

**Ecole des Mines  
de Paris**

---

Contrat INERIS/ARMINES no. CNS 003 27 84

Rapport final

**ÉTUDE SUR LA RÉALISATION DE CARTOGRAPHIES DE LA QUALITÉ  
DE L'AIR DANS LES ZONES PEU/PAS COUVERTES PAR LES RÉSEAUX  
DE STATIONS FIXES À L'AIDE DE MÉTHODES GÉOSTATISTIQUES**

Chantal de FOUQUET

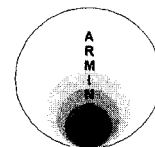
-----

Rapport N-9/03/G

Juin 2003



**CENTRE DE GEOSTATISTIQUE**  
35 rue Saint-Honoré  
77305 FONTAINEBLEAU (France)  
Tél. : 33-1 64 69 47 81 Fax : 33-1 64 69 47 05  
<http://cg.ensmp.fr>



1. Compléments sur la concentration saisonnière ou annuelle en NO <sub>2</sub> sur l'agglomération de Mulhouse. ....	2
<b>1.1. Introduction</b> .....	2
<b>1.2. Relations entre moyennes saisonnières</b> .....	3
<b>1.3. Estimation monovariante de la concentration en NO<sub>2</sub></b> .....	6
<b>1.4. Estimation de la moyenne annuelle par les moyennes saisonnières</b> ....	22
<b>1.5. Conclusion</b> .....	23
<b>1.6. Annexes. NO<sub>2</sub> sur l'agglomération de Mulhouse.</b> .....	26
2. Conclusions sur la cartographie du NO <sub>2</sub> .....	32
<b>2.1. Maxima saisonniers</b> .....	32
<b>2.2. NO<sub>2</sub> routier de la vallée de la Thur</b> .....	32
<b>2.3. NO<sub>2</sub> urbain à Mulhouse et Montpellier</b> .....	33

**Le présent rapport comporte deux parties :**

- **un complément sur l'étude de la concentration saisonnière ou annuelle en NO<sub>2</sub> sur l'agglomération de Mulhouse, avec des figures annexes.**
- **le rapport final de la convention INERIS/ARMINES CNS 003 27 84**

# **1. Compléments sur la concentration saisonnière ou annuelle en NO<sub>2</sub> sur l'agglomération de Mulhouse.**

## **1.1. Introduction**

Le présent rapport complète la "Méthodologie de cartographie de la concentration annuelle en NO<sub>2</sub> sur l'agglomération de Mulhouse" (avril 2003, troisième rapport d'avancement ; convention Ineris-Armines, CNS 003 27 84), désignée dans la suite comme 3ème rapport CG.

Dans cette étude, il apparaissait que les écarts entre les différents modèles avec dérive externe pour la moyenne saisonnière ou annuelle restaient faibles, par comparaison aux erreurs expérimentales en validation croisée. Ces résultats de validation croisée sont mieux corrélés aux concentrations mesurées pour la moyenne hivernale que pour la moyenne estivale.

On se propose ici

- de comparer les estimations multivariées au krigeage monovarié de la concentration en NO<sub>2</sub>, pour préciser l'intérêt d'incorporer ou non des variables auxiliaires dans l'estimation des moyennes saisonnières ou annuelles.

- de mentionner la possibilité d'un cokrigeage de la moyenne annuelle par les moyennes saisonnières. Bien que ce cokrigeage ne soit pas directement faisable en l'état actuel du logiciel Isatis, il présenterait l'intérêt d'utiliser au mieux l'information disponible (toutes les moyennes hivernales ainsi que toutes les moyennes estivales - une extension aux mesures par quinzaine étant envisageable), et d'assurer une parfaite cohérence entre les trois estimations : moyenne hivernale, moyenne estivale et moyenne "annuelle". Lorsque les modèles multivariés diffèrent pour la moyenne hivernale et estivale, le calcul de la variance de krigeage devra être repris.

Dans tout le document, les moyennes saisonnières sont calculées sur 3 quinzaines, et les moyennes annuelles sur les 6 quinzaines. Les stations comportant de 1 à 3 tubes sont traitées ensemble. Rappelons que la moyenne hivernale est disponible pour 64 tubes, la moyenne estivale pour 62, l'intersection pour la moyenne annuelle comportant 52 tubes.

## 1.2. Relations entre moyennes saisonnières

Dans la suite, les termes de concentrations fortes ou faibles sont relatifs aux histogrammes présentés ; ils ne font pas référence à des valeurs guide.

### Moyennes hivernales

Le nuage de corrélation des moyennes saisonnières (figure 1.) montre que les tubes du mode des faibles concentrations hivernales présentent en été également une faible concentration. Ces tubes sont plutôt situés sur le pourtour de l'agglomération, excepté vers le Nord, où les concentrations hivernales sont fortes ou moyennes.

Ces fortes concentrations hivernales se situent principalement au centre de l'agglomération, ainsi que plus localement, au Nord. Pour ces tubes, les concentrations estivales sont fortes ou moyennes (voir les classes de couleur sur les histogrammes).

### Moyennes estivales

Le mode des fortes valeurs estivales, repérées par des carrés rouges sur la figure 2., est regroupé dans la tranche 20-28  $\text{mg}/\text{m}^3$ , avec une valeur voisine de 35  $\text{mg}/\text{m}^3$ . Il comprend en particulier un ensemble de tubes s'écartant de la majorité des points du nuage de corrélation des moyennes hiver-été, caractérisés par un surplus local de la concentration estivale.

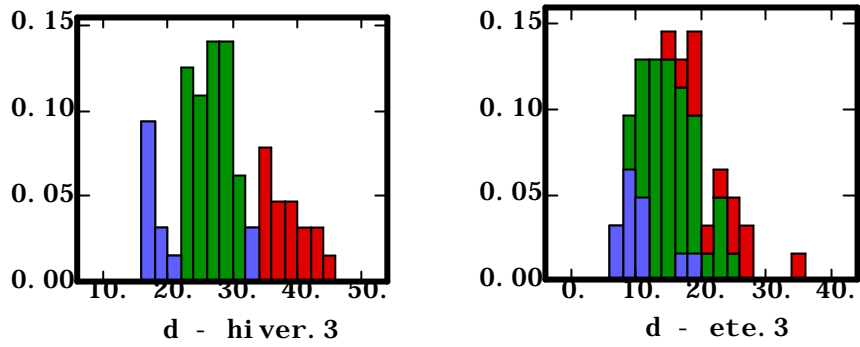
Les fortes mesures estivales se trouvent approximativement entre les ordonnées  $y=5287$  à  $y=5295$  environ, ainsi qu'en un tube plus en Nord. Elles présentent une extension Est-Ouest marquée, et sont moins étendues dans la direction Nord-Sud que les fortes moyennes hivernales. Cette extension est-ouest peut expliquer la structure peu marquée observée sur le variogramme de la moyenne estivale dans cette direction.

Les tubes de fortes concentrations estivales recouvrent l'ensemble des valeurs hivernales intermédiaires à fortes (donc toutes, sauf le mode des faibles concentrations hivernales).

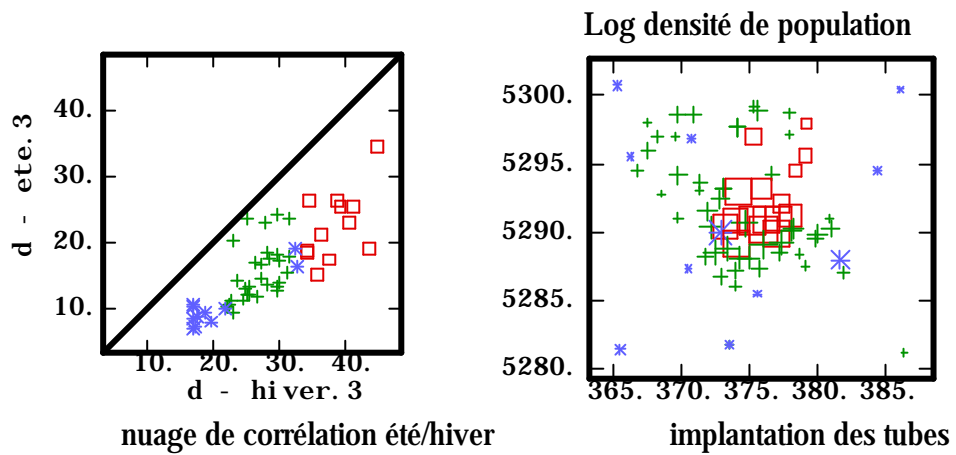
Vers le nord, deux tubes rapprochés appartiennent l'un, au mode des fortes moyennes hivernales, et l'autre, au mode des fortes moyennes estivales.

**L'implantation et par suite la structure spatiale des concentrations, en particulier des valeurs fortes, diffèrent entre l'hiver et l'été.**

**Les faibles moyennes hivernales donnent de faibles moyennes estivales. Les fortes moyennes hivernales donnent des moyennes estivales plutôt fortes à fortes, alors que les fortes moyennes estivales correspondent à des moyennes hivernales intermédiaires à fortes.**



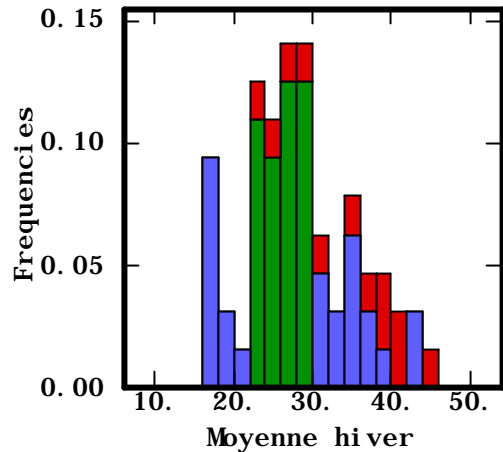
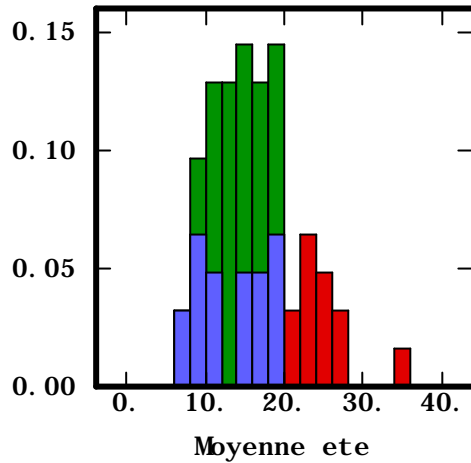
**a) histogrammes des moyennes saisonnières**



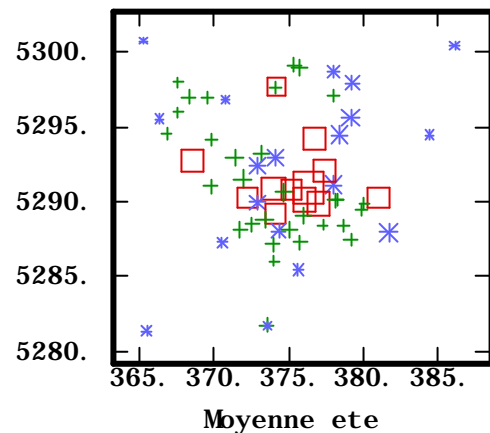
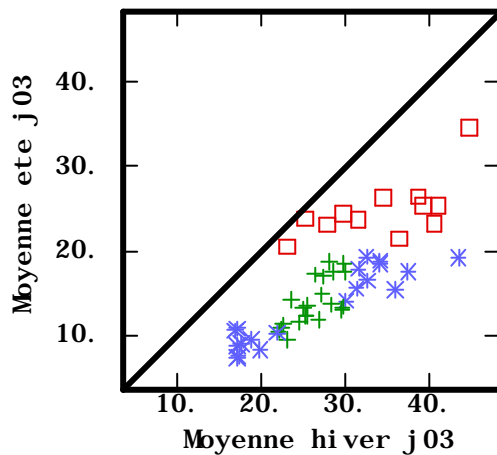
**b) Nuage de corrélation des moyennes saisonnières et implantation des tubes**

*Sur la carte d'implantation, les symboles sont proportionnels au logarithme translaté de la densité de population. La bissectrice est tracée sur le nuage de corrélation des moyennes saisonnières.*

**Figure 1. Classes de concentrations hivernales en NO2**



**a) histogrammes des moyennes saisonnières**



**b) Nuage de corrélation des moyennes saisonnières et implantation des tubes**

Sur la carte d'implantation, les symboles sont proportionnels à la moyenne estivale de NO<sub>2</sub>. La bissectrice est tracée sur le nuage de corrélation des moyennes saisonnières.

**Figure 2. Classes de concentration des moyennes saisonnières de NO<sub>2</sub>**

Remarque : dans une population d'histogramme unimodal et présentant une corrélation spatiale, un suréchantillonnage local peut engendrer un mode supplémentaire sur l'histogramme des concentrations. Bien que la densité d'implantation des stations soit localement plus forte, au centre de l'agglomération ainsi que vers le nord et le nord-est, les modes sur les histogrammes ne semblent pas résulter de ces resserrements locaux des tubes.

## **1.3. Estimation monovariante de la concentration en NO<sub>2</sub>**

### **1.3.1. Etude variographique monovariante de NO<sub>2</sub>**

Les variogrammes expérimentaux directionnels des moyennes saisonnières et annuelles de NO<sub>2</sub> indiquent une légère anisotropie pour la moyenne hivernale, la variabilité suivant les directions Est-Ouest et à 135 degrés étant plutôt supérieure aux petites distances, et inférieures aux plus grandes, à la variabilité suivant les directions Nord-Sud et à 45 degrés (figure 3.), ce que confirme la carte variographique (figure 4.). Recalculés avec les 52 stations comportant des mesures pour les 6 quinzaines, les variogrammes directionnels de la concentration hivernale présentent des anisotropies moins marquées (figures 3.).

Si les paliers restent inférieurs, les fluctuations du variogramme sont nettement plus fortes pour la moyenne estivale ou annuelle que pour la moyenne hivernale : suivant la direction Est-Ouest, le variogramme de la moyenne estivale semble pratiquement péritique. La forte variabilité aux premiers pas de ce variogramme peut s'expliquer, comme indiqué précédemment, par l'extension Est-Ouest des fortes concentrations estivales.

La variabilité spatiale de la concentration en NO<sub>2</sub> varie donc selon la saison. Mais le nombre de stations informées exhaustivement durant les 2 saisons ne permet pas une comparaison vraiment précise de ces structures. Dans la suite, nous retenons une modélisation simplifiée en négligeant les anisotropies des moyennes saisonnières ou annuelles, mal identifiables à partir des 52 tubes "annuels".

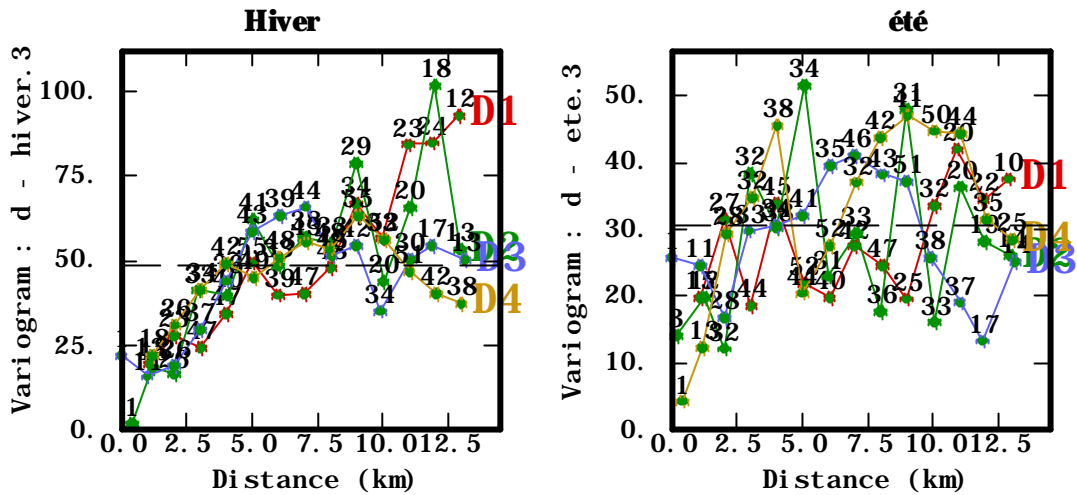
#### **Corrélation des moyennes saisonnières de NO<sub>2</sub>**

Les principales différences dans la structure spatiale des concentrations saisonnières proviennent de deux groupes de tubes (figure 5.):

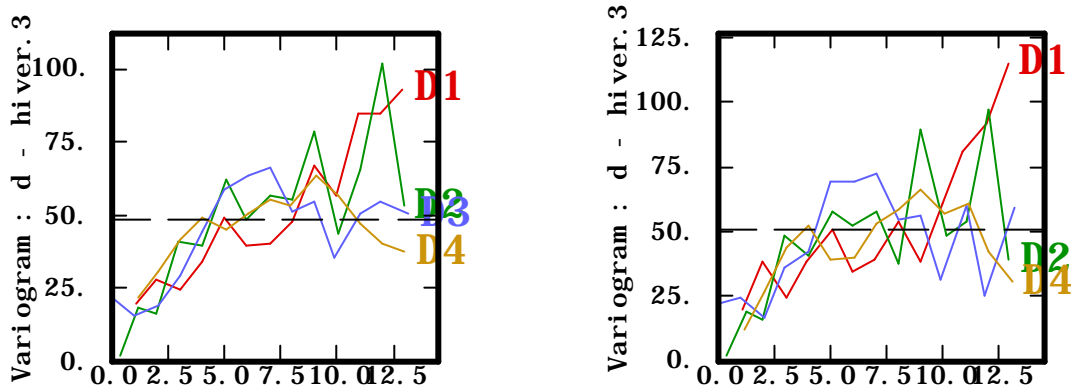
- trois tubes implantés vers le Nord, et figurés par des étoiles bleues, ont des concentrations estivales moyennes et des concentrations hivernales fortes, et par suite des écarts saisonniers plus importants, supérieurs à 30 ng/m<sup>3</sup>.

- certains tubes appartenant au mode des fortes concentrations estivales correspondent à des concentrations hivernales recouvrant un large intervalle de valeurs intermédiaires. Certains de ces tubes, figurés par des carrés rouges, sont situés vers l'ordonnée  $y=5290$  dans la partie centrale de l'agglomération, ainsi qu'en périphérie vers le Nord et le Nord-Ouest (entre les ordonnées  $y=5290$ . et 5295, ainsi que l'un des tubes du doublon au Nord) ; ils présentent des écarts plus réduits entre les moyennes saisonnières. Par rapport à la majorité des points du nuage, la moyenne estivale est ici majorée d'environ 5 ng/m<sup>3</sup>.

Supposons la moyenne hivernale fortement corrélée à l'une des variables explicatives, en particulier pour les fortes concentrations ; comme les tubes présentant des valeurs fortes diffèrent entre les deux saisons, alors cette variable explicative sera moins bien corrélée aux moyennes estivales, en particulier pour à leurs fortes concentrations. La différence d'implantation des fortes concentrations suivant les saisons explique la différence de liaison entre ces moyennes saisonnières et les variables auxiliaires.



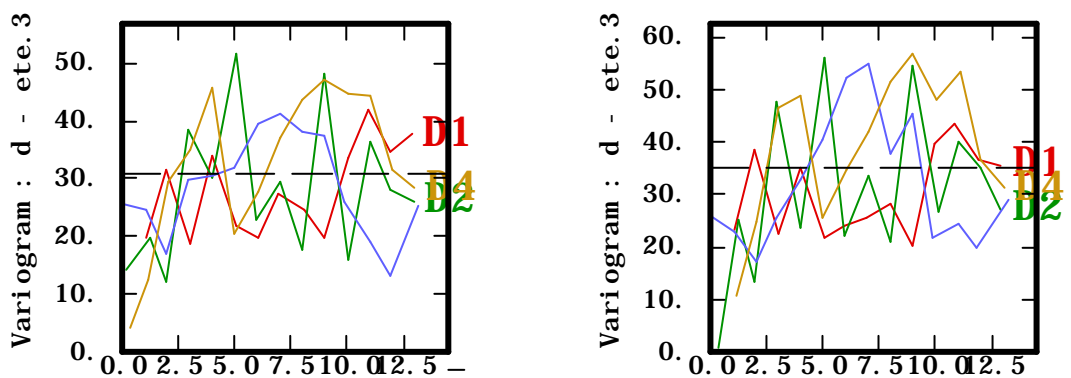
a) moyennes saisonnières, tous tubes. Le nombre de couples par point du variogramme expérimental est reporté.



64 tubes hivernaux

52 tubes annuels

b) moyenne hivernale



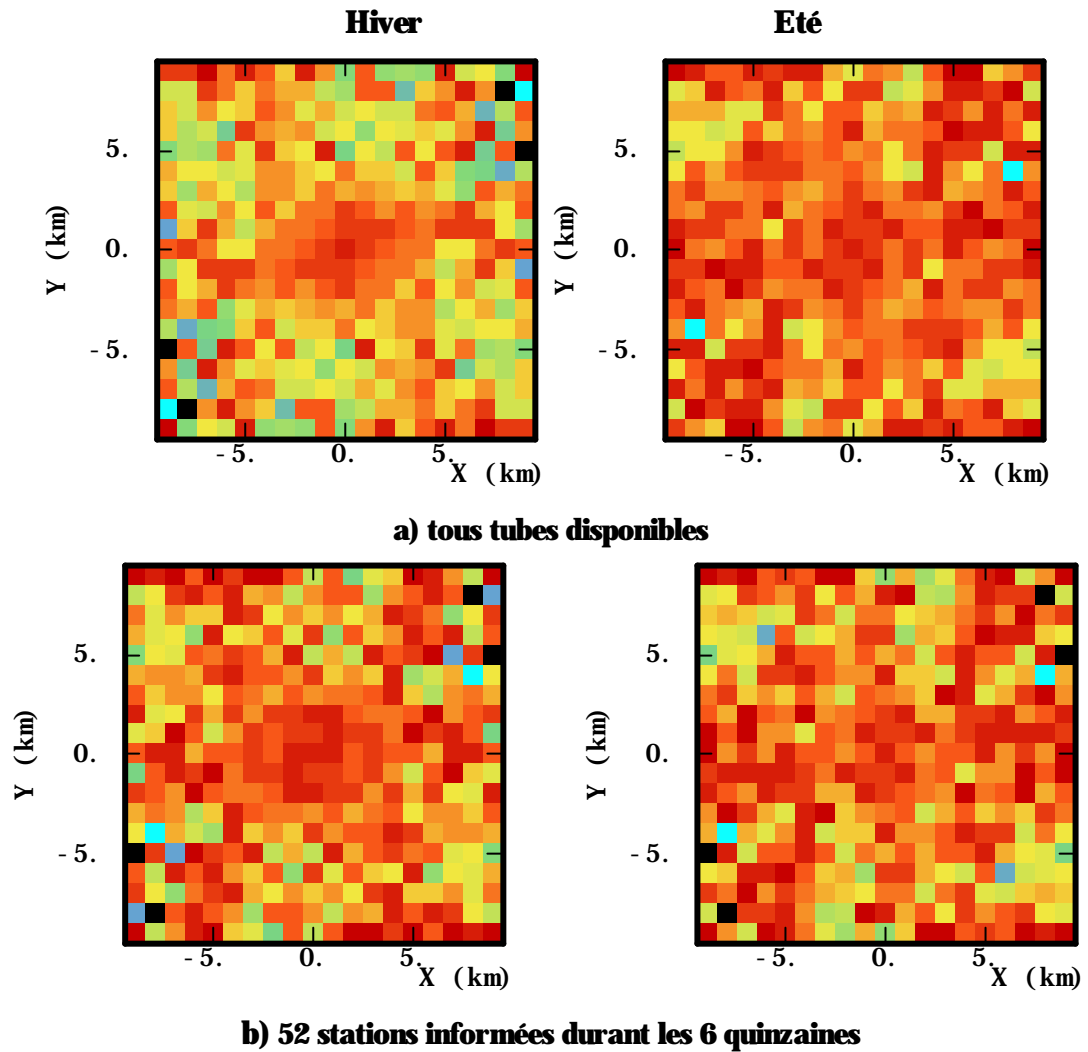
62 tubes estivaux

52 tubes annuels

c) moyenne estivale

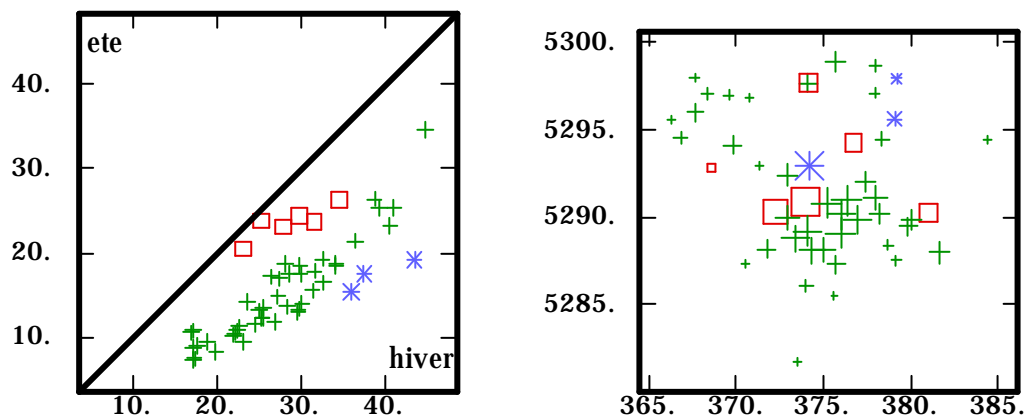
D1 : Est-Ouest, axe Ox ; D2 : 45 degrés ; D3 : Nord-Sud, axe Oy ; D4 : 135degrés.

**Figure 3. Variogrammes de la moyenne saisonnière de NO2 (pas : 1km).**



Le rouge indique les valeurs faibles.

**Figure 4. Carte variographique des moyennes saisonnières (pas : 1 km)**



Sur la carte d'implantation (à droite), le tracé est proportionnel au logarithme translaté de la densité de population.

**Figure 5. Comparaison des moyennes saisonnières.**

### 1.3.2. Modélisation monovariante

La détermination des anisotropies saisonnières étant incertaine, nous ajustons les variogrammes omnidirectionnels des moyennes saisonnières et annuelle (voir le variogramme expérimental de la concentration annuelle en figure 6.), d'abord séparément les uns des autres sans rechercher nécessairement de structure commune (figure 7. ; ces variogrammes sont calculés avec un nombre de tubes différents selon les saisons). L'effet de pépite ajusté est supérieur pour la moyenne estivale ( $12 \text{ (mg/m}^3\text{)}^2$ ) à celui de la moyenne hivernale ( $7 \text{ (mg/m}^3\text{)}^2$ ), alors même que, pour les 52 tubes communs, moyenne et variance hivernales ( $m=28.3 \text{ mg/m}^3$ ,  $s^2 = 50.4 \text{ (mg/m}^3\text{)}^2$ ) sont supérieures aux valeurs estivales ( $m=16.2 \text{ mg/m}^3$ ,  $s^2 = 35.0 \text{ (mg/m}^3\text{)}^2$ ). Pour la moyenne annuelle, l'effet de pépite ajusté est nul. La portée de la structure d'échelle plurikilométrique est supérieure en hiver (6.5km), et inférieure en été ou pour l'année (4.7km et 4.5km, respectivement). La non stationnarité spatiale des variogrammes de la moyenne hivernale ou annuelle à l'échelle d'observation est prise en compte par une composante linéaire.

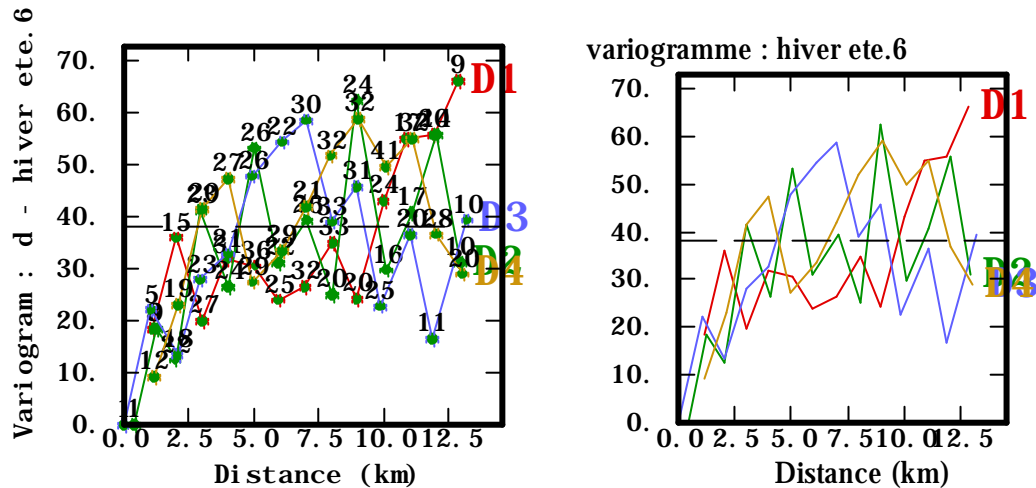
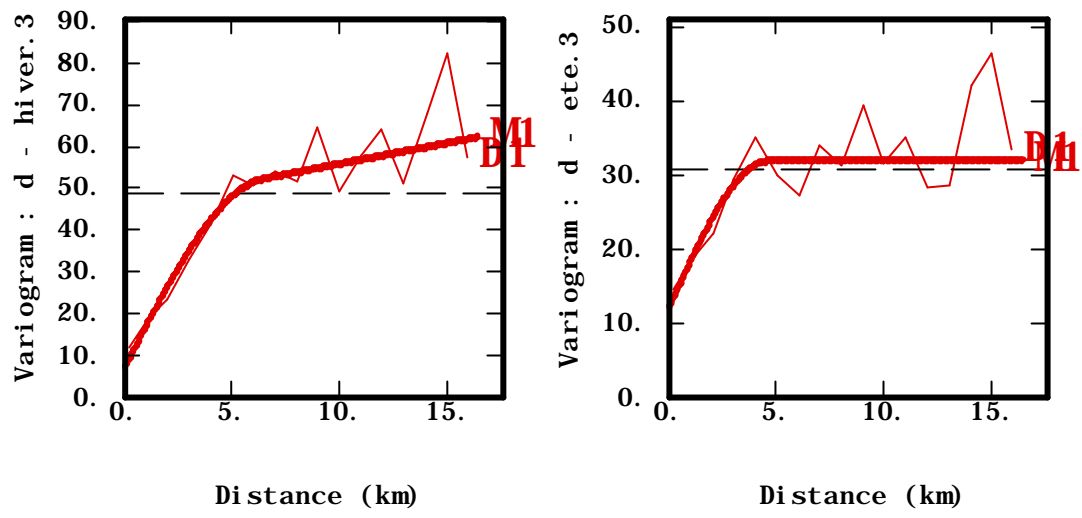
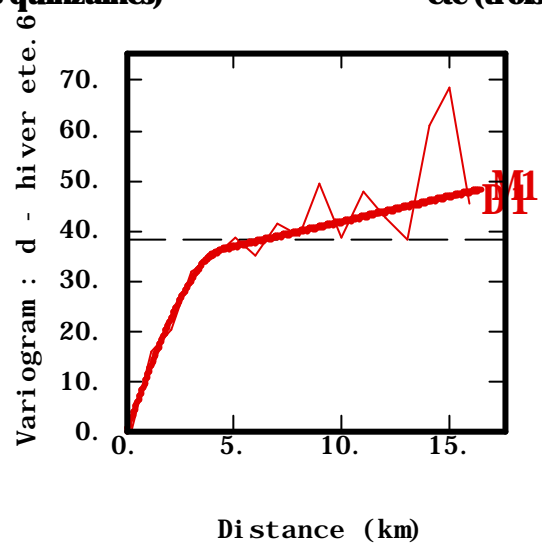


Figure 6. Variogrammes de la moyenne annuelle de NO2 (52 tubes. Pas : 1 km).



**hiver (trois quinzaines)**

**été (trois quinzaines)**



**annuel (six quinzaines)**

- Modèle NO2 hiver : 7.+ 39. Sphérique (6.5km) + 1 lin(km)
  - Modèle NO2 été : 12. + 20. Sphérique(4.7km)
  - Modèle NO2 annuel : 32 Sphérique(4.5km)+1lin(km)
- Pas du calcul des variogrammes expérimentaux : 1km

**Figure 7. Ajustement des variogrammes omnidirectionnels de NO2 saisonnier ou annuel.**

### 1.3.3. Validation croisée

Le critère de comparaison est la validation croisée usuelle. Comme les tubes sont plus nombreux car plus resserrés au centre de l'agglomération, les statistiques traduisent principalement la qualité de la modélisation dans cette zone. Pondérer chaque tube, en attribuant un poids plus important aux tubes situés en périphérie, permettrait éventuellement de corriger les statistiques d'erreur d'estimation, en atténuant l'influence prépondérante de la zone urbaine dense.

Remarque : les statistiques de validation croisée ne sont ainsi pas nécessairement "objectives", car elles dépendent de l'implantation des données. Si certaines zones sont suréchantillonnées par rapport à d'autres, ces statistiques peuvent indiquer la qualité du modèle de façon préférentielle pour ces zones.

Les résultats de validation croisée, obtenus avec un même voisinage à 10km, sont présentés en annexe aux figures 16. à 18., ainsi qu'au tableau 1., où sont rappelés les "meilleurs" résultats obtenus précédemment dans les modèles avec dérive externe (cf. 3ème rapport d'avancement).

L'apport des variables "explicatives" utilisées en dérive externe apparaît plus marqué l'hiver, aussi bien selon le coefficient de corrélation entre valeur mesurée  $Z$  et valeur estimée  $Z^*$ , que selon la moyenne et la variance des erreurs d'estimation. L'intérêt des variables explicatives est moindre pour l'été, en particulier pour la moyenne des erreurs quadratiques. Quant-à la moyenne annuelle, la validation croisée donne des résultats analogues pour les modèles avec ou sans dérive.

Pour l'estimation monovariante, le coefficient de corrélation le plus élevé entre valeur mesurée et validation croisée s'observe pour la moyenne annuelle. Pour la moyenne estivale, le nuage de corrélation ( $Z, Z^*$ ) est redressé par rapport à la première bissectrice, ce qui traduit une dispersion des estimations bien inférieure à celle des concentrations mesurées, autrement dit, un fort effet lissant de l'estimation. Dans ce cas, couper sur l'estimation  $Z^*(x)$  au lieu de la concentration "réelle"  $Z(x)$  pour délimiter les endroits où la concentration estivale dépasse un seuil  $s$ , peut conduire à de nombreuses erreurs de sélection : c'est "l'effet d'information", d'autant plus marqué que le nuage de corrélation ( $Z, Z^*$ ) s'éloigne de la première bissectrice et que le seuil  $s$  prend des valeurs fortes ou faibles par rapport à la moyenne.

Dans la suite, nous détaillons les deux questions suivantes :

- peut-on apprécier l'apport effectif des variables auxiliaires pour l'estimation de la moyenne annuelle ou saisonnière, en particulier pour les concentrations estivales, estimées avec une précision relative médiocre ?

- l'estimation de la moyenne annuelle est-elle améliorée si on la calcule comme la moyenne des estimations saisonnières, ce qui permettrait d'utiliser les variables "explicatives" pour la moyenne hivernale, ainsi que l'ensemble des tubes respectivement disponibles pour l'estimation des moyennes saisonnières ? De plus, la cohérence entre moyennes saisonnières et annuelle serait ainsi assurée.

période	modèle	r (Z,Z*)	erreur		erreur relative		err.standardisée	
			m	s2	m	s2	m	s2
hiver 60 (64)	NO2	.76	0.27	19.6	0.04	0.031	0.03	0.81
	g linéaire population, bâti dense	.83	0.12	14.4	0.02	0.026	-	-
été 59 (62)	NO2	.62	0.15	18.3	0.08	0.092	0.02	0.79
	g sphérique 3km pop., bâti dense, NOx	.67	0.13	17.0	0.05	0.088	-	-
annuel 49 (52)	NO2 annuel	.81	0.27	12.6	0.05	0.040	0.03	0.69
	g linéaire population, bâti dense	.81	0.05	12.5	0.03	0.032	-	-
	g linéaire NOx	.82	0.21	11.8	0.03	0.035	-	-
	<b>NO2 hiver + été</b>	<b>.80</b>	<b>0.18</b>	<b>13.6</b>	<b>0.05</b>	<b>0.041</b>		
	<b>DE hiver + été</b>	<b>.85</b>	<b>0.07</b>	<b>10.20</b>	<b>0.03</b>	<b>0.030</b>		

**En gras** : moyenne des estimations hivernales et estivales, respectivement en krigeage monovariante ou avec dérive externe.

Les résultats avec dérive externe sont repris du 3ème rapport d'avancement.

Erreur :  $Z^*(x) - Z(x)$  ; erreur relative :  $\frac{Z^*(x) - Z(x)}{Z(x)}$  ; erreur standardisée  $\frac{Z^*(x) - Z(x)}{s_k(x)}$

r (Z,Z\*) : coefficient de corrélation entre la mesure et le résultats de validation croisée

m : moyenne, s2 : variance, g : variogramme

L'effectif des données estimées en validation croisée est reporté. La différence avec le nombre total de mesures disponibles indiqué entre parenthèses provient d'un nombre insuffisant de données dans le voisinage de krigeage, pour la réestimation de certains tubes.

Pour les modèles avec dérive externe, les statistiques des erreurs standardisées ne sont pas reportées, le palier des variogrammes n'ayant pas été calé.

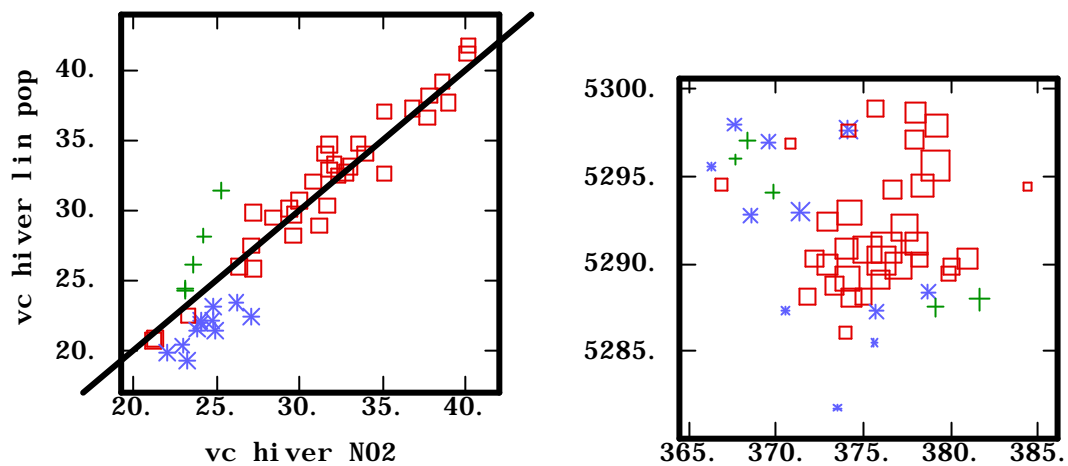
Voisinage de krigeage : 10km

**Tableau 1. Validation croisée. Comparaison des moyennes saisonnières ou annuelle en modèle monovariante et avec dérive externe.**

### 1.3.3.a. Comparaison des validations croisées pour NO2 hivernal

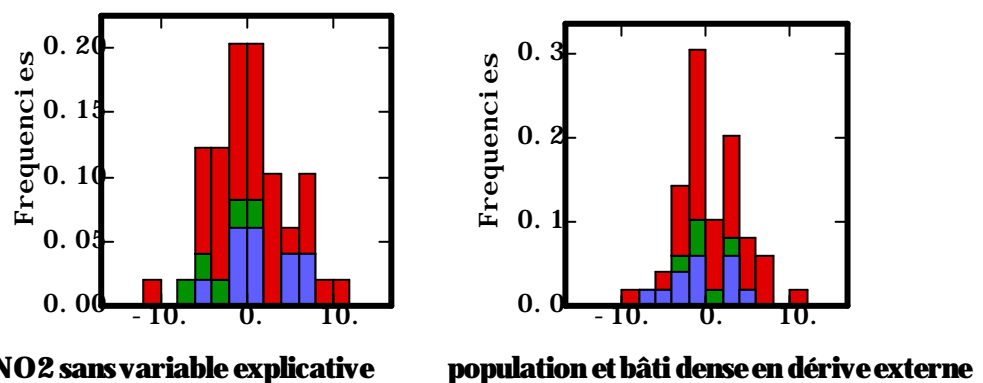
Les histogrammes des erreurs de validation croisée avec ou sans variable explicative (logarithme translaté du bâti dense et de la densité de population) confirment des erreurs localement plus importantes pour le krigeage monovarié (figure 8.). Les deux erreurs les plus fortes en valeur absolue (une positive et l'autre négative) se retrouvent respectivement aux mêmes tubes pour les deux estimations (le tube le plus à l'est pour cette saison, ainsi que celui situé au nord-ouest du point  $x=380, y=5295$ ).

Les écarts les plus importants *entre les deux estimations* sont localisés en périphérie de l'agglomération, notamment au Nord-Ouest et au Sud-Est. Pour ces 16 tubes (étoiles bleues ou croix vertes sur la figure, selon que l'estimation multivarié est inférieure ou supérieure à l'estimation monovarié), les statistiques d'erreur de validation croisée indiquent une moyenne des erreurs de  $m=-0.80$  pour NO2 monovarié et  $m=0.20$  en dérive externe, et une variance des erreurs de 16.7 pour NO2 monovarié et 10.5 en dérive externe.



Sur le nuage de corrélation, la droite indique la première bissectrice.

#### a) Nuage de corrélation des estimations et implantation des tubes



#### b) histogramme des erreurs de validation croisée $Z^*-Z$ (hiver)

Etoiles bleues : la validation croisée monovarié est inférieure à celle avec dérive externe ; croix vertes : c'est l'inverse ; carrés rouges : résultats plutôt analogues pour les deux modèles.

**Figure 8. NO2 hivernal. Comparaison des validations croisées avec et sans variable auxiliaire (49 tubes de la validation croisée annuelle).**

La dérive externe permettrait effectivement de “recaler” le modèle dans les zones peu échantillonnées en NO<sub>2</sub>, ce qui diminue la variance des erreurs de validation croisée.

Remarque : en un tube, deux erreurs intermédiaires, mais de signe opposés, peuvent se traduire par un écart important entre les deux modèles. Par ailleurs, un écart faible entre les deux modèles, peut correspondre à de fortes erreurs de validation croisée, analogues pour les deux modèles.

### **1.3.3.b. NO<sub>2</sub> estival**

Les erreurs relatives sont ici plus importantes. Le modèle avec dérive externe donne un coefficient de corrélation entre validation croisée et mesures aux tubes légèrement supérieur, et une variance des erreurs de validation croisée légèrement inférieure, à ceux du modèle monovarié. La variance de l’erreur de validation croisée est supérieure en été à celle de la “meilleure” validation croisée hivernale (mais les tubes ne sont pas tout-à-fait les mêmes).

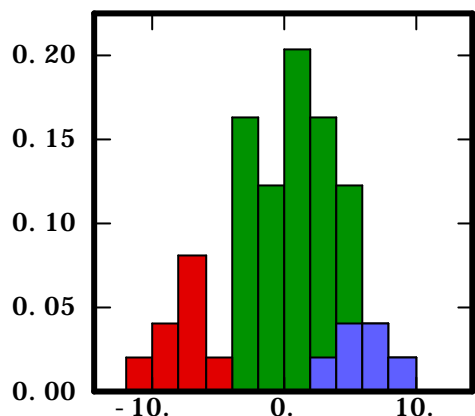
La majorité des fortes erreurs se retrouvent aux mêmes tubes pour les deux modèles, quelques-unes étant cependant corrigées dans le modèle avec dérive externe (figure 9.). C’est en particulier le cas pour les deux “carrés rouges”, situés autour de  $y=5290$  et entre  $x=370$ . et  $375$ , sous-estimés en validation croisée monovarié mais pas dans le modèle avec dérive externe. Le repérage des fortes erreurs de validation croisée sur le nuage de corrélation de ces estimations montre que les deux modèles peuvent fournir en certains tubes des valeurs proches, mais erronées.

De fortes erreurs de validation croisée se retrouvent, aussi bien en dérive externe que pour le NO<sub>2</sub> monovarié, pour certains tubes implantés dans la tranche Est-Ouest comprise entre  $y=5290$ . et  $y=5295$ km ou à son voisinage. Dans cette tranche, certaines concentrations fortes sont sous-estimées en validation croisée à cause de concentrations voisines moyennes ; à proximité immédiate, des valeurs intermédiaires sont surestimées en validation croisée, à de ces fortes concentrations. Cet effet s’observe également pour le doublon au Nord, qui présente deux fortes erreurs de validation croisée de signes opposés.

### **1.3.3.c. NO<sub>2</sub> annuel**

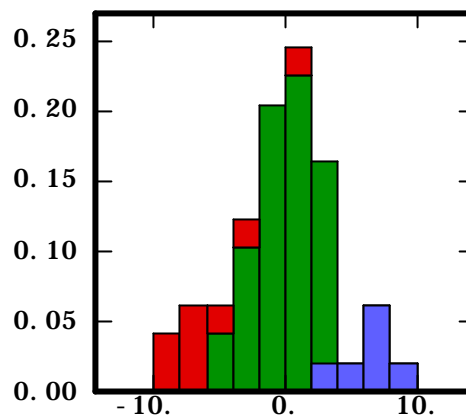
Les fortes erreurs de validation croisée se retrouvent aux mêmes tubes, pour l’estimation monovarié ou pour les deux modèles avec dérive externe. De forts écarts entre le modèle monovarié et celui avec les émissions de NO<sub>x</sub> en dérive externe se trouvent au centre et au sud de l’agglomération (figure 10.). Supprimer des statistiques le tube à l’est présentant une réestimation systématiquement médiocre fournit des résultats légèrement favorables au modèle avec NO<sub>x</sub> en dérive externe.

Il n’apparaît pas de critère décisif en faveur de l’un des modèles, NO<sub>2</sub> monovarié ou l’un des modèles avec dérive externe



vc ete NO2 - ete

**NO2 sans variable explicative**

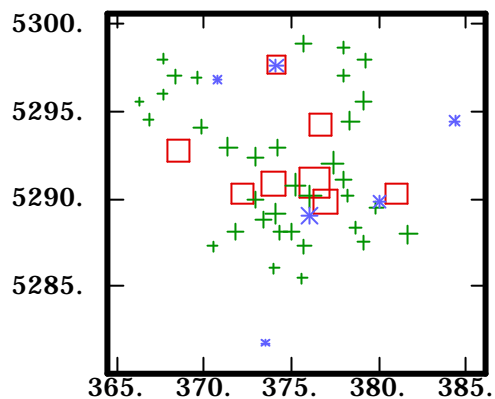
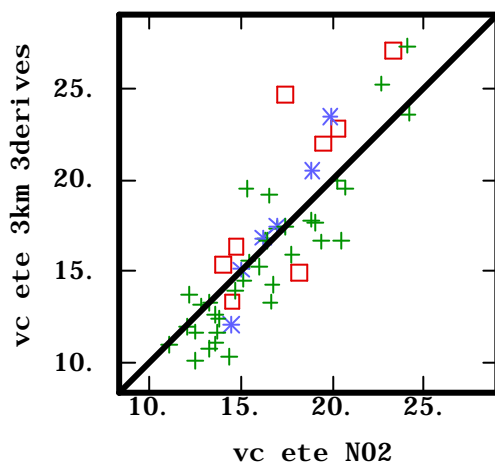


vc ete 3km 3der - ete

**population et bâti dense en dérive externe**

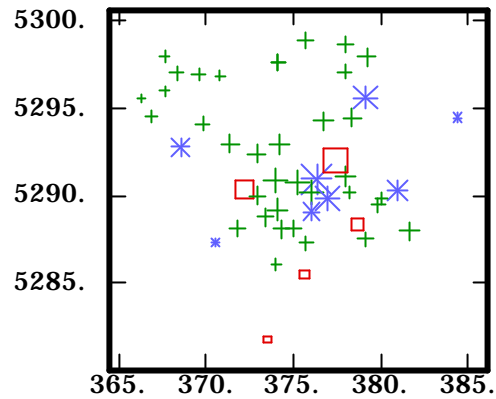
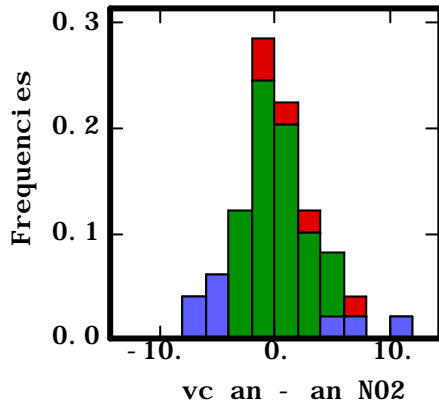
*Validation croisée monovariante : carrés rouges,  $Z^*$  sous-estime fortement  $Z$ ; étoiles bleues,  $Z^*$  surestime fortement  $Z$ . Les tubes repérés se correspondent dans les quatre graphiques.*

**a) histogramme des erreurs de validation croisée  $Z^* - Z$**

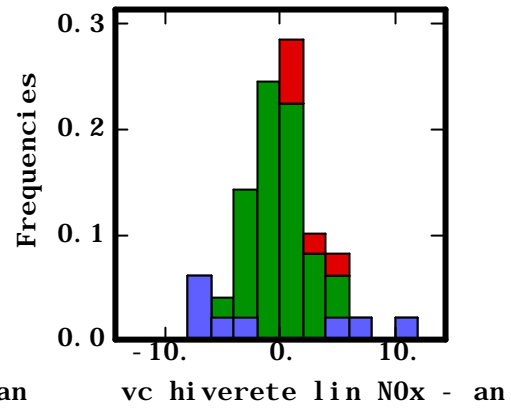
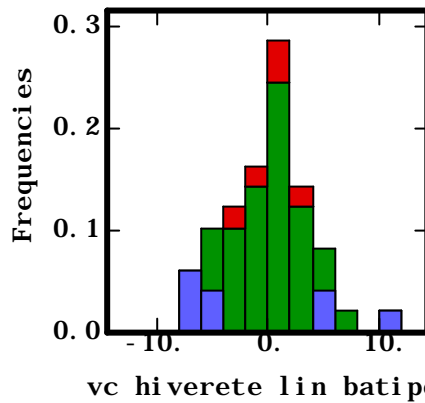


**b) Nuage de corrélation des estimations et implantation des tubes repérés**

**Figure 9. NO2 estival. Comparaison des validations croisées avec et sans variable auxiliaire (49 tubes de la validation croisée annuelle).**



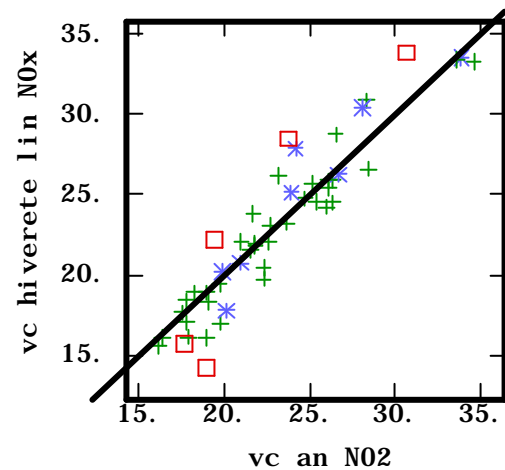
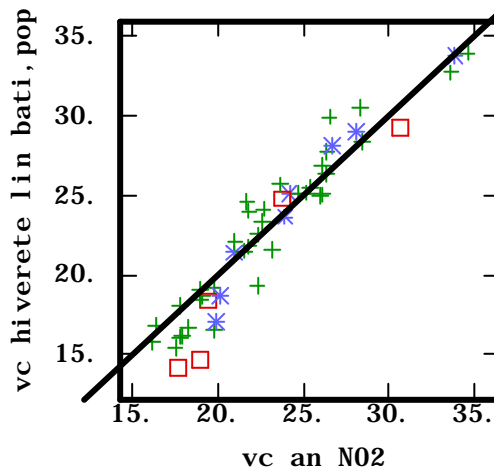
**Estimation monovariante de NO2 annuel**



**Population et bâti dense en dérive externe**

**Emissions de NOx en dérive externe**

**a) histogramme des erreurs de validation croisée  $Z^* - Z$  et repérage des tubes**



**Population et bâti dense en dérive externe**

**Emissions de NOx en dérive externe**

*Carrés rouges : forts écarts entre le modèle monovariante et l'un des modèles avec dérive externe; étoiles bleues : fortes erreurs absolues d'estimation (dans le même sens pour les trois modèles).*

**b) Nuage de corrélation des estimations. Abscisse : estimation monovariante de NO2**

*Les tubes repérés se correspondent dans les différents graphiques.*

**Figure 10. NO2 annuel. Comparaison des estimations avec et sans variable auxiliaire.**

### 1.3.3.d. Discussion

L'examen des erreurs de validation croisée confirme les statistiques globales du tableau 1. (avec les réserves faites sur l'objectivité de ce critère). La densité de population, le bâti, ou les émissions de NOx (cf. 3ème rapport CG) améliorent l'estimation pour la moyenne hivernale, pour laquelle cette amélioration porte principalement sur les zones en périphérie (cf. implantation des tubes repérés, figure 8.), qui comportent moins de tubes. Pour cartographier NO2 sur une zone peu couverte par les tubes, le modèle avec dérive se montre utile.

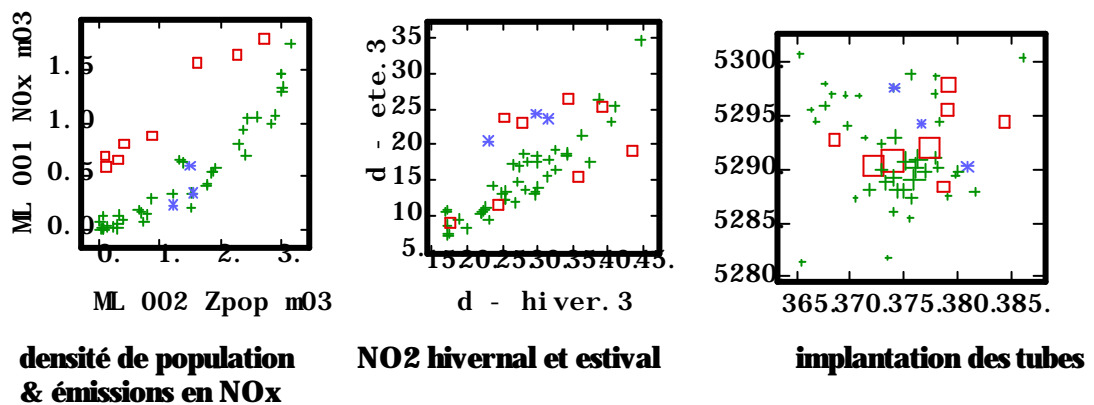
Les fortes concentrations estivales sont mal ré-estimées. Ceci est à relier au variogramme quasi pépitique dans la direction est-ouest, correspondant à un fort contraste de concentrations dans cette zone de valeurs fortes. Au contraire, les "fortes" valeurs de la concentration hivernale forment une tache plus étendue, ce qui permet une meilleure estimation croisée d'une de ces concentrations par les mesures voisines.

Les corrélations entre les variables "explicatives" et les moyennes saisonnières ou annuelles expliquent-elles ces résultats?

Le nuage de corrélation entre densité de population et émissions de NOx montre un groupe de tubes (repérés par des carrés rouges) correspondant à des émissions en NOx plutôt fortes, relativement à la densité de population. Il serait intéressant de vérifier si fortes émissions correspondent à des zones plus industrielles ou résidentielles (Figure 11.). Certains de ces tubes appartiennent au groupe des "teneurs estivales plus fortes" précédemment discuté, mais ce n'est pas systématique.

Réciproquement, ces tubes à "teneurs estivales plus fortes" n'apparaissent pas de façon caractéristique sur les nuages de corrélation des variables explicatives, ce qui montre l'inaptitude de ces dernières à expliquer les différences de structure spatiale des moyennes hivernales et estivales.

La présence de plusieurs populations sur les nuages de corrélation explique l'apport différencié des variables "explicatives" sur les moyennes de NO2. Mais ces écarts entre variables explicatives portant sur un nombre assez réduit de tubes, leur influence sur les résultats de validation croisée reste faible.



**Figure 11. Fortes émissions en NOx relativement à la densité de population.**

### **1.3.3.e. Retour sur les liaisons entre variables auxiliaires aux tubes**

La figure 12. récapitule, pour les 75 tubes instrumentés, les relations entre variables auxiliaires, en logarithme translaté : densité de population, bâti dense et émissions de NOx.

Les classes de forte densité de population, repérées par des carrés rouges, recouvrent les classes de valeurs intermédiaires ou fortes du bâti dense. Si les maxima de population et de bâti dense se correspondent, le nuage de corrélation reste cependant très dispersé, notamment pour les valeurs moyennes ou fortes.

Les valeurs fortes de la densité de population sont liées à des émissions plutôt fortes ou moyennes, mais ne correspondent pas à toutes les valeurs fortes de NOx. Ce nuage de corrélation densité de population-émissions de NOx indique une population de tubes, mentionnée précédemment, caractérisée pour une densité de population plutôt faible, par des émissions plutôt fortes. Dans cette population, le sous-groupe correspondant à une faible densité de population est repéré ici par des étoiles bleues ; pour ces tubes le bâti dense prend des valeurs faibles ou moyennes. Il pourra être utile de vérifier, dans ces deux cas, si ces fortes émissions de NOx non liées à la densité de population sont liées au bâti industriel ou lâche.

Malgré la présence de deux populations différentes sur leur nuage, pour les trois variables auxiliaires retenues, les plus fortes corrélations s'observent entre densité de population et émissions de NOx. Le bâti dense fournit donc une information différente de celle apportée par les deux autres variables (voir aussi l'ACP figure 22 du premier rapport d'avancement, entre moyennes saisonnières et variables auxiliaires).

Ceci est cohérent avec les variables retenues en dérive : lorsque 2 variables sont présentes, le bâti en fait partie, car il apporte une information différente par rapport aux deux autres Variables auxiliaires.

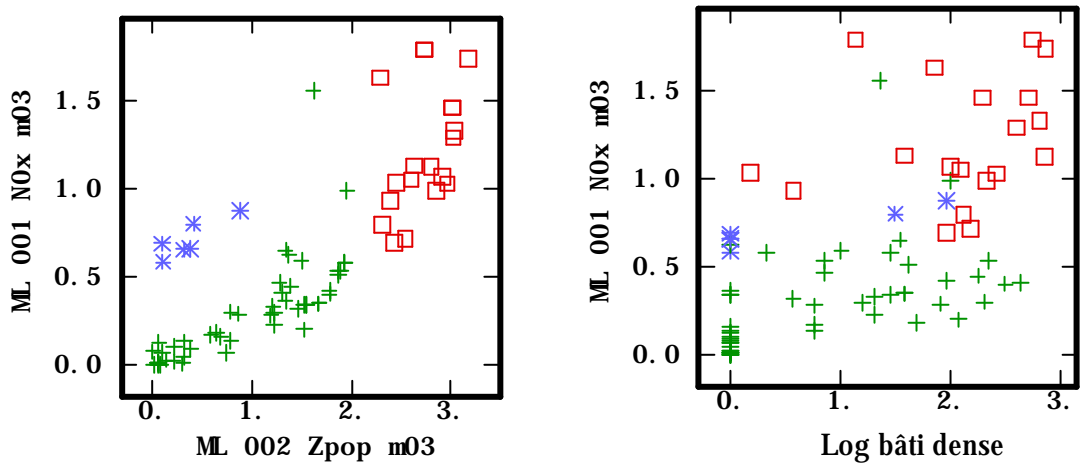
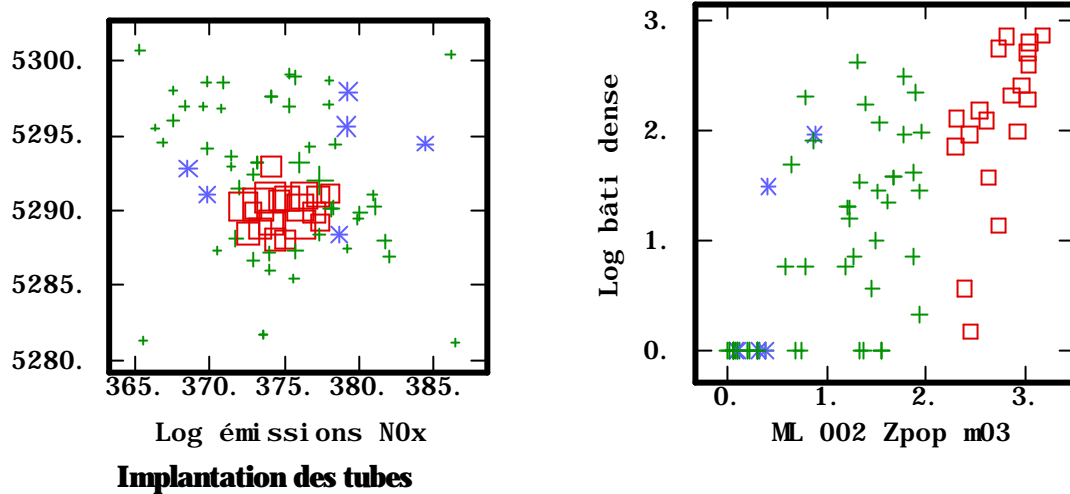
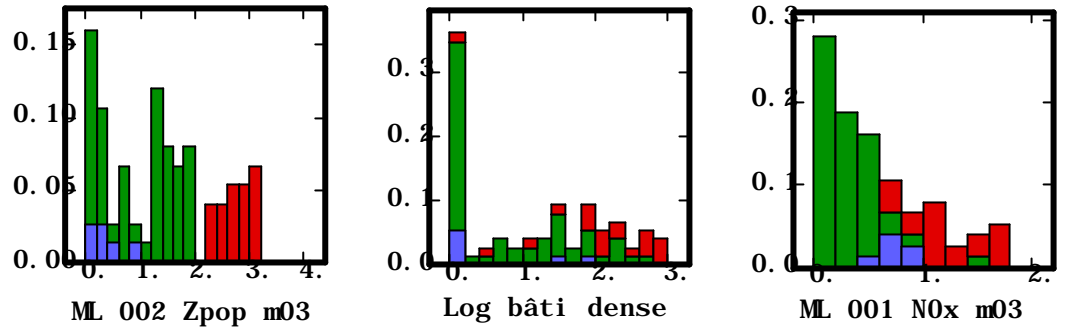
### **1.3.3.f. Retour sur les liaisons entre variables auxiliaires et NO2**

Sur la figure 13. sont repérées les fortes concentrations respectivement pour la moyenne hivernale et pour la moyenne estivale. Ces corrélations ont été examinées en détail dans le troisième rapport d'avancement. On note que les valeurs relativement fortes des coefficients de corrélation entre les variables auxiliaires et la moyenne estivale sont partiellement dus à un seul tube, correspondant au maximum de la concentration estivale

Les plus fortes concentrations de la moyenne hivernale sont liées aux fortes valeurs du bâti dense ; la réciproque est cependant fautive : à de fortes valeurs du bâti dense sont associées également des valeurs moyennes de la concentration hivernale.

Sauf en un tube, aux fortes émissions de NOx sont associées de fortes concentrations estivales. Mais certaines fortes concentrations estivales s'observent en des points d'émissions moyennes en NOx. NOx seul n'explique donc pas toutes les fortes concentrations estivales. Ceci est cohérent avec l'estimation de la moyenne estivale avec NOx en dérive, mais avec des résultats de validation croisée assez médiocres.

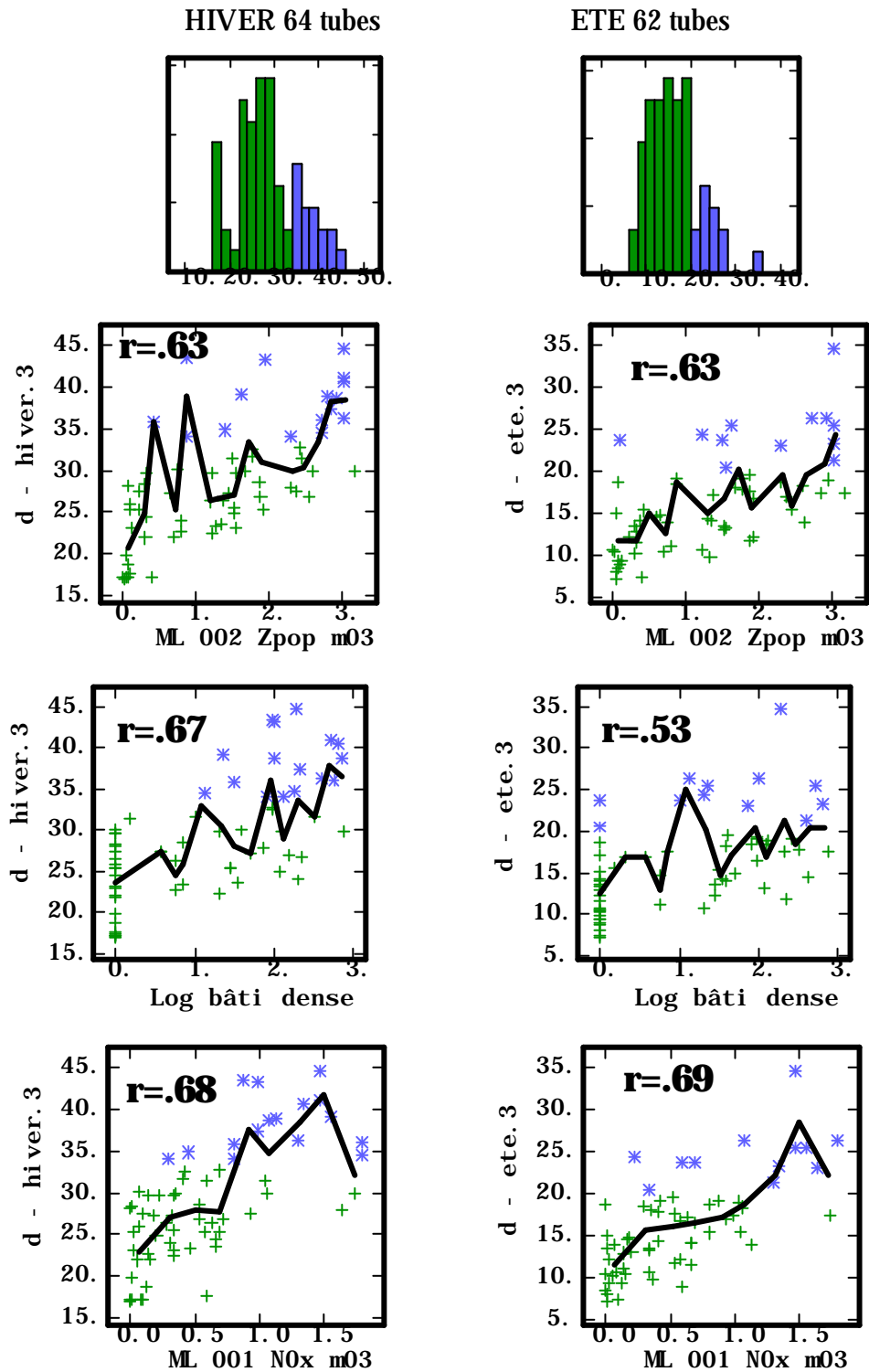
Pour la densité de population ainsi que pour le bâti dense, la dispersion des nuages de corrélation avec la moyenne estivale est très forte : à une même classe de variables explicatives, correspondent des valeurs très différentes de la concentration estivale. Peut-être le bâti lâche permettrait-il de représenter localement des émissions liées à la circulation automobile ?



Corrélations des Log translétés, pour 75 sites

population, bâti dense	.75
population, émissions NOx	.80
bâti dense, émissions NOx	.62

**Figure 12. Relations entre trois variables auxiliaires**



**Figure 13. Liaisons entre variables auxiliaires et moyennes saisonnières.**

## 1.4. Estimation de la moyenne annuelle par les moyennes saisonnières

L'apport des variables explicatives est démontré pour la moyenne hivernale, mais pas pour la moyenne annuelle. Améliore-t-on l'estimation de la moyenne annuelle, en la calculant comme la moyenne des estimations saisonnières, ce qui permettrait

- de reporter sur la moyenne annuelle le gain de précision fourni par les variables auxiliaires sur la moyenne hivernale ;
- par cokrigage, d'utiliser toutes les mesures de moyennes saisonnières, au lieu des seules mesures disponibles simultanément en hiver et en été, comme c'est le cas pour le krigeage "direct" de la moyenne annuelle ?

Laissant de côté la question du calcul de la variance de l'erreur d'estimation, nous effectuons la moyenne des (estimations par) validations croisées hivernales et estivales, et comparons cette "validation croisée" de la moyenne annuelle aux mesures "annuelles". Cette nouvelle validation croisée de la moyenne annuelle est calculée pour les estimations monovariées de NO<sub>2</sub> d'une part, puis pour les modèles avec dérive externe d'autre part (population et bâti dense en dérive en hiver, auxquels se rajoutent les émissions de NO<sub>x</sub> pour l'été), la moyenne hivernale et la moyenne estivale étant obtenues par validation croisée à partir de 64 et de 62 tubes respectivement.

En krigeage monovarié, l'estimation de la moyenne annuelle par la moyenne des estimations saisonnières n'apporte pas d'amélioration aux résultats de validation croisée (cf. tableau 1. en début de ce rapport, lignes en gras). Pour les modèles saisonniers avec dérive externe, les résultats de validation croisée de la moyenne annuelle s'améliorent sensiblement, pour le coefficient de corrélation comme pour la moyenne. En pratique, l'estimation de la moyenne annuelle par cokrigage des estimations saisonnières, nécessite une modélisation cohérente, utilisant par exemple les mêmes variables en dérive et un modèle variographique bivariable du type modèle linéaire de corégionalisation. Dans le calcul précédent, le variogramme est linéaire pour la moyenne hivernale, et sphérique de portée 3km, pour la moyenne estivale. Le cokrigage de la moyenne annuelle par les moyennes saisonnières nécessite d'adapter les conditions de non-biais.

Une autre solution, consiste à calculer les régressions, linéaires ou non, des moyennes saisonnières par rapport aux variables explicatives (variables non nécessairement identiques pour les deux saisons), et à ajuster un modèle variographique bivariable des résidus. *Des tests sont donc encore nécessaires pour choisir "un bon" modèle.*

Nous ne développons pas cette modélisation, qui nécessite de reprendre la comparaison de l'apport des différentes variables explicatives sur la précision de l'estimation des moyennes saisonnières *et annuelle*, et des différents modèles variographiques. *Ce cokrigage*, qui permet d'utiliser l'ensemble des moyennes saisonnières pour l'estimation de la moyenne annuelle, nous semble *une voie à développer*. Une extension consisterait à cokriquer les moyennes saisonnières (et donc la moyenne annuelle) par les mesures des quinzaines.

A titre indicatif, l'ajustement d'un modèle linéaire de corégionalisation des moyennes saisonnières et annuelle de NO<sub>2</sub>, cohérent avec la "relation de fermeture"

$NO_2^{AN} = \frac{1}{2} \gamma NO_2^{HIVER} + NO_2^{ETE}$ , est reporté en annexe. La modélisation sera évidemment à reprendre pour un cokrigage utilisant les variables explicatives.

## 1.5. Conclusion

La carte d'implantation des tubes, ainsi que le nuage de corrélation des concentrations saisonnières, indiquent une localisation différente des fortes concentrations respectivement hivernales et estivales, alors que pour ces deux saisons, les faibles concentrations sont situées principalement en périphérie, excepté vers le Nord de l'agglomération.

Les variogrammes des concentrations saisonnières confirment ces différences de structure spatiale. Les variables explicatives se corrélaient mieux avec la moyenne hivernale, ce qui est cohérent avec de meilleurs résultats de validation croisée pour cette saison, lorsque les variables explicatives sont introduites en dérive externe. La concentration estivale proportionnellement plus forte pour un ensemble de tubes n'a pu être expliquée par des caractéristiques particulières des variables explicatives communes à ces tubes. La forte dispersion des nuages de corrélation entre moyenne estivale et variables explicatives explique le faible apport de l'estimation multivariable pour cette saison.

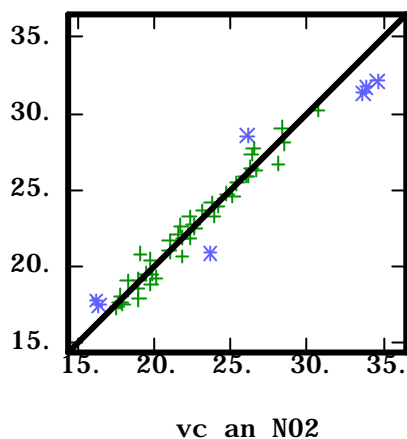
Il apparaît cependant utile, pour utiliser l'ensemble des mesures saisonnières (voire toutes les mesures par quinzaines), et pour assurer une bonne cohérence des estimations saisonnières et hivernales, de cokriquer la moyenne annuelle par les moyennes saisonnières.

Nous n'avons pas cherché à ajuster un "meilleur modèle" pour ce cokrigeage. Les résultats présentés dans le 3ème rapport d'avancement indiquent une sensibilité des modèles (pour les variables explicatives notamment) au choix des tubes. Il semble alors plus utile d'ajuster un modèle tenant compte des récentes campagnes de mesures, incluant d'éventuels tubes supplémentaires.

*Il est également important de préciser le critère de choix d'un "meilleur modèle".* Comme les tubes sont resserrés au centre de l'agglomération, la validation croisée classique privilégie la bonne adaptation du modèle sur ces tubes, ce qui peut se justifier en relation avec la densité de population supérieure dans ces zones. Le critère serait différent si l'on souhaite donner le même "poids" à des zones de même superficie. Un critère privilégiant la bonne restitution des concentrations les plus fortes est également envisageable.

Des résultats préliminaires, non détaillés ici, semblent indiquer une amélioration de la corrélation entre les concentrations saisonnières de NO<sub>2</sub> et l'occupation des sols, fournie à maille 200m, lorsque les pourcentages des différents types de bâti sont cumulés.

Par ailleurs, la question de l'influence du support des variables explicatives, et notamment du bâti, et de son éventuelle optimisation reste en suspens.



N02/Mulhouse tubes mai 03

- Variable #1 : vc an NO2
- Variable #2 : vc hiver + vc ete

Nb. samples :	49	49
Minimum :	16.1614	17.3358
Maximum :	34.6085	32.1188
Mean :	23.0237	22.9413
Std. Dev. :	4.48788	4.17081
Coef of Var. :	0.194924	0.181803

Correlations :	1	0.974579
	0.974579	1

Regression Lines :

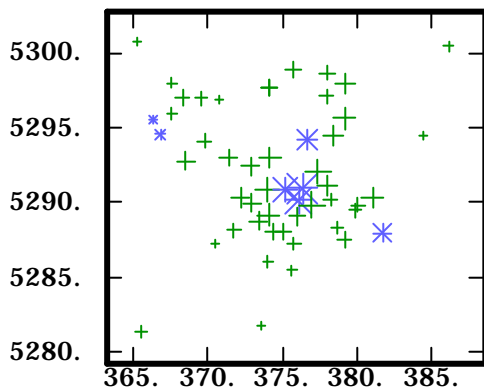
$$V2 = (0.905724) * V1 + (2.088224)$$

$$V1 = (1.048668) * V2 + (-1.034155)$$

vc an NO2 : krigeage de la moyenne annuelle de NO2 (validation croisée, estimation monovariante).

vc hiver + vc été : moyenne des validations croisées de la concentration saisonnière de NO2, chacune des estimations saisonnières étant monovariante.

**a) nuage de corrélation des deux estimations (validation croisée). Abscisse : estimation de la moyenne annuelle, ordonnée : moyenne des estimations saisonnières.**



N02/Mulhouse tubes mai 03

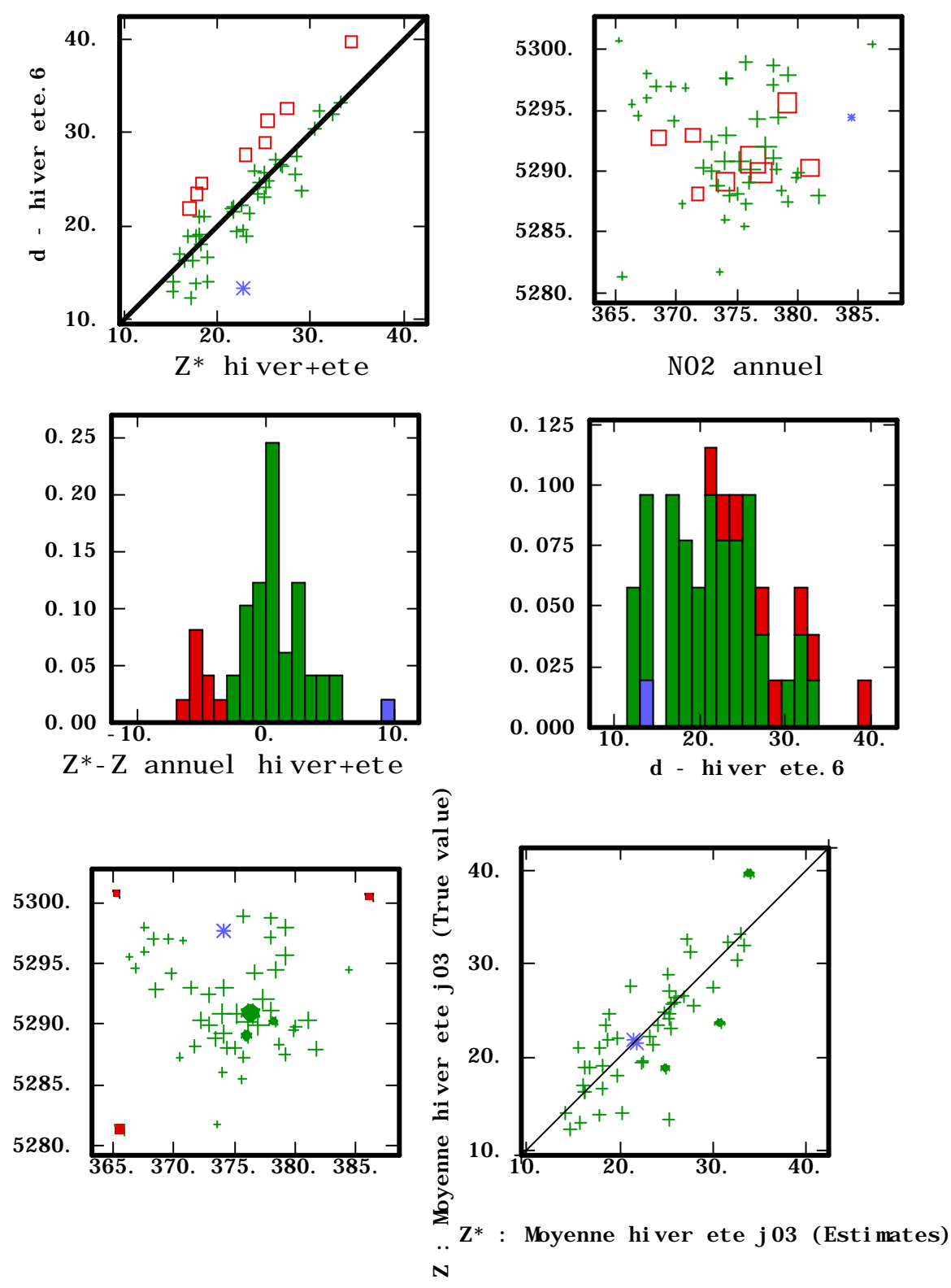
- Variable #1 : d - hiver ete.6

Nb. samples :	52
Minimum :	12.2312
Maximum :	39.6695
Mean :	22.2641
Std. Dev. :	6.17941
Coef of Var. :	0.277551

Tubes repérés par des \* : les valeurs faibles se situent à l'ouest, les valeurs fortes au centre, et les deux tubes intermédiaires respectivement vers le nord et au sud-est.

**b) Implantation des tubes repérés sur le nuage de corrélation.**

**Figure 14. Validation croisée de l'estimation monovariante de la moyenne annuelle de NO2 : krigeage annuel et moyenne des deux krigeages saisonniers**

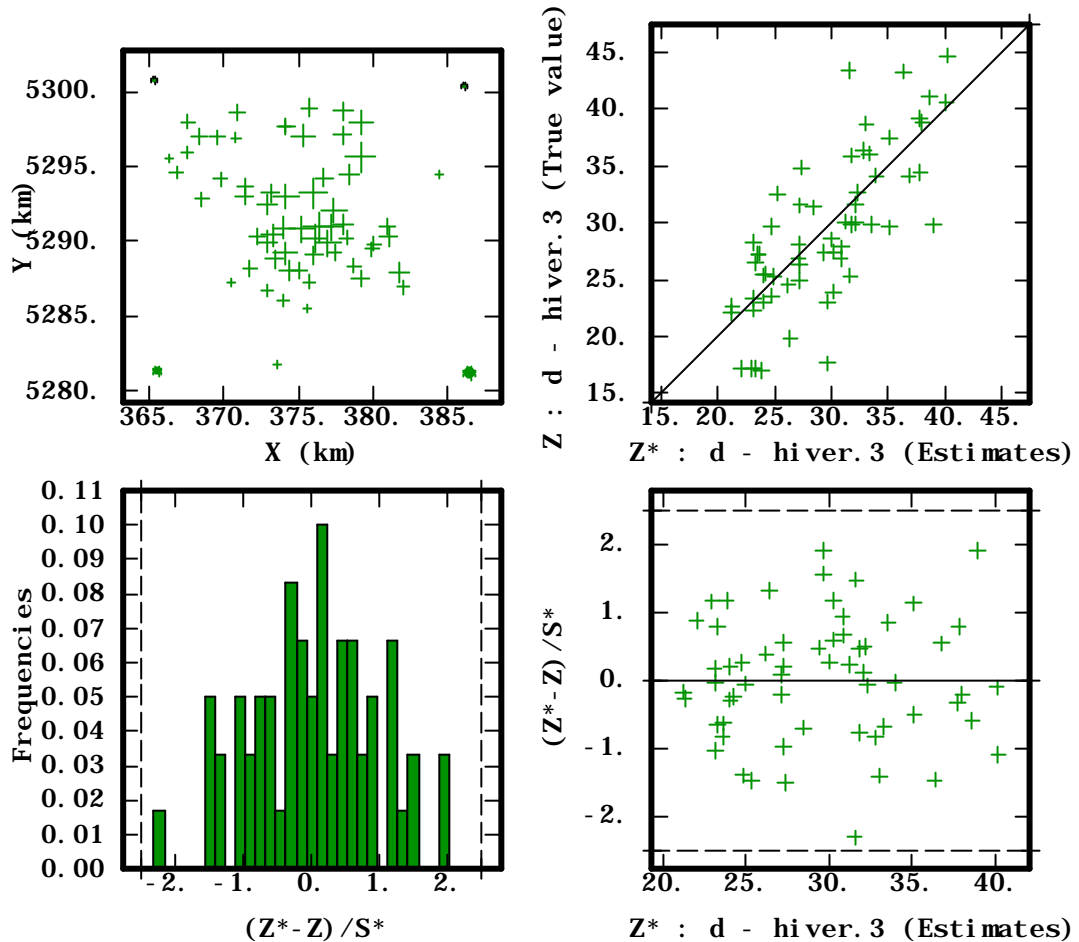


Les tubes en rouge ne sont pas ré-estimés, par manque de données en nombre suffisant dans le voisinage. Le couple de tubes distants de 3m au Nord est indiqué en bleu.

**Figure 15. NO2 "annuel". Validation croisée pour les modèles de référence : population, bâti dense et NOx en dérive, variogramme linéaire.**

## 1.6. Annexes. NO2 sur l'agglomération de Mulhouse.

### 1.6.1. Validation croisée monovariante : krigeage de NO2



N02/Mulhouse tubes mai03

- Variable #1 : d - hiver.3

Standard Parameter File for Model: N02 hiver

Standard Parameter File for Neighborhood: Isotrope rayon 10000

Cross validation statistics based on 60 test data

	Mean	Variance
Error	0.26733	19.62856
Std. Error	0.02502	0.81496

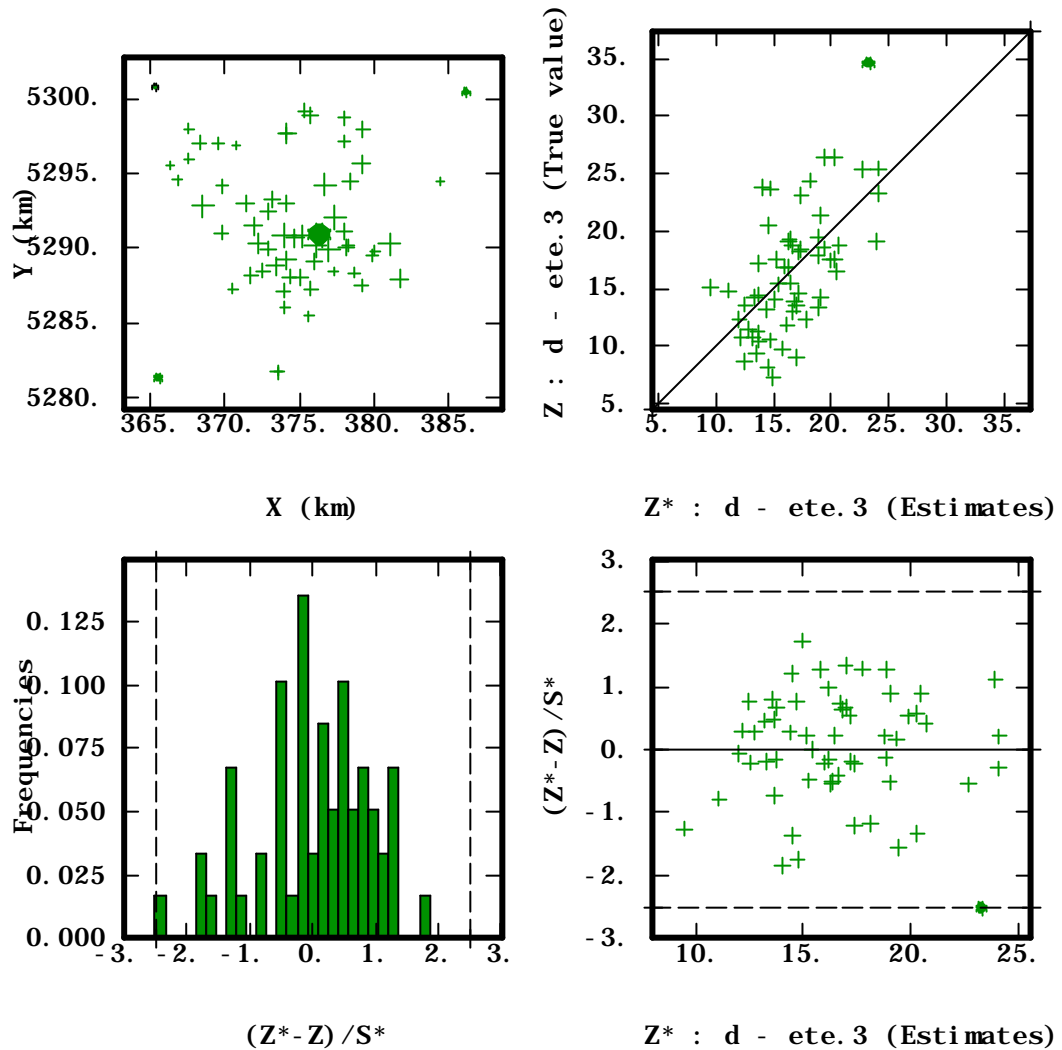
Cross validation statistics based on 60 robust data

	Mean	Variance
Error	0.26733	19.62856
Std. Error	0.02502	0.81496

A data is robust when its Standardized Error lies between -2.500000 and 2.500000

Isatis

**Figure 16. Validation croisée de NO2 hivernal. Modèle monovariante, isotrope.**



Isatis

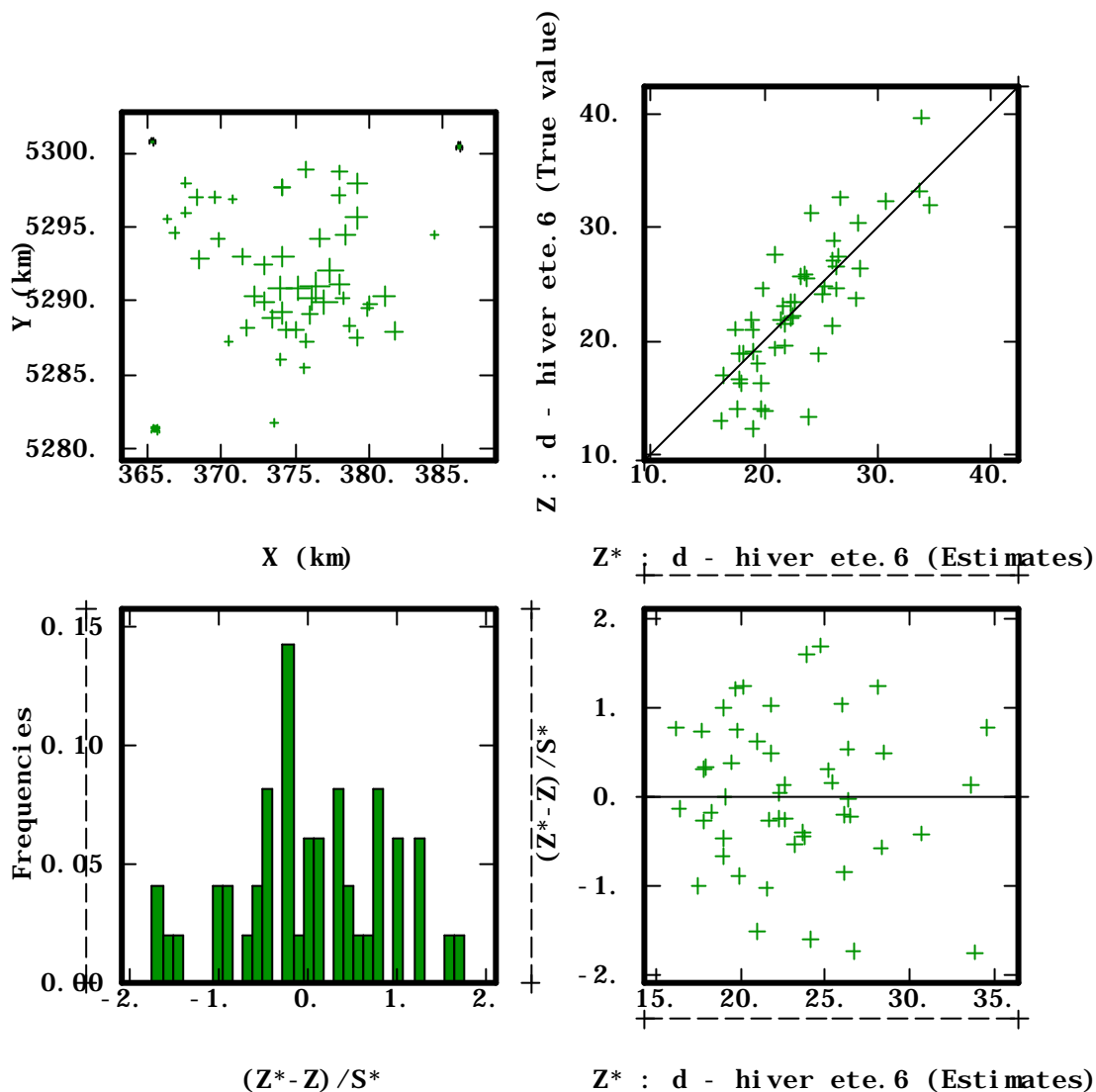
```

N02/Mulhouse tubes mai03
- Variable #1 : d - ete.3
Standard Parameter File for Model: N02 ete
Standard Parameter File for Neighborhood: Isotrope rayon 10000
Cross validation statistics based on 59 test data
      Mean      Variance
Error      0.14624    18.31832
Std. Error  0.02392     0.79384

Cross validation statistics based on 58 robust data
      Mean      Variance
Error      0.34429    16.32008
Std. Error  0.06752     0.69538

A data is robust when its Standardized Error
lies between -2.500000 and 2.500000
  
```

**Figure 17. Validation croisée de NO2 estival. Modèle monovarié, isotrope.**



Isatis

N02/Milhouse tubes mai03  
 - Variable #1 : d - hiver etc.6  
 Standard Parameter File for Model: N02 an  
 Standard Parameter File for Neighborhood: Isotrope rayon 10000  
 Cross validation statistics based on 49 test data

	Mean	Variance
Error	0.26550	12.60675
Std. Error	0.02775	0.69121

Cross validation statistics based on 49 robust data

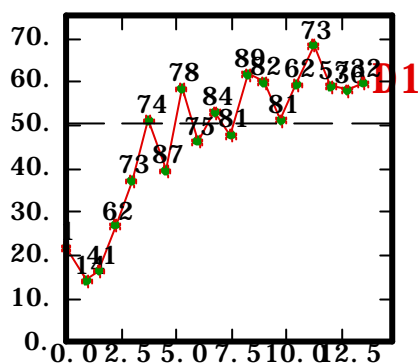
	Mean	Variance
Error	0.26550	12.60675
Std. Error	0.02775	0.69121

A data is robust when its Standardized Error lies between -2.50000 and 2.50000

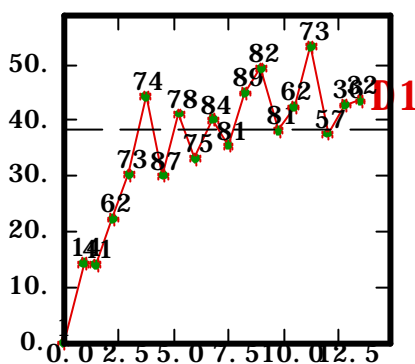
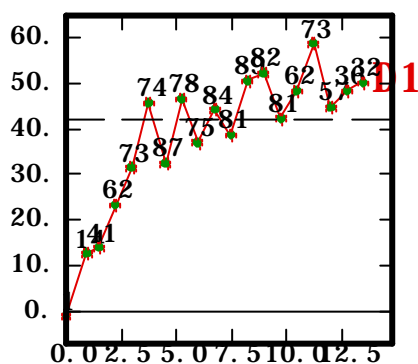
Figure 18. Validation croisée de N02 annuel. Modèle monovariante, isotrope.

## 1.6.2. Variogrammes des moyennes saisonnières et annuelles

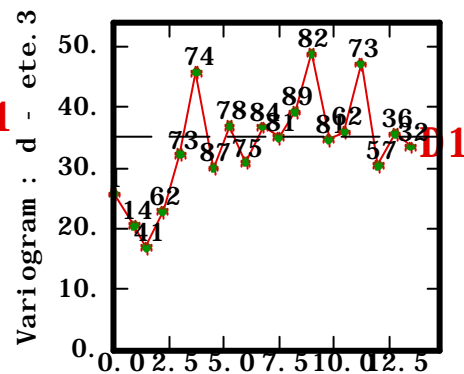
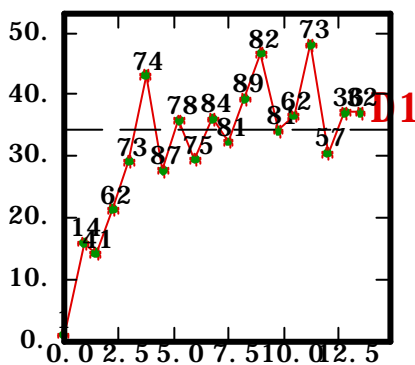
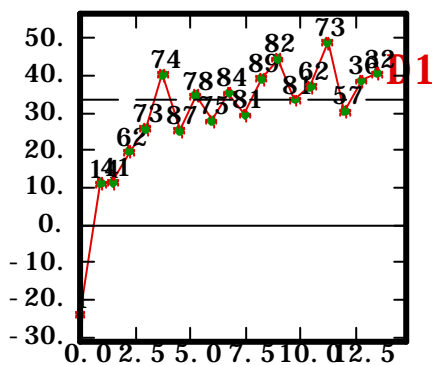
HIVER



ANNUEL



ETE



Distance (km)

Distance (km)

Distance (km)

Isatis

N02/Milhouse tubes mai03(52 stations annuelles)

- Variable #1 : d - etc. 3
- Variable #2 : d - hiver etc. 6
- Variable #3 : d - hiver. 3

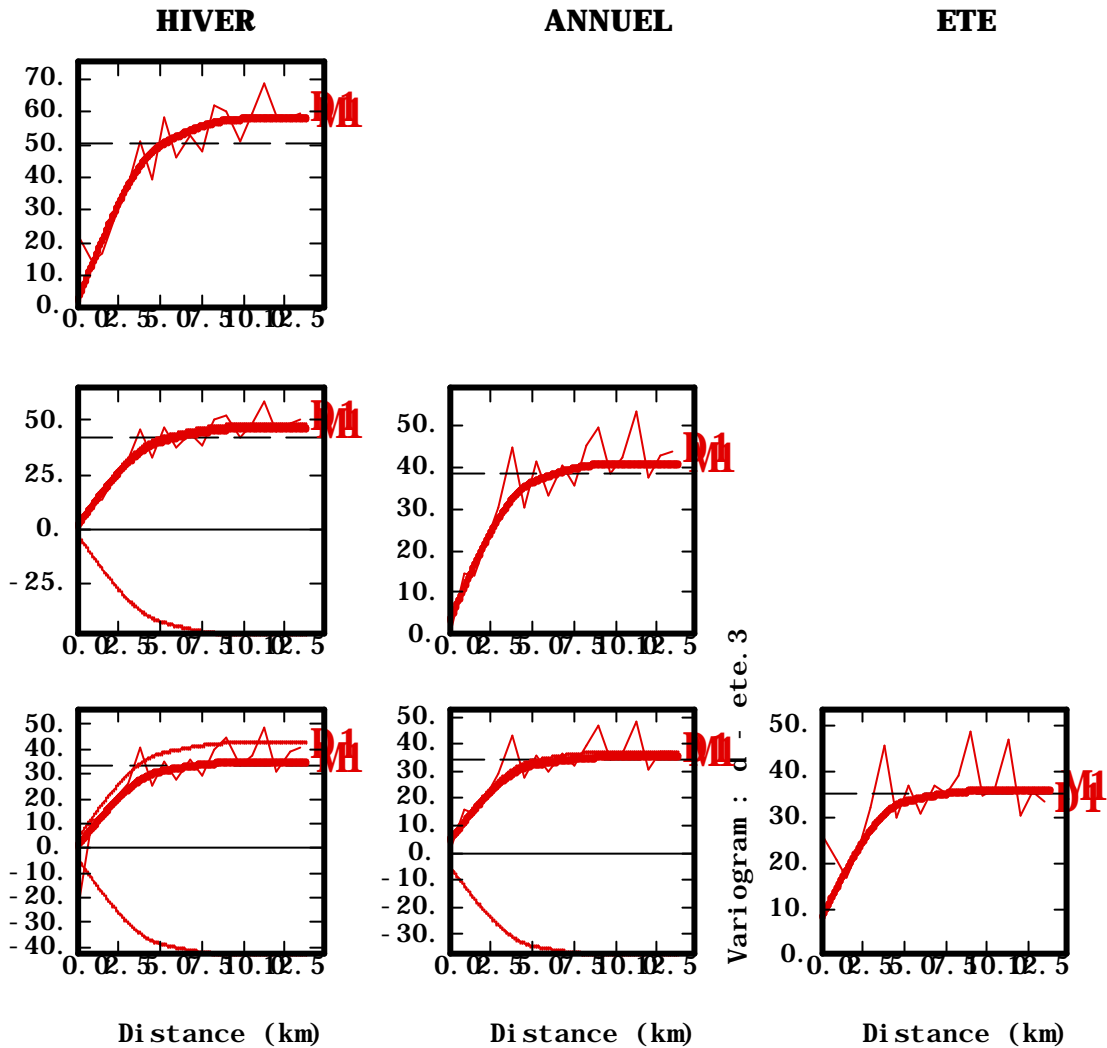
Experimental Variogram(s) : 1 direction(s)

D1 - Az= 0.00, Ay= 0.00, Ax= 0.00

Angular tolerance = 90.00

Lag = 0.75km, Count = 19 lags, Tolerance = 50.00%

**Figure 19. variogrammes omnidirectionnels simples et croisés des moyennes saisonnières et annuelles (52 tubes annuels, pas : 750m)**



N02/Mulhouse tubes mai03(52 stations annuelles)

- Variable #1 : d - etc. 3
- Variable #2 : d - hiver. 6
- Variable #3 : d - hiver. 3

Experimental Variogram(s) : 1 direction(s)

D1 - Az= 0.00, Ay= 0.00, Ax= 0.00

Angular tolerance = 90.00

Lag = 0.75km, Count = 19 lags, Tolerance = 50.00%

S1 - Nugget effect

8	4.5	1
4.5	3	1.5
1	1.5	2

S2 - Spherical - Range = 5.00km

20	21	22
21	23.5	26
22	26	30

S3 - Spherical - Range = 10.00km

8	10	12
10	14.5	19
12	19	26

**Figure 20. Ajustement des variogrammes simples et croisés des moyennes en NO2, compatible avec la relation de fermeture**

Soient  $g_{\text{HIVER}}$ ,  $g_{\text{ETE}}$ ,  $g_{\text{HIVER ETE}}$  les modèles variographiques simples et croisés des moyennes saisonnières en NO<sub>2</sub>. Ce modèle détermine le variogramme simple de la moyenne annuelle, ainsi que les variogrammes croisés entre la moyenne annuelle et les moyennes saisonnières. En effet, la relation de “fermeture”  $NO_2^{\text{AN}} = \frac{1}{2} NO_2^{\text{HIVER}} + NO_2^{\text{ETE}}$  implique les relations suivantes entre les variogrammes

- moyenne annuelle :  $g_{\text{AN}} = \frac{1}{4} g_{\text{HIVER}} + g_{\text{ETE}} + \frac{1}{2} g_{\text{HIVER,ETE}}$

- croisé annuel, hiver :  $g_{\text{AN HIVER}} = \frac{1}{2} g_{\text{HIVER}} + g_{\text{HIVER ETE}}$

- croisé annuel, été :  $g_{\text{AN ETE}} = \frac{1}{2} g_{\text{ETE}} + g_{\text{HIVER ETE}}$

L’ajustement de la figure 20., effectué en modèle linéaire de corégionalisation, est compatible avec ces conditions.

Contrairement aux ajustements effectués indépendamment pour les moyennes saisonnières et la moyenne annuelle, nous avons retenu

- un effet de pépite pour l’ensemble des variogrammes simples et croisés ;
- un schéma sphérique de portée 5km comme structure intermédiaire commune ;
- un schéma sphérique de portée 10km comme structure commune à “grande” portée

## **2. Conclusions sur la cartographie du NO<sub>2</sub>**

L'étude de la concentration moyenne saisonnière ou annuelle en NO<sub>2</sub> menée dans le cadre de la présente convention a donné lieu aux rapports d'avancement suivants :

[1] Pierre Chauvet, octobre 2002. Examen préliminaire des données de Mulhouse. Premier rapport d'avancement, N-15/02/G. ENSMP, Fontainebleau.

[2] Serge Séguret, février 2003. Estimation du dioxyde d'azote routier dans la vallée de la Thur. Deuxième rapport d'avancement, N1/03/G. ENSMP, Fontainebleau.

[3] Chantal de Fouquet, avril 2003. Méthodologie de cartographie de la concentration annuelle de NO<sub>2</sub> sur l'agglomération de Mulhouse. Troisième rapport d'avancement N6/03/G. ENSMP, Fontainebleau.

[4] Gaëlle Le Loc'h, juin 2003. Etude exploratoire du dioxyde d'azote sur l'agglomération de Montpellier. Quatrième rapport d'avancement, ENSMP, Fontainebleau.

### **2.1. Maxima saisonniers**

Le cas de la vallée de la Thur se différencie de ceux des agglomérations de Mulhouse et de Montpellier à la fois par les caractéristiques du milieu (une route suivant un fond de vallée d'une part, deux agglomérations d'autre part), et par l'implantation des mesures (transects le long de la route, espacés d'environ 500m pour la vallée de la Thur, couverture de l'agglomération avec des resserrement locaux pour Mulhouse et Montpellier). Les mesures saisonnières recouvrent 6 quinzaines consécutives pour la vallée de la Thur, trois à Mulhouse et quatre à Montpellier. Avec ces supports temporels différents, les mesures des concentrations saisonnières atteignent approximativement 60 ng/m<sup>3</sup> en été et 50 en hiver dans la vallée de la Thur, alors que pour les deux agglomérations, le maximum est atteint en hiver, avec 45 ng/m<sup>3</sup> (maximum estival : 35 ng/m<sup>3</sup>) pour Mulhouse, et 37 ng/m<sup>3</sup> (maximum estival : 23 ng/m<sup>3</sup>) pour Montpellier.

Dans les deux agglomérations, sans doutes à cause des installations de chauffage ainsi que de l'activité industrielle, les maximum et les moyennes sont supérieurs en hiver. A proximité immédiate de la route, les concentrations sont supérieures en été, mais dès que la distance à la route dépasse quelques dizaines de mètres, les moyennes hivernales deviennent supérieures aux moyennes estivales, avec un écart d'une dizaine de ng/m<sup>3</sup> au-delà d'une cinquantaine de mètres à la route (cf. rapport [2], figure 56 p 47).

Les résultats sont récapitulés séparément pour les vallées de la Thur et pour les deux agglomérations.

### **2.2. NO<sub>2</sub> routier de la vallée de la Thur**

La concentration en NO<sub>2</sub> diminue vers le fond de la vallée et croît fortement à proximité immédiate de la route. La modélisation retenue est du type dérive + résidu, la moyenne (dérive) et la variance des résidus changeant selon la saison. La dérive conserve la même forme pour l'été et l'hiver, le paramétrage étant modifié, avec un gradient latéral plus

marqué en été. Quelle que soit la distance à la route, l'écart-type des résidus est supérieur en été (cf. [2], figure 56 p 47). A même densité de mesures, cela implique une moins bonne précision pour l'estimation des concentrations estivales. Aux distances pour lesquelles la moyenne estivale est inférieure à la moyenne hivernale, la précision relative également est plus médiocre en été qu'en hiver.

Des points considérés comme singuliers ont été exclus de la modélisation. Pour un calcul d'exposition réaliste, il conviendrait d'en tenir compte pour améliorer la modélisation longitudinalement, suivant la route.

La modélisation de la dérive ne tient pas compte des variables auxiliaires comme la densité de population ou l'occupation des sols. En effet, perpendiculairement à la route, le fort gradient de concentrations s'explique par la distance à la route, donc à l'aide d'une des coordonnées du repère local ( $x$ , distance à la route). Sur les transects, les tubes sont localement très resserrés relativement à la grille informée pour les variables auxiliaires. Plusieurs tubes d'un même transect reçoivent donc la même valeur (par "migration"), ou des valeurs voisines (par krigeage) pour les variables auxiliaires, qui ne peuvent donc pas expliquer la variabilité de la concentration.

Le long de la route, l'abscisse curviligne traduit l'avancement dans la vallée, ce qui explique une partie de la variabilité des concentrations (la diminution moyenne de NO<sub>2</sub> suivant la coordonnée  $y$  croissante), ainsi que celle des variables explicatives. Cependant, des tubes isolés avec de faibles concentrations ont été écartés de la modélisation. Il serait intéressant de vérifier si certaines variables explicatives, ou des caractéristiques locales de la route, apparaissent corrélées aux résidus des concentrations ([NO<sub>2</sub>]-dérive) en ces points supposés "anomaliques".

Une présentation synthétique des résultats, ainsi que des perspectives de développement, sont données p 48 et suivantes du rapport [2].

Par ailleurs, disposant de six quinzaines par saison, nous proposons pour un éventuel complément d'étude d'examiner le caractère représentatif ou non, de trois quinzaines parmi les six.

### **2.3. NO<sub>2</sub> urbain à Mulhouse et Montpellier**

Les mesures par analyseurs n'ont pas été examinées dans cette étude. Elles seront indispensables pour préciser la variabilité temporelle. Actuellement, on ne dispose pas encore des données pour caler un variogramme spatio-temporel, en vue de proposer des schéma d'échantillonnage éventuellement mieux adaptés.

Les variables auxiliaires fournies ne coïncident pas pour les deux agglomérations, la type de bâti manquant pour Montpellier. Par ailleurs, le nombre de tubes disponibles simultanément en été et en hiver reste réduit, en particulier pour Montpellier. Il sera donc nécessaire de valider les conclusions de l'étude.

Les moyennes hivernales sont supérieures aux estivales. Il apparaît des différences marquées entre quinzaines d'une même saison, certaines quinzaines estivales ou hivernales étant en

revanche assez bien corrélées à celles de la saison opposées (voir les coefficients de corrélation pour Montpellier, [4], p13 tableaux 7 et 8 ; pour Mulhouse, [3], tableau 2, p13 et figure 3, p10. Les coefficients de corrélation ne tiennent pas compte des écarts sur les moyennes, différentes entre les saisons). Les nuages entre mesures par quinzaines indiquent une corrélation linéaire, mais font aussi apparaître différentes “populations”, c’est-à-dire des groupes de stations présentant un comportement différencié.

Sur les ACP entre variables auxiliaires et moyennes saisonnières, ces dernières apparaissent regroupées (Montpellier, [4], figure 31 p 36 ; Mulhouse [1], figure 22), alors que les moyennes par quinzaines ne se répartissent pas clairement suivant les saisons (Mulhouse, [3], figure 4 p11 ; Montpellier, [4], figure 25 p 32). Ceci indique des variations à court terme de la concentration, dues par exemple aux conditions météorologiques. Un examen détaillé des différentes populations de tubes identifiées sur les nuages de corrélation du NO<sub>2</sub> par quinzaine permettrait de mieux expliquer ces comportements.

Les changements importants du NO<sub>2</sub> entre quinzaines successives ont pour conséquences qu’un modèle saisonnier ou annuel de la concentration serait inadapté à l’estimation par quinzaine ; au minimum, il serait nécessaire d’adapter le paramétrage.

La représentativité des moyennes sur trois ou quatre quinzaines comme estimateur des moyennes saisonnières n’a pas été examinée.

La dispersion des mesures par quinzaines, et les liens entre quinzaines pour des saisons opposées, fournissent un coefficient de corrélation assez marqué pour Mulhouse entre moyennes saisonnières, de l’ordre de 0.80. Si les ACP montrent des corrélations plutôt analogues entre moyennes saisonnières et variables auxiliaires, ces corrélations avec les variables auxiliaires restent cependant assez faibles (Mulhouse, [1], figure 22; Montpellier [4], figure 31, p36). Mieux que les ACP, les nuages de corrélation montrent que NO<sub>2</sub> est moins lié aux variables explicatives disponibles en été qu’en hiver. A Montpellier, où deux hivers successifs ont été instrumentés, des écarts de corrélation notables apparaissent entre mesures hivernales et variables auxiliaires, mais l’implantation des tubes diffère largement d’une année à l’autre (voir les coefficients de corrélation avec les facteurs des ACP des variables auxiliaires, [4], tableau 19 p 40) ; ces résultats demandent donc une confirmation.

La qualité des modèles en vue d’une cartographie de la moyenne saisonnière ou annuelle de la concentration a été examinée suivant la technique de “validation croisée classique”, par suppression de chaque donnée successivement, et réestimation de cette valeur par les autres mesures. Comme les tubes sont plus resserrés localement, notamment au centre de l’agglomération pour Mulhouse, les statistiques de validation croisée “privilégient” ces zones : les points y sont proportionnellement plus nombreux, et en présence d’une structure spatiale, la proximité des données tend à fournir des estimations plus satisfaisantes (cet effet est corrigé par le calcul des erreurs standardisées).

Avant de décider de la qualité d’un modèle, il est nécessaire de préciser le critère de comparaison, et en particulier la zone concernée.

Les modèles retenus pour l’estimation saisonnière sont du type dérive externe, ou dérive et résidus pour Montpellier. Pour les deux agglomérations, les résultats de validation croisée indiquent une forte dispersion du nuage valeur mesurée - valeur réestimée pour la moyenne

estivale, l'estimation étant de meilleure qualité pour la moyenne hivernale (avec un calage du nuage valeur estimée - valeur mesurée sur la première bissectrice, malgré une dispersion encore importante de ce nuage). Pour Mulhouse, ainsi que sous réserve de confirmation pour Montpellier (cf. respectivement la première partie du présent rapport et [4], tableau 28 p 51), l'apport des variables auxiliaires est faible pour l'estimation des moyennes estivales. La prise en compte des variables auxiliaires améliore l'estimation des moyennes hivernales.

L'apport respectif des différentes variables explicatives (densité de population, émissions, type de bâti) n'apparaît cependant pas encore très clairement. Plusieurs modèles utilisant deux variables simultanément ont été retenus. Comme on l'a discuté précédemment, le choix d'un modèle "optimal" suppose de préciser le critère (qui n'est pas nécessairement le résultat de validation croisée) ainsi que l'extension de la zone à estimer.

Pour Mulhouse, l'estimation de la moyenne annuelle par la moyenne des estimations des variables saisonnières améliore sensiblement les résultats de validation croisée. Le cokrigeage de la moyenne annuelle par les mesures saisonnières permet en particulier d'utiliser les mesures disponibles en un site durant une saison seulement. Ce cokrigeage permet également de reporter sur la moyenne annuelle l'apport des variables explicatives pour la moyenne hivernale. Enfin, il assure la cohérence des estimations saisonnières et annuelles.

Pour les deux agglomérations, une fois précisé le critère de choix d'un "bon" modèle (valable pour le centre de l'agglomération, en extrapolation en périphérie ?), on pourra s'intéresser à l'influence du support de régularisation des variables auxiliaires, et examiner une éventuelle corrélation différée spatialement entre variables auxiliaires et concentration mesurée (cf. [1], en conclusion). La définition d'un "meilleur" schéma d'échantillonnage dépend également de la zone à cartographier ainsi que du modèle retenu.

Nous renvoyons par ailleurs aux conclusions des rapports [3] et [4] pour les applications spécifiques à Mulhouse ou Montpellier.