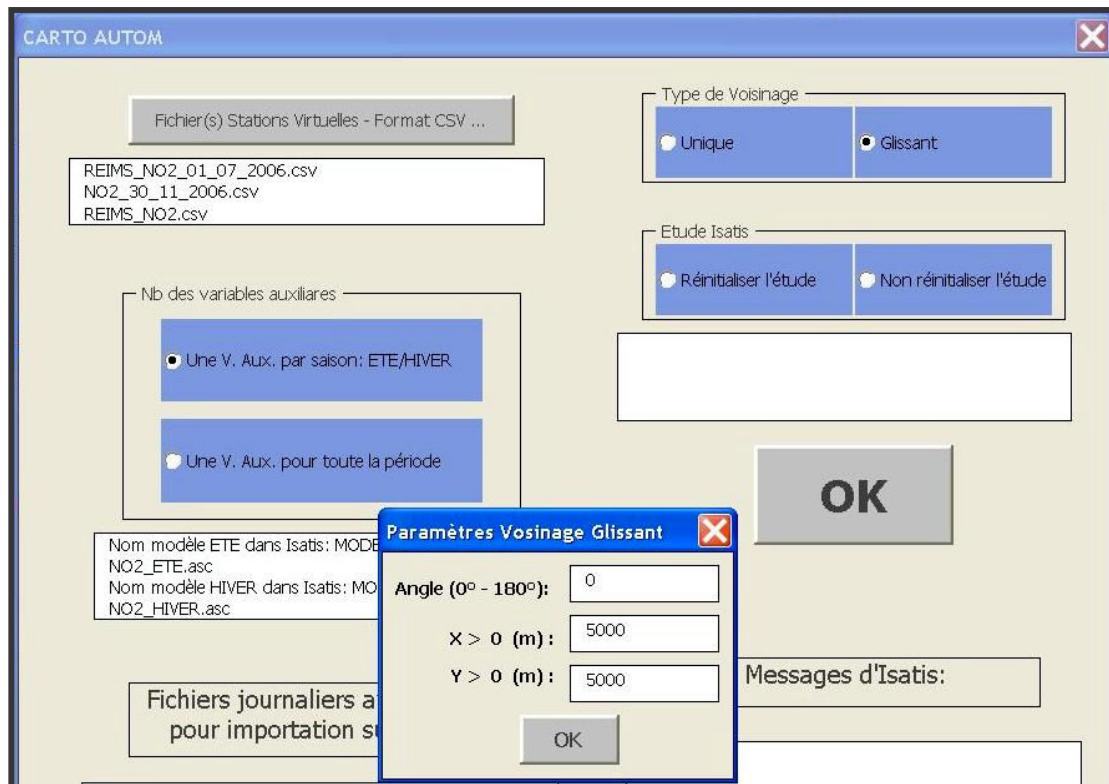


Figure 48 : Boîte de dialogue « Paramètres Voisinage Glissant »

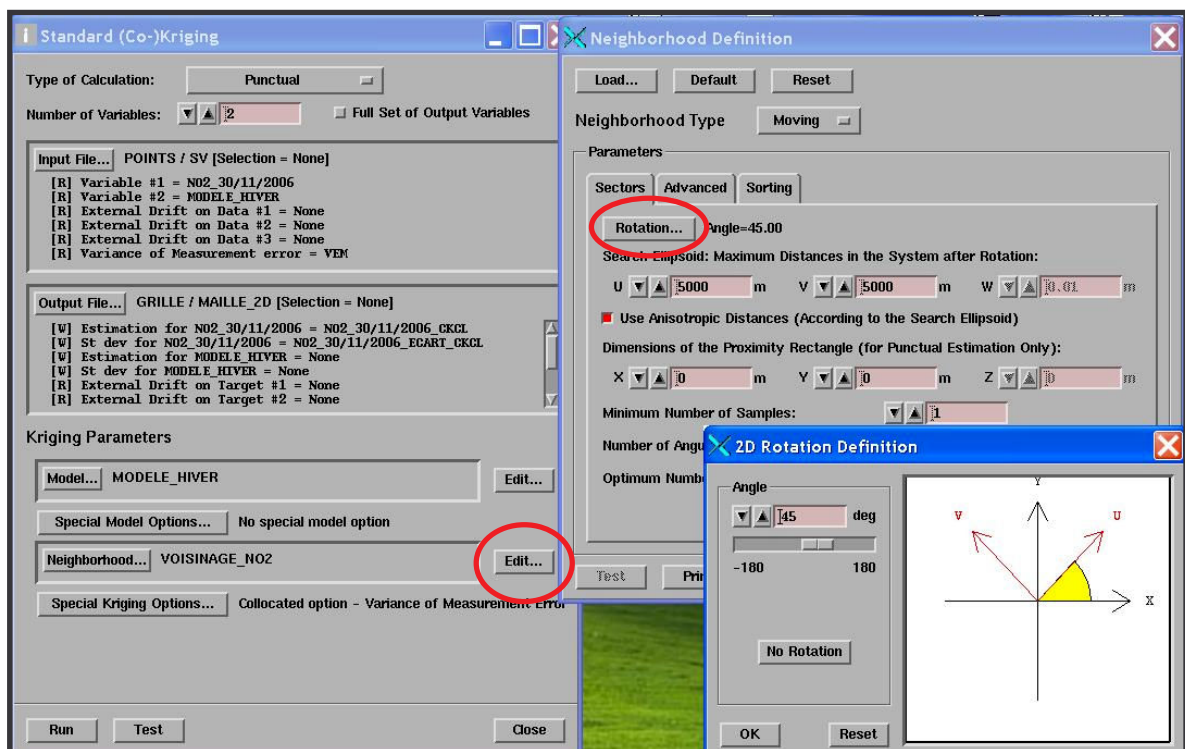


Figure 49 : Ajustement de la direction du voisinage dans Isatis

Les paramètres « X(m) » et « Y(m) » correspondent aux distances en mètre du rayon de recherche dans l'axe X et Y. Si les deux valeurs rentrées sont identiques, le rayon de recherche de points voisin représentera un cercle. Dans le cas où les valeurs rentrées sont différentes, une ellipse orientée dans la direction définie par l'angle de rotation sera utilisée.

Par exemple dans le cas de la Figure 48, l'utilisateur a défini un voisinage d'estimation correspondant à une ellipse de rayon 5000m dans la direction de E45°N et 2500m dans la direction N45°W.

On rappelle que les paramètres du voisinage d'estimation peuvent être définis en fonction de l'anisotropie géométrique de la pollution par NO<sub>2</sub>. Celle-ci peut être détectée lors de l'étude variographique préalable des campagnes par tubes passifs disponibles.

L'utilisateur doit obligatoirement saisir une valeur numérique dans les trois zones de texte de la boîte de dialogue : « Paramètres Voisinage Glissant », si un ou plusieurs paramètres ont été laissés vides, le message d'erreur «Entrez les paramètres du voisinage » apparaîtra lors de la sélection du bouton « OK ».

Par ailleurs, les paramètres ont été fixés à zéro (0) par défaut, cependant les distances nulles (= 0) ne sont pas acceptées. Si on oublie de changer la valeur zéro des distances, l'erreur « Une distance de voisinage n'est peut pas être nulle» apparaîtra lors de la sélection du bouton « OK ».

L'utilisateur doit porter une attention particulière aux valeurs de paramètres rentrées, il faut par exemple s'assurer de l'existence des points de données (stations) dans le rayon de recherche défini, dans le cas contraire, l'erreur « ERROR DURING FILE CLASSIFICATION (SORTING). SEARCH ELLIPSOID OR SORTING PARAMETERS TOO SMALL » se présentera lors de l'estimation sur Isatis. Cette erreur provoque la fermeture d'Isatis lorsque ce logiciel est utilisé en mode batch.

## 1.2.4 ETUDE ISATIS

Dans Isatis, le chargement des données et la réalisation des interpolations s'effectuent au sein d'une étude (nom « Study » dans le logiciel). Une étude est l'espace de travail qu'Isatis réserve dans le disque dur afin d'enregistrer les données, les paramètres choisis (variogrammes, modèles, voisinage, etc.) et les nouvelles variables créées lors des opérations du logiciel (p. ex. estimations par cokrigage).

Lors du lancement d'Isatis en mode batch, il est nécessaire de rajouter dans les paramètres le nom de l'étude où les opérations demandées seront effectuées.

Une étude Isatis doit donc être créée avant l'utilisation du programme de cartographie automatique.

### 1.2.4.1 CREATION D'UNE ETUDE SUR ISATIS

Pour créer une étude sur Isatis, il faut d'abord faire apparaître la fenêtre « Data File manager » à l'aide du menu « File », option « Data File manager » (cf. Figure 50).

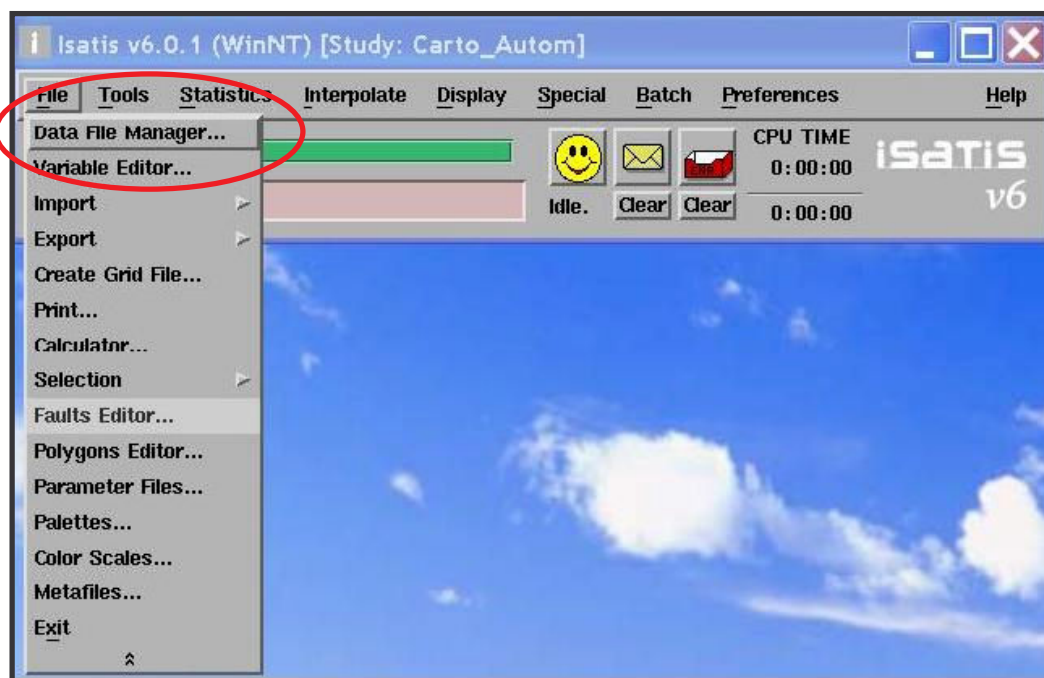


Figure 50 : Sélection du sub-menu « Data File Manager » du menu « File » d'Isatis

Ensuite l'option « Create » du menu « Study » de cette fenêtre permet de définir le nom de l'étude et l'emplacement dans le disque dur.

Le nom et l'emplacement de l'étude peuvent être choisis par l'utilisateur. La Figure 51 montre la création sur Isatis d'une étude nommée : « Carto\_Autom », avec un emplacement dans le disque dur choisi automatiquement par le logiciel.

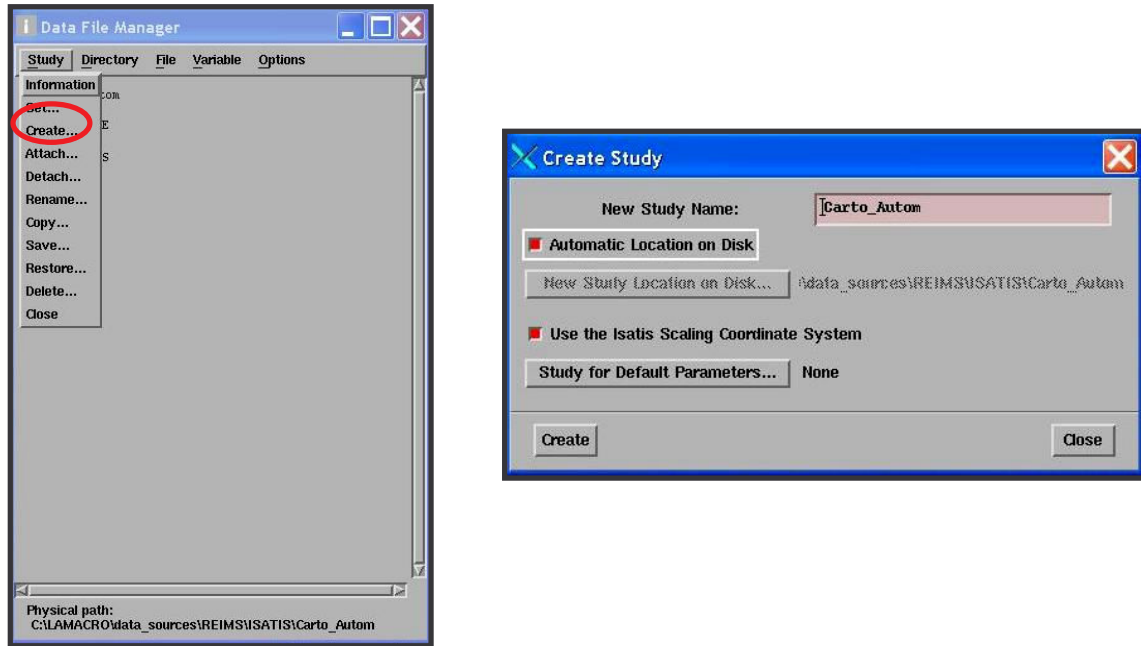


Figure 51 : Sélection du sub-menu « Create » du menu « Study » de la fenêtre « Data File Manager » d'Isatis

L'étude ainsi créée doit aussi contenir les modèles de variogramme qui seront utilisés lors des estimations par cokrigage colocalisé. Ces modèles ont dû être choisis lors de l'analyse des campagnes par tubes passifs disponibles.

#### 1.2.4.2 IMPORTATION ET EXPORTATION DES MODELES SUR ISATIS

Les modèles peuvent être exportés ou importés à l'aide des outils disponibles dans l'option « Variogram Model » du menu « Parameters Files » (cf. Figure 52 et Figure 53).

Le bouton « ASCII Export... » permet de créer un fichier texte avec une extension « mvario » contenant les paramètres du modèle ; le bouton « ASCII Import... » permet ensuite d'importer ce fichier et donc le modèle dans une autre étude Isatis.

Le fichier « mvario » peut être modifié à l'aide d'un éditeur de texte (p. ex. Notepad). Le Tableau 37 montre un exemple des 9 premières lignes de ce type de fichier.

La troisième ligne « Pset Name » représente le nom qu'Isatis donnera au modèle lors de l'importation du fichier, il s'agit du nom qui doit être rentré lors du choix de la variable auxiliaire dans le cadre « Nb des variables auxiliaires » de l'interface du programme « Cartographie Automatique » (cf. 1.2.2).

Les lignes 4 à 9 renseignent sur la localisation dans la structure des données d'Isatis des variables auxquelles fait référence le modèle : Dossier, Répertoire, Nom de la sélection et Nom des variables.

Afin d'éviter des problèmes de compatibilité lors de l'estimation par cokrigage, il est conseillé de copier les lignes 4 à 9 dans tous les fichiers « mvario » qui seront utilisés par la programme de cartographie automatique et de décocher la case « Check Variable Names Consistency in Variogram Models » de l'onglet « Miscellaneous » du sub-menu « Study Environment » du menu Preferences d'Isatis (cf. Figure 54).

Le Tableau 36 affiche un exemple du message d'erreur produit par isatis lorsque les noms des variables à estimer ne correspondent pas aux informations fournies par les lignes 4 à 9 du fichier « mvario » importé.

```
The variables referenced in the Parameters File MODELE_AN
do not match some of the variables declared in the Data File
List of the variables in the Data File:
- NO2_30/11/2006 (Not referenced)
- MODELE_AN (Not referenced)
List of the variables in the Parameter File:
- TOTO_NO2
- TOTO_Var_AUX
The process is interrupted because you asked to
check variable names consistency in variogram models
in Preferences/Study Environment/Miscellaneous
```

*Tableau 36 : Un message d'erreur d'Isatis*

Cette erreur provoque la fermeture d'Isatis lorsque ce logiciel est utilisé en mode batch. Par ailleurs les lignes 8 et 9 donnent l'ordre des variables dans le modèle, la première variable du modèle doit toujours correspondre aux concentrations de NO<sub>2</sub> et la seconde à la variable auxiliaire, l'utilisateur doit vérifier que cet ordre est toujours respecté, sinon on risque d'utiliser la corrélation spatiale de la variable auxiliaire pour estimer les concentrations de NO<sub>2</sub> (inversion des variables dans le modèle).

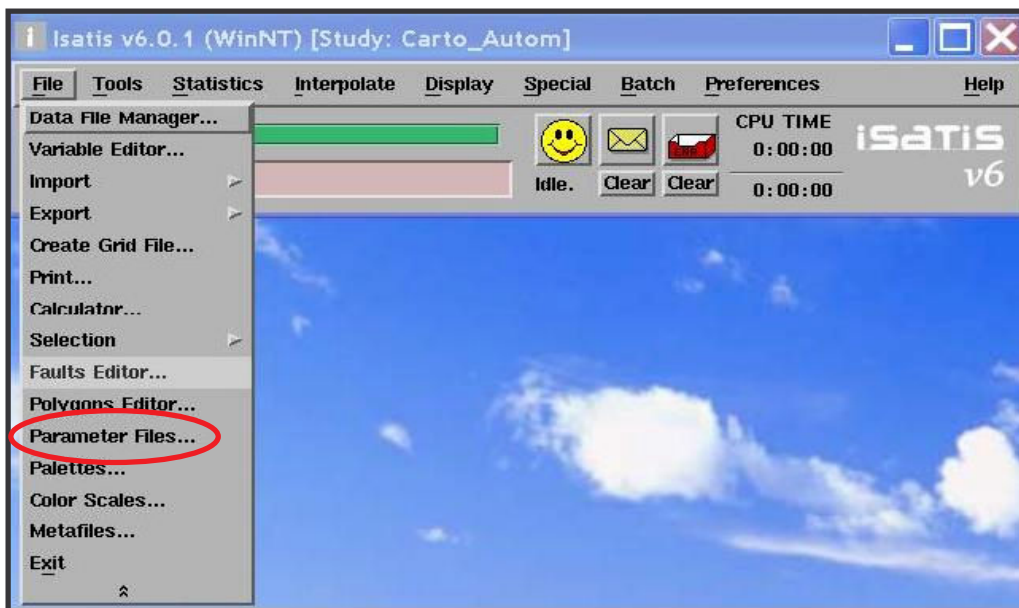


Figure 52 : Sélection du sub-menu « Parameter Files » du menu « File » d'Isatis

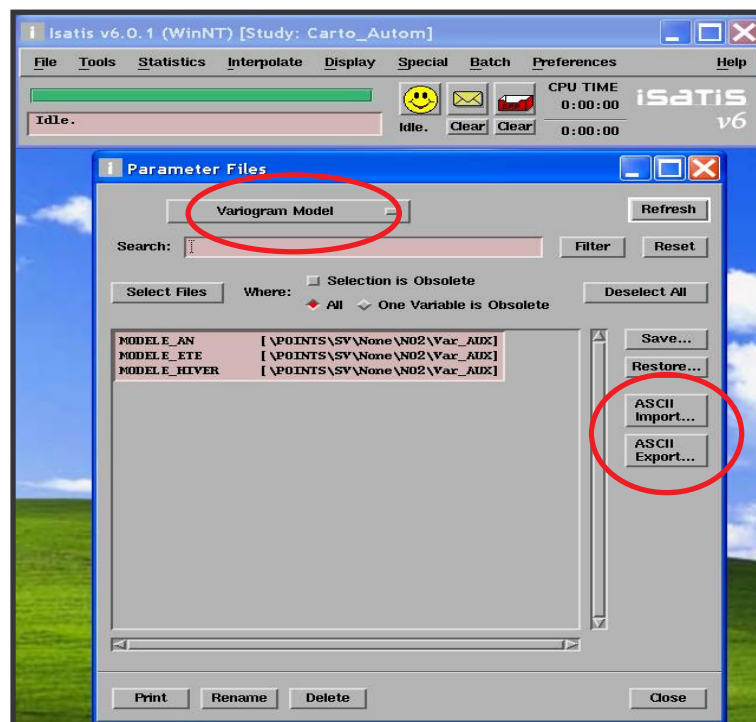


Figure 53 : Boutons permettant d'importer ou exporter les modèles au format Ascii (Option « Variogram Model » de la fenêtre « Parameters File » d'Isatis)

# Ligne	Contenu
1	Version = 521
2	PSet Type = Variogram Model
3	PSet Name = MODELE_AN
4	Directory = POINTS
5	File = SV
6	Selection = None
7	Number of Variables = 2
8	Variable = NO2
9	Variable = Var AUX

Tableau 37: Exemple des 9 premières lignes du fichier texte représentant un modèle du variogramme d'Isatis

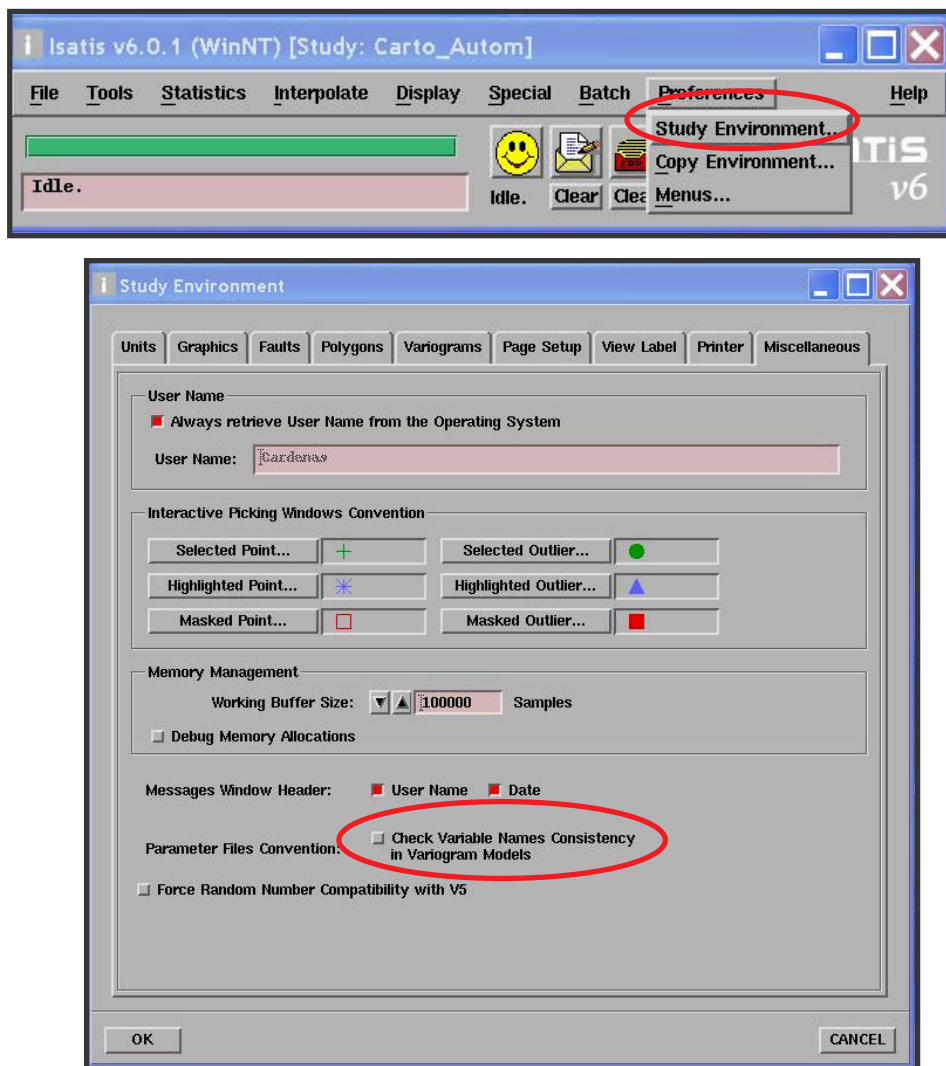


Figure 54 : Désélection de l'option « Check Variable Names Consistency in Variogram Models » de l'onglet « Miscellaneous » de la fenêtre « Study Environment » (menu « Preferences » d'Isatis)



### 1.2.4.3 SAUVEGARDE D'UNE ETUDE SUR ISATIS

Pour effectuer les estimations, le programme de cartographie automatique demande de rentrer un BackUp de l'étude Isatis contenant les modèles de variogrammes.

Ce fichier peut être créé à l'aide du sub-menu « Save... » du menu « Study » de la fenêtre « Data File Manager » d'Isatis.

La Figure 55 montre un exemple de création d'un fichier « Carto\_Autom.backup » qui est une copie de secours de l'étude Carto\_Autom créée dans la section 1.2.4.1.

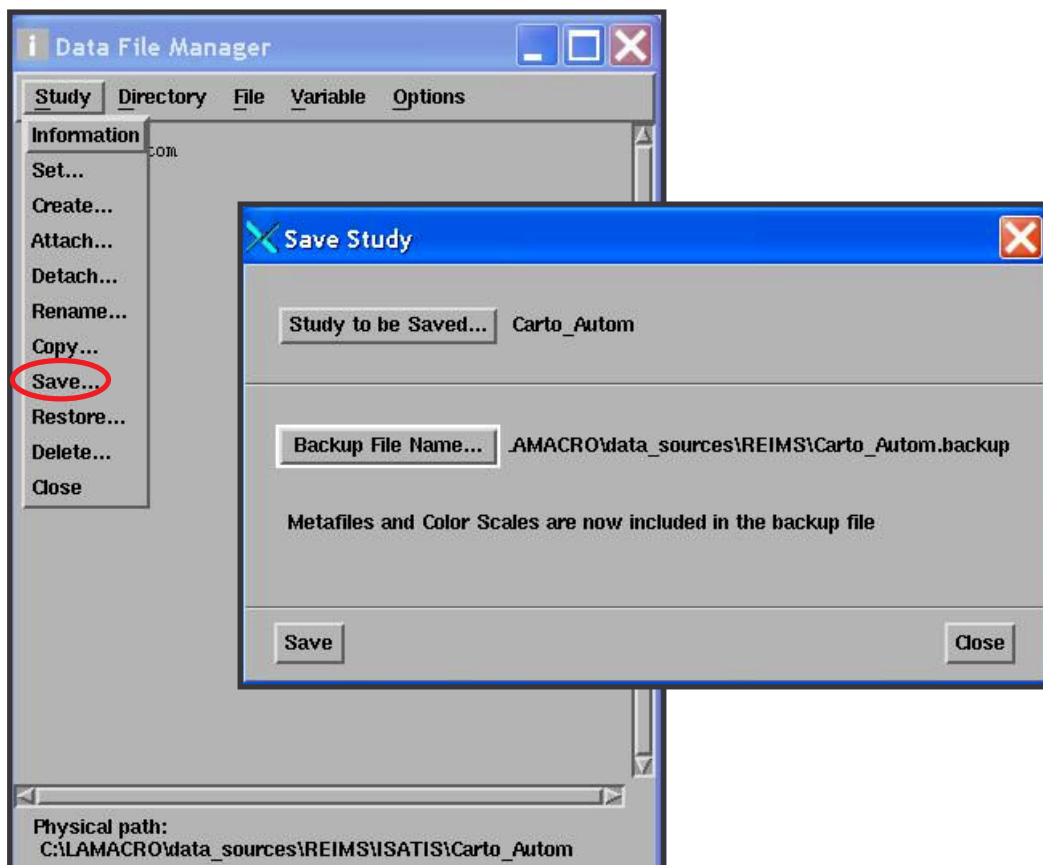


Figure 55 : Sauvegarde d'une étude sur Isatis

Le nom et l'emplacement du fichier « backup » peuvent être choisis par l'utilisateur.

### 1.2.4.4 CHOIX D L'OPTION D'INITIALISATION DE L'ETUDE ISATIS

Dans l'interface du programme de « Cartographie Automatique », le cadre dénommé « Etude Isatis » permet de rentrer le nom de l'étude Isatis contenant les modèles de variogrammes. Cette étude sera utilisée par Isatis pour charger les variables et enregistrer les estimations réalisées.

L'interface propose de choisir parmi deux options : la réinitialisation de l'étude ou la continuation d'une étude déjà existante. La première option doit être choisie lors d'une première utilisation du programme ou lorsqu'on veut effacer et récrire les estimations réalisées sur l'étude Isatis utilisée.

La seconde option permet d'enregistrer les estimations réalisées dans une étude déjà existante, cette option peut être intéressante lorsque l'on veut comparer a posteriori sur Isatis les résultats des estimations des jours différents.

Par exemple dans une agglomération où l'on veut appliquer la méthode de cartographie automatique, la première option doit être employée lors de la cartographie de la première journée. Ensuite si l'on veut conserver toutes les cartographies des pollutions journalières dans la même étude Isatis, la seconde option doit être choisie.

Lorsque l'utilisateur sélectionne l'option : « Réinitialiser l'étude », on accède à la boîte de dialogue « Ouvrir un fichier » (cf. Figure 56), laquelle permet de chercher et charger le fichier BACKUP de l'étude Isatis. Il s'agit du fichier créé comme il a été expliqué dans la section 1.2.4.3.

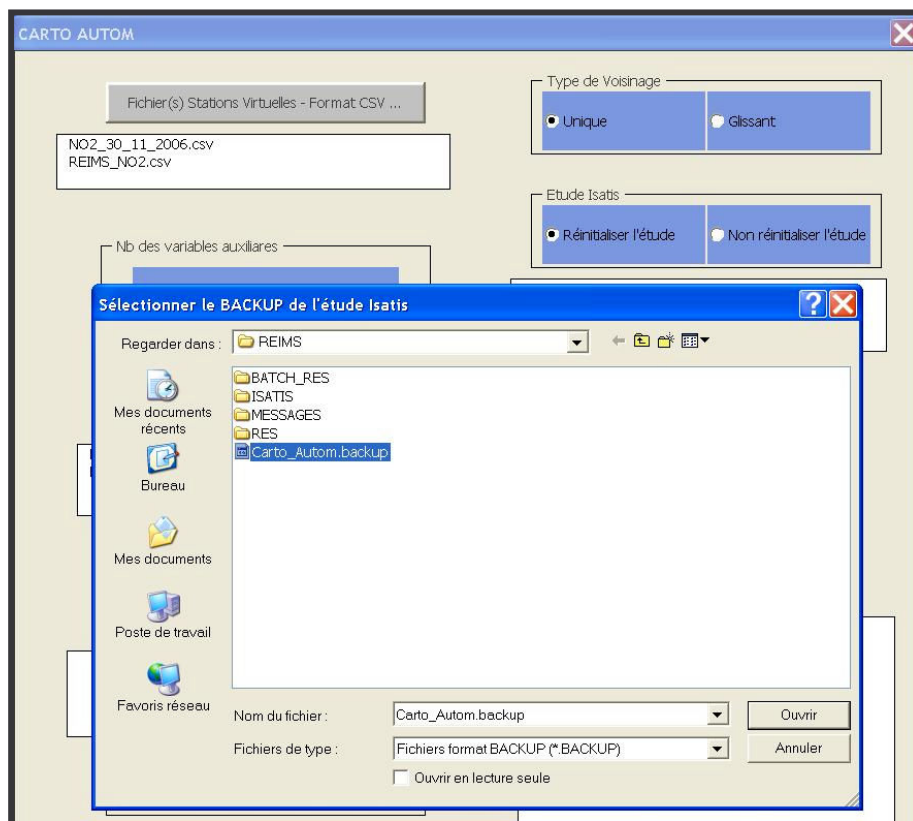


Figure 56 : Fenêtre « Sélectionner le BACKUP de l'étude Isatis »

Le bouton « ouvrir » de cette boîte fait apparaître la boîte de dialogue « Nom Etude Isatis » où l'utilisateur est invité à rentrer le nom que l'étude portera sur Isatis. Le nom par défaut est « Carto\_Autom » (cf. Figure 57).

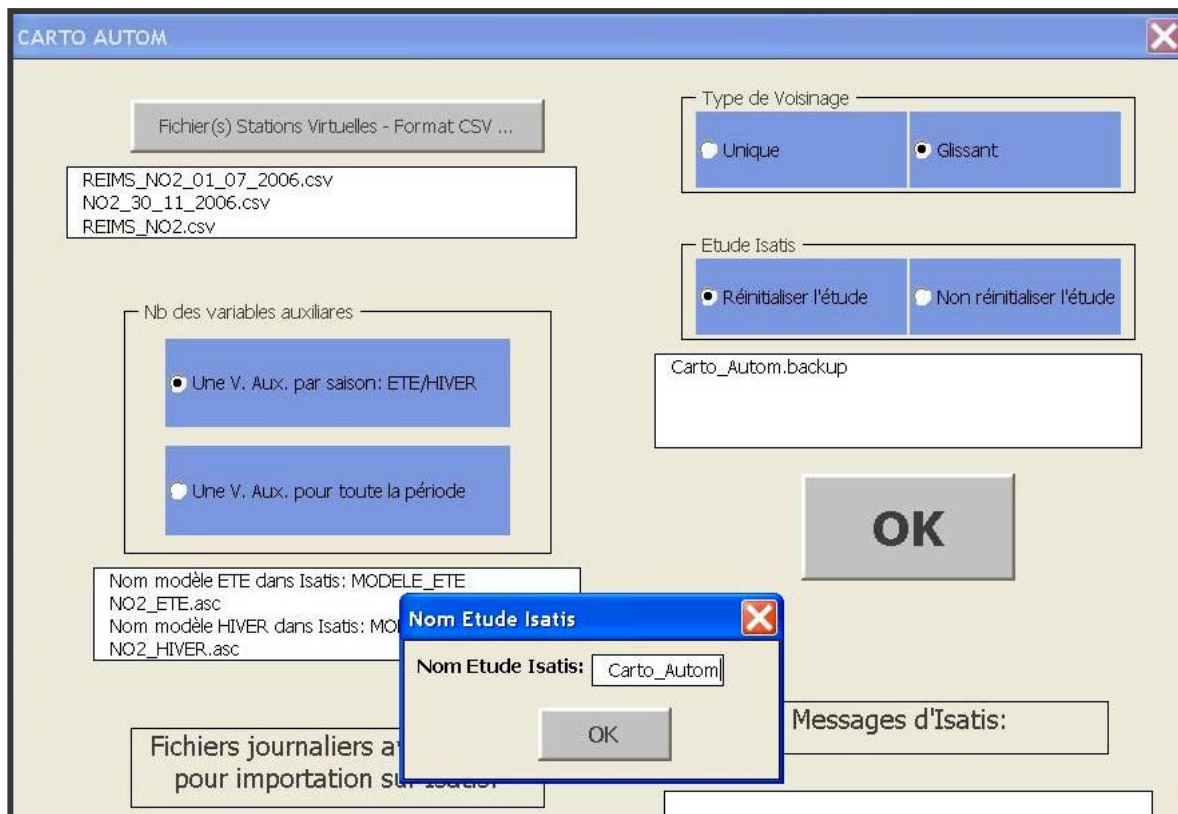


Figure 57 : Boîte de dialogue : « Nom Etude Isatis »

La sélection du bouton « OK » permet d'afficher sur l'interface, dans la boîte de dialogue localisée dans la partie inférieure des cases à cocher, le nom du fichier BACKUP avec le chemin d'accès suivi du nom de l'étude Isatis.

Si l'option « Non Réinitialiser l'étude » a été choisie, seulement la boîte de dialogue « Nom Etude Isatis » sera affichée lors de la sélection de la case à cocher. Le nom de l'étude rentré doit donc correspondre à une étude existante, sinon Isatis produira une erreur et les estimations ne s'effectueront pas.

### 1.2.5 EXECUTION DU PROGRAMME

Le bouton « OK » localisé au milieu à droite de l'interface permet l'exécution du programme. Dans le cas où une information est manquante, un message « Attention » apparaît indiquant à l'utilisateur de rentrer cette information (cf. Figure 58).

La première action réalisée lors de l'exécution est l'analyse des fichiers CSV. Si un fichier ne respecte pas les formats prédéfinis, des messages indiquant le type d'erreur apparaissent, comme indiqué dans la section 1.2.1.

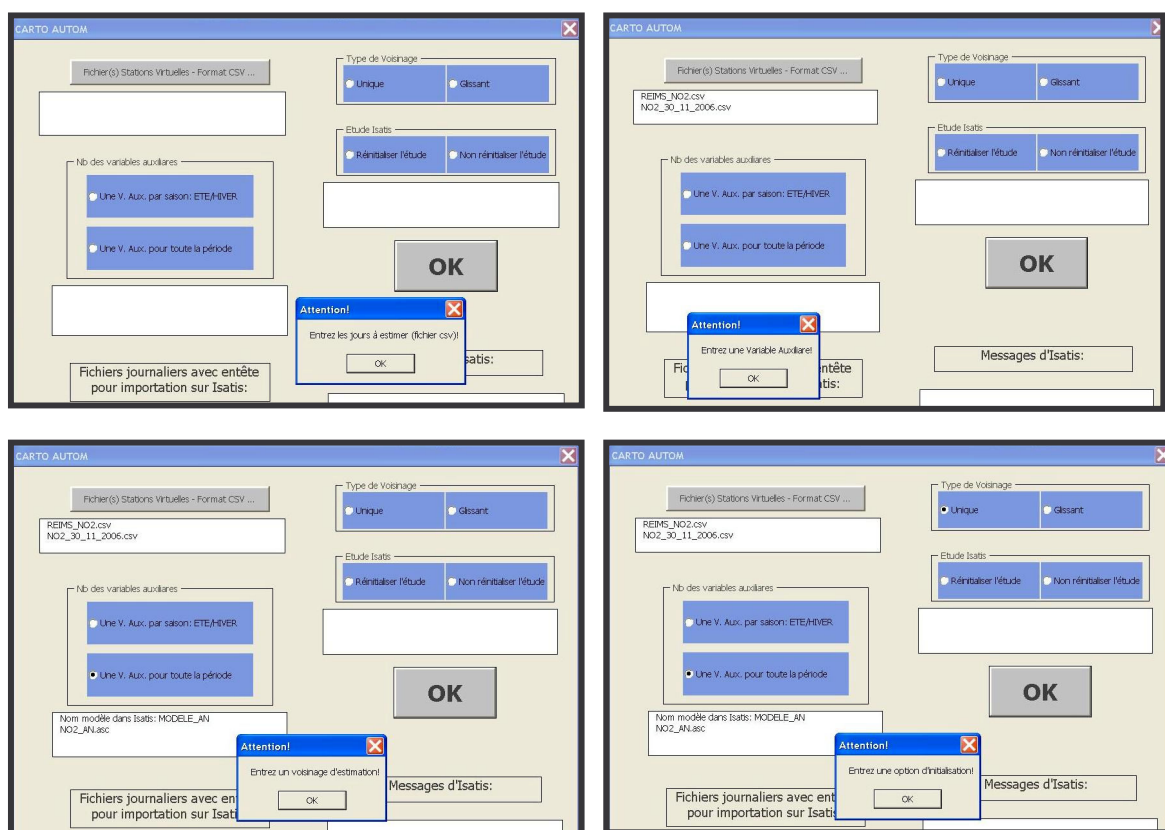


Figure 58 : Les 4 boîtes de dialogue « Attention » qu'apparaissent lors de la sélection du bouton « OK » de l'interface principale dans le cas où une information est manquante

### 1.2.5.1 CREATION DES FICHIERS AVEC ENTETE

Après l'analyse des informations, le programme insère d'abord un entête (« header » sur Isatis) à chaque fichier CSV valable, ensuite rajoute l'expression « \_hd » dans la dernière partie du nom du fichier et l'enregistre dans le même répertoire du fichier CSV original.

Les noms des fichiers ainsi créés sont affichés avec le chemin d'accès dans la boîte de dialogue dénommée : « Fichiers journaliers avec entête pour importation sur Isatis :» de la partie inférieure gauche de l'interface (cf. Figure 59)

L'entête inséré contient les paramètres nécessaires à l'importation du fichier sur Isatis. Celui-ci diffère selon l'option d'initialisation de l'étude sur Isatis.

Dans le cas d'une réinitialisation d'étude, l'entête du premier fichier CSV à traiter doit contenir les caractéristiques de tous les 9 champs du fichier. L'Annexe 8 affiche un exemple d'entête d'un fichier CSV chargé de réinitialiser une étude sur Isatis.

S'il ne s'agit pas du premier fichier à importer sur Isatis, l'entête ne comportera que les renseignements sur les concentrations de NO<sub>2</sub>. L'Annexe 9 montre un exemple d'entête d'un fichier CSV qui sera importé dans une étude existante.

Si un fichier avec le même nom existe dans le répertoire cible, celui-ci est écrasé et remplacé par le nouveau fichier CSV créé.

Les fichiers CSV finals peuvent être consultés a posteriori, ils peuvent être utiles par exemple lors d'une utilisation manuelle du logiciel Isatis.

#### 1.2.5.2 CREATION DES FICHIERS IJNL

Après la création des fichiers CSV avec entête, le programme procède à l'élaboration de ceux qui permettent le lancement d'Isatis en mode batch. Pour être reconnus par Isatis, ces dernières doivent avoir l'extension « IJNL ». Ils ressemblent toutes les opérations qu'Isatis doit effectuer pour mener à bien les estimations demandées.

Les fichiers IJNL sont stockés dans un sous-répertoire spécialement créé à cette fin. Celui-ci est nommé « BATCH\_RES » et est localisée dans le répertoire où se trouve le fichier CSV original.

Si le sous-répertoire est déjà présent, le programme insère les fichiers IJNL dans celui-ci. Dans le cas où un fichier avec le même nom existerait dans le répertoire « BATCH\_RES », celui-ci serait écrasé et remplacé par le nouveau fichier IJNL créé.

Les instructions contenant le fichier IJNL dépendent également du type d'initialisation de l'étude Isatis, voici la façon de les identifier:

- Le fichier chargé d'initialiser une étude porte comme nom celui du fichier CSV avec l'expression « \_Batch\_init.ijnl » dans la dernière partie du nom.
- Pour les fichiers qui seront lancées dans une étude déjà existante, l'expression « \_Batch\_run.ijnl » est rajoutée dans la dernière partie du nom du fichier.

Les noms des fichiers « IJNL » créés sont également affichées avec le chemin d'accès sur l'interface dans la boîte de dialogue « Fichiers batch pour le lancement d'Isatis : » (cf. Figure 59). Ces fichiers peuvent être consultés a posteriori, ils peuvent être utiles par exemple lors d'une utilisation manuelle du logiciel Isatis.

Un exemple des premières lignes d'un fichier du type « BATCH\_INIT.IJNL » est affiché dans l'Annexe 10. L'aide en ligne d'Isatis (touche « F1 » de l'écran lorsque l'on travaille sur le logiciel) fourni une description détaillée de la syntaxe des instructions des fichiers « batch ». Dans l'Annexe 11 se trouve également un exemple des premières lignes d'un fichier du type « BATCH\_RUN.IJNL ».

Le fichier d'initialisation comporte des instructions d'effacement, de création et de restauration d'une étude sur Isatis, le fichier « RUN » commence directement par l'importation des données sur une étude existante.

### 1.2.5.3 LANCEMENT AUTOMATIQUE D'ISATIS

Après la création des fichiers de lancement d'Isatis, l'interface affiche une boîte de dialogue renseignant le nombre total de fichiers IJNL créés et demandant à confirmer le lancement d'Isatis (cf. Figure 59).

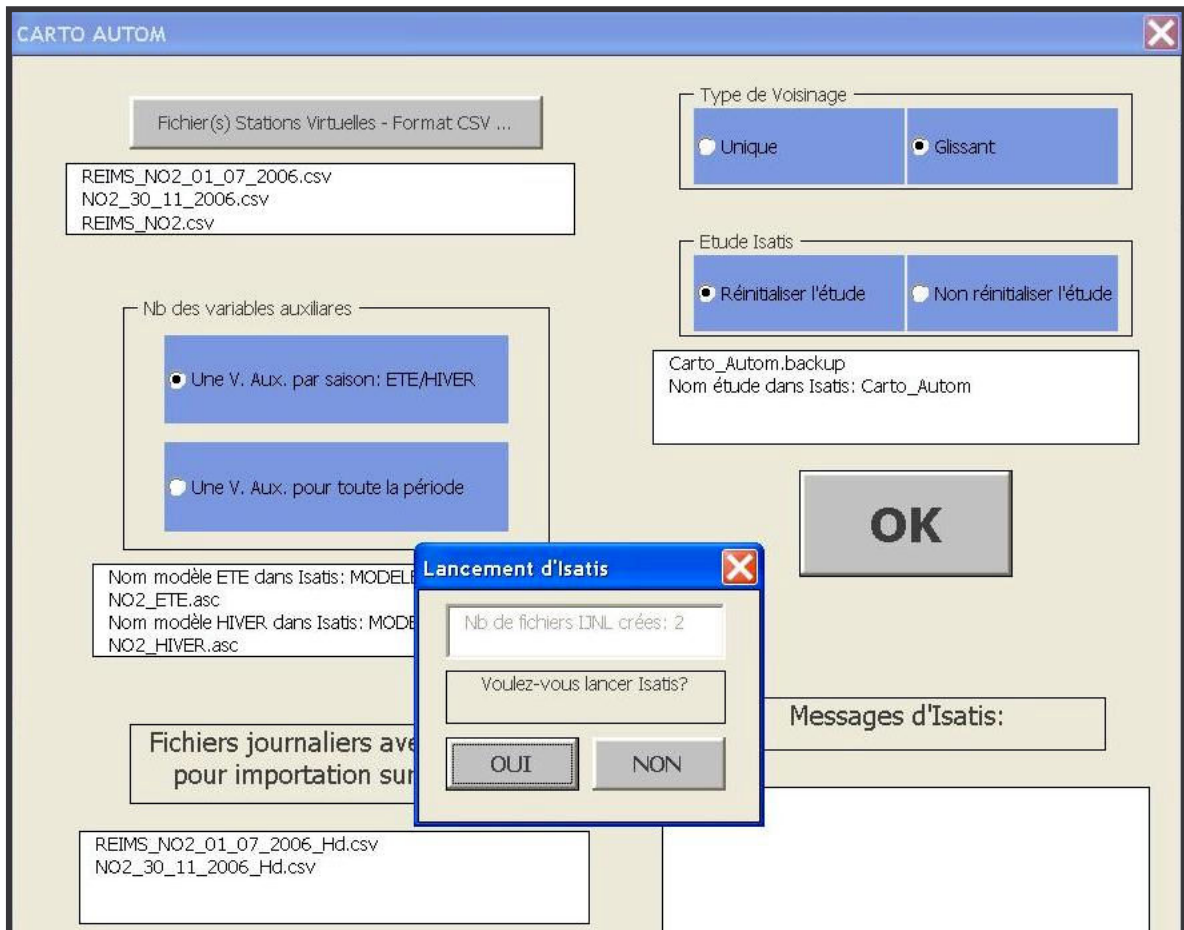


Figure 59 : Boîte de dialogue « Lancement d'Isatis »

Le bouton « NON » fait apparaître une autre boîte qui donne la possibilité de relancer ou de sortir de l'interface (cf. Figure 60).

Si l'on souhaite lancer automatiquement les interpolations et que l'on dispose de la version Windows d'Isatis, l'utilisateur peut sélectionner le bouton « OUI ». Dans ce cas le logiciel cherche automatiquement le fichier exécutable, s'il ne le trouve pas, une boîte de dialogue permettant de le sélectionner manuellement apparaîtra à l'écran (cf. Figure 61).

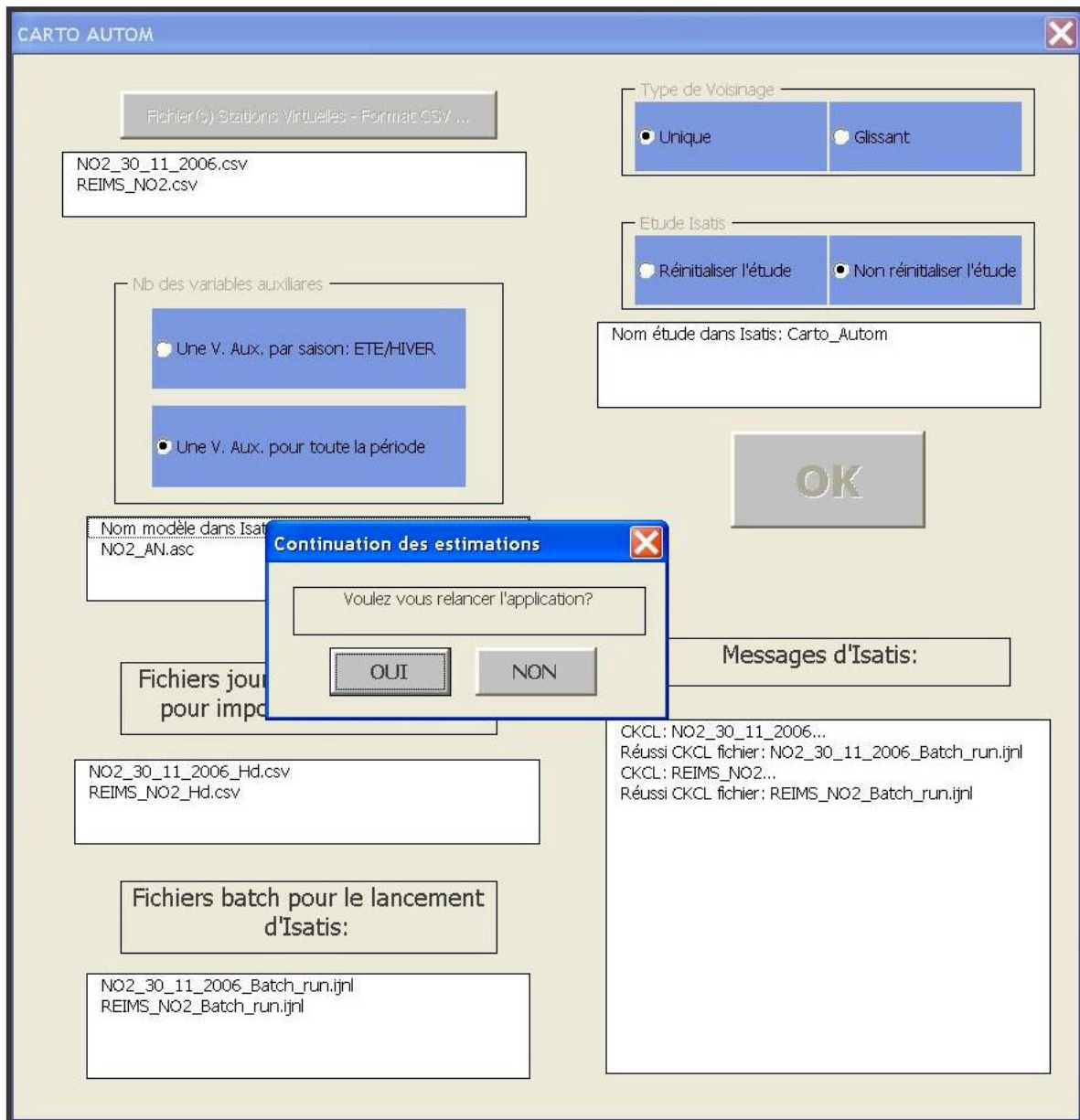


Figure 60 : Boite de dialogue : « Continuation des estimations »

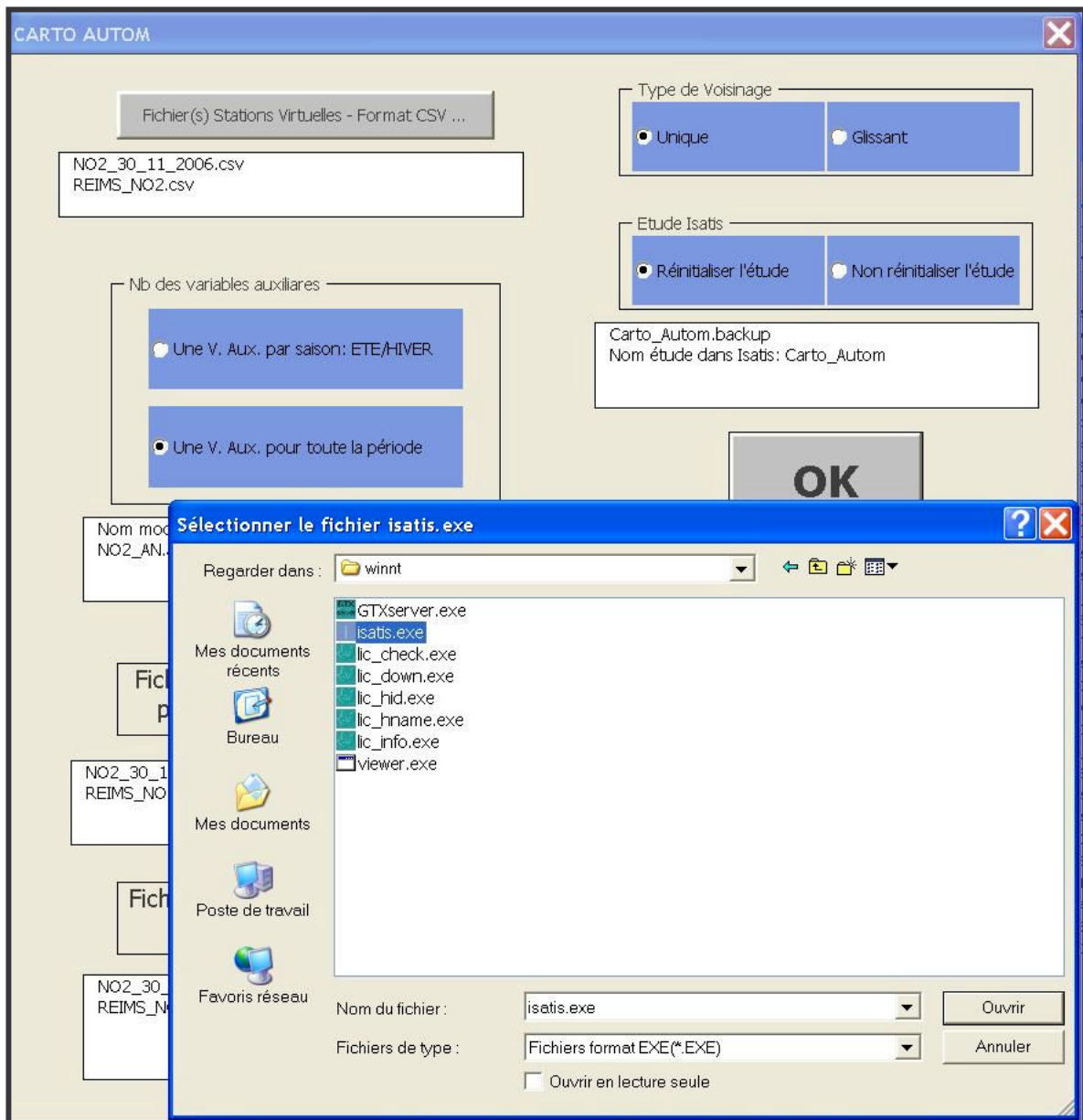


Figure 61 : Boite de dialogue « Sélectionner le logiciel isatis.exe »

Si l'utilisateur sort de cette fenêtre sans sélectionner un fichier, le message : «Le chemin d'accès à Isatis est incorrect, Voulez-vous répéter ? » apparaît (cf. Figure 62).



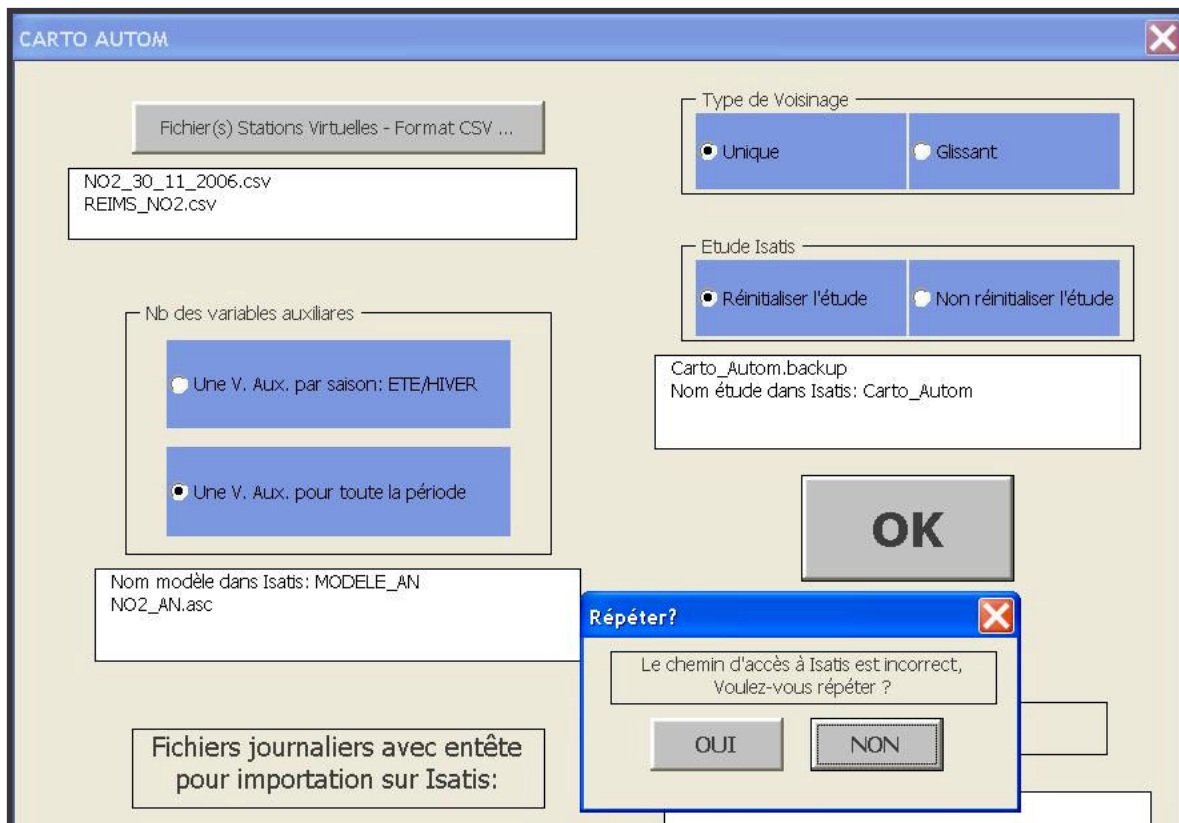


Figure 62 : Boîte de dialogue : Répéter ?

Le bouton « OUI » de la boîte « Répéter ? » permet de sélectionner le fichier exécutable d'Isatis (cf. ; Figure 61), le bouton « NON » affiche la boîte de dialogue de la Figure 60, demandant à relancer ou à sortir de l'interface.

Si le logiciel Isatis est trouvé, deux actions sont réalisées :

- D'abord un fichier « carto\_autom\_isatis.ini » avec une section « [Logiciel Isatis] » et une clé « Chemin d'accès= » est créé dans la racine du disque dur (« C:/ »). Le programme remplit cette variable avec le chemin d'accès du logiciel Isatis, par exemple :

« [Logiciel Isatis] »

« Chemin d'accès=D:\Logiciels\Geovariances\Isatis\bin\winnt\isatis.exe »

L'objectif de ce fichier INI est de ne pas demander cette information à chaque lancement de l'interface. Lors de la prochaine exécution du programme, ce fichier INI est lu et le logiciel Isatis est exécuté automatiquement.

- Ensuite le programme lance le premier fichier IJNL, il attend la finalisation des actions sur Isatis, puis lance le fichier suivant et ainsi de suite jusqu'à l'exécution de tous les fichiers IJNL créés.

Pendant la réalisation des interpolations sur Isatis, quelques messages s'affichent dans la boîte de dialogue « Messages d'Isatis » localisée dans la partie inférieure droite de l'interface.

Le premier message « CKCL : *NOM\_FICHER\_CSV...* » indique que l'interface a lancé une estimation par cokrigage colocalisé et qu'elle attend les résultats d'Isatis (cf. Figure 63 et Figure 64).

Si les estimations ont été réalisées le message : « CKCL Réussi (*Nom du fichier batch lancé*) » est affiché (cf. Figure 63), si une erreur empêchant la réalisation des estimations s'est produit sur Isatis, le message : « CKCL Echec (*Nom du fichier batch lancé*) » apparaît (cf. Figure 64).

Les sous-répertoires RES et MESSAGES sont créés par l'interface afin d'y stocker les résultats et les messages d'Isatis. Si les estimations sont réalisées, l'utilisateur est invité à consulter dans le répertoire RES les fichiers produits par Isatis.

Deux fichiers texte l'un contenant les messages d'information produits par Isatis et l'autre le temps de calcul de chaque opération, sont enregistrés dans le répertoire MESSAGES. Il s'agit des messages qui s'affichent dans les fenêtres « Isatis Messages » et « CPU TIME » lors de l'utilisation manuelle du logiciel.

Le nom du première fichier crée correspond au nom du fichier CSV original plus l'expression « *\_isatis.txt* » dans la dernière partie, l'Annexe 13 montre un exemple des premières lignes de ce fichier texte. Pour le deuxième fichier l'expression « *\_isatis\_cpu\_timer.txt* » a été rajouté, l'Annexe 14 montre un exemple de ce fichier texte.

Dans le cas où une estimation n'est pas réalisée, les messages d'erreur d'Isatis sont consignés dans le premier fichier texte. Pour les interpréter il est conseillé de consulter les documents de référence fournis par « GEOVARIANCES » société éditeur du logiciel : l'aide en ligne, le support et les notices d'utilisation.

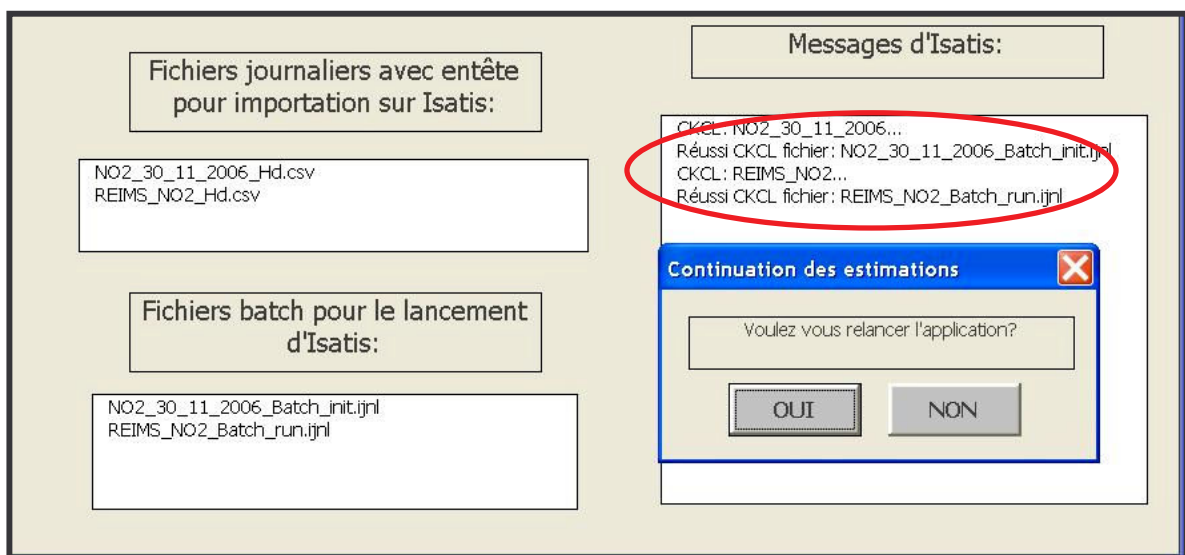


Figure 63 : Exemples des messages lorsque les estimations ont été réalisées par Isatis

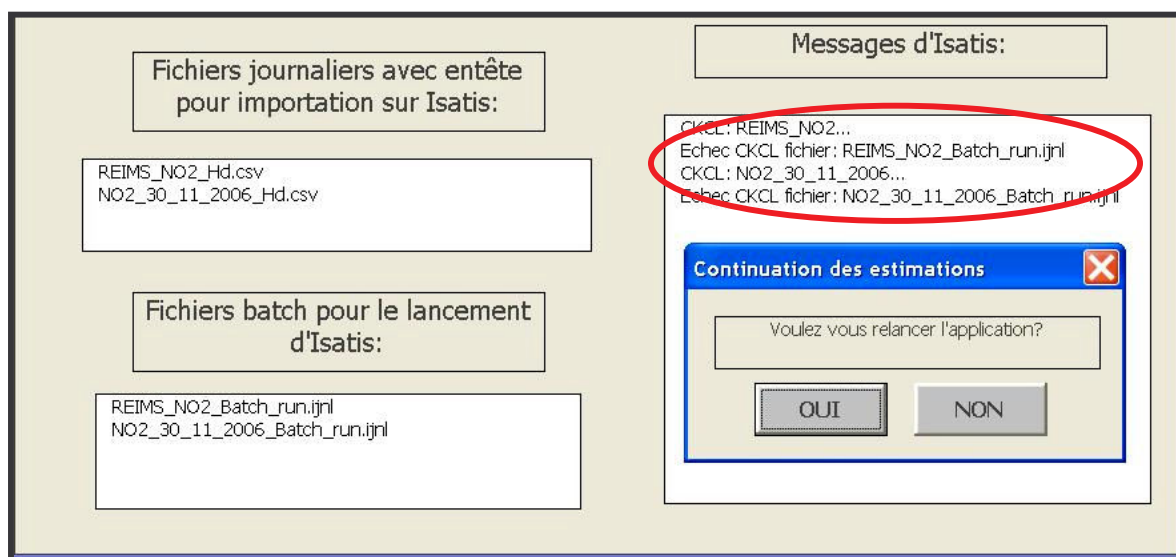


Figure 64 : Exemples des messages lorsque les estimations n'ont pas été réalisées par Isatis

La durée des opérations sur Isatis dépend de l'étendu du domaine d'étude et de la taille de la maille d'estimation. Par exemple, pour le cas d'étude sur Reims, l'exécution d'Isatis dure 14 secondes (cf. l'Annexe 14). Si on effectue des estimations sur une maille très fine, par exemple dans le cas où l'on appliquerait la méthode d'imbrication d'échelle pour calculer la pollution à proximité des grandes routes (cf. 6), le temps de calcul sera plus important.

Pendant qu'Isatis tourne en mode batch, l'interface reste bloquée jusqu'à la finalisation des estimations du dernier fichier IJNL lancée. Cependant l'utilisateur peut effectuer d'autres opérations sur l'ordinateur (p. ex. lancer une autre application ou consulter les fichiers déjà créés par Isatis avec l'Explorateur Windows).

Lorsque les calculs sont finis, la boîte de dialogue « Continuation des estimations » s'affiche, donnant la possibilité de relancer ou de sortir de l'interface (cf. Figure 63 et Figure 64).

#### 1.2.5.4 EXPORTATION DES RESULTATS D'ISATIS

Après chaque estimation réussie, Isatis enregistre quatre fichiers dans le sous-répertoire « RES ». Le nom des fichiers est composé du nom du fichier initial (sans l'extension « CSV ») avec l'expression « ...\_CKCL » dans la dernière partie du nom.

Les fichiers présentent les caractéristiques suivantes :

- Une image TIFF représentant une carte des concentrations de NO<sub>2</sub> calculées sur la zone d'étude. L'échelle des couleurs « arc-en-ciel inversée » d'Isatis à été choisi (bleu pour les faibles concentrations et rouge pour les fortes concentrations). L'échelle s'adapte automatiquement aux valeurs minimales et maximales de chaque estimation.

De la même façon que le lancement du fichier exécutable d'Isatis, l'affichage de cette image a été aussi automatisé. Le programme vérifié automatiquement l'existence d'un programme exécutable associé à l'extension TIFF, s'il ne le trouve pas, une boîte de dialogue permettant de le sélectionner est affichée (Figure 65), si le logiciel d'affichage d'images est renseigné, deux actions sont réalisées :

- D'abord, dans le fichier « carto\_autom\_isatis.ini » une nouvelle section « [Logiciel exécution fichiers TIFF] » est créée et la clé « Chemin d'accès= » est renseignée avec le chemin d'accès du logiciel d'images, par exemple :

« [Logiciel exécution fichiers TIFF] »

« Chemin d'accès=C:\Program Files\IrfanView\i\_view32.exe »

Lors de la prochaine exécution du programme, ce fichier INI est lu et l'affichage de l'image est réalisé automatiquement.

- Ensuite le programme affiche automatiquement chaque TIFF qui est exporté par Isatis (Figure 66).
- Une grille SIG (extension ASC) représentant les concentrations de NO<sub>2</sub> calculées sur la zone d'étude. Pour visualiser cette grille, l'on doit disposer d'un logiciel SIG (ArcView, MapInfo, etc.). Ce fichier peut être aussi réimporté sur Isatis car les coordonnées sont incluses. Un exemple des premières lignes d'un fichier ASC est montré dans l'Annexe 12.
  - Deux tables (extension CSV) contenant les principales statistiques des résultats. Ces fichiers sont accessibles à partir d'un logiciel de traitement de texte (NOTEPAD, WORD, XEMACS, etc.) ou par un tableur (EXCEL, etc.). Les formules utilisées pour le calcul des statistiques sont affichées dans le tableau de l'Annexe 15.
    - Le premier fichier comprend quelques statistiques calculées d'une part sur l'ensemble des stations (virtuelles et fixes) et d'autre part sur l'ensemble des stations fixes seulement. Deux lignes et 28 colonnes forment ce fichier. La première ligne est réservée aux noms des champs et la seconde aux valeurs numériques des statistiques. La dernière partie du nom de ce fichier est « ...\_STATS\_ALL\_CKCL.CSV », une description détaillée des champs est renseignée dans l'Annexe 16.
    - L'autre fichier comporte quelques statistiques permettant de comparer l'estimation avec la concentration mesurée dans chaque station fixe. Ce fichier est composé de 16 colonnes, et le nombre de lignes est égal au nombre de stations fixes présentes dans la zone d'étude plus la première ligne qui est réservée au nom des champs. Ce fichier peut être aussi réimporté sur Isatis car les coordonnées sont incluses. La dernière partie du nom de ce fichier est « ...\_STATS\_PAR\_SF\_CKCL.CSV », une description détaillée des champs est renseignée dans l'Annexe 17.

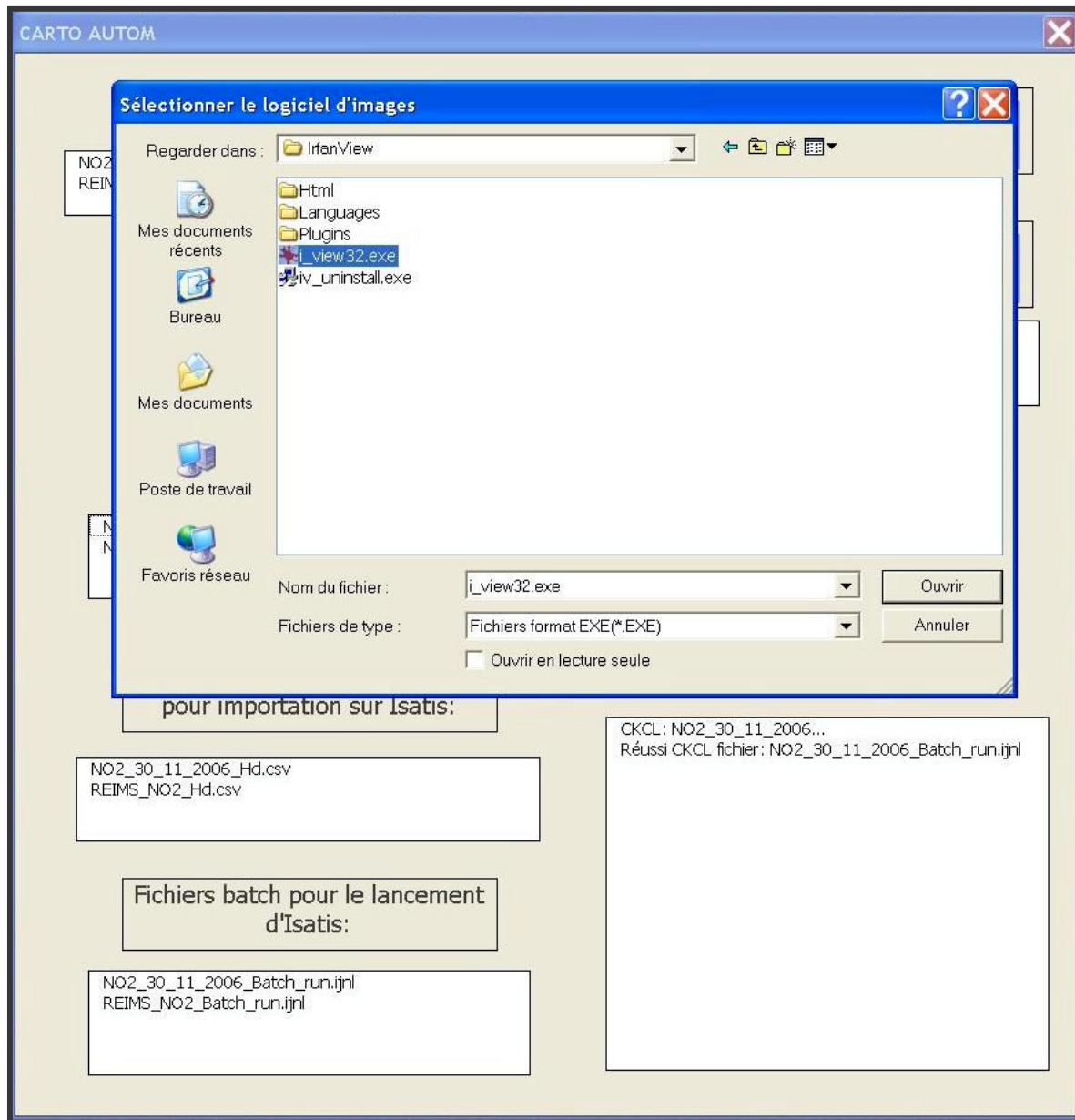


Figure 65 : Boîte de dialogue « Sélectionner le logiciel d'images »

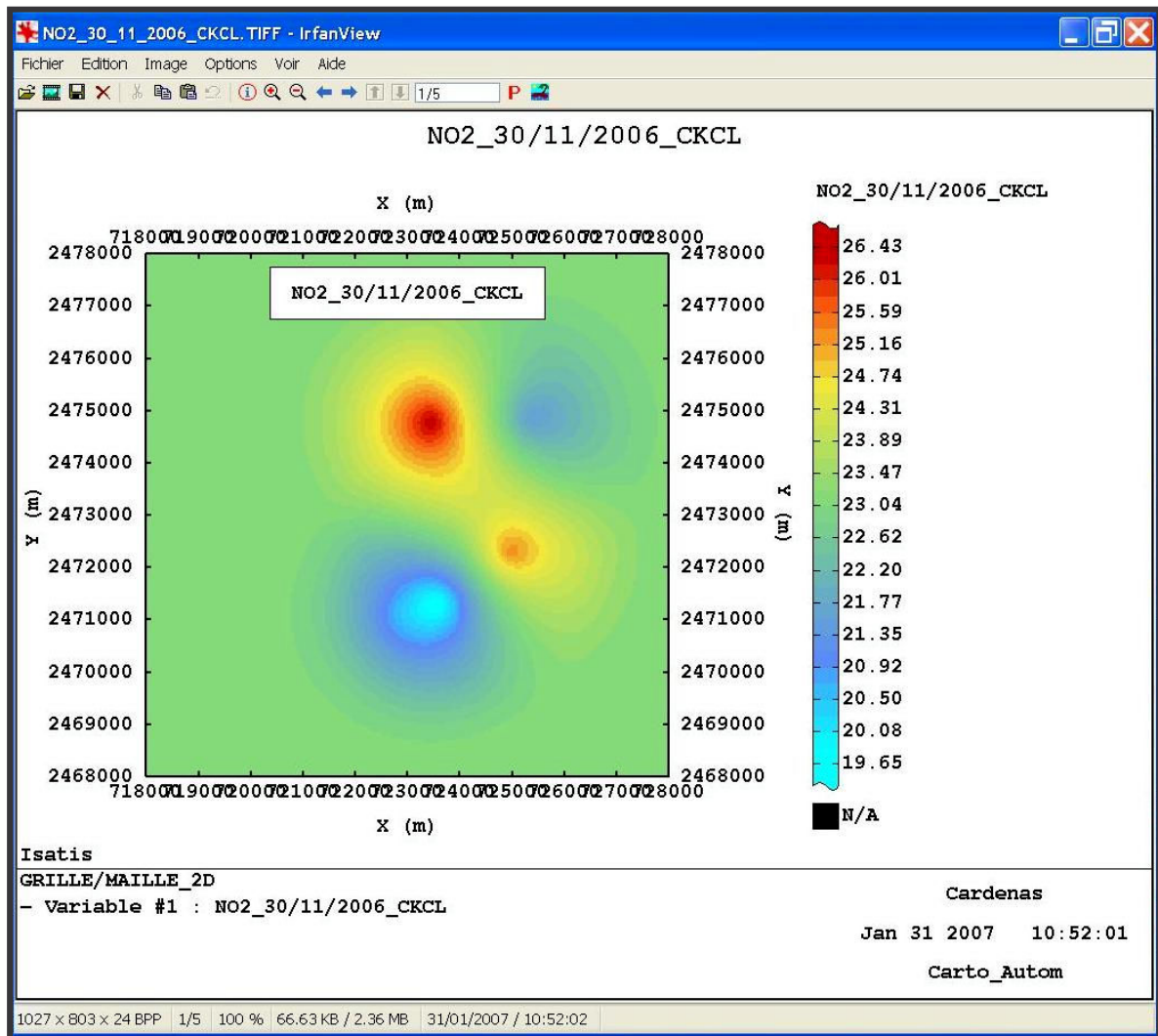


Figure 66 : Exemple d'affichage automatique d'un fichier TIFF

## Annexe 3 : Format des fichiers outil Excel: Corrélation Tubes- Stations

- Fichier « tubes.xls »

NOM	X	Y	23/12_07/01	07_20/01	20/01_03/02	03_17/02
1	722850	2474714	26	43	36	40
2	723104	2474790	36	47	49	44
3	723450	2474768	30	44	39	37
4	723758	2474707	31	46	40	43
5	724094	2474808	30	40	40	38
6	724282	2474816	31	43	40	40
8	724024	2474384	27	40	40	37
9	723826	2474385	29	43	43	41
10	723409	2474506	36	45	47	45

Le fichier doit comporter l'entête correspondante :

NOM, X, Y, Campagne n°1,....., campagne n°n.

Descriptif :

COL1 : NOM, nom du site de prélèvement de l'échantillonneur passif

COL2 : X, position géographique x du site

COL3 : Y, position géographique y du site

COL4 : 23/12\_07/01, valeur de la concentration pour le prélèvement

- Fichier « stations.xls »

NOM	X	Y	23/12_07/01	07_20/01	20/01_03/02	03_17/02
SF1	725174	2474799	26	43	37	37
SF2	722747	2471002	16	40	27	34
SF3	724769	2472098	28	49		35
SF4	723376	2474728	34	56	45	48
SF6	723494	2471203	15	36	27	31
SF7	725838	2473231	12	40	28	27

Le fichier doit comporter l'entête correspondante :  
 NOM, X, Y, Campagne n°1,....., campagne n°n.

Descriptif :

- COL1 : NOM, nom de la station de mesure  
 COL2 : X, position géographique x de la station de mesure  
 COL3 : Y, position géographique y de la station de mesure  
 COL4 : 23/12\_07/01, valeur de la concentration moyenne mesurée



- Fichier « virtuelles\_param.csv »

```
NOM;X;Y;VEM;SEL_SF;SFREF;M;B;RHO
SF1;725174;2474799;N/A;1.00;N/A;N/A;N/A;N/A
SF2;722747;2471002;N/A;1.00;N/A;N/A;N/A;N/A
SF3;724769;2472098;N/A;1.00;N/A;N/A;N/A;N/A
SF4;723376;2474728;N/A;1.00;N/A;N/A;N/A;N/A
SF6;723494;2471203;N/A;1.00;N/A;N/A;N/A;N/A
SF7;725838;2473231;N/A;1.00;N/A;N/A;N/A;N/A
1;722850;2474714;11.19;0.00;SF2;0.81;11.67;0.95
4;723758;2474707;19.75;0.00;SF2;0.70;18.68;0.94
6;724282;2474816;11.34;0.00;SF1;0.81;8.56;0.94
8;724024;2474384;46.26;0.00;SF3;0.59;12.54;0.94
9;723826;2474385;22.38;0.00;SF6;0.78;16.67;0.91
16;723722;2474154;55.55;0.00;SF3;0.50;18.60;0.94
17;724004;2474184;31.59;0.00;SF2;0.60;18.77;0.91
...
```

Le fichier comprend deux parties :

- la partie haute concerne les stations fixes
- la partie basse concerne les sites de prélèvement.

Descriptif :

COL1 : NOM, nom de la station ou du site d'échantillonnage passif  
 COL2 : X, position géographique x de la station ou du site  
 COL3 : Y, position géographique y de la station ou du site  
 COL4 : VEM, variance de l'erreur d'estimation  
 « N/A » pour une station de mesure (car VEM non calculée)  
 COL5 : SEL\_SF, cette variable vaut « 0 » pour les sites d'échantillonnage  
 et « 1 » pour les stations de mesures  
 COL6 : SREF, indique le nom de la station de mesure corrélée au site  
 d'échantillonnage, « N/A » pour une station de mesure.  
 COL7 : M, pente, « N/A » pour une station de mesure  
 COL8 : B ordonnée à l'origine, « N/A » station de mesure  
 COL9 : RHO, coefficient de corrélation ( r ), « N/A » station de mesure

- Fichier d'entrée « NO2\_23\_12\_2001.XLS »

NOM	X	Y	NO2_23_12_2001
SF1	725174	2474799	26
SF2	722747	2471002	16
SF3	724769	2472098	28
SF4	723376	2474728	34
SF6	723494	2471203	15
SF7	725838	2473231	12

- COL1 : NOM, nom de la station  
COL2 : X, position géographique x de la station  
COL3 : Y, position géographique y de la station  
COL4 : NO2\_23\_12\_2001, concentrations pour la période

*Annexe 4 : Application de la méthode de cokrigage temporel pour l'estimation de la moyenne annuelle de la concentration de NO<sub>2</sub> dans l'agglomération lilloise*

Il faut d'abord noter qu'en matière de pollution atmosphérique on ne considère généralement que deux saisons : l'hiver, du 1<sup>er</sup> Octobre au 31 Mars, et l'été, du 1<sup>er</sup> Avril au 30 Septembre.

Pour appliquer la méthode, Il faut disposer de plusieurs campagnes de mesure par tubes passifs sur des périodes représentatives de chaque saison :

- Pour l'hiver : au moins trois campagnes de mesure (p. ex. en janvier et février).
- Pour l'été : au moins trois campagnes de mesure (p. ex. en juin et juillet).

On calcule d'abord la moyenne des campagnes de chaque saison pour obtenir la moyenne saisonnière, puis on vérifie s'il y a une forte corrélation entre les mesures des deux saisons (coefficient de corrélation  $\geq 0.6$ ).

Comme la corrélation ne peut se calculer que sur les échantillons représentés dans les deux saisons, il faut avoir un nombre considérable d'échantillons communs aux deux saisons. Si la corrélation est forte et le nombre d'échantillons communs suffisant, on peut donc effectuer un cokrigage entre les deux moyennes saisonnières.

Si on dispose de l'information auxiliaire, on peut chercher à construire une variable auxiliaire bien corrélée à chacune des moyennes saisonnières, de cette façon les estimations seront améliorées. Les informations auxiliaires sont les plus souvent : la densité de population, le cadastre d'émissions NO<sub>x</sub>, la distance à la route, la densité du bâti, les conditions météorologiques, la présence d'ozone, 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 nécessaire de construire une variable auxiliaire spécifique à chaque saison.

Ensuite on effectue le cokrigage de chaque moyenne saisonnière à l'aide de la moyenne de la saison opposée et de la variable auxiliaire si elle a été préalablement définie.

Un avantage de la méthode est le fait que l'on peut conserver (et donc exploiter) les sites n'étant pas mesurés que dans une saison. D'ailleurs on peut aussi calculer la variance d'estimation de chaque cokrigage, cette variable est une mesure qualitative de la qualité de l'estimation de chaque moyenne saisonnière. Le cokrigage suppose l'ajustement de deux variogrammes simples (un pour chaque saison) et du variogramme croisé.

Finalement la moyenne arithmétique de deux cokrigages saisonnières est calculé, la variable résultante est la moyenne annuelle du NO<sub>2</sub>. La variance d'estimation de la moyenne annuelle ne peut pas être calculée de la même façon car la variance n'est pas une variable additive.

Pour déterminer la qualité des résultats, une validation croisée des deux moyennes saisonnières sur les échantillons est effectuée, ensuite on calcule la

moyenne de deux estimations par validation croisée et on trace le nuage de corrélation avec la moyenne annuelle des mesures. P. ex. dans l'étude du LCSQA 2004 le coefficient de corrélation de ce nuage est de 0.97, le coefficient calculé sur les résultats de la validation croisée par krigeage en dérive externe sur la moyenne annuelle a été plus faible : 0.82, ces résultats démontrent l'amélioration apportée par le cokrigeage.

Le Tableau 38 affiche les statistiques des campagnes par tubes passifs menées 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 des 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	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	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
<b>Moyenne été 2003</b>	<b>22-mai-03</b>	<b>3-juil-03</b>	<b>137</b>	<b>15.2</b>	<b>59.07</b>	<b>25.91</b>	<b>6.89</b>	<b>47.47</b>	<b>0.266</b>

Nb Q	Date Début	Date Fin	Nb Tubes	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	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
<b>Moyenne hiver 2003-2004</b>	<b>8-janv-04</b>	<b>18-févr-04</b>	<b>136</b>	<b>21.12</b>	<b>50.43</b>	<b>34.08</b>	<b>5.36</b>	<b>28.7</b>	<b>0.157</b>

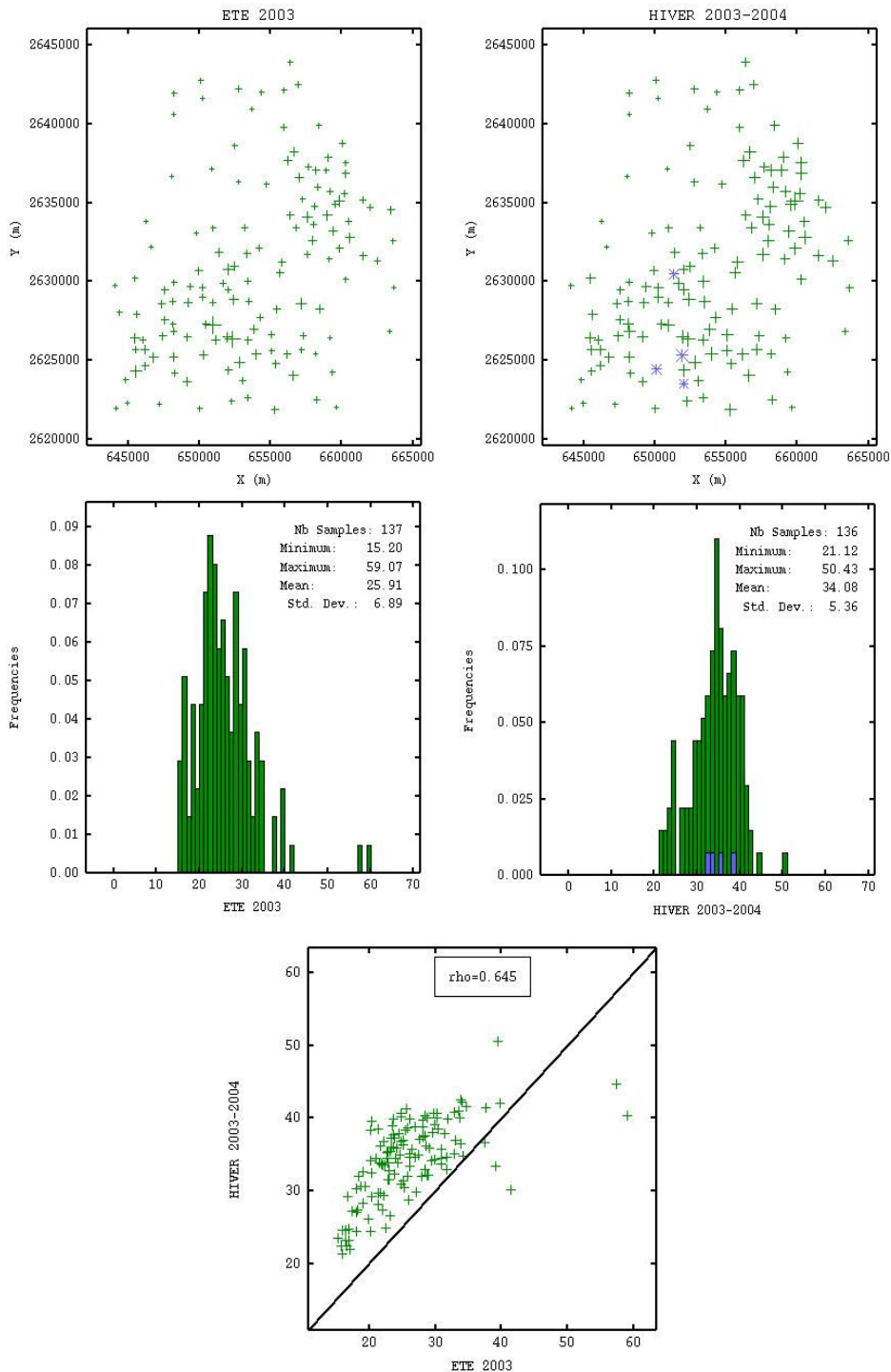
2003-2004	Date Début	Date Fin	Nb Tubes	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
<b>Moyenne Annuelle</b>	<b>22-mai-03</b>	<b>18-févr-04</b>	<b>132</b>	<b>18.46</b>	<b>50.96</b>	<b>29.93</b>	<b>5.62</b>	<b>31.58</b>	<b>0.187</b>

Tableau 38 : Statistiques des campagnes par tubes passifs menée dans l'année 2003-2004 et des moyennes saisonnières et annuelle

Pour calculer les deux moyennes saisonnières on n'a retenu que les échantillons où il y a au moins deux campagnes de mesure sur la saison. Ensuite la moyenne arithmétique des moyennes saisonnières est calculée (sur 132 points de mesure dans ce cas).

La localisation géographique et les principales statistiques de deux moyennes saisonnières sont présentées sur les cartes de la Figure 67. On observe qu'il existe une corrélation entre les deux saisons, le coefficient de corrélation (non

spatial) est de 0.645, le variogramme croisée permettra de vérifier la corrélation spatiale entre les deux saisons.



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 67 : Cartes et statistiques des deux moyennes saisonnières**

La fonction auxiliaire calculée en 2005 [4] est conservée, il s'agit de la somme des deux variables normées : la densité d'émissions et la densité de population (cf. Figure 68)

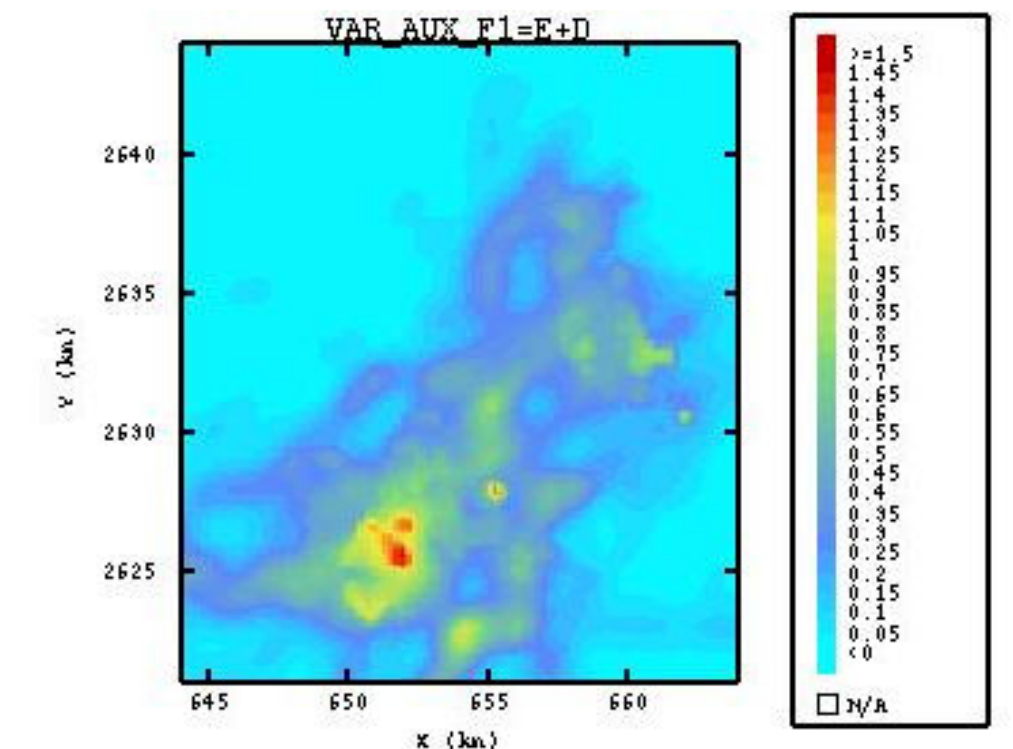


Figure 68 : Carte de la variable auxiliaire

Dans ce cas d'étude, une même fonction auxiliaire pour les estimations de deux moyennes saisonnières est retenue : les nuages de la Figure 69 montrent que les coefficients de corrélation entre la variable auxiliaire et les deux concentrations moyennes sont similaires : 0.623 pour la saison estivale et 0.617 pour la saison hivernale.

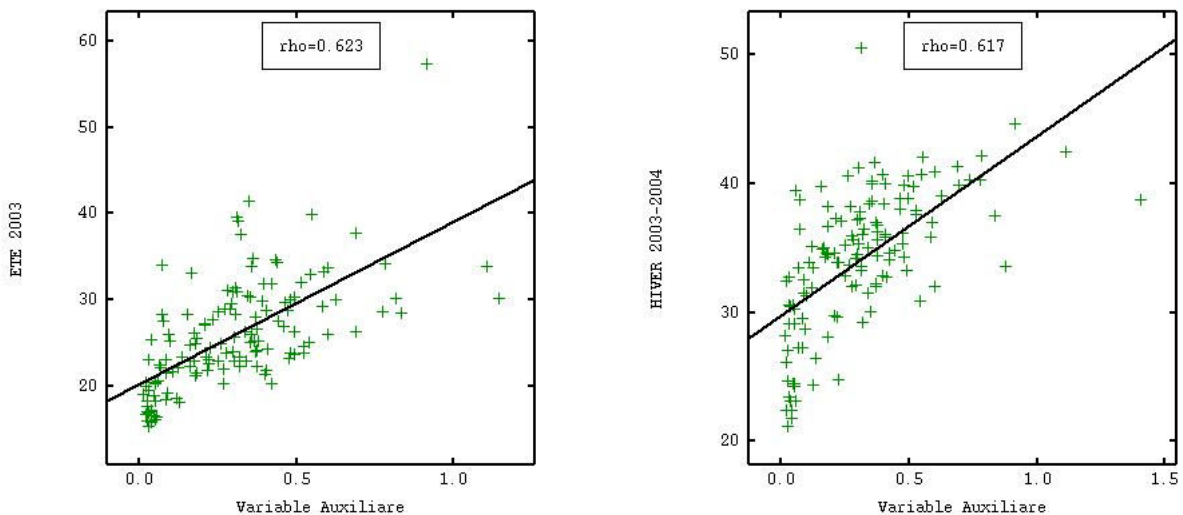


Figure 69 : Nuage de corrélation entre les deux moyennes saisonnières et la variable auxiliaire

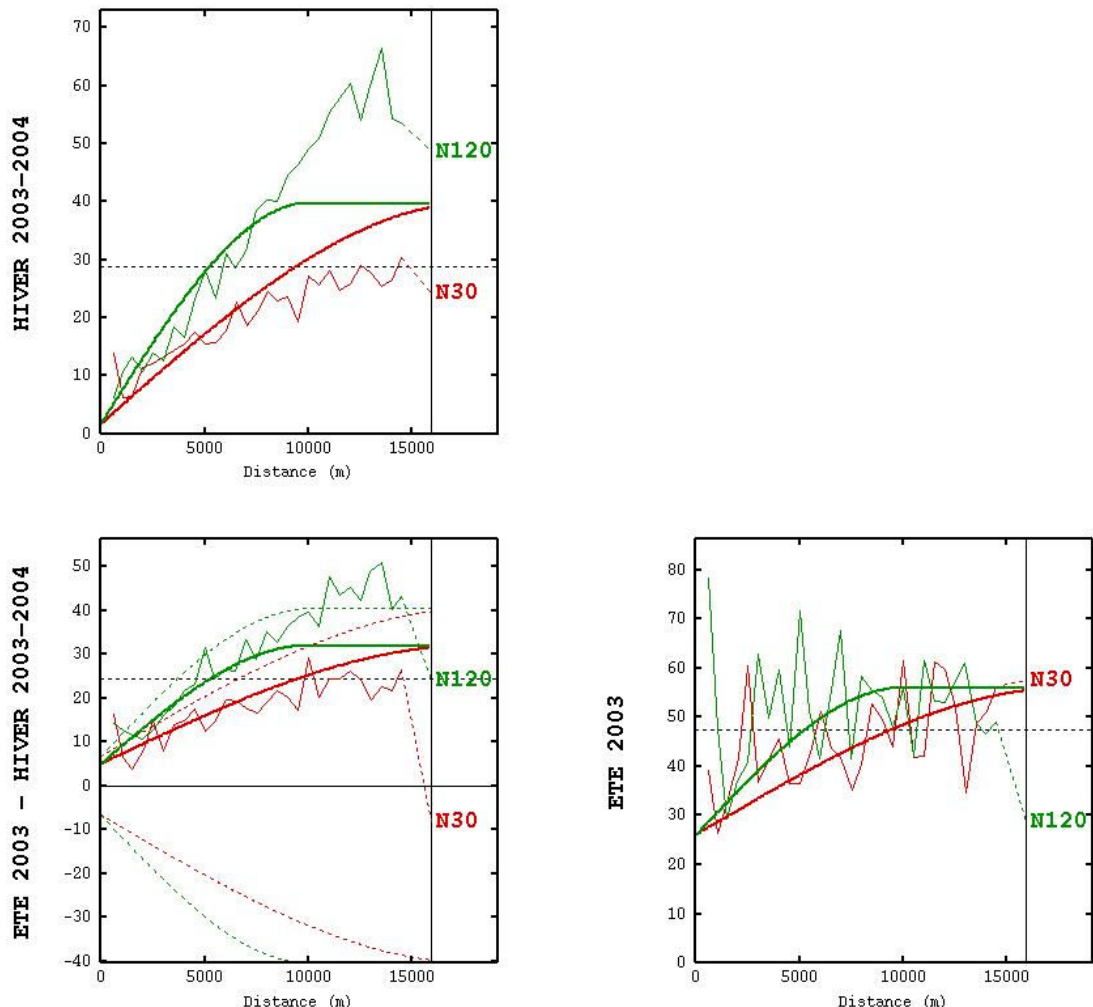
Les modèles bivariés ajustés aux variogrammes expérimentaux des concentrations saisonnières et du variogramme croisée sont affichées dans la Figure 70. Il s'agit d'un effet de pépite et d'une structure sphérique.

Dans la région Lilloise, la pollution par dioxyde d'azote ne présente pas une très bonne corrélation spatiale dans la saison estivale: variogramme expérimentale très irrégulier et modèle ajusté présentant une forte composante pépétique.

Ce différence de comportement entre l'été et l'hiver, qui a été aussi détecté sur d'autres études du LCSQA, peut s'expliquer par des conditions purement climatiques (en été, le rayonnement solaire favorise les phénomènes d'advection, propices à la dispersion des polluants).

Les résultats des estimations réalisées sur chaque saison par validation croisée sont montrées dans les Figure 71 et Figure 72. La méthode du cokrigage entre moyennes saisonnières avec comme dérive externe la fonction auxiliaire préalablement définie est utilisée. La Figure 73 montre également le nuage de corrélation calculé sur les tubes passifs entre la moyenne des estimations saisonnières et la moyenne annuelle des mesures.

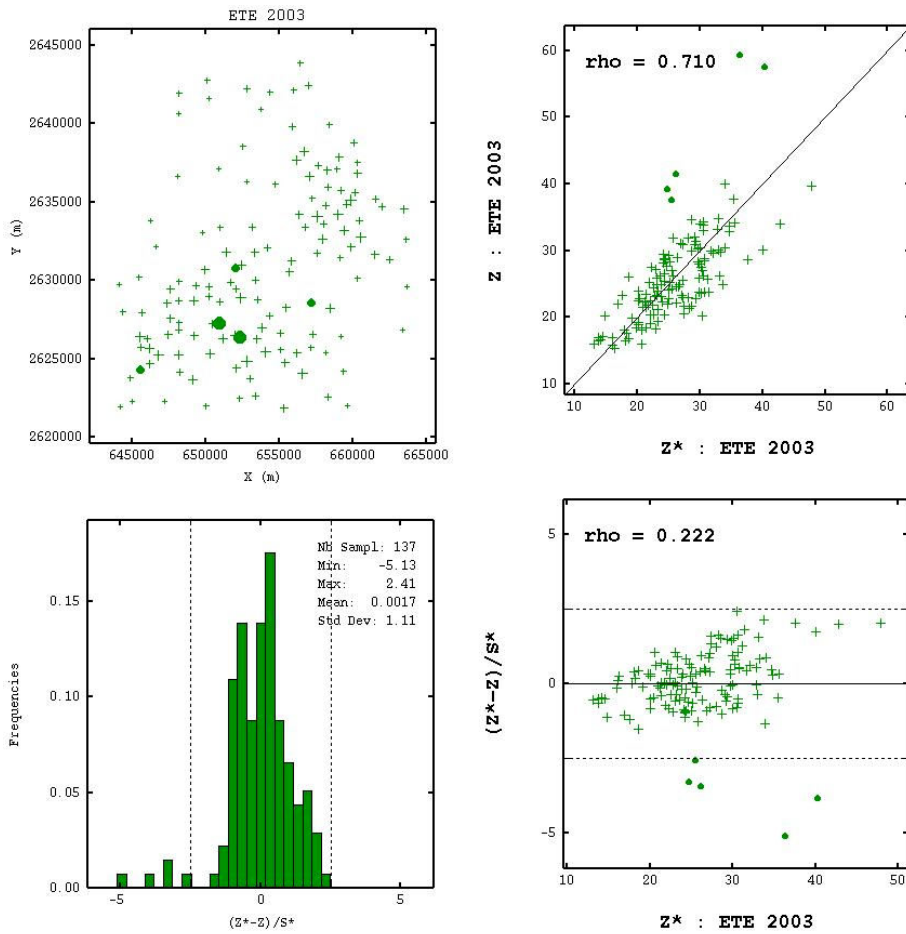
Ces graphiques confortent l'utilisation du cokrigage pour l'estimation de la moyenne annuelle : nuage très serré, points proches de la bissectrice et forte corrélation (coefficient de 0.946).



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

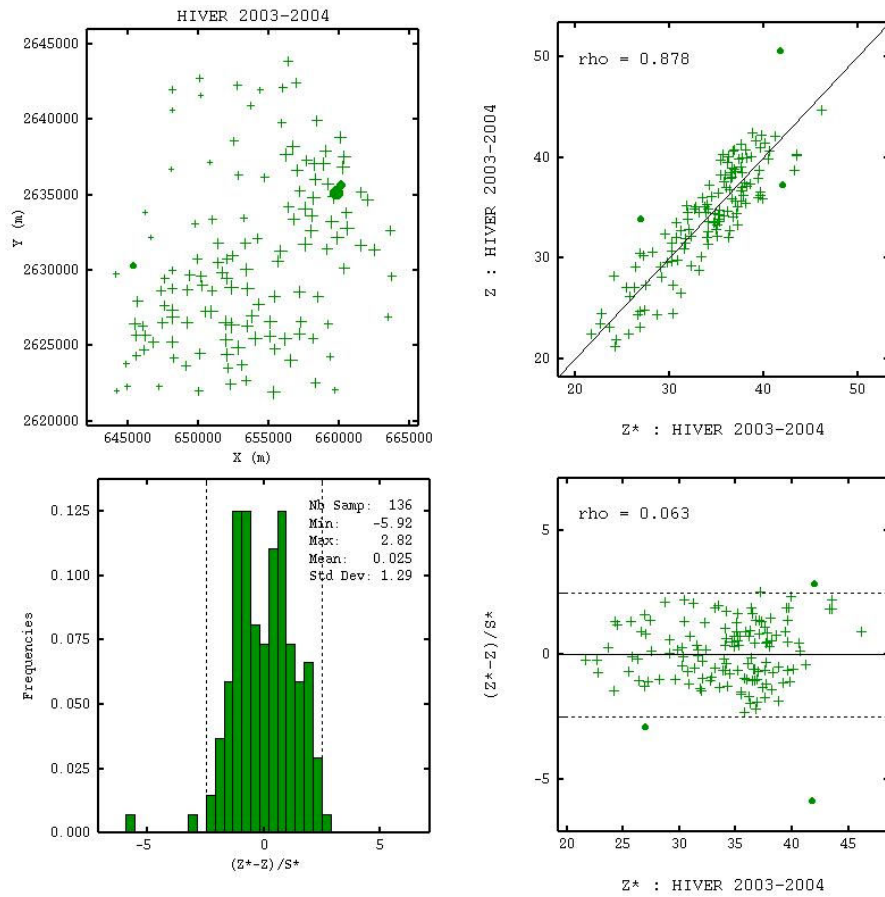
Figure 70 : Modèle des variogrammes simples et croisés des moyennes saisonnières





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

Figure 71 : Validation croisée de la moyenne estivale



Variable (Nb=136)	Moyenne	Variance
Erreur	0.1	6.62
Erreur Réduite	0.025	1.67

Figure 72 : Validation croisée de la moyenne hivernale

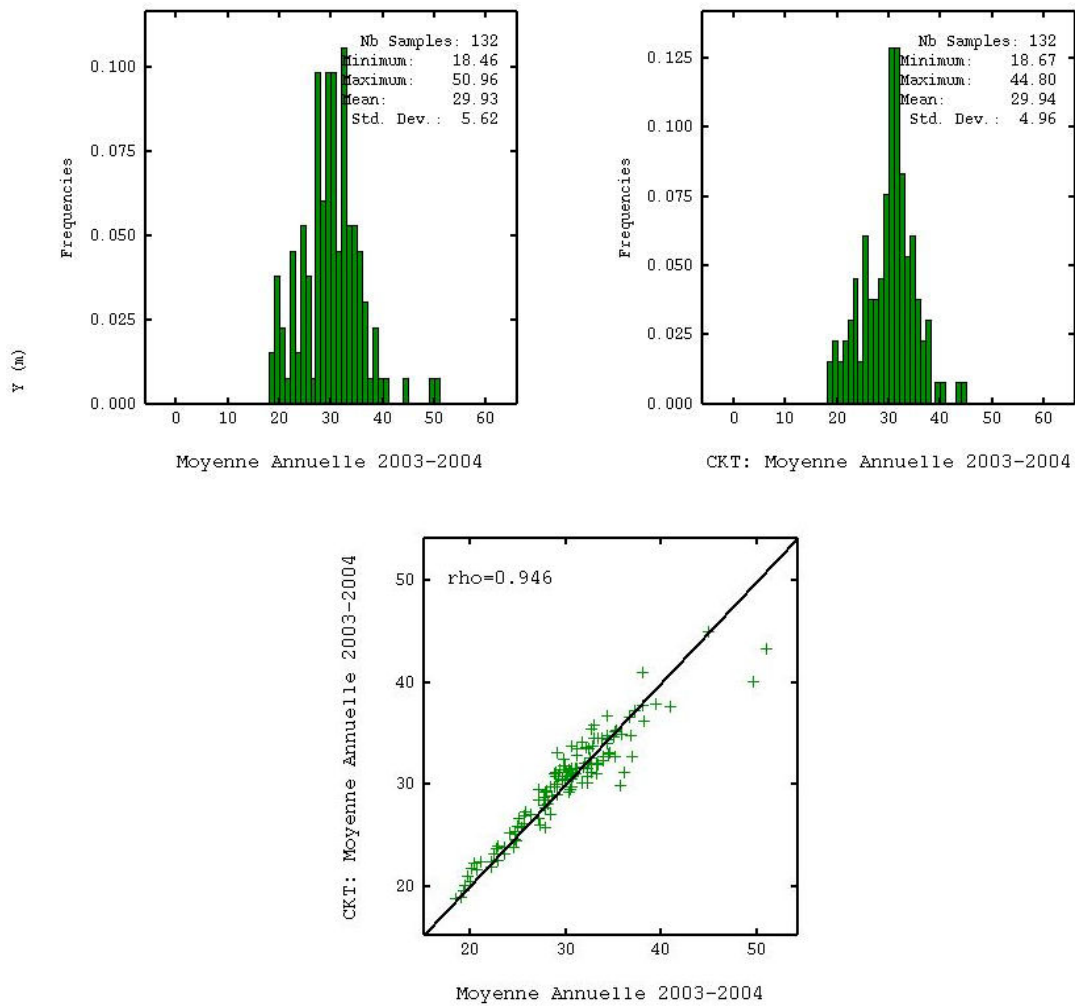


Figure 73 : Histogrammes et nuages de corrélation de l'estimation de la moyenne annuelle et de la moyenne annuelle statistique, aux tubes (année 2003-2004)

Annexe 5 : Modèles choisis pour la méthode de la cartographie automatique journalière de NO2 dans l'agglomération lilloise

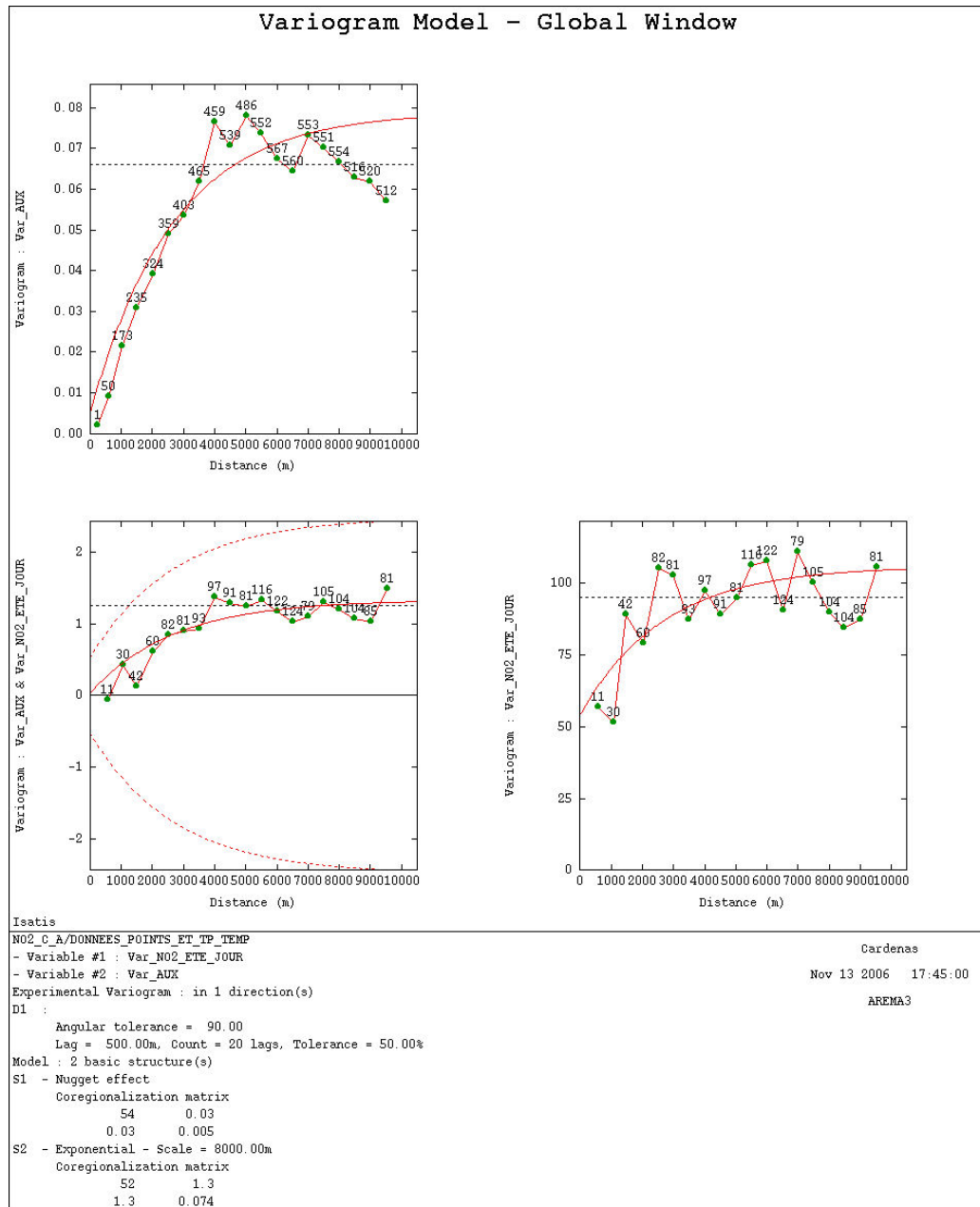


Figure 74 : Modèle croisé entre un jour de la période estivale 2003-2004 et la variable auxiliaire F1K

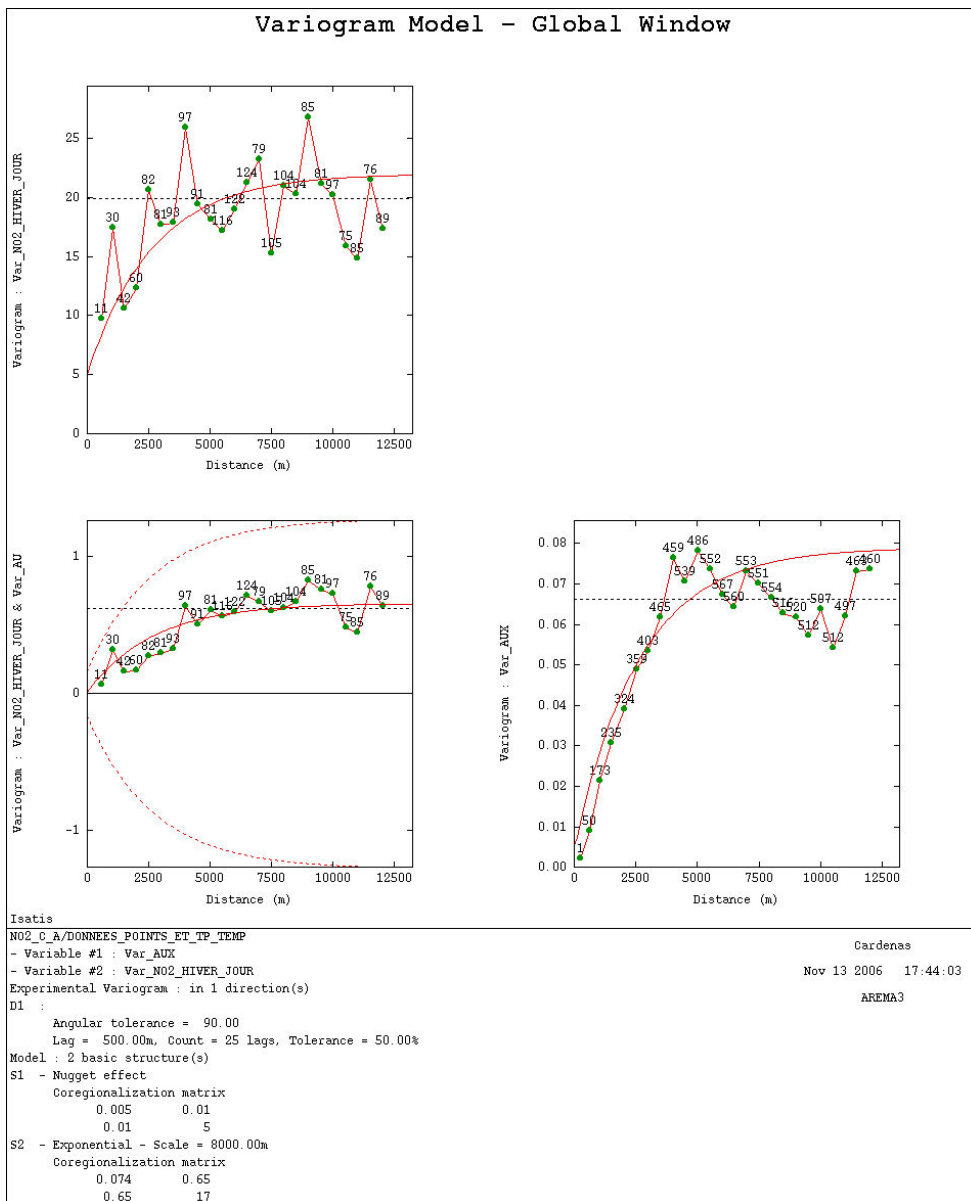
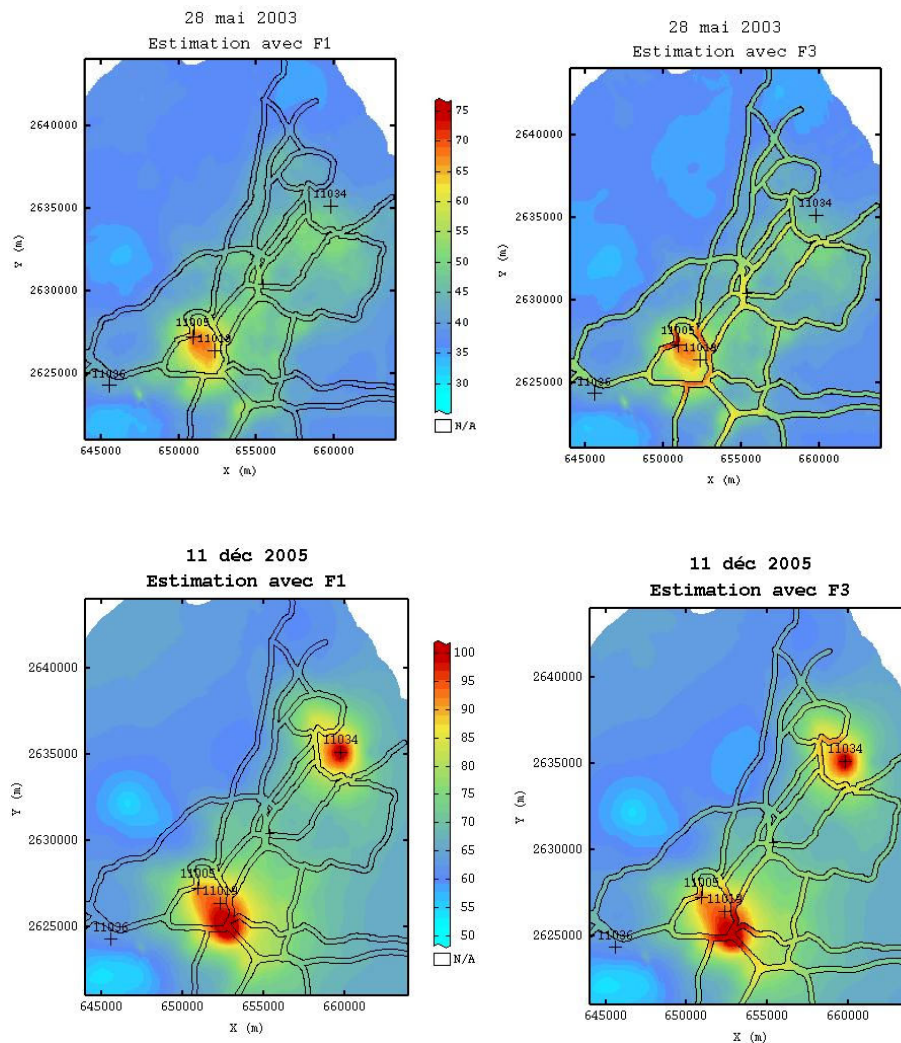


Figure 75 : Modèle croisé entre un jour de la période hivernale 2003-2004 et la variable auxiliaire F1K

Annexe 6 : Exemple de deux estimations avec les fonctions F1 et F3 réalisées par CKJA dans l'agglomération lilloise



A gauche : Estimations avec F1

A droite : Estimations avec F3

En haut : Journée du 28 mai 2003

En bas : Journée du 11 décembre 2005

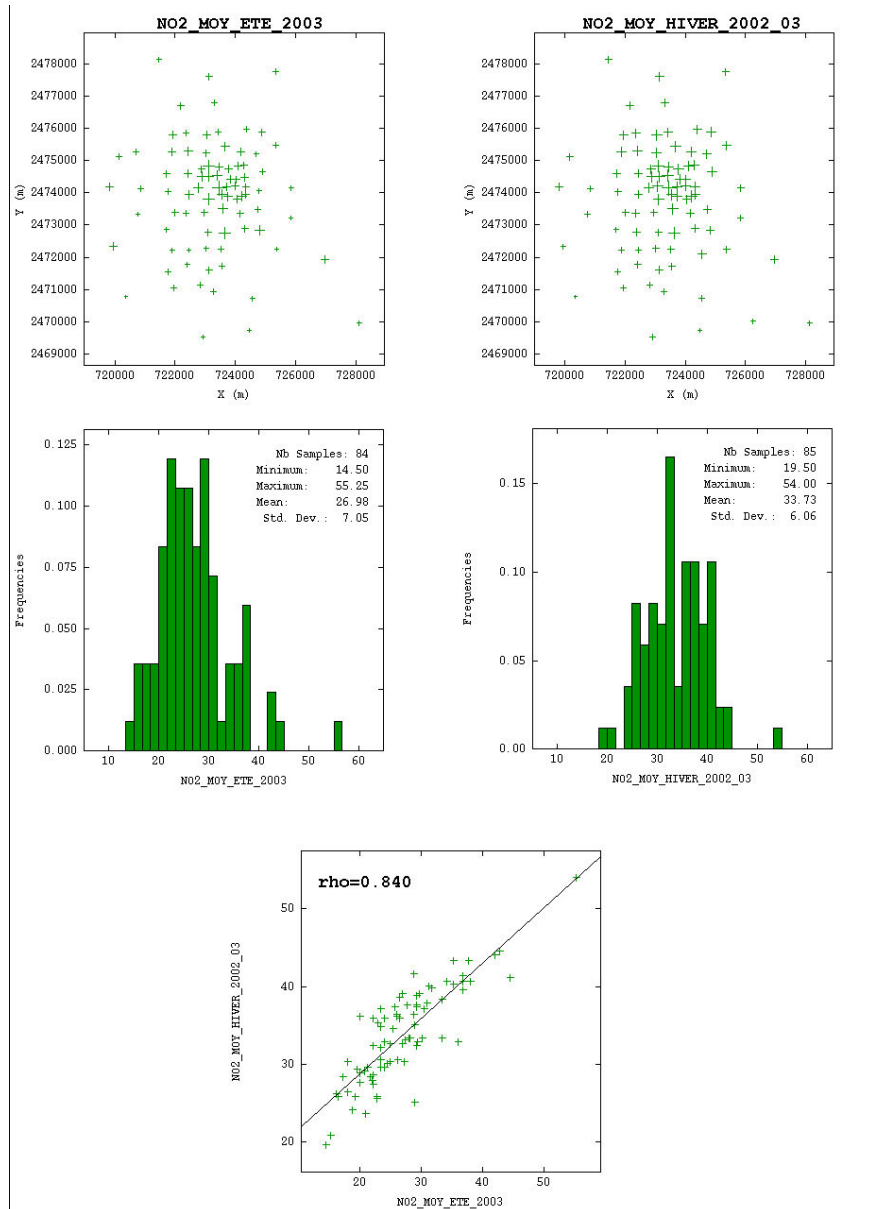
Sur les cartes:

Bande de 150m autour des principales routes

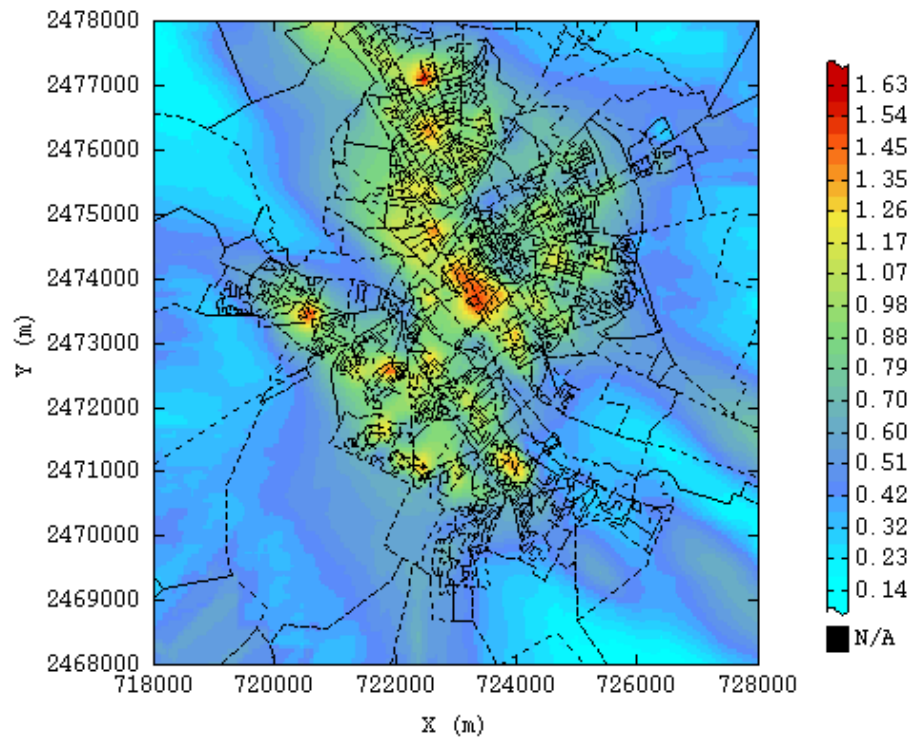
Localisation géographique et code AASQA des 4 stations de trafic

Figure 76 : Exemple de deux estimations avec les fonctions F1 et F3 réalisées par CKJA

*Annexe 7 : Principales figures et statistiques de l'application de la méthode de cokrigage temporel pour l'estimation de la moyenne annuelle de la concentration de NO2 dans l'agglomération rémoise*



*Figure 77 : Cartes et statistiques des deux moyennes saisonnières (agglomération rémoise, année 2002-2003)*



Lignes pointillées : recoupage par îlots dans l'agglomération rémoise

Figure 78 : Carte de la variable auxiliaire

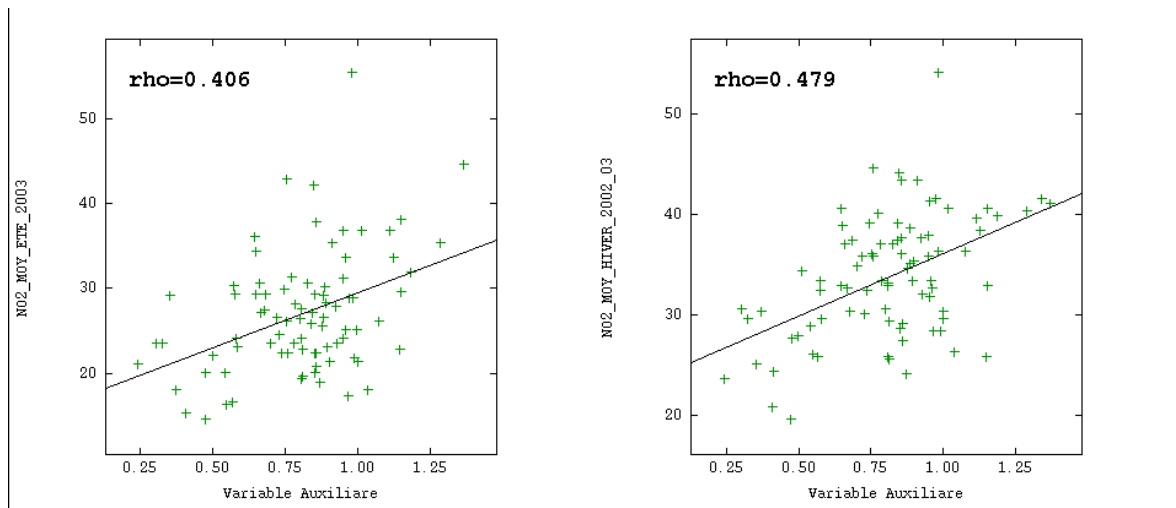


Figure 79 : Nuages de corrélations entre les deux moyennes saisonnières et la variable auxiliaire sélectionné dans l'agglomération rémoise (année 2002-2003)



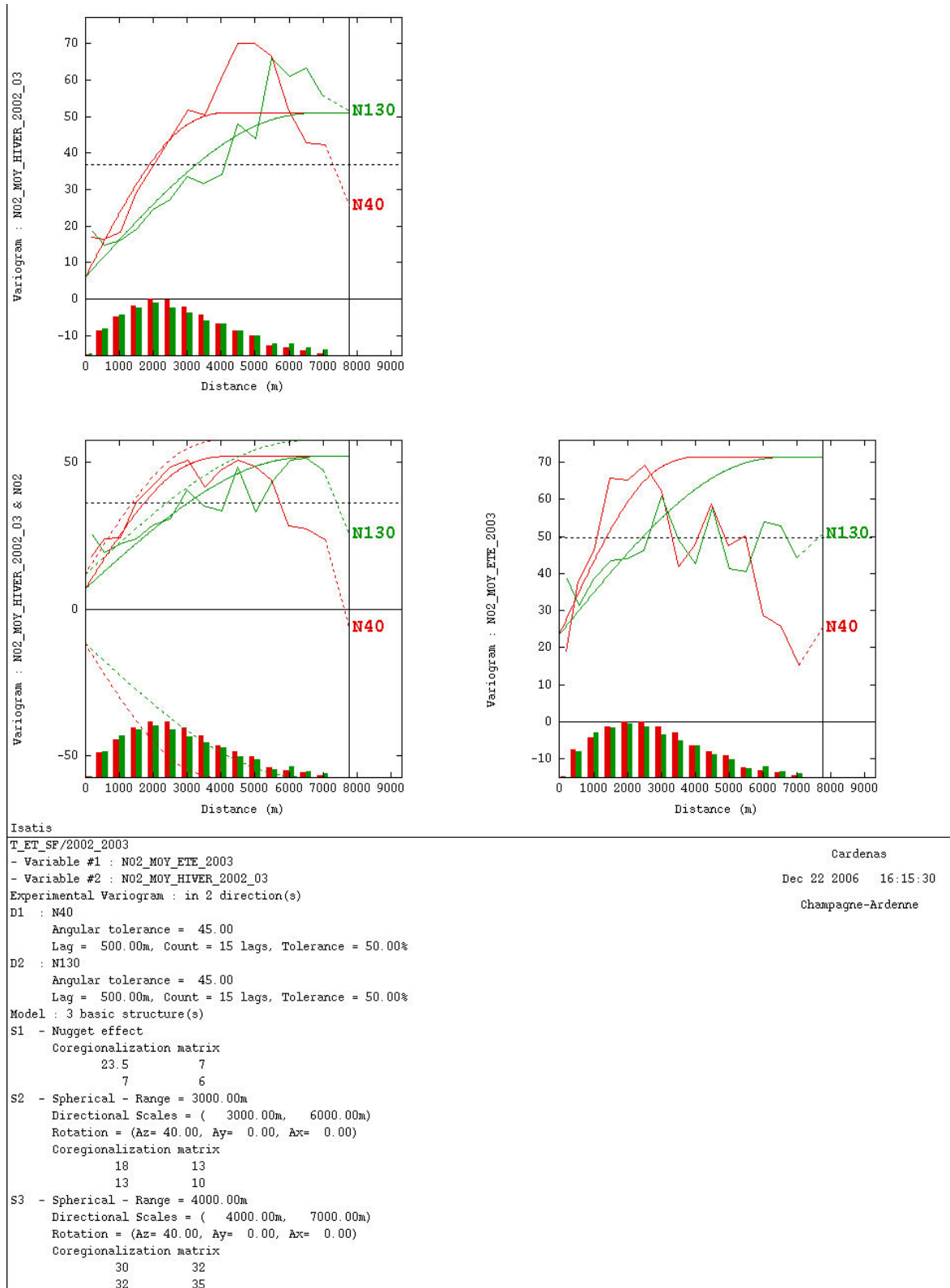


Figure 80 : Modèles des variogrammes simples et croisés des moyennes saisonnières (agglomération rémoise, année 2002-2003)

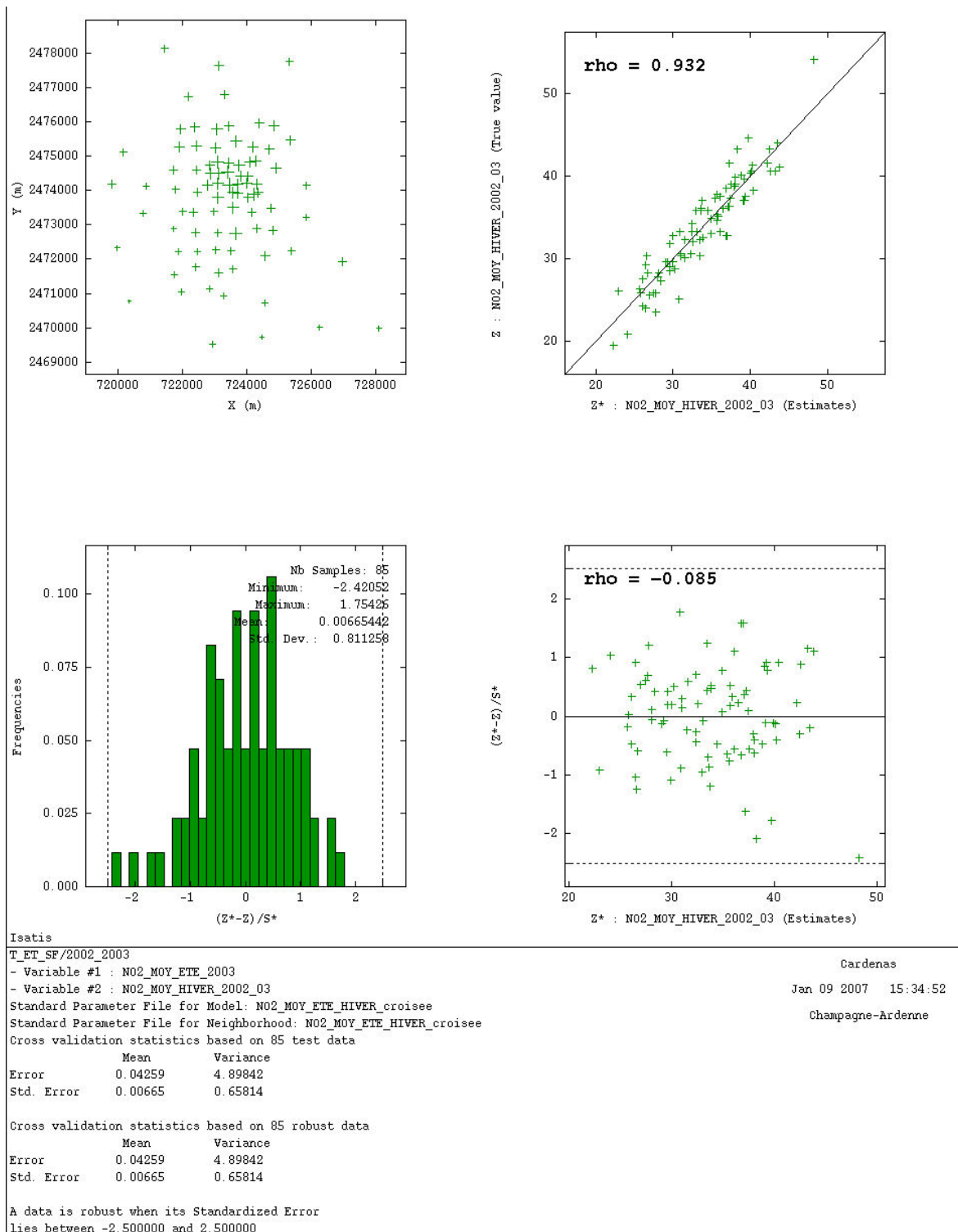


Figure 81 : Validation croisée de la moyenne hivernale 2002-2003, agglomération rémoise

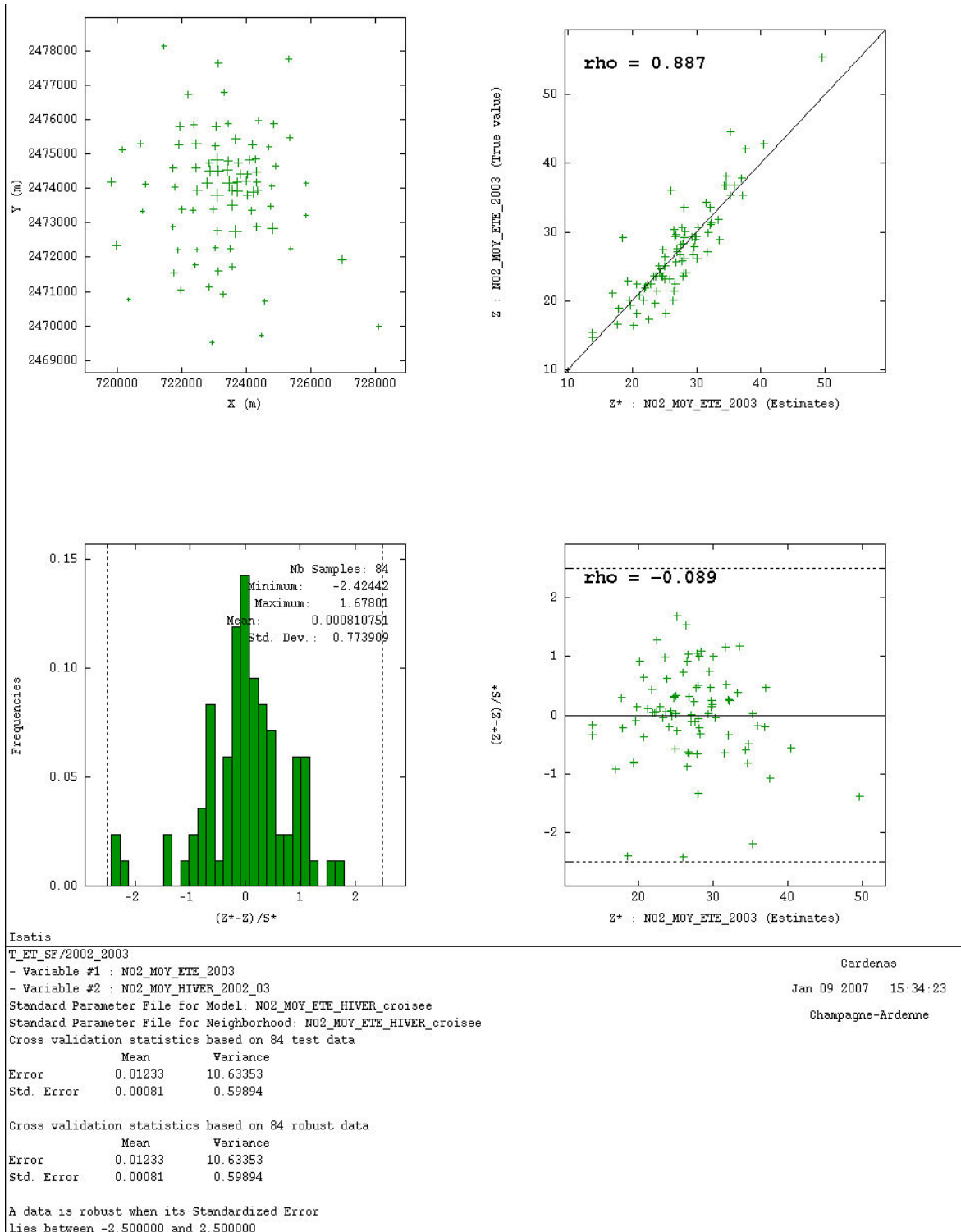


Figure 82 : Validation croisée de la moyenne estivale 2003, agglomération rémoise

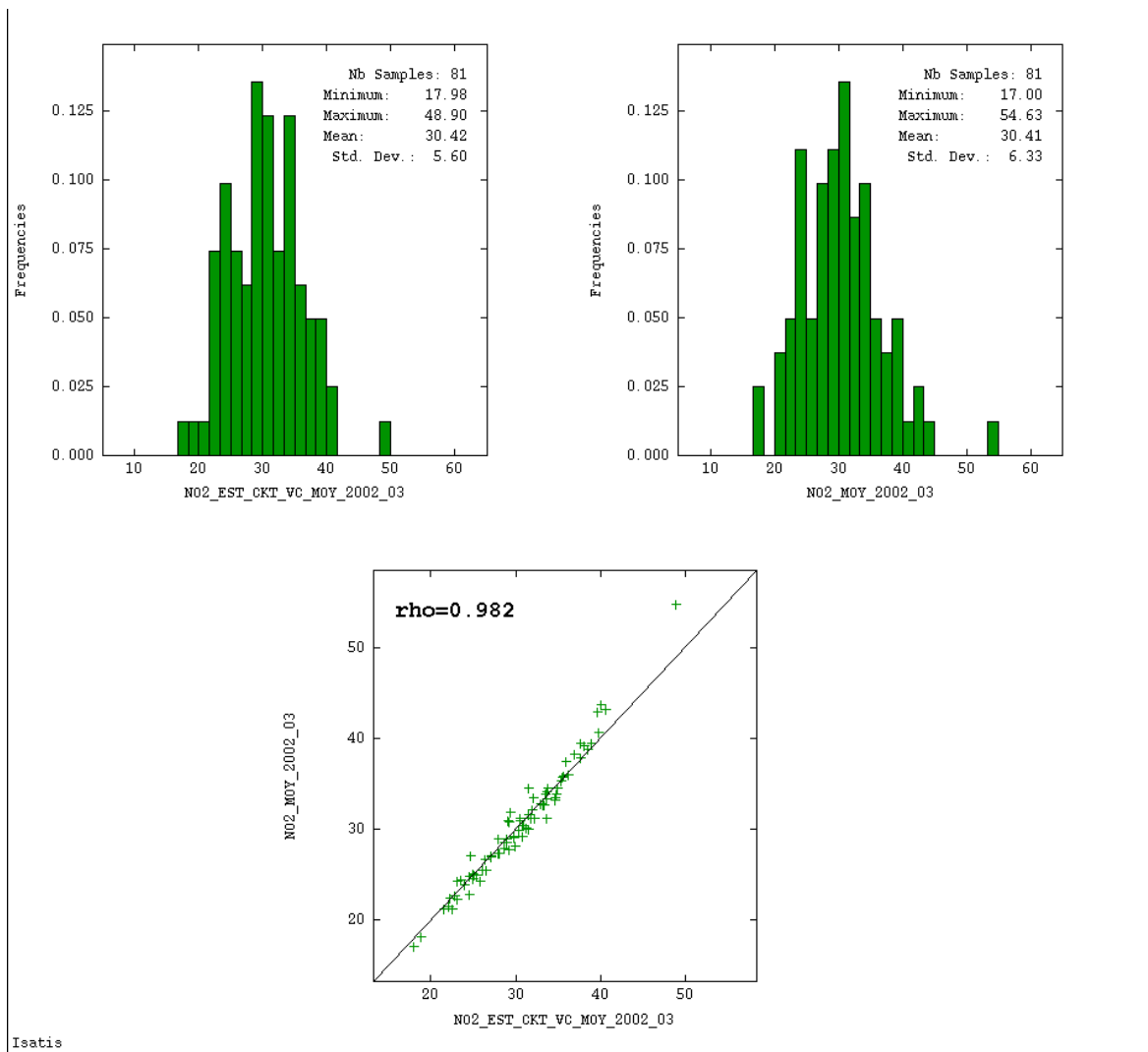
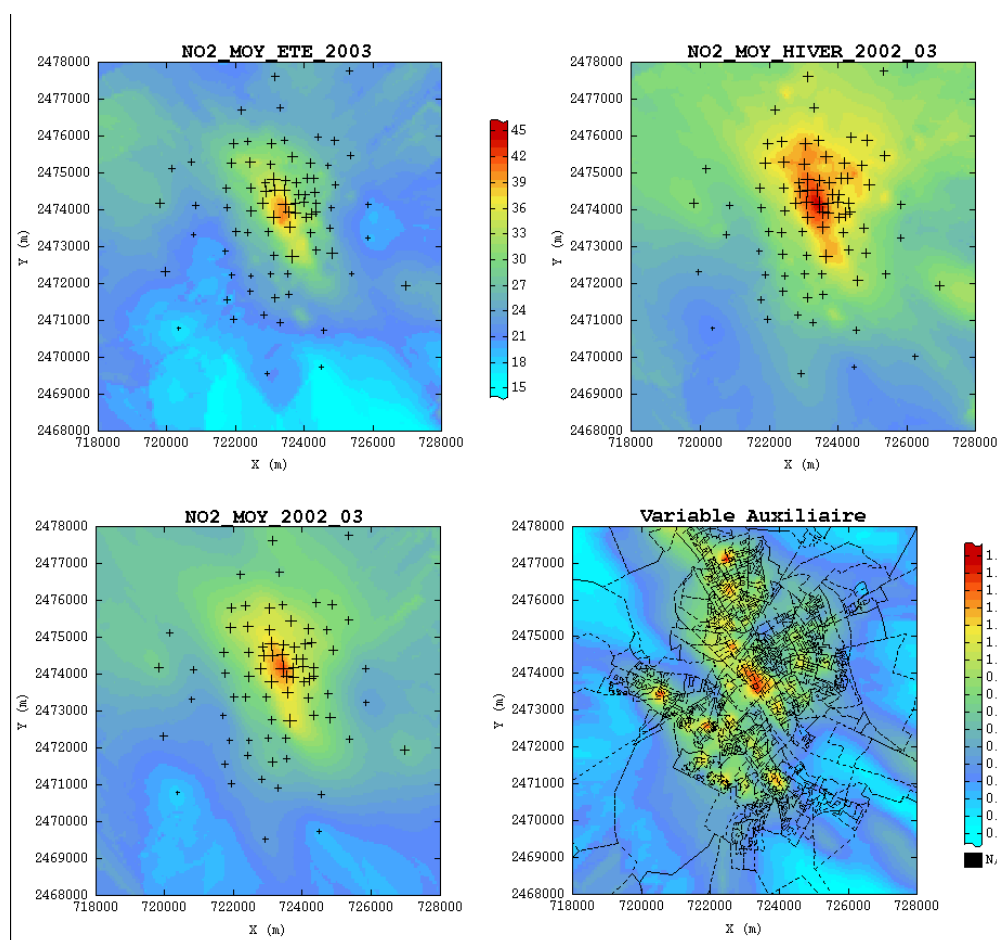


Figure 83 : Histogrammes et nuages de corrélation de l'estimation aux tubes de la moyenne annuelle et de la moyenne annuelle statistique, (agglomération rémoise, année 2002-2003)

La figure ci-joint montre les résultats des estimations des moyennes saisonnières et annuelle calculées par cokrigage temporel pour l'année 2002-2003 dans l'agglomération rémoise.



En haut : Moyennes saisonnières

En bas à gauche : Moyenne annuelle

La taille des croix est proportionnelle aux concentrations mesurées aux tubes passifs

En Bas à droite : Variable auxiliaire et découpage en îlots dans l'agglomération rémoise

Figure 84 : Cartes des moyennes saisonnières et annuelle des concentrations estimées de NO<sub>2</sub> et de la variable auxiliaire (agglomération rémoise, année 2002-2003)

CKT	Nb nœuds	Min	Max	Moy	$\sigma$	$\sigma^2$ ( $\mu\text{g}/\text{m}^3$ ) <sup>2</sup>	CV
Moyenne hiver 2002-2003	10 000	18.2	42.3	25.6	4.1	16.5	0.16
Moyenne été 2003	10 000	14.1	40.5	22.9	3.9	15.4	0.17
Moyenne Annuelle	10 000	20.6	45.0	28.7	4.5	20.6	0.16

Tableau 39 : Statistiques des estimations des moyennes saisonnières et annuelles de NO<sub>2</sub> sur la région rémoise (année 2002-2003)

## Annexe 8 : Exemple de fichier CSV avec entête (réinitialisation d'une étude isatis)

```

#
# structure=free
#
# csv_file=Y, csv_sep=";", csv_dec="."
# nskip=1
#
# field=1, type=alpha , name="NOM"
# field=2 , type=xg , name="X";
#   f_type=Decimal , f_length=10 , f_digits=2, unit="m"
# field=3 , type=yg , name="Y";
#   f_type=Decimal , f_length=10 , f_digits=2, unit="m"
# field=4 , type=alpha , name="SFREF"
# field=5 , type=numeric , name="M", ffff="N/A" ;
#   bitlength=32 , unit="" ;
#   f_type=Decimal , f_length=10 , f_digits=2
# field=6 , type=numeric , name="B", ffff="N/A" ;
#   bitlength=32 , unit="" ;
#   f_type=Decimal , f_length=10 , f_digits=2
# field=7 , type=numeric , name="NO2_30/11/2006", ffff="N/A" ;
#   bitlength=32 , unit="" ;
#   f_type=Decimal , f_length=10 , f_digits=2
# field=8 , type=numeric , name="VEM", ffff="N/A" ;
#   bitlength=32 , unit="" ;
#   f_type=Decimal , f_length=10 , f_digits=2
# field=9 , type=numeric , name="SEL_SF", ffff="N/A" ;
#   bitlength=1 , unit="" ;
#   f_type=Decimal , f_length=10 , f_digits=2
NOM;X;Y;SFREF;M;B;NO2_30/11/2006;VEM;SEL_SF
1;722850;2474714;N/A;N/A;N/A;28.05;8.22;0
2;723758;2474707;N/A;N/A;N/A;32.94;9.15;0
3;724282;2474816;N/A;N/A;N/A;25.59;6.96;0
4;724024;2474384;N/A;N/A;N/A;28.01;16.68;0
...

```

Annexe 9 : Exemple de fichier CSV avec entête (importation dans une étude Isatis existante)

```
#
# structure=free
#
# csv_file=Y, csv_sep=";", csv_dec="."
# nskip=1
#
# field=7 , type=numeric , name="NO2", ffff="N/A" ;
#  bitlength=32 , unit="" ;
#  f_type=Decimal , f_length=10 , f_digits=2
NOM;X;Y;SFREF;M;B;NO2;VEM;SEL_SF
1;722850;2474714;N/A;N/A;N/A;28.05;8.22;0
2;723758;2474707;N/A;N/A;N/A;32.94;9.15;0
3;724282;2474816;N/A;N/A;N/A;25.59;6.96;0
4;724024;2474384;N/A;N/A;N/A;28.01;16.68;0
5;723826;2474385;N/A;N/A;N/A;31.18;13.11;0
6;723722;2474154;N/A;N/A;N/A;31.7;25.55;0
7;724004;2474184;N/A;N/A;N/A;30.96;4.23;0
8;724045;2473764;N/A;N/A;N/A;30.13;23.1;0
9;723730;2473875;N/A;N/A;N/A;33.64;4.42;0
...
```

## Annexe 10 : Exemple des premières lignes d'un fichier BATCH\_INIT.IJNL

```

%VAR CHM = C:\LAMACRO\data_sources\REIMS\MESSAGES
%VAR CHFB =C:\LAMACRO\data_sources\REIMS\Carto_Autom.backup
%VAR CHR = C:\LAMACRO\data_sources\REIMS\RES
%VAR CHFSV = C:\LAMACRO\data_sources\REIMS\Moyenne_annuelle_NO2_2006_Hd.csv
%VAR CHFG = C:\LAMACRO\data_sources\REIMS\NO2_AN.asc
%VAR NO2J = NO2_30/11/2006
%VAR SEL = SEL_SF
%VAR VEM = VEM
%VAR MODEL = MODELE_AN
%VAR VAUX = MODELE_AN
%VAR VOISINAGE = Unique
%VAR X = 0
%VAR Y = 0
%VAR ANGLE = 0
%VAR NO2 = NO2_30/11/2006_CKCL
%VAR NO2E = NO2_30/11/2006_ECART_CKCL
%VAR ST = Carto_Autom
%VAR DP = POINTS
%VAR FSV = SV
%VAR DG = GRILLE
%VAR FM2D = MAILLE_2D
%VAR FMV = MAILLE_VERTICAL
%REDIRECT MSG C:\LAMACRO\data_sources\REIMS\MESSAGES\Moyenne_annuelle_NO2_2006_isatis.txt
%REDIRECT CPU C:\LAMACRO\data_sources\REIMS\MESSAGES\Moyenne_annuelle_NO2_2006_isatis_cpu_timer.txt
%VAR CHST = C:\LAMACRO\data_sources\REIMS\ISATIS\Carto_Autom
%VAR CHRGEST = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL.ASC
%VAR NOM = NOM
%VAR CHRTIFF = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL.TIFF
%VAR CHRCSV1 = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL_STATS_ALL.CSV
%VAR CHRCSV2 = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL_STATS_PAR_SF.CSV
%VAR CHSTTOTO = C:\LAMACRO\data_sources\REIMS\ISATIS\TOTO
#
***** Bulletin Name *****      =B= Data File Manager
***** Bulletin Version *****     =N= 510
File manager Action                 =A= Create Study
New Study Name                      =A= TOTO
Automatic Location on Disk          =A= N
New Study Location on Disk          =A= $CHSTTOTO
Use Isatis Coordinate System        =A= Y
Study for Default Parameters        =A= None
***** Bulletin Name *****      =B= Data File Manager
#
***** Bulletin Version *****     =N= 510
File manager Action                 =A= Set Current Study
Old Study Name                      =A= TOTO
***** Bulletin Name *****      =B= Data File Manager
***** Bulletin Version *****     =N= 510
File manager Action                 =A= Delete Study
Old Study Name                      =A= $ST
#
***** Bulletin Name *****      =B= Data File Manager
***** Bulletin Version *****     =N= 510
File manager Action                 =A= Restore Study
Backup File Name                    =A= $CHFB
New Study Name                      =A= $ST
Automatic Location on Disk          =A= N

```



```

New Study Location on Disk      =A= $CHST
#
***** Bulletin Name *****  =B= Data File Manager
**** Bulletin Version ****     =N= 510
File manager Action            =A= Delete Study
Old Study Name                 =A= TOTO
#
***** Bulletin Name *****  =B= ASCII File Import
**** Bulletin Version ****     =N= 510
Input ASCII Filename           =A= $CHFSV
Format described in data file  =A= Y
Header File                    =A= $CHFSV
File Creation Mode             =A= Create a New File
Directory Name                 =A= $DP
Header Points File Name        =A= None
File Name                      =A= $FSV
#
***** Bulletin Name *****  =B= ArcView Import
**** Bulletin Version ****     =N= 331
Arcview File Type              =A= GRID
GRID File Type                 =A= Ascii format
GRID Data File Name            =A= $CHFG
GRID Isatis Directory Name     =A= $DG
GRID Isatis File Name          =A= $FM2D
GRID Isatis Variable Name      =A= $VAUX
SHAPEFILE Shapefile Name      =A= None
SHAPEFILE Import As           =A=
SHAPEFILE Polygons Use Field for Name =A= N
SHAPEFILE Polygons Name Field  =A=
SHAPEFILE Polygons Different Levels =A= N
SHAPEFILE Polygons Group by Name =A= N
SHAPEFILE Polygons Add Thickness =A= Y
SHAPEFILE Points One Sample Per Shape =A= N
SHAPEFILE Unit                 =A= Meter
SHAPEFILE Isatis Directory Name =A= None
SHAPEFILE Isatis File Name     =A= None
#
...

```

## Annexe 11 : Exemple des premières lignes d'un fichier BATCH\_RUN.IJNL

```

%VAR CHM = C:\LAMACRO\data_sources\REIMS\MESSAGES
%VAR CHFB = None
%VAR CHR = C:\LAMACRO\data_sources\REIMS\RES
%VAR CHFSV = C:\LAMACRO\data_sources\REIMS\Moyenne_annuelle_NO2_2006_Hd.csv
%VAR CHFG = C:\LAMACRO\data_sources\REIMS\NO2_AN.asc
%VAR NO2J = NO2_30/11/2006
%VAR SEL = SEL_SF
%VAR VEM = VEM
%VAR MODEL = MODELE_AN
%VAR VAUX = MODELE_AN
%VAR VOISINAGE = Unique
%VAR X = 0
%VAR Y = 0
%VAR ANGLE = 0
%VAR NO2 = NO2_30/11/2006_CKCL
%VAR NO2E = NO2_30/11/2006_ECART_CKCL
%VAR ST = Carto_Autom
%VAR DP = POINTS
%VAR FSV = SV
%VAR DG = GRILLE
%VAR FM2D = MAILLE_2D
%VAR FMV = MAILLE_VERTICAL
%REDIRECT MSG C:\LAMACRO\data_sources\REIMS\MESSAGES\Moyenne_annuelle_NO2_2006_isatis.txt
%REDIRECT CPU C:\LAMACRO\data_sources\REIMS\MESSAGES\Moyenne_annuelle_NO2_2006_isatis_cpu_timer.txt
%VAR CHST = C:\LAMACRO\data_sources\REIMS\ISATIS\Carto_Autom
%VAR CHRGEST = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL.ASC
%VAR NOM = NOM
%VAR CHRTIFF = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL.TIFF
%VAR CHRCSV1 = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL_STATS_ALL.CSV
%VAR CHRCSV2 = C:\LAMACRO\data_sources\REIMS\RES\Moyenne_annuelle_NO2_2006_CKCL_STATS_PAR_SF.CSV
%VAR CHSTTOTO = C:\LAMACRO\data_sources\REIMS\ISATIS\TOTO
#
***** Bulletin Name *****      =B= ASCII File Import
**** Bulletin Version ****        =N= 510
Input ASCII Filename              =A= $CHFSV
Format described in data file     =A= Y
Header File                       =A= $CHFSV
File Creation Mode                =A= Add Variables in an Existing File
Directory Name                    =A= $DP
Header Points File Name           =A= None
File Name                         =A= $FSV
#
***** Bulletin Name *****      =B= ArcView Import

```

```
***** Bulletin Version *****      =N= 331
Arcview File Type                     =A= GRID
GRID File Type                        =A= Ascii format
GRID Data File Name                   =A= $CHFG
GRID Isatis Directory Name            =A= $DG
GRID Isatis File Name                 =A= $FM2D
GRID Isatis Variable Name             =A= $VAUX
SHAPEFILE Shapefile Name              =A= None
SHAPEFILE Import As                   =A=
SHAPEFILE Polygons Use Field for Name =A= N
SHAPEFILE Polygons Name Field         =A=
SHAPEFILE Polygons Different Levels   =A= N
SHAPEFILE Polygons Group by Name     =A= N
SHAPEFILE Polygons Add Thickness     =A= Y
SHAPEFILE Points One Sample Per Shape =A= N
SHAPEFILE Unit                        =A= Meter
SHAPEFILE Isatis Directory Name       =A= None
SHAPEFILE Isatis File Name            =A= None
#
...
```

## Annexe 12 : Exemple des premières lignes d'un fichier ASC (« grille SIG »)

```
ncols 100
nrows 100
xllcenter 718050
yllcenter 2.46805e+006
cellsize 100
nodata_value 0
14.4696 14.7408 15.0809 15.4109 15.5398 15.867 16.2679 16.5768 16.8655
17.1214 17.991 18.1102 18.2543 18.4872 18.6481 18.9589 19.5446 19.9015
20.4998 20.947 21.5322 22.2373 22.8456 23.4386 23.9468 24.5352 24.9509
25.0696 25.0529 24.8191 24.4594 24.0062 23.5473 23.153 22.888 22.6525
22.3277 21.9566 21.6173 21.2944 20.9755 20.6925 20.3601 20.0451
19.7941 19.7282 19.676 19.614 19.395 19.2175 19.0303 18.8876 18.6675
18.5389 18.2612 18.0023 17.7474 17.6909 17.2879 17.1115 16.8168
16.6422 16.4555 16.5078 16.4088 16.4064 16.2638 16.1499 16.0082
15.9302 15.8468 15.8793 15.5541 15.5313 15.3676 15.0499 15.0034
14.8876 14.7512 14.9162 15.0569 15.099 14.8338 14.7514 14.7245 14.759
14.9209 15.664 15.7051 16.2205 15.8684 15.9567 16.0474 16.1376 16.2421
16.3538 16.4765 16.6213 16.757 16.8925
14.3405 14.5173 14.8694 15.2835 15.6676 16.0594 16.3249 16.7488 17.116
17.8653 18.0885 18.4777 18.58 18.7238 18.8395 19.1107 19.4203 19.9158
20.4154 20.7384 21.2353 21.8547 22.4347 23.0007 23.4586 23.9994
24.3896 24.5643 24.6752 24.6243 24.4235 24.1119 23.7481 23.3832
23.1284 22.9336 22.7023 22.3827 22.0515 21.7632 21.3671 21.09 20.7216
20.2724 19.9866 20.1142 20.2291 20.2052 20.0361 19.8073 19.6141
19.4665 19.2523 19.0817 18.7634 18.5633 18.1854 17.9651 17.7173
17.4121 17.1503 16.8916 16.7605 16.6993 16.6014 16.6027 16.4682
16.4075 16.2696 16.146 16.0262 16.0096 15.7193 15.7182 15.5623 15.2817
15.1433 14.8937 14.8141 14.9498 14.9999 14.7219 14.7256 14.644 14.5521
14.7149 14.8645 15.5827 16.0393 16.1049 15.7573 15.8267 15.93 16.0462
16.1781 16.3425 16.5006 16.6593 16.8306 17.0101
...
```

Annexe 13 : Exemple des premières lignes d'un fichier contenant les messages d'isatis (« \_isatis.txt »)

```
$$$$>>> RUNNING BULLETIN: Data File Manager
File Manager Action: Set Current Study
Setting Study: TOTO...
$$$$>>> END OF RUN...
$$$$>>> RUNNING BULLETIN: Data File Manager
File Manager Action: Delete Study
$$$$>>> END OF RUN...
$$$$>>> RUNNING BULLETIN: Data File Manager
File Manager Action: Restore Study
Extracting ./ ...
Extracting DESCR ...
Extracting DEF/ ...
Extracting DEF/DEF2600 ...
Extracting DEF/DEF1000 ...
Extracting DEF/DEF2670 ...
Extracting DEF/envIRON.def ...
Extracting GTX/ ...
Extracting GTX/DESCR ...
Extracting TMP/ ...
Extracting PARAM/ ...
Extracting PARAM/DVARIO/ ...
Extracting PARAM/MVARIO/ ...
Extracting PARAM/MVARIO/PSET.0 ...
Extracting PARAM/MVARIO/PSET.1 ...
Extracting PARAM/MVARIO/PSET.2 ...
Extracting PARAM/NEIGH/ ...
Extracting PARAM/NEIGH/PSET.0 ...
Extracting PARAM/MHISTO/ ...
Extracting PARAM/MCORREL/ ...
Extracting PARAM/PCAMOD/ ...
Extracting PARAM/PROP/ ...
Extracting CSCALE/ ...
Extracting CSCALE/CSCA.0 ...
Extracting CSCALE/CSCA.1 ...
Extracting CSCALE/CSCA.2 ...
Extracting CSCALE/CSCA.3 ...
Extracting CSCALE/CSCA.4 ...
Extracting CSCALE/CSCA.5 ...
Extracting CSCALE/CSCA.6 ...
Extracting CSCALE/CSCA.7 ...
...
```

Annexe 14 : Exemple des premières lignes d'un fichier contenant les temps de calcul d'Isatis (« \_isatis\_cpu\_timer.txt »)

```

0 0:00:00 isatis
0 0:00:00 Data File Manager
0 0:00:00 Data File Manager
0 0:00:00 Data File Manager
0 0:00:00 ASCII File Import
1 0:00:01 isatis
0 0:00:00 Data File Manager
0 0:00:00 Data File Manager
0 0:00:00 Data File Manager
0 0:00:00 ASCII File Import
0 0:00:00 ArcView Import
0 0:00:00 Migrate Grid to Point
0 0:00:00 Calculator
11 0:00:11 Standard (Co-)Kriging
0 0:00:00 Cross-validation
0 0:00:00 Calculator
1 0:00:01 Calculator
0 0:00:00 Create Grid File
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
1 0:00:01 Copy Points Statistics to Grid
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 Calculator
0 0:00:00 Calculator
0 0:00:00 Calculator
0 0:00:00 Copy Points Statistics to Grid
0 0:00:00 New Display Page
1 0:00:01 ArcView Export
0 0:00:00 ASCII File Export
0 0:00:00 ASCII File Export
14 0:00:14 isatis

```

Remarque : dans cet exemple, la durée de l'exécution d'Isatis a été de 14 secondes, dont 11 pour le cokrigage.

## Annexe 15 : Formules pour le calcul des statistiques produites par Isatis

#	Nom de la statistique	Formule	# Ligne, Annexe
1	Erreur	$(Z-Z^*)$	10, Annexe 17
2	Moyenne de l'erreur	$[\Sigma(Z-Z^*)]/n$	19, Annexe 16
3	Erreur carrée	$(Z-Z^*)^2$	11, Annexe 17
4	Moyenne de l'erreur carrée	$[\Sigma(Z-Z^*)^2]/n$	20, Annexe 16
5	Variance de l'erreur	$\{[\Sigma(Z-Z^*)^2]/n\} - \{[\Sigma(Z-Z^*)]/n\}^2$	22, Annexe 16
6	Valeur absolue de l'erreur	$\{\text{Racine carrée} [(Z-Z^*)^2]\}$	12, Annexe 17
7	Racine carrée de la moyenne de l'erreur carrée (RMSE)	Racine carrée $\{[\Sigma(Z-Z^*)^2]/n\}$	21, Annexe 16
8	Erreur relative	$(Z-Z^*)/Z$	15, Annexe 17
9	Moyenne de l'erreur relative	$\{\Sigma[(Z-Z^*)/Z]\}/n$	23, Annexe 16
10	Erreur relative carrée	$[(Z-Z^*)/Z]^2$	16, Annexe 17
11	Moyenne de l'erreur relative carrée	$\{\Sigma[(Z-Z^*)/Z]^2\}/n$	24, Annexe 16
12	Variance de l'erreur relative	$\{\Sigma[(Z-Z^*)/Z]^2\}/n - (\{\Sigma[(Z-Z^*)/Z]\}/n)^2$	25, Annexe 16
13	Erreur réduite	$(Z-Z^*)/\sigma$	13, Annexe 17
14	Moyenne de l'erreur réduite	$\{\Sigma[(Z-Z^*)/\sigma]\}/n$	26, Annexe 16
15	Erreur réduite carrée	$[(Z-Z^*)/\sigma]^2$	14, Annexe 17
16	Moyenne de l'erreur réduite carrée	$\{\Sigma[(Z-Z^*)/\sigma]^2\}/n$	27, Annexe 16
17	Variance de l'erreur réduite	$\{\Sigma[(Z-Z^*)/\sigma]^2\}/n - (\{\Sigma[(Z-Z^*)/\sigma]\}/n)^2$	28, Annexe 16
18	Moyenne de l'erreur absolue	$(\Sigma \{\text{Racine carrée} [(Z-Z^*)^2]\})/n$	29, Annexe 16

Abréviations utilisées dans ce tableau :

- Z: Mesure
- Z<sup>\*</sup>: Estimation
- σ: Ecart -type d'estimation
- n: Nombre des points
- Σ : Symbole représentant la somme de « n » éléments

**Annexe 16 : Description détaillée des champs du fichier:**  
**«...\_STATS\_ALL\_CKCL.CSV »**

<i>Champ</i>	<i>Nom du Champ</i>	<i>Description</i>
1	NO2_vs_NO2_CKCL_SF_VC_RHO	$\rho$ entre les [NO2] mesurées et estimées par VC dans les SF
2	NO2_vs_VAUX_SF_RHO	$\rho$ entre les [NO2] mesurées dans les SF et la V. Aux.
3	NO2_vs_VAUX_RHO	$\rho$ entre les [NO2] des stations (SF et SV) et la V. Aux.
4	NO2_NB	Nombre de stations (SV et SF) disponibles
5	NO2_MIN	Valeur minimale des [NO2] des stations (SV et SF)
6	NO2_MAX	Valeur maximale des [NO2] des stations (SV et SF)
7	NO2_MOY	Valeur moyenne des [NO2] des stations (SV et SF)
8	NO2_ECART	Ecart- type des [NO2] des stations (SV et SF)
9	NO2_SF_NB	Nombre de SF disponibles
10	NO2_SF_MIN	Valeur minimale des [NO2] mesurées dans les SF
11	NO2_SF_MAX	Valeur maximale des [NO2] mesurées dans les SF
12	NO2_SF_MOY	Valeur moyenne des [NO2] mesurées dans les SF
13	NO2_SF_ECART	Ecart- type des [NO2] mesurées dans les SF
14	NO2_CKCL_SF_VC_NB	Nombre des stations fixes disponibles pour la VC
15	NO2_CKCL_SF_VC_MIN	Valeur minimale des [NO2] estimées par VC dans les SF
16	NO2_CKCL_SF_VC_MAX	Valeur maximale des [NO2] estimées par VC dans les SF
17	NO2_CKCL_SF_VC_MOY	Valeur moyenne des [NO2] estimées par VC dans les SF
18	NO2_CKCL_SF_VC_ECART	Ecart- type des [NO2] estimées par VC dans les SF
19	NO2_CKCL_ERREUR_SF_VC_MOY	Moyenne de l'erreur par VC dans les SF (ligne 2 de l'Annexe 15)
20	NO2_CKCL_ERREUR_CARREE_SF_VC_MOY	Moyenne de l'erreur carrée par VC dans les SF (ligne 4 de l'Annexe 15)
21	NO2_CKCL_RMSE_SF_VC_MOY	Racine de la moyenne de l'erreur carrée par VC dans les SF (ligne 7 de l'Annexe 15)
22	NO2_CKCL_ERREUR_SF_VC_VAR	Variance de l'erreur par VC dans les SF (ligne 5 de l'Annexe 15)
23	NO2_CKCL_ERREUR_REL_SF_VC_MOY	Moyenne de l'erreur relative par VC dans les SF (ligne 9 de l'Annexe 15)
24	NO2_CKCL_ERREUR_REL_CARREE_SF_VC_MOY	Moyenne de l'erreur relative carrée par VC dans les SF (ligne 11 de l'Annexe 15)
25	NO2_CKCL_ERREUR_REL_SF_VC_VAR	Variance de l'erreur relative par VC dans les SF (ligne 12 de l'Annexe 15)
26	NO2_CKCL_ERREUR_RED_SF_VC_MOY	Moyenne de l'erreur réduite par VC dans les SF (ligne 14 de l'Annexe 15)
27	NO2_CKCL_ERREUR_RED_CARREE_SF_VC_MOY	Moyenne de l'erreur réduite carrée par VC dans les SF (ligne 16 de l'Annexe 15)
28	NO2_CKCL_ERREUR_RED_SF_VC_VAR	Variance de l'erreur réduite par VC dans les SF (ligne 17 de l'Annexe 15)
29	NO2_CKCL_ERREUR_ABS_SF_VC_MOY	Moyenne de l'erreur absolue par VC dans les SF (ligne 18 de l'Annexe 15)

Abréviations utilisées dans ce tableau :

Concentration NO2 : [NO2]

Stations Fixes : SF

Variable Auxiliaire : V. Aux.

Coefficient de corrélation :  $\rho$

Stations Virtuelles : SV

Validation Croisée : VC



**Annexe 17 : Description détaillée des champs du fichier:  
«...\_STATS\_PAR\_SF\_CKCL.CSV »**

<b>Champ</b>	<b>Nom du Champ</b>	<b>Description</b>
1	SN+ Sample Number (READONLY)+	Numéro de la SF dans la structure d'Isatis
2	X	Coordonnée X (vient du fichier CSV original)
3	Y	Coordonnée Y (vient du fichier CSV original)
4	NOM	Nom ou numérotation de la SF (vient du fichier CSV original)
5	NO2	Mesure de la [NO2] (vient du fichier CSV original)
6	VEM	Mesure de la VEM (vient du fichier CSV original)
7	« Nom de la Variable auxiliaire »	Valeur de la Variable auxiliaire
8	NO2_CKCL_SF_VC	[NO2] estimée par validation croisée
9	NO2_CKCL_EST_ECART_SF_VC	Ecart- type d'estimation de la [NO2] par validation croisée
10	NO2_CKCL_ERREUR_SF_VC	Erreur par VC dans la SF (ligne 1 de l'Annexe 15)
11	NO2_CKCL_ERREUR_CARREE_SF_VC	Erreur par VC carrée dans la SF (ligne 3 de l'Annexe 15)
12	NO2_CKCL_ERREUR_ABS_SF_VC	Valeur absolu de l'Erreur par VC dans la SF (ligne 6 de l'Annexe 15)
13	NO2_CKCL_ERREUR_RED_SF_VC	Erreur réduite par VC dans la SF (ligne 13 de l'Annexe 15)
14	NO2_CKCL_ERREUR_RED_CARREE_SF_VC	Erreur réduite carrée par VC dans la SF (ligne 15 de l'Annexe 15)
15	NO2_CKCL_ERREUR_REL_SF_VC	Erreur relative par VC dans la SF (ligne 8 de l'Annexe 15)
16	NO2_CKCL_ERREUR_REL_CARREE_SF_VC	Erreur relative carrée par VC dans la SF (ligne 10 de l'Annexe 15)

Abréviations utilisées dans ce tableau :

- Concentration NO2 : [NO2]
- Stations Fixes : SF
- Validation Croisée : VC