



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



### Adaptation des plans d'échantillonnage spatial aux objectifs de campagne

ANDRE WROBLEWSKI et LAURE MALHERBE

Novembre 2005





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.**

# **ECOLE DES MINES DE DOUAI**

## **DEPARTEMENT CHIMIE ET ENVIRONNEMENT**

### **ADAPTATION DES PLANS D'ECHANTILLONNAGE SPATIAL AUX OBJECTIFS DE CAMPAGNE**

ANDRE WROBLEWSKI et ESPERANZA PERDRIX

**Novembre 2005**



## Adaptation des plans d'échantillonnage spatial aux objectifs des campagnes

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

**Thème : Traitements numériques**

**Convention 05000051**

**Financée par la Direction des Préventions des Pollutions et des Risques (DPPR)**

**NOVEMBRE 2005**

**Personnes ayant participé à l'étude**

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

## Résumé de l'étude EMD/INERIS 2005

### ADAPTATION DES PLANS D'ECHANTILLONNAGE SPATIAL AUX OBJECTIFS DE CAMPAGNE

**André Wroblewski (EMD)**

[wroblewski@ensm-douai.fr](mailto:wroblewski@ensm-douai.fr) ; tel : + 33 (0)327712611

**Laure Malherbe (INERIS)**

[Laure.Malherbe@ineris.fr](mailto:Laure.Malherbe@ineris.fr); tel : + 33 (0)3 44 55 62 18

Dans cette étude, intitulée « Adaptation des plans d'échantillonnage spatial aux objectifs de campagne », nous tentons d'apporter une réponse à la question du dimensionnement d'une campagne de mesure en fonction de son objectif.

Dans cette phase, il est proposé de s'intéresser plus particulièrement aux objectifs de surveillance suivants :

- surveillance des zones rurales non couvertes par les stations de mesures (couverture hors zone agglomération) ;
- surveillance des villes de 10 000 à 100 000 habitants.

Des recommandations pour l'implantation d'un nouveau site de mesure seront fournies dans le guide rédigé en 2006. L'expérience des AASQA indique qu'en général, ce cas d'étude peut se rattacher aux objectifs précédents, les campagnes d'échantillonnage par tubes étant ponctuellement complétées de mesures automatiques par moyens mobiles.

Compte tenu de la répartition spatiale des concentrations, plus ou moins homogène selon les composés chimiques, il convient de mener cette étude polluant par polluant. Dans le programme réalisé en 2005, nous abordons le cas du dioxyde d'azote.

La méthodologie appliquée repose sur l'exploitation de données issues de simulations déterministes ou de campagnes de mesure. La démarche consiste d'une part, à tester différentes tailles de maille pour un maillage régulier et d'autre part (cas des données simulées), à expérimenter différents types de maille régulière. L'échantillonnage est simulé en sélectionnant tout ou partie des données initiales.

Afin d'évaluer l'impact des différents plans d'échantillonnage (distingués selon le type et la taille de la maille régulière) et l'effet d'une dégradation du maillage, un scénario de référence est défini comme base de comparaison. Pour chaque maillage testé, cette comparaison porte sur les paramètres statistiques du jeu de données sélectionné, le variogramme expérimental et la qualité du variogramme modélisé. On confronte également la carte de krigeage à l'estimation de référence.

L'analyse des différents scénarios nous conduit à deux pistes de conclusion.

- L'étude de la structure du variogramme expérimental en fonction du nombre de points fait apparaître, pour chaque expérience, une nette modification de structure correspondant à un seuil de dégradation de l'information.

- La comparaison des cartes de krigeage et d'écart-types de krigeage est, elle aussi, riche d'enseignements. Cependant les résultats diffèrent suivant le type de secteur étudié :
  - pour une zone hors agglomération, l'évolution de la qualité du krigeage subit une brusque dégradation avec la diminution du nombre de points; un optimum semble se dégager pour une structure expérimentale de 100 points. Cela correspond à une densité d'échantillonnage de 2,3 points pour 100 km<sup>2</sup> ou 1 point pour 43,5 km<sup>2</sup>. La supériorité d'une géométrie de maille par rapport à une autre n'a pu être établie.
  - pour une zone urbaine, les résultats de krigeage indiquent qu'avec des données en nombre assez limité mais bien réparties dans l'espace, on peut obtenir une estimation fiable des concentrations de fond. Pour les deux agglomérations étudiées, ce nombre se situe entre 40 et 50, soit 1 point par maille de 1 à 1,5 km de côté.

Ces résultats doivent être confirmés par un complément d'étude sur d'autres jeux de données afin de pouvoir généraliser les préconisations en terme de structure de maillage, de densité d'échantillonnage et de nombre de points. Ce rapport constitue un document préliminaire ; il fournit une présentation détaillée de la méthodologie développée pour l'étude des maillages et de son application à différents jeux de données. A partir de ce travail et des résultats à venir, un guide de bonnes pratiques sera rédigé en 2006.

# SOMMAIRE

<b>Résumé de l'étude EMD/INERIS 2005</b> .....	<b>4</b>
<b>GLOSSAIRE</b> .....	<b>10</b>
<b>CONTEXTE ET OBJET</b> .....	<b>11</b>
<b>METHODOLOGIE</b> .....	<b>11</b>
<b>COUVERTURE HORS ZONE AGGLOMERATION</b> .....	<b>13</b>
<b>I. PRÉSENTATION DU JEU DE DONNÉES INITIAL</b> .....	<b>15</b>
Méthodologie de sélection des données.....	16
<b>II. ETUDE DE LA CAMPAGNE DU 10 ET 25 JUILLET 2002</b> .....	<b>20</b>
II.1 Etude du jeu de données de référence et établissement de la carte de référence. ....	20
• II.1.1 Etude du variogramme expérimental .....	20
• Modèle ref .....	21
II.2. Etude de l'échantillonnage sur une grille de 17x17 carré au pas de 4km ....	23
• II.2.1. Etude du variogramme expérimental .....	23
• II.2.2. Ajustement du variogramme modélisé et test de validation croisée.....	23
• Modèle car17x17 .....	24
• II.2.3. Etablissement de la carte.....	25
II.3 Etude de l'échantillonnage sur une grille de 17x17 losange au pas de 4km	26
• II.3.1. Etude du variogramme expérimental .....	26
• II.3.2. Ajustement du variogramme modélisé et test de validation croisée.....	26
• Modèle los17x17 .....	27
• II.3.3. Etablissement de la carte.....	28
II.4. Comparaison des cartes et discussion .....	29
<b>III. ETUDE DE LA CAMPAGNE DU 3 AU 17 AVRIL 2003</b> .....	<b>31</b>
III.1 Etude du jeu de données de référence et établissement de la carte de référence. ....	31
• III.1.1 Etude du variogramme expérimental .....	31
• Modèle ref .....	32
III.2. Etude de l'échantillonnage sur une grille de 17x17 carre au pas de 4km ...	33

• III.2.1. Etude du variogramme expérimental .....	33
• III.2.2. Ajustement du variogramme modélisé et test de validation croisée....	34
• Modèle car17x17 .....	35
• III.2.3. Etablissement de la carte.....	36
III.3 Etude de l'échantillonnage sur une grille de 17x17 losange au pas de 4km	36
• III.3.1. Etude du variogramme expérimental .....	36
• III.3.2. Ajustement du variogramme modélisé et test de validation croisée....	37
• Modèle los17x17.....	38
• III.3.3. Etablissement de la carte.....	39
III.4. Comparaison des cartes et discussion .....	40
<b>IV DISCUSSION.....</b>	<b>42</b>
<b>ETUDE D'UNE VILLE DE MOINS DE 100 000 HABITANTS : CAS DE CHERBOURG-OCTEVILLE.....</b>	<b>44</b>
<b>I. INTRODUCTION .....</b>	<b>44</b>
I.1. Présentation du jeu de données initial .....	44
I.2. Méthodologie de sélection des données.....	45
<b>II. ETUDE DU JEU DE DONNÉES INITIAL ET ÉTABLISSEMENT DE LA CARTE DE RÉFÉRENCE.....</b>	<b>48</b>
II.1. Etude du variogramme expérimental .....	48
II.2. Ajustement du variogramme modélisé et test de validation croisée.....	48
• Modèle 1 .....	49
• Modèle 2.....	49
II.3.Choix du meilleur modèle et établissement de la carte de référence.....	51
• Toutes les données.....	51
• Toutes les données.....	51
<b>III. ETUDE DE LA SÉLECTION DE DONNÉES N°1 (MAILLES DE 1 KM).....</b>	<b>54</b>
III.1. Etude du variogramme expérimental .....	54
III.2. Ajustement du variogramme modélisé et test de validation croisée.....	54
• Modèle 1 .....	55
• Modèle 2.....	55
• Sélection 1 (1 km).....	55
III.3. Test de validation absolue menée sur les données complémentaires de la sélection 1 .....	56
• Données complémentaires de la sélection 1 .....	57
• Données complémentaires de la sélection 1 .....	57



III.4.Choix du meilleur modèle et établissement de la carte de la sélection 1 .....	58
• Sélection 1 .....	59
• Sélection 1 .....	59
<b>IV. ETABLISSEMENT DE LA CARTE À PARTIR DE LA SÉLECTION N°2.....</b>	<b>62</b>
IV.1. Etude du variogramme expérimental.....	62
IV.2. Ajustement du variogramme modélisé et test de validation croisée .....	62
IV.3. Test de validation absolue mené sur les données complémentaires de la sélection 2 .....	63
IV.4. Etablissement de la carte pour la sélection 2 .....	64
• Sélection 2 .....	65
<b>V COMPARAISON DES CARTES ET DISCUSSION .....</b>	<b>67</b>
<b>CAS DE LA ROCHELLE.....</b>	<b>72</b>
<b>B.I. DESCRIPTION DU CAS.....</b>	<b>72</b>
B.I.1. Présentation du jeu de données initial.....	72
B.I.2. Méthodologie de sélection des données .....	72
<b>B.II. ETUDE DU JEU DE DONNÉES INITIAL ET ÉTABLISSEMENT DE LA CARTE DE RÉFÉRENCE .....</b>	<b>74</b>
B.II.1. Etude du variogramme expérimental.....	75
B.II.2. Modélisation du variogramme et validation croisée .....	75
B.II.3. Etablissement de la carte de référence .....	77
<b>B.III. ETUDE DE LA SÉLECTION DE DONNÉES N°1 (MAILLE DE 1 KM).....</b>	<b>78</b>
B.III.1. Etude du variogramme expérimental.....	79
B.III.2. Modélisation du variogramme et test de validation croisée .....	79
B.III.3. Test de validation absolue sur le complémentaire de la sélection N°1 ....	80
B.III.4. Etablissement de la carte pour la sélection 1 .....	82
<b>B.IV. ETUDE DE LA SÉLECTION DE DONNÉES N°2 (MAILLE DE 1.5 KM).....</b>	<b>82</b>
B.IV.1. Etude du variogramme expérimental .....	82
B.IV.2. Modélisation du variogramme et test de validation croisée.....	83
B.IV.3. Test de validation absolue sur le complémentaire de la sélection 2.....	84
B.IV.4. Etablissement de la carte pour la sélection 2 .....	85
<b>B.V. ETUDE DE LA SÉLECTION DE DONNÉES N°3 (MAILLE DE 2 KM).....</b>	<b>86</b>
B.V.1. Etude du variogramme expérimental .....	86
B.V.2. Modélisation du variogramme et test de validation croisée.....	87

B.V.3. Test de validation absolue sur le complémentaire de la sélection 3.....	89
B.V.4. Etablissement de la carte pour la sélection 3.....	90
B.V.5. Test de validation absolue sur l'intersection des complémentaires.....	90
<b>B.VI. COMPARAISON DES ESTIMATIONS ET DISCUSSION.....</b>	<b>92</b>
B.IV.1. Influence du nombre de données sur le variogramme expérimental.....	92
B.VI.2. Influence sur l'estimation du nombre de données de krigeage pour un modèle établi de variogramme .....	92
• B.VI.2.1. Estimation sur un jeu de données indépendant .....	92
• B.VI.2.2. Estimation sur la grille .....	93
B.VI.3. Effet du nombre de données expérimentales sur la qualité des cartes ...	94
<b>CONCLUSIONS ET PERSPECTIVES .....</b>	<b>96</b>
<b>REFERENCES .....</b>	<b>99</b>
<b>LISTE DES ANNEXES.....</b>	<b>100</b>

## GLOSSAIRE

Dans le texte, sont désignés par *validation croisée* et *validation absolue* les tests suivants :

### ***Validation croisée :***

Elle a pour objet de contrôler l'adéquation entre le variogramme modélisé et les données expérimentales qui ont servi à construire ce modèle.

Elle consiste à éliminer temporairement un site de mesure et à y estimer la concentration par krigeage, à l'aide des données restantes et du modèle de variogramme. On dispose ainsi d'une concentration estimée  $Z_i^*$  et d'une concentration mesurée  $Z_i$ . Cette opération est répétée pour chacun des points d'échantillonnage, ce qui permet de calculer des statistiques sur l'erreur  $Z^*-Z$ .

### ***Validation absolue :***

Elle permet d'évaluer la précision de l'estimation.

Elle consiste à estimer les concentrations sur un jeu de données indépendant (i.e. non utilisées dans le calcul du variogramme) et à comparer les concentrations calculées et mesurées.

## CONTEXTE ET OBJET

La définition et la mise en place des PSQA (Plans de Surveillance de la Qualité de l'Air), ainsi que le développement d'outils d'évaluation de la qualité de l'air complémentaires à la mesure (analyse objective, géostatistique, modélisation déterministe...) ont engendré un véritable besoin de méthodes de référence pour l'élaboration de plans d'échantillonnage.

Les questions soulevées dans ce contexte concernent aussi bien les critères de dimensionnement de campagnes de mesure que les bonnes pratiques de traitement de données. Il convient de les traiter en fonction de l'objectif des campagnes.

Dans cette étude, prévue sur deux ans, nous abordons la première question, à savoir le dimensionnement d'une campagne de mesure selon son objectif. L'aspect spatial de l'échantillonnage est ici examiné. Le but est de tester des plans d'échantillonnage spatial pour différents objectifs de surveillance, et de déterminer les stratégies les plus adéquates.

Il est proposé de s'intéresser :

- à la surveillance des zones rurales non couvertes par les stations de mesures (couverture hors zone agglomération) et par extension, à la surveillance à l'échelle régionale ;
- à la surveillance des villes de 10 000 à 100 000 habitants ;
- à l'implantation de stations fixes.

Les résultats de ces travaux pourront être considérés comme un complément technique du *Guide des aides de l'ADEME en matière d'équipements pour la surveillance*, qui fournit à ce sujet de premières orientations [1].

Compte tenu de la répartition spatiale des concentrations, plus ou moins homogène selon les composés chimiques, cette analyse doit être menée polluant par polluant (dioxyde d'azote, ozone, benzène). Dans le programme 2005, nous abordons le cas du dioxyde d'azote, polluant réglementé faisant l'objet d'une surveillance régulière par campagnes de mesure. Le cas des autres polluants sera traité en 2006.

D'autre part, deux types de plans d'échantillonnage sont largement utilisés : les plans d'échantillonnage réguliers et les plans d'échantillonnage stratifiés [2]. Le présent document est consacré à l'évaluation de schémas réguliers pour les zones rurales et les villes de moins de 100 000 habitants.

L'élargissement de l'étude aux plans stratifiés et à l'ensemble des objectifs et polluants précédemment mentionnés donnera lieu en 2006 à un guide de recommandations.

## METHODOLOGIE

L'étude se compose de deux parties, la première traite de la surveillance du NO<sub>2</sub> en zone rurale, la seconde de la surveillance du NO<sub>2</sub> dans une ville de moins de 100000 habitants.

Pour la surveillance en zone rurale (1<sup>ère</sup> partie), notre travail s'appuie sur des données de simulation obtenues sur la région Centre en collaboration avec LIG'AIR

(66x66km<sup>2</sup> au nord d'Orléans, voir travaux de l'EMD 2002-2003 [3-4]). Le domaine étudié s'étend sur environ 4400 km<sup>2</sup>.

Pour la surveillance des villes de 10 000 à 100 000 habitants (2<sup>ème</sup> partie), l'étude repose sur des données de campagnes de mesure effectuées dans les agglomérations de Cherbourg et de La Rochelle (domaines d'environ 125 km<sup>2</sup> et 230 km<sup>2</sup>).

La méthodologie adoptée consiste à définir différentes tailles et si possible, différentes géométries de maille d'échantillonnage, et à évaluer l'influence du maillage sur la qualité de l'estimation.

#### Choix des tailles de maille :

Le choix des tailles de maille est dicté par les caractéristiques des jeux de données exploités, notamment le nombre de points de mesure (ou de points simulés) et la surface couverte, et par la technique de traitement des données. Dans cette étude, il s'agit d'établir des cartographies par estimation géostatistique. Comme un variogramme ne peut pas être construit raisonnablement avec moins de 30 points [4], la taille maximale de maille est définie de façon que l'aire d'étude soit couverte par un minimum d'environ 30 points, répartis le plus régulièrement possible.

#### Choix du type de maille régulière :

Deux types de mailles régulières sont considérés dans la 1<sup>ère</sup> partie (données simulées) :

- maille carrée ;
- maille losange.

Du fait de la moindre densité de points, les jeux de données expérimentaux de la 2<sup>ème</sup> partie ne permettent pas une étude aussi fine ; on se limite une maille carrée.

#### Méthode de comparaison :

Afin de comparer les différents plans d'échantillonnage (distingués selon le type et la taille de la maille régulière), il est nécessaire de définir un scénario de référence.

Puisque généralement, le nombre de tubes passifs par campagne<sup>1</sup> de mesure n'excède pas 200 ou 250 points, ce scénario correspond :

- soit à la totalité des points de mesure disponibles, dans le cas de données expérimentales ;
- soit à une sélection régulière de 256 points dans le jeu de données simulées.

A partir du jeu de données de référence, on calcule le variogramme de référence, les paramètres statistiques de référence (minimum, 25<sup>ème</sup> percentile, médiane, 75<sup>ème</sup> percentile, maximum, moyenne, écart-type, variance) et la carte krigée de référence, obtenue par krigeage ordinaire ponctuel [4].

Ensuite, pour chaque maillage testé, on compare :

- les paramètres statistiques du jeu de données sélectionné par le maillage avec les paramètres statistiques de référence.
- graphiquement, le variogramme expérimental avec le variogramme expérimental de référence.

---

<sup>1</sup> On entend ici par « campagne » une période de 14 jours consécutifs, pendant laquelle les échantillonneurs passifs de dioxyde d'azote sont exposés.

La qualité du variogramme modélisé est appréciée

- par une validation croisée (krigeage sur les points sélectionnés par le maillage testé) ;
- par une validation absolue (krigeage sur les points non sélectionnés par le maillage dans le cas d'un jeu de données expérimentales, sur les points décalés d'1/2 maille de simulation par rapport au jeu de référence, dans le cas des données simulées).

Dans ce second test, l'impact du nombre et de la localisation des données d'entrée du krigeage n'est pas différencié de l'effet lié à la dégradation du variogramme. C'est pourquoi la validation absolue est répétée en utilisant le variogramme de référence.

Enfin, on procède aux calculs d'estimation par krigeage sur l'ensemble du domaine (même grille, même voisinage de krigeage dans tous les calculs). Les cartes obtenues sont comparées à la carte estimée de référence.

Tous les calculs sont effectués avec le logiciel Isatis (Géovariances, v 5.0.2)

**1<sup>ERE</sup> PARTIE :**  
**COUVERTURE HORS ZONE AGGLOMERATION**

## I. PRESENTATION DU JEU DE DONNEES INITIAL

Le jeu de données étudié a été élaboré au cours de l'étude LCSQA 2003 « Cartographie de la pollution atmosphérique en zone rurale influencée au moyen d'un modèle déterministe de dispersion des polluants »<sup>2</sup>. Il est constitué des données de concentration en dioxyde d'azote obtenues par échantillonnage passif lors de deux campagnes de mesure, d'une durée de 14 jours chacune, ainsi que des résultats de modélisation déterministe correspondant à la simulation de ces campagnes sur une partie nord de la région Centre.

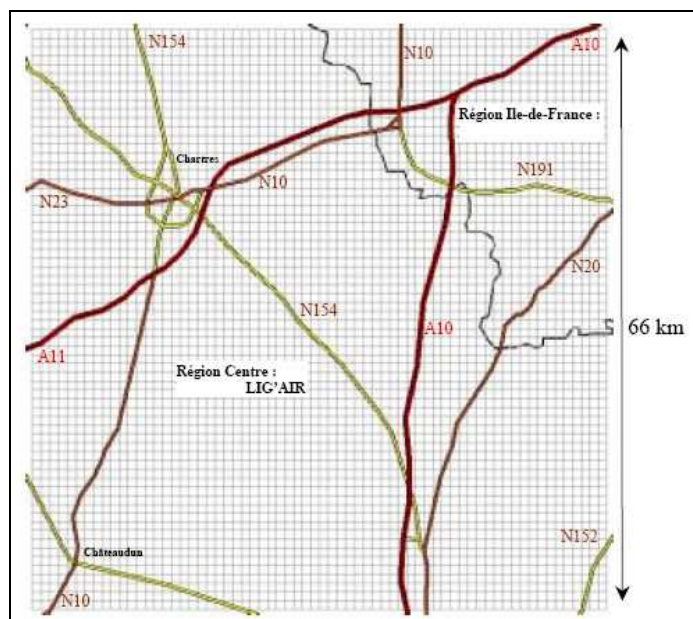


Figure 1 – Zone d'étude, nord de la région centre, avec maillage de modélisation.

Une campagne de mesure des teneurs en  $\text{NO}_2$  par échantillonnage passif sur une période de quinze jours entre les 10 et 25 juillet 2002 a été effectuée par l'EMD en partenariat avec LIG'Air. 63 tubes ont été répartis sur un domaine composé de 36 mailles de 11 km sur 11 km. Les simulations déterministes ont été réalisées avec Transchim (CORIA) sur un domaine de 66x66 km<sup>2</sup> avec une maille de 1 km.

La deuxième campagne de mesure des teneurs en  $\text{NO}_2$  par échantillonnage passif sur une période de quinze jours a eu lieu du 3 au 17 avril 2003. 102 tubes ont été répartis sur un domaine composé de 36 mailles de 11 km sur 11 km avec un raffinement le long des axes routiers. Les simulations déterministes ont été réalisées avec Transchim (CORIA) sur un domaine de 66x66 km<sup>2</sup> avec une maille de 1 km.

Principales caractéristiques du jeu de données étudié :

	Campagne du 10 au 25 juillet 2002	Campagne du 3 au 17 avril 2003
Moyenne (en $\mu\text{g}/\text{m}^3$ )	12,48	18.25
Maximum (en $\mu\text{g}/\text{m}^3$ )	40.56	29.77
Minimum (en $\mu\text{g}/\text{m}^3$ )	5.91	10.76

<sup>2</sup> LCSQA-EMD, 2003. *Cartographie de la pollution atmosphérique en zone rurale influencée au moyen d'un modèle déterministe de dispersion des polluants*. Convention 115/2003.



Les données exploitées dans cette partie sont les résultats des simulations calées préalablement sur les campagnes de mesure. Complétant les données de ces campagnes, ils représentent en effet une plus grande densité d'information spatiale. On suppose qu'ils fournissent une image fidèle de la réalité.

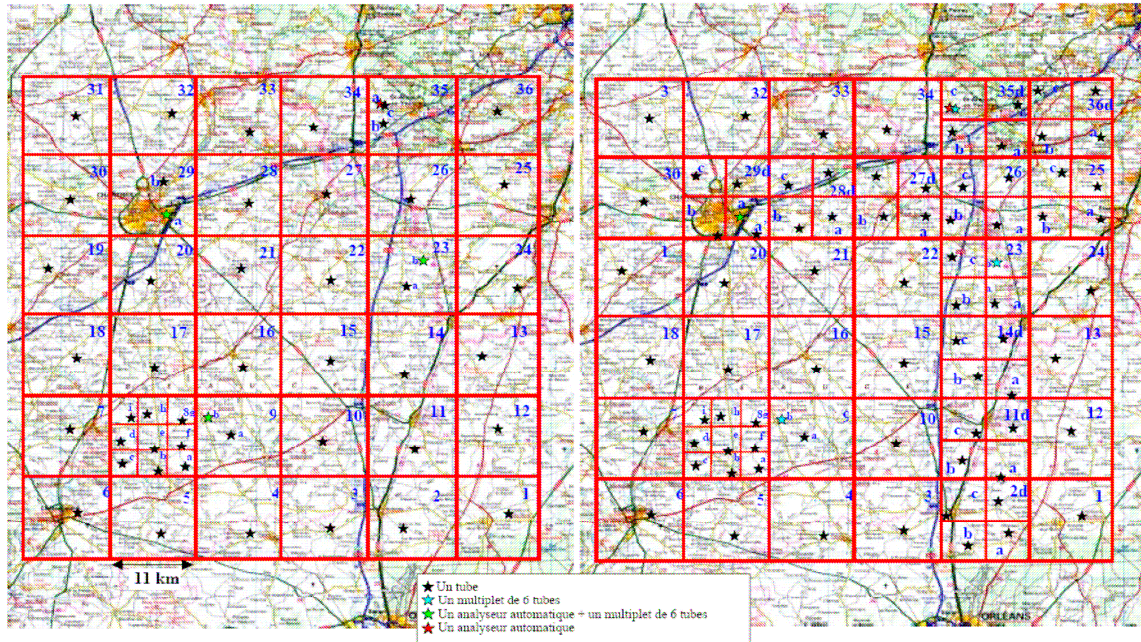


Figure 2 : échantillonnage spatial de la première (à gauche) et la deuxième (à droite) campagne de mesures

### Méthodologie de sélection des données

L'objectif est de travailler sur les données de modélisation déterministe, afin de pouvoir tester de manière systématique des méthodes de dégradation de maillage. Dans cet objectif, deux premiers maillages ont été établis :

- un premier maillage de 33x33 au pas de 2km, appelé maillage de référence,
- un second maillage de 33x33 au pas de 2km, décalé d'une demi-maille par rapport au précédent et appelé maillage de validation absolue (voir figure n°3.). Les points de ce maillage ne sont jamais sélectionnés comme points d'échantillonnage dans l'élaboration des cartographies mais ils constituent un jeu de données indépendant pour la validation.

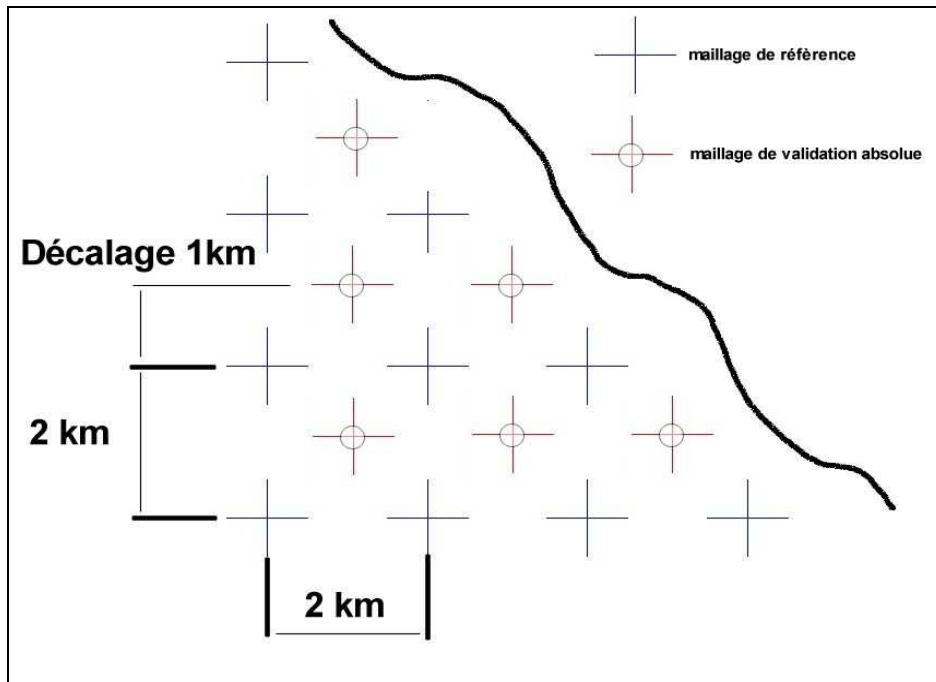


Figure n°3 : Maillage de référence

A partir du maillage de référence nous avons élaboré 8 maillages dégradés. Les quatre premiers maillages ont été dégradés de manière régulière à partir du maillage de référence :

- maillage de référence 33x33 (1089 pts) au pas de 2km,
- maillage de 17x17 carré (289 pts) au pas de 4km,
- maillage de 11x11 carré (121 pts) au pas de 6km,
- maillage de 9x9 carré (81 pts) au pas de 8km,
- maillage de 7x7 carré (49 pts) au pas de 10km,
- maillage de 17x17 losange (281 pts) au pas de 4km,
- maillage de 11x11 losange (121 pts) au pas de 6km,
- maillage de 9x9 losange (77 pts) au pas de 8km,
- maillage de 7x7 losange (49 pts) au pas de 10km.

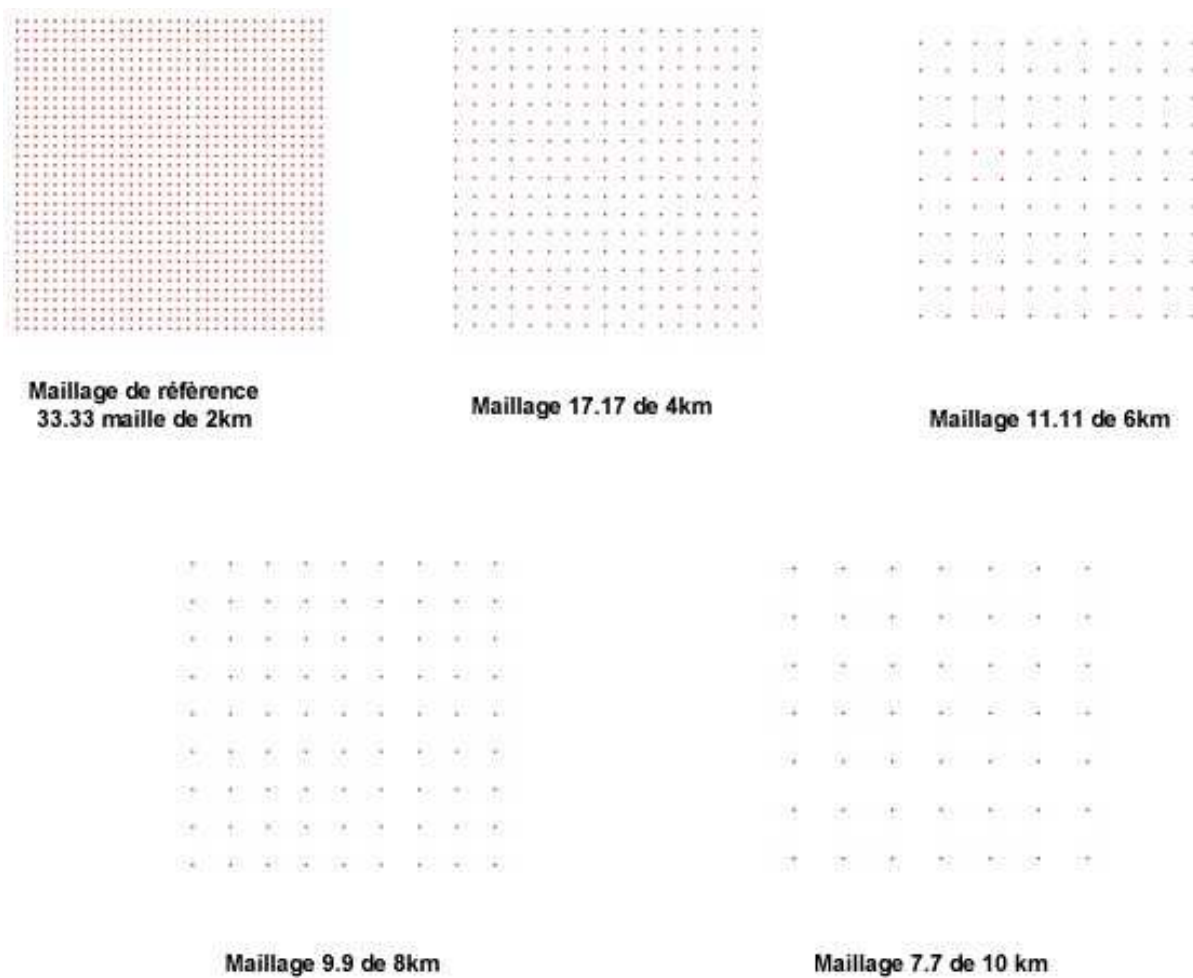


Figure n°4 : différents maillages carrés utilisés

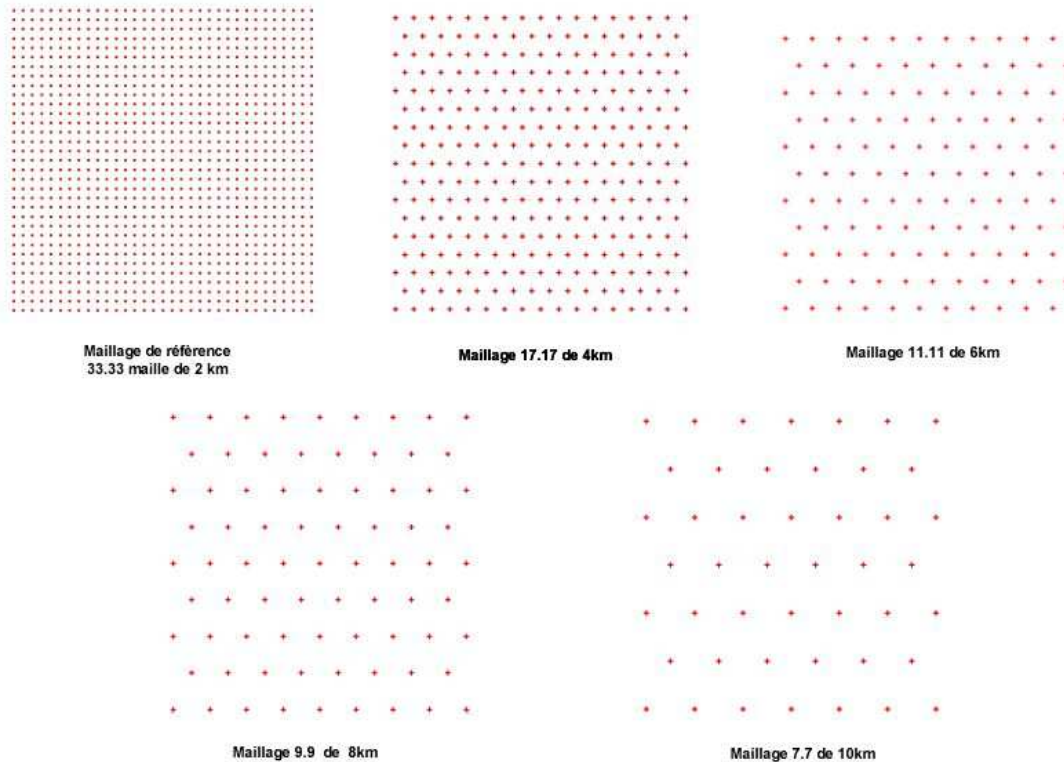


Figure n°5 : différents maillages « losanges » utilisés

Les données de chaque maillage sont exploitées selon une démarche de géostatistique, comme s'il s'agissait d'un jeu de données expérimental. Les étapes classiques d'une cartographie géostatistique sont ainsi mises en œuvre : analyse exploratoire des données et modélisation de la structure spatiale, validation du modèle, estimation par krigeage.

Pour chacun des maillages testés, les variogrammes expérimentaux ont été ainsi déterminés. Nous avons ensuite modélisé un variogramme puis effectué la validation croisée. L'utilisation d'un jeu de données de référence a permis de procéder à une validation absolue. Enfin pour chaque maillage nous avons effectué un krigeage aux points des données de validation absolue en utilisant le variogramme de référence.

Ces cartographies sont comparées entre elles afin de déterminer des critères de sélection de maillages pour la réalisation de campagne d'échantillonnage. Nous présentons dans la suite du rapport, de manière complète, le cas de référence et le cas du maillage 17x17. Les autres résultats sont détaillés en annexe du rapport.

## II. ETUDE DE LA CAMPAGNE DU 10 AU 25 JUILLET 2002

### II.1 Etude du jeu de données de référence et élaboration de la carte de référence.

#### II.1.1 ETUDE DU VARIOGRAMME EXPERIMENTAL

Le variogramme retenu est omnidirectionnel, avec 10 pas de 500 m chacun.

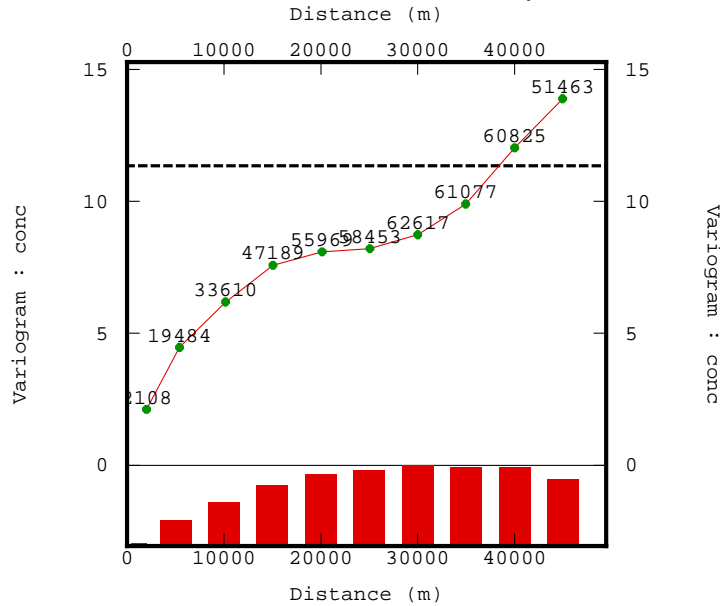
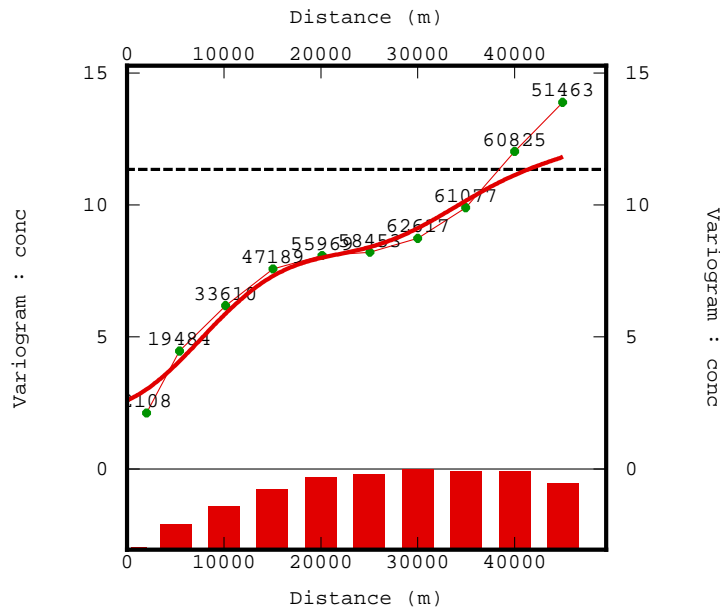


Figure n°6 : Variogramme expérimental (NO<sub>2</sub> en µg/m<sup>3</sup>)

#### II.1.2 AJUSTEMENT DU VARIOGRAMME MODELISE ET TEST DE VALIDATION CROISEE :

Le modèle de variogramme ajusté est composé d'une structure pépitique (palier=2.6) additionnée d'une structure sinus cardinal et d'une structure linéaire.



**Modèle : Pépité(2.6)+ sincard(75000 ;1.9)+lineaire(41500 ;6.7)**

Figure n°7 : Variogramme modélisé (NO<sub>2</sub> en µg/m<sup>3</sup>)



Afin d'apprécier la qualité du modèle, nous l'avons soumis au test de validation croisée. Ce test consiste à calculer, au moyen du modèle, une estimation de la concentration en polluant, pour chaque point de l'espace où celle-ci a été mesurée.

**Tableau 1 – Statistiques sur l'erreur d'estimation obtenues par validation croisée sur toutes les données**

<i>MODELE REF</i>	moyenne	variance
Erreur	-0,0096	3,23
Erreur standardisée	-0,0064	1,02

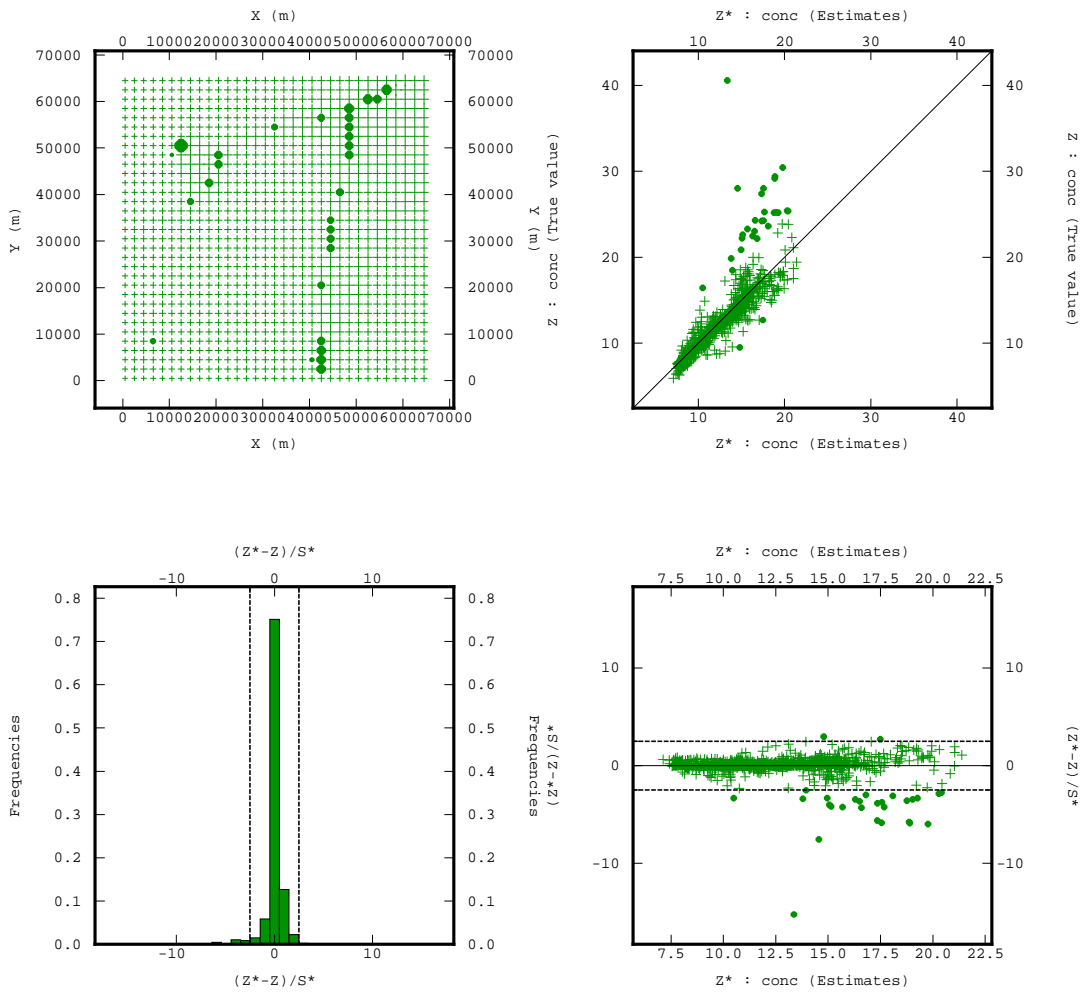


Figure n°8 : Résultats de validation croisée pour le modèle de référence

La plupart des données sont correctement réestimées. La sous-estimation de certaines valeurs s'explique par leur situation de proximité.

### II.1.3. Etablissement de la carte de référence

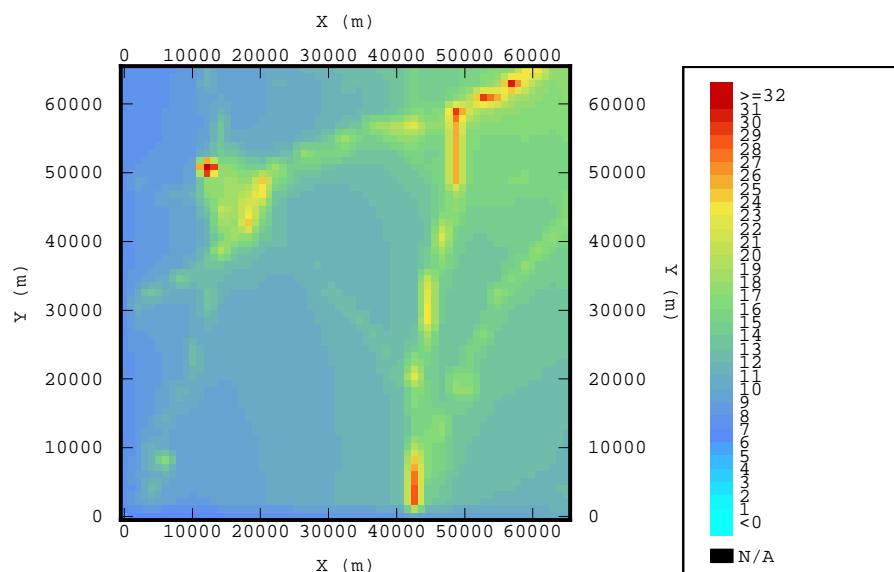


Figure n°9 : Estimation par krigeage ponctuel, à partir de toutes les données, sur une grille de 2km de coté.

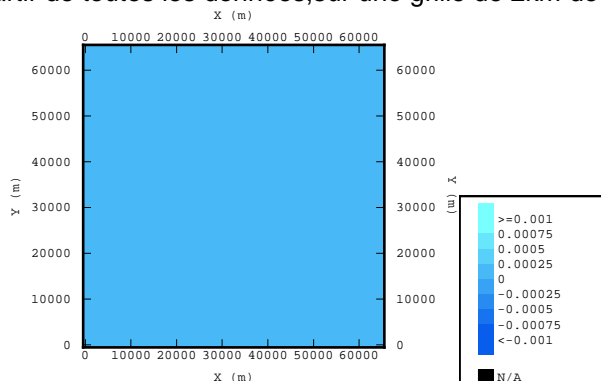


Figure n°10 : Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, sur une grille de 2km de coté.

A partir de ce cas de référence l'impact de la dégradation du maillage d'échantillonnage a été étudié. D'abord sur un maillage carré, ensuite sur un maillage losange.

## II.2. Etude de l'échantillonnage sur une grille de 17x17 carré au pas de 4km

### II.2.1. ETUDE DU VARIOGRAMME EXPERIMENTAL

Le variogramme retenu est omnidirectionnel, avec 10 pas de 500 m chacun.

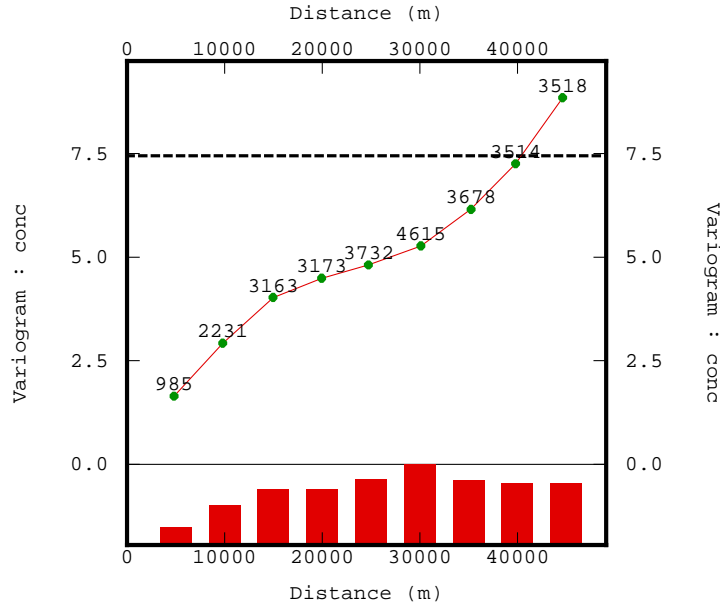
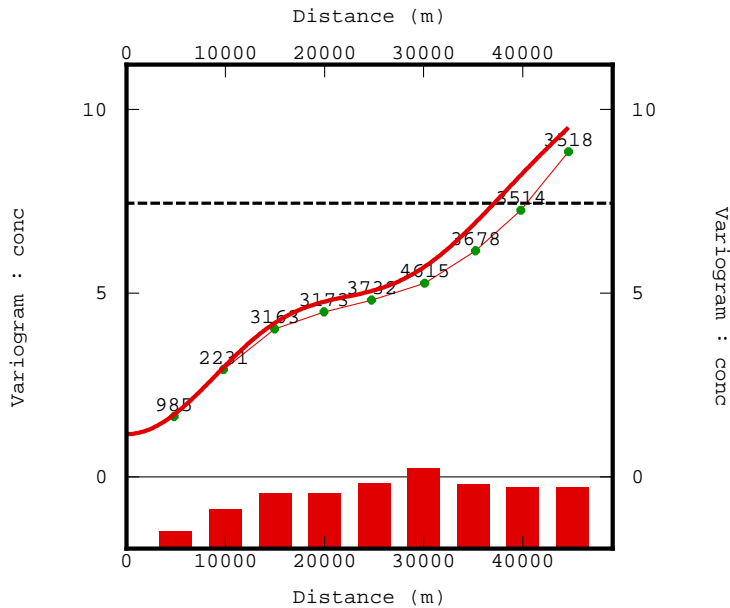


Figure n°11 : Variogramme expérimental (NO<sub>2</sub> en µg/m<sup>3</sup>)

### II.2.2. AJUSTEMENT DU VARIOGRAMME MODELISE ET TEST DE VALIDATION CROISEE

Le modèle de variogramme ajusté est composé d'une structure pépitique (palier=1.5) additionnée d'une structure sinus cardinal et d'une structure puissance.



**Modèle : Pépité(1.5)+ sincard(1600 ;3.5)+puissance(41000 ;1 ;6)**

Figure n°12 : Variogramme modélisé (NO<sub>2</sub> en µg/m<sup>3</sup>)



Le modèle a été soumis au test de validation croisée dont les résultats sont consignés ci-dessous.

**Tableau 2 – Statistiques sur l’erreur d’estimation obtenues par validation croisée sur toutes les données**

MODELE CAR17X17	moyenne	variance
Erreur	0,0047	3.298
Erreur standardisée	-0,0006	0.985

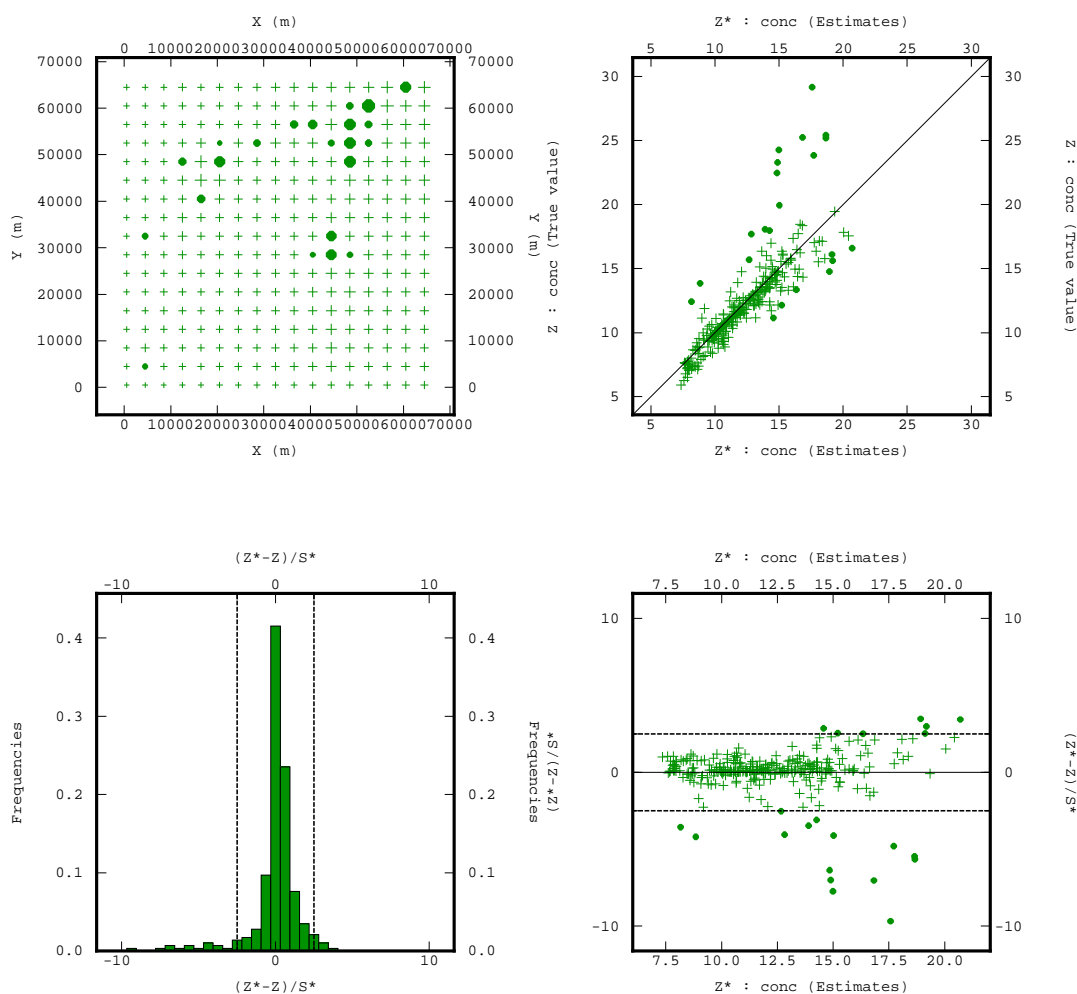


Figure n°13 : Résultats de validation croisée pour le modèle carre de 17x17

La plupart des données restent correctement réestimées, on retrouve toujours la sous-estimation de certaines valeurs dues à la proximité des axes automobiles.

### II.2.3. ÉTABLISSEMENT DE LA CARTE

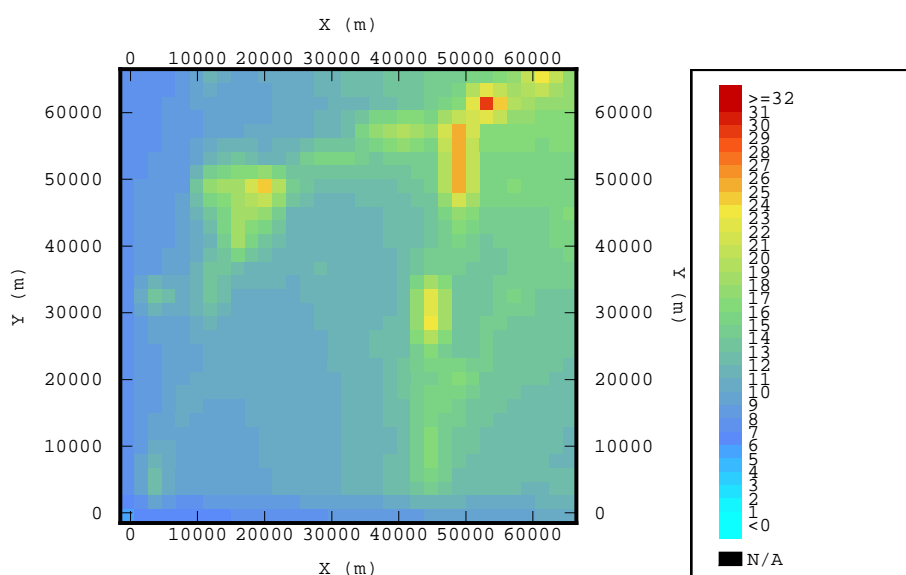


Figure n°14 : Estimation par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de coté.

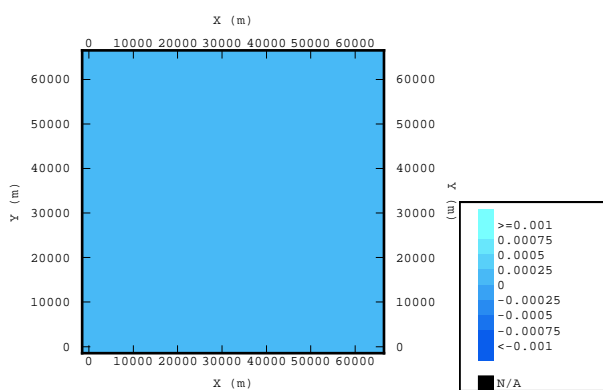


Figure n°15 : Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de coté.

## II.3 Etude de l'échantillonnage sur une grille de 17x17 losange au pas de 4km

### II.3.1. ETUDE DU VARIOGRAMME EXPERIMENTAL

Le variogramme retenu est omnidirectionnel, avec 10 pas de 500 m chacun.

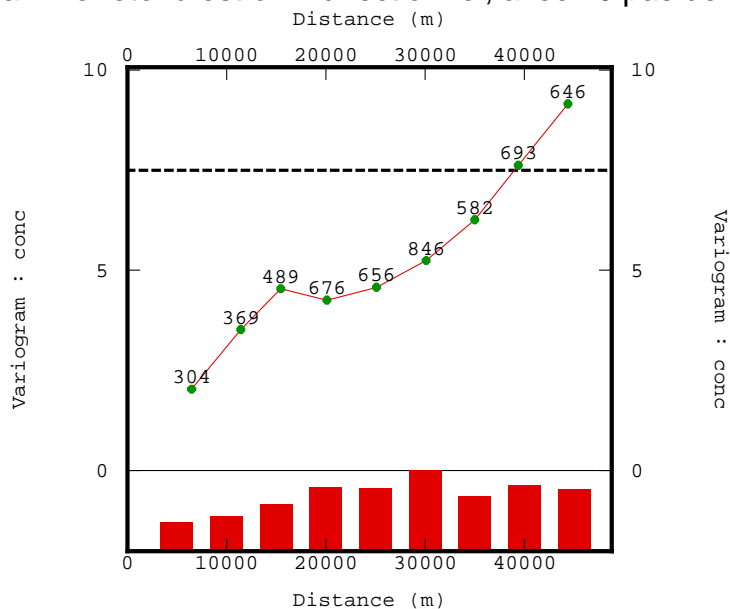
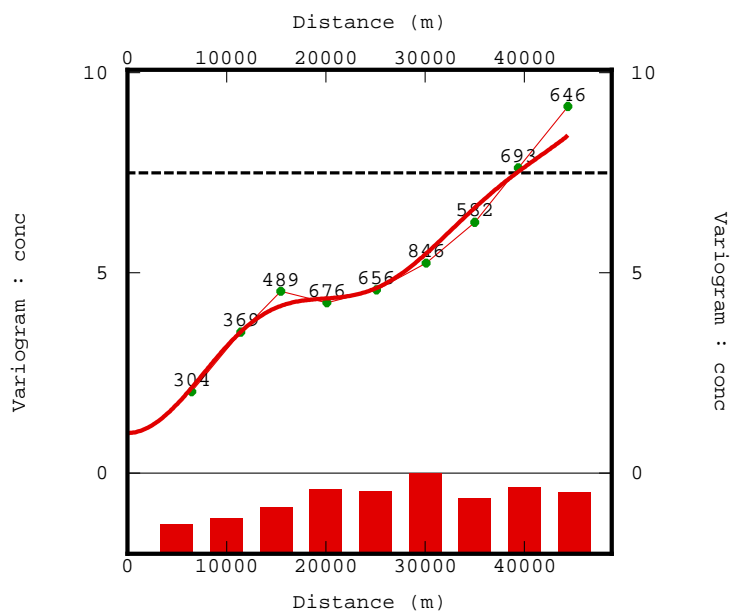


Figure n°16 : Variogramme expérimental (NO<sub>2</sub> en µg/m<sup>3</sup>)

### II.3.2. AJUSTEMENT DU VARIOGRAMME MODELISE ET TEST DE VALIDATION CROISEE

Le modèle de variogramme ajusté est composé d'une structure pépitique (palier=0.5) additionnée d'une structure sinus cardinal et d'une structure puissance.



**Modèle : Pépité(0.5)+ sincard(42800 ;1)+puissance(10000 ;1 ;1.6)**

Figure n°17 : Variogramme modélisé (NO<sub>2</sub> en µg/m<sup>3</sup>)

Le modèle a été soumis au test de validation croisée dont les résultats sont consignés ci-dessous.

**Tableau 3 – Statistiques sur l’erreur d’estimation obtenues par validation croisée sur toutes les données**

<i>MODELE LOS17X17</i>	moyenne	variance
Erreur	0,01237	4.709
Erreur standardisée	0,0051	3.670

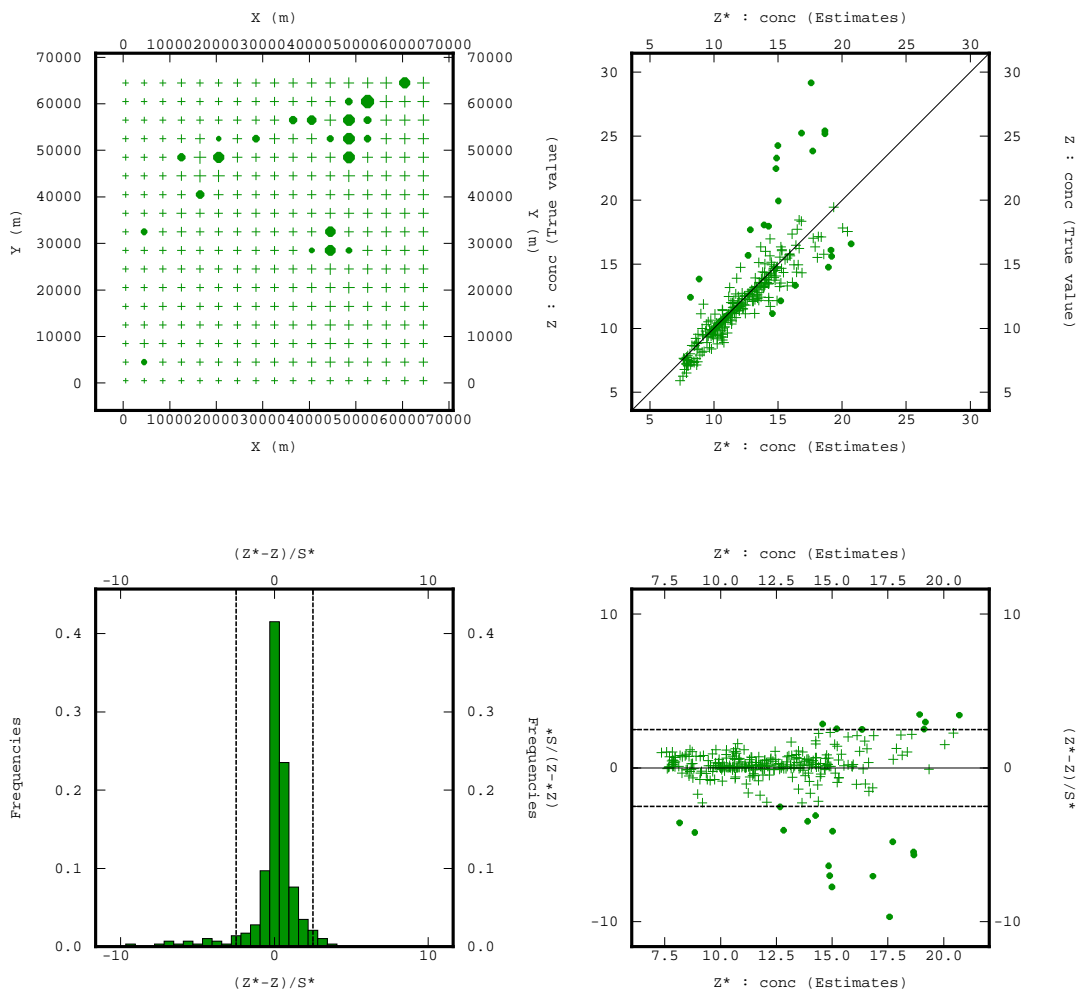


Figure n°18 : Résultats de validation croisée pour le modèle losange 17x17

Le choix d’un maillage différent influe peu, à ce niveau de dégradation, sur la réestimation des données. La sous-estimation de certaines valeurs s’explique toujours par leur situation de proximité.

### II.3.3. ÉTABLISSEMENT DE LA CARTE

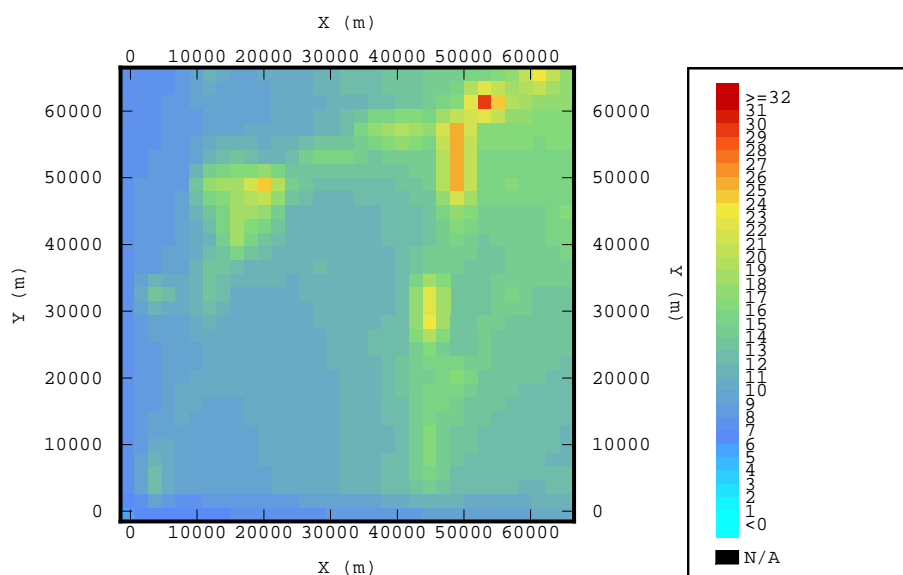


Figure n°19 : Estimation par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de coté.

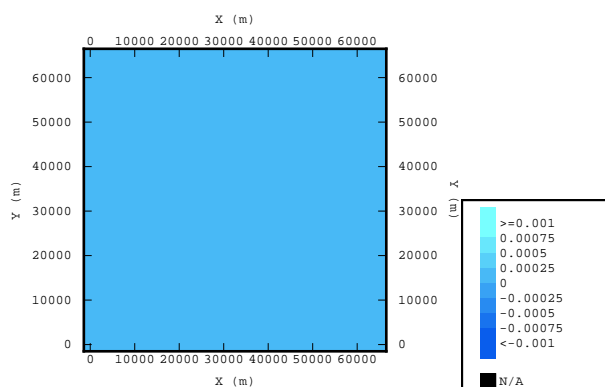


Figure n°20 : Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de coté.

## II.4. Comparaison des cartes et discussion

### Influence du nombre de données sur le variogramme expérimental

Le tableau de synthèse des résultats des différents cas de dégradations traités est établi :

maillages	ref	car_17	car_9	car_7	los_17	los_9	los_7
nombre de points	1089	289	81	49	281	77	49
Min ( $\mu\text{g}/\text{m}^3$ )	5.91	5.91	5.91	5.91	5.91	5.91	5.91
Max ( $\mu\text{g}/\text{m}^3$ )	40.56	29.16	25.42	19.43	28.01	25.25	16.43
Moyenne ( $\mu\text{g}/\text{m}^3$ )	12.43	12.25	11.93	11.3	12.34	11.96	11.79
ecart-type ( $\mu\text{g}/\text{m}^3$ )	3.47	3.5	3.73	2.99	3.53	3.35	3.15
Variance ( $\mu\text{g}/\text{m}^3$ )	12.06	12.24	13.9	8.94	12.46	11.23	9.95
Validation croisée							
erreur							
moy	-0.0096	0.00473	0.0166	0.03745	0.01237	0.02267	0.39313
var	3.23	3.298	3.019	0.987	4.70916	4.21041	1.16542
erreur std.							
moy	-0.0064	-0.00055	0.0042	0.019	0.00506	0.00586	0.32467
var	1.02	0.985	1.335	0.718	3.67046	1.17449	0.87146
validation absolue							
error							
moy	0.044	-0.066	-0.1807	-0.6581	0.0278	-0.3371	-0.3169
var	1.83	2.76	4.712	4.49	1.9676	4.6478	3.9043
erreur std.							
moy	0.025	-0.055	-0.145	-0.8641	0.0284	-0.1802	-0.3032
var	0.58	1.98	2.984	7.5306	1.9943	1.3301	3.5562

Tableau 4 : synthèse de validation des différents schémas de dégradation résultats

On peut noter que la variation de l'erreur de validation croisée évolue peu au cours du processus de dégradation. En revanche l'erreur de validation absolue se dégrade régulièrement avec saut significatif à partir d'un schéma de dégradation de 81 points (grille 9x9).

A partir des différentes restitutions de la carte de données expérimentales, nous avons établi un graphe (figure n°21) de synthèse des différents variogrammes expérimentaux.

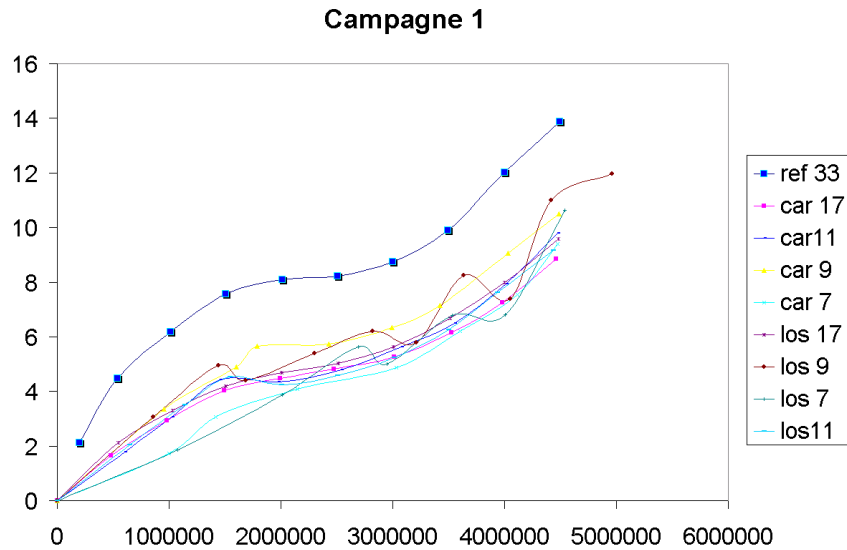


Figure n°21 : Comparaison des différents variogrammes

L'analyse de ce graphique montre un net changement de structure du variogramme avec la diminution du nombre de points, elle apparaît clairement à partir de la grille 9x9 (car 9) pour le maillage carré et (los 9) pour le maillage losange.

Une analyse de la variance des erreurs d'estimation confirme ce résultat.

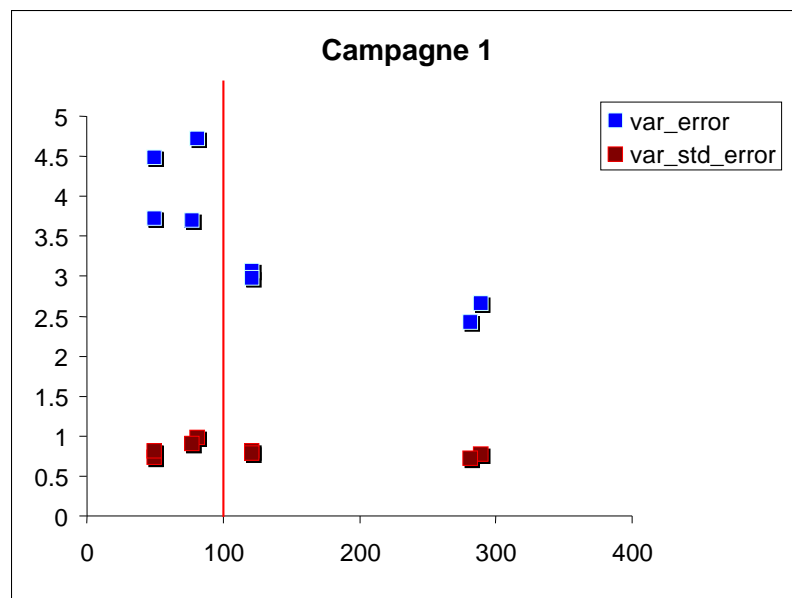


Figure n°22 : variance de l'erreur de krigeage en fonction du nombre de points de mesure pour les maillages carré

On remarque une dégradation significative de la variance de l'erreur de krigeage à partir d'un seuil, évalué pour ce jeu de données expérimental aux alentours de 100 points.

### III. ETUDE DE LA CAMPAGNE DU 3 AU 17 AVRIL 2003

#### III.1 Etude du jeu de données de référence et élaboration de la carte de référence.

##### III.1.1 ETUDE DU VARIOGRAMME EXPERIMENTAL

Le variogramme retenu est omnidirectionnel, avec 10 pas de 500 m chacun.

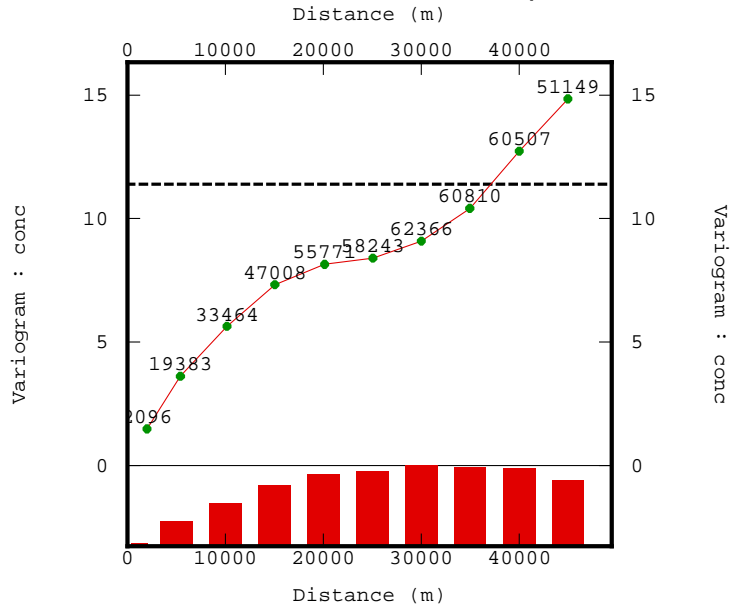
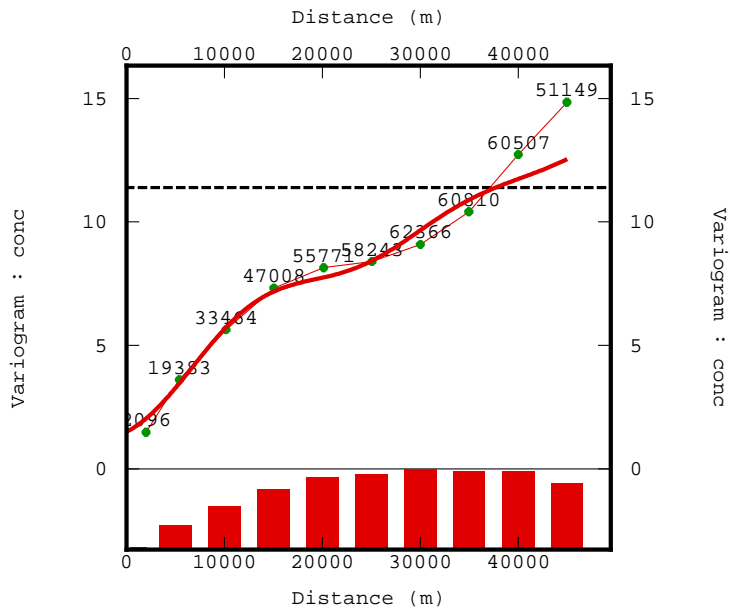


Figure n°23 : Variogramme expérimental (NO<sub>2</sub> en µg/m<sup>3</sup>)

##### III.1.2 AJUSTEMENT DU VARIOGRAMME MODELISE ET TEST DE VALIDATION CROISEE

Le modèle de variogramme ajusté est composé d'une structure pépitique (palier=1.5) additionnée d'une structure sinus cardinal et d'une structure linéaire.



Modèle : Pépité(1.5)+ sincard(66000 ;2.2)+lineaire(45000 ;9)

Figure n°24 : Variogramme modélisé (NO<sub>2</sub> en µg/m<sup>3</sup>)



Comme précédemment, le modèle de variogramme est soumis au test de validation croisée dont les résultats sont présentés ci-après :

**Tableau 5 – Statistiques sur l'erreur d'estimation obtenues par validation croisée sur toutes les données**

<i>MODELE REF</i>	moyenne	variance
Erreur	0.00422	1.3658
Erreur standardisée	0.0304	0.706

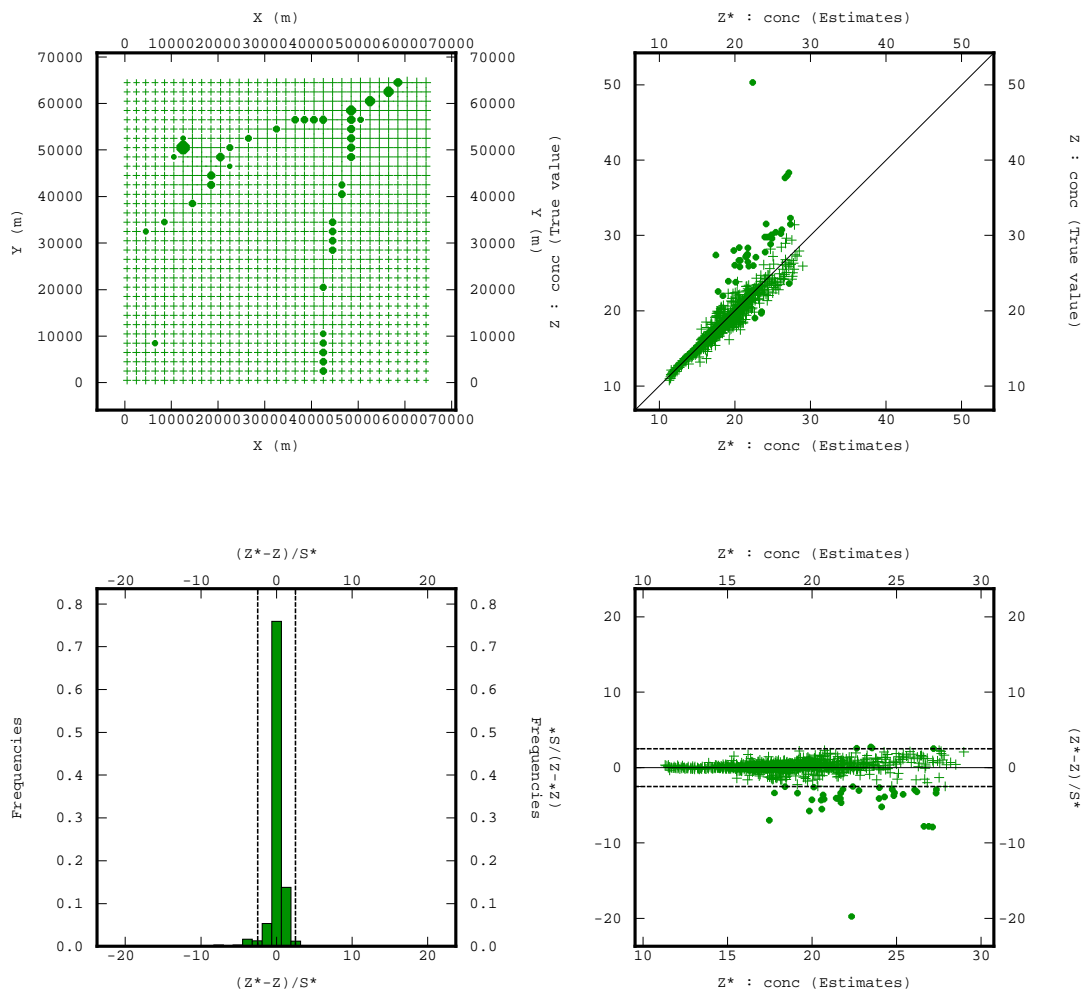


Figure n°25 : Résultats de validation croisée pour le modèle de référence

Sur cette deuxième campagne la plupart des données restent correctement réestimées. On retrouve toujours la sous-estimation de certaines valeurs dues à la proximité des axes automobiles.

### III.1.2. Etablissement de la carte de référence

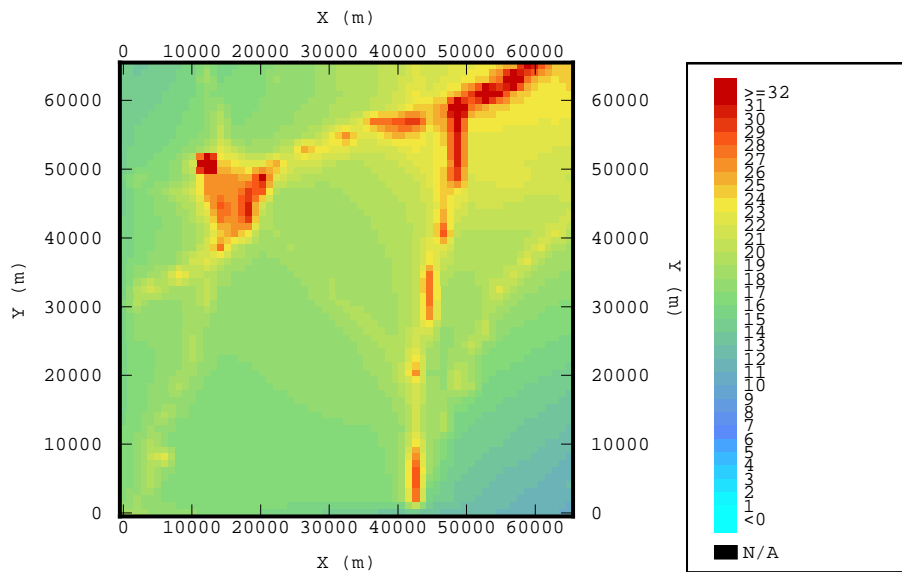


Figure n°26 : Estimation par krigeage ponctuel, à partir de toutes les données, sur une grille de 2km de côté.

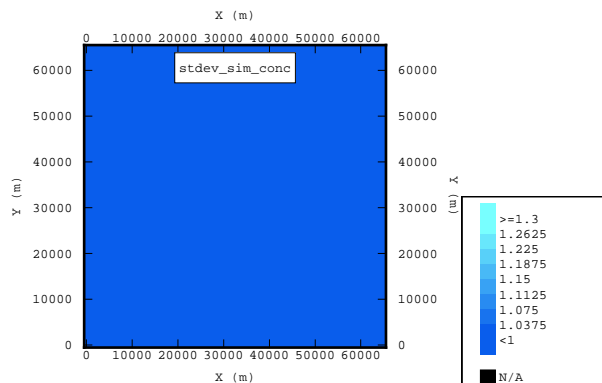


Figure n°27: Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, sur une grille de 2km de côté.

A partir de ce cas de référence l'impact de la dégradation du maillage d'échantillonnage est étudié, d'abord sur un maillage carré, puis sur un maillage losange.

## III.2. Etude de l'échantillonnage sur une grille de 17x17 carre au pas de 4km

### III.2.1. ETUDE DU VARIOGRAMME EXPERIMENTAL

Le variogramme retenu est omnidirectionnel, avec 10 pas de 500 m chacun.

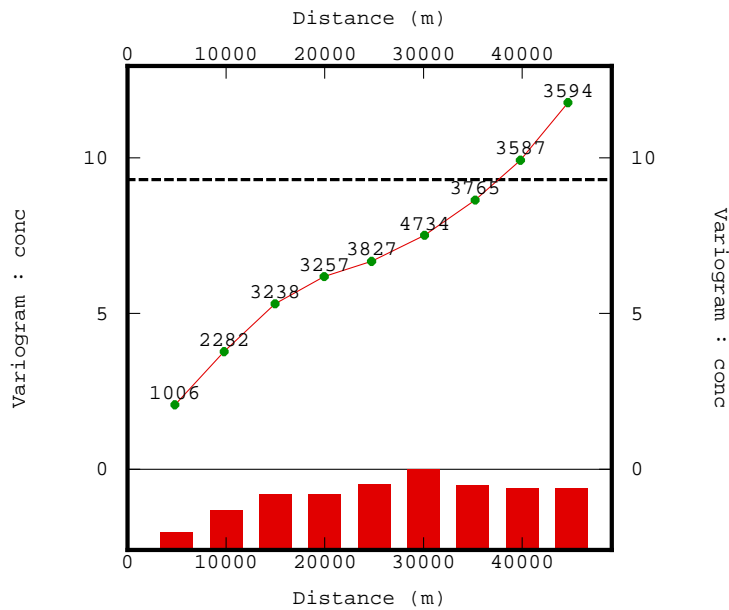
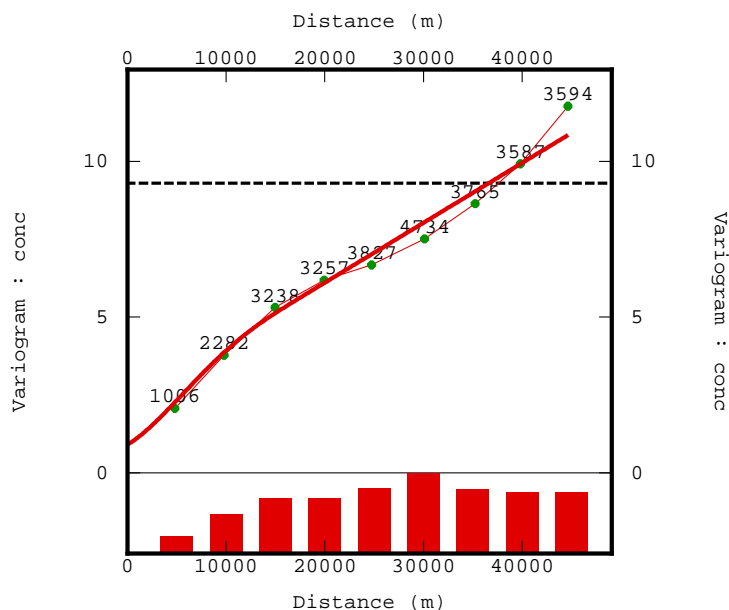


Figure n°28 : Variogramme expérimental (NO<sub>2</sub> en µg/m<sup>3</sup>)

### III.2.2. AJUSTEMENT DU VARIOGRAMME MODELISE ET TEST DE VALIDATION CROISEE

Le modèle de variogramme ajusté est composé d'une structure pépitique (palier=0.9) additionnée d'une cubique et d'une fonction puissance.



**Modèle : Pépité(0.9)+ cub(18600 ;1.35)+puissance(1;13000;2.5)**

Figure n°29 : Variogramme modélisé (NO<sub>2</sub> en µg/m<sup>3</sup>)

Le modèle est soumis au test de validation croisée dont les résultats sont les suivants :

**Tableau 6 – Statistiques sur l’erreur d’estimation obtenues par validation croisée sur toutes les données**

MODELE CAR17X17	moyenne	variance
Erreur	-0,0326	2.0012
Erreur standardisée	-0,0261	1.2921

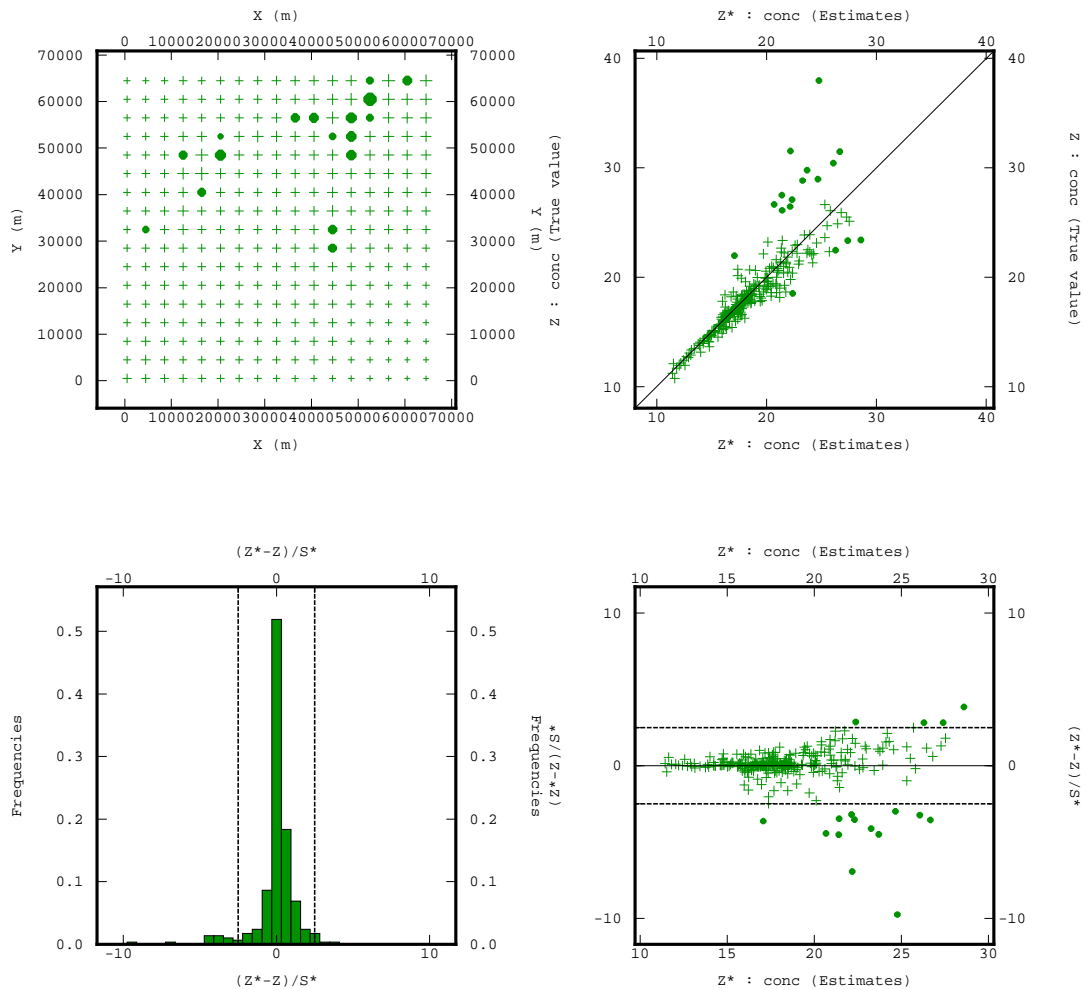


Figure n°30 : Résultats de validation croisée pour le modèle carré de 17x17

La dégradation du maillage n’a pas beaucoup d’influence, sur la réestimation des données. Il reste certaines valeurs sous-estimées du fait de leur proximité à certain gros émetteurs.

### III.2.3. ÉTABLISSEMENT DE LA CARTE

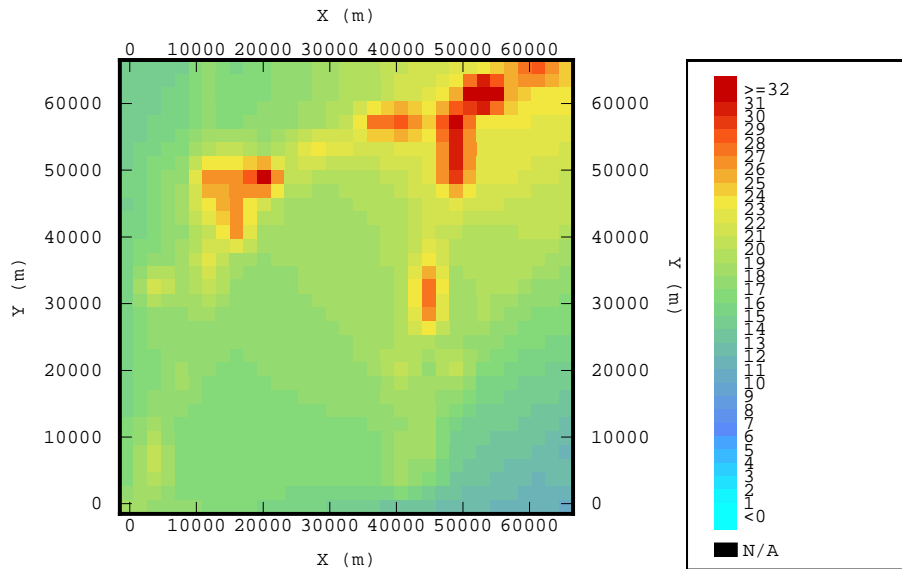


Figure n°31 : Estimation par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de coté.

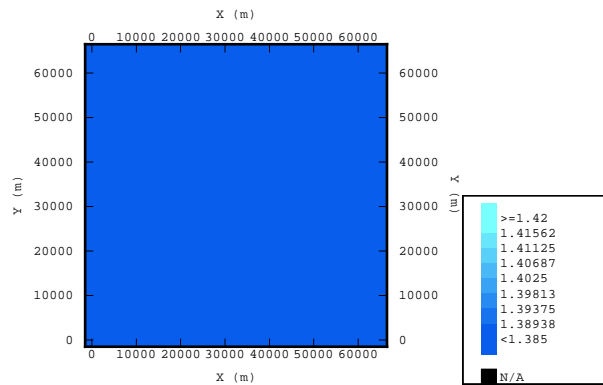
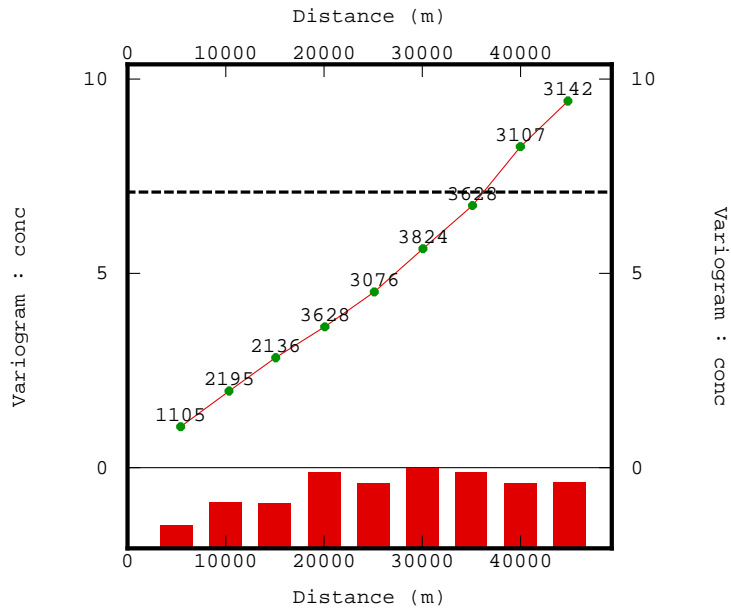


Figure n°32 : Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de coté.

## III.3 Etude de l'échantillonnage sur une grille de 17x17 losange au pas de 4km

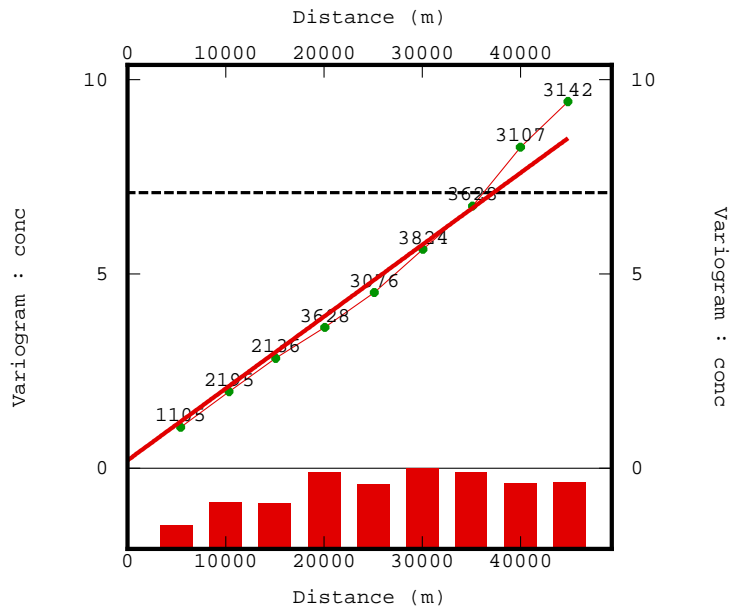
### III.3.1. ETUDE DU VARIOGRAMME EXPERIMENTAL

Le variogramme retenu est omnidirectionnel, avec 10 pas de 500 m chacun.

Figure n°33 : Variogramme expérimental (NO<sub>2</sub> en µg/m<sup>3</sup>)

### III.3.2. AJUSTEMENT DU VARIOGRAMME MODELISE ET TEST DE VALIDATION CROISEE

Le modèle de variogramme ajusté est composé d'une structure pépitique (palier=0.2) additionnée d'une structure linéaire.



**Modèle : Pépité(0.2)+ 1<sup>er</sup>order(10000 ;1.85)**

Figure n°34 : Variogramme modélisé (NO<sub>2</sub> en µg/m<sup>3</sup>)

Le modèle est soumis au test de validation croisée dont les résultats sont consignés ci-après :

#### **Tableau 7 – Statistiques sur l'erreur d'estimation obtenues par validation croisée sur toutes les données**

<i>MODELE LOS17X17</i>	moyenne	variance
Erreur	0,0345	1.51174
Erreur standardisée	0,0446	2.5093

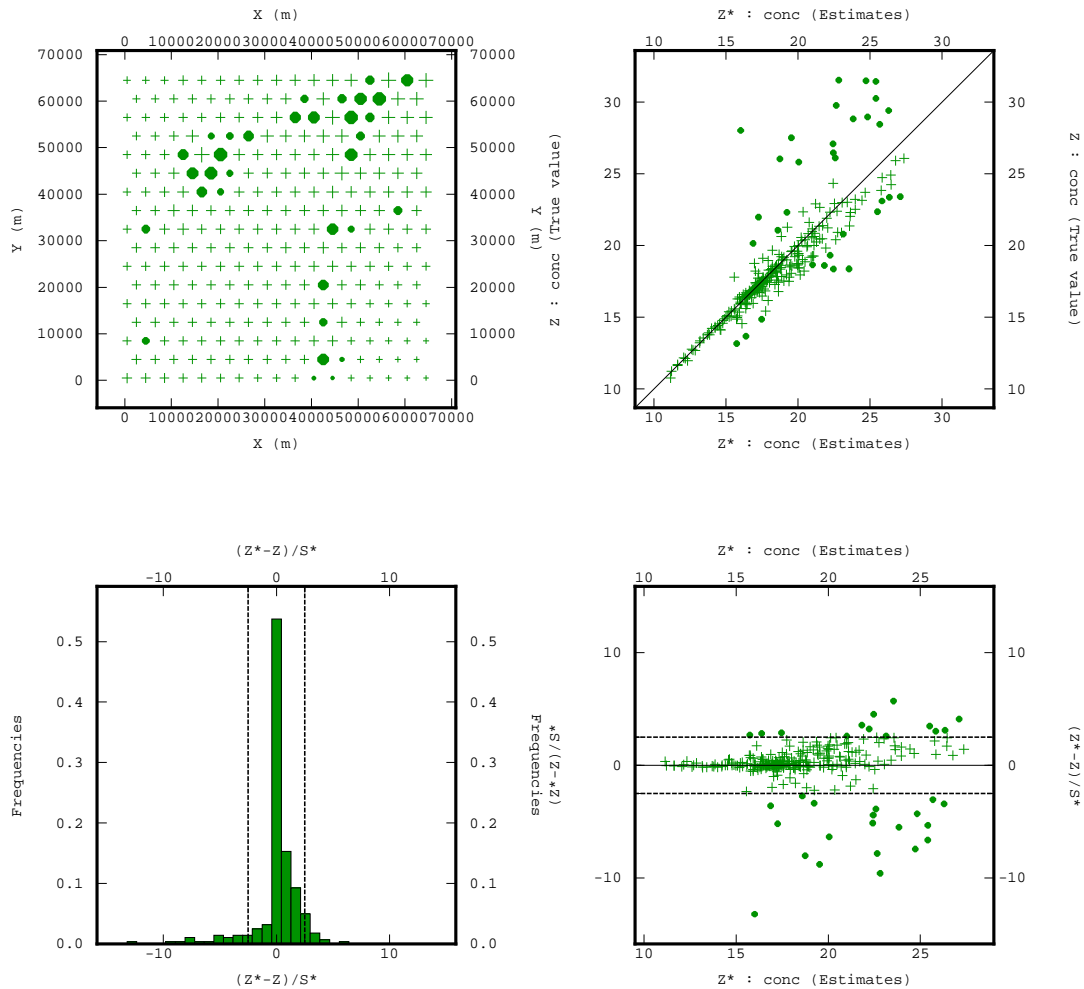


Figure n°35 : Résultats de validation croisée pour le modèle losange 17x17

De nouveau, dans cette deuxième campagne, le choix d'une structure de maillage différente n'a pas beaucoup d'influence, sur la réestimation des données. Il reste certaines valeurs sous-estimées du fait de leur proximité à certain gros émetteurs.

### III.3.3. ÉTABLISSEMENT DE LA CARTE

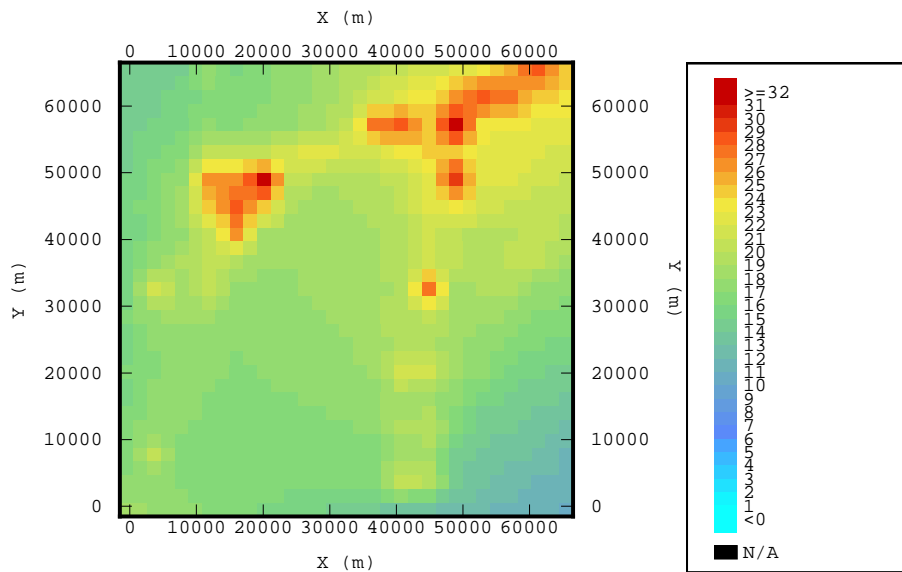


Figure n°36 : Estimation par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de côté.

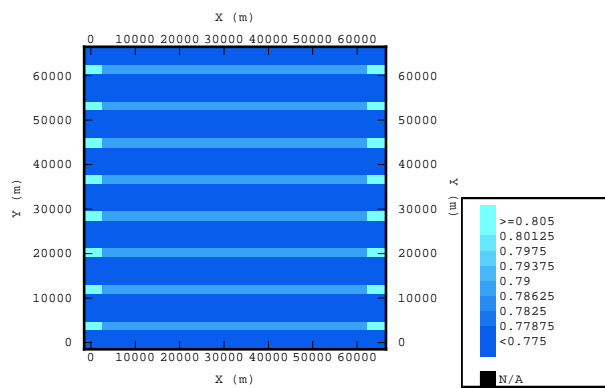


Figure n°37 : Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, sur une grille de 4km de côté.



### III.4. Comparaison des cartes et discussion

#### Influence du nombre de données sur le variogramme expérimental

A partir des différents cas de dégradations traités, le tableau suivant, qui synthétise les résultats obtenus, est établi :

	ref	car_17	car_9	car_7	los_17	los_9	los_7
nombre de points	1089	289	81	49	281	77	49
Min ( $\mu\text{g}/\text{m}^3$ )	10.76	10.76	10.76	11.24	10.76	10.76	11.24
Max ( $\mu\text{g}/\text{m}^3$ )	32.31	27.5	20.99	29.4	31.52	29.77	29.4
Moyenne ( $\mu\text{g}/\text{m}^3$ )	18.56	18.14	16.99	17.92	18.58	18.25	18.08
ecart-type ( $\mu\text{g}/\text{m}^3$ )	3.65	3.66	3.83	3.15	3.65	3.46	3.27
Variance ( $\mu\text{g}/\text{m}^3$ )	11.38	9.3	5.0583	9.93	13.33	11.96	10.67
Validation croisée							
Erreur							
moy	0.00182	0.00614	0.00345	0.02481	0.00795	0.01854	0.0343
var	2.7167	3.02322	3.18741	3.27754	3.52218	3.05048	3.08989
Erreur std.							
moy	0.00063	0.0022	0.0022	0.00729	0.00414	0.00824	0.01449
var	1.3536	1.64412	3.33237	1.13414	4.23693	2.5566	2.27579
validation absolue							
Erreur							
moy	0.0422	-0.0326	-0.1059	-0.41	0.0345	-0.172	-0.2877
var	1.3658	2.0012	3.3061	4.3761	1.51174	3.1175	4.018
Erreur std.							
moy	0.0304	-0.0261	-0.1286	-0.258	0.0446	-0.1828	-0.1748
var	0.706	1.2921	5.0503	1.6828	2.5093	3.474	1.4395

Tableau 8 : synthèse de validation des différents schémas de dégradation résultats

Les résultats de cette deuxième campagne confirment ceux précédemment acquis. La variation de l'erreur de validation croisée évolue peu au cours du processus de dégradation. En revanche l'erreur de validation absolue se dégrade régulièrement avec saut significatif à partir d'un schéma de dégradation de 81 points (grille 9x9).

A partir des différentes restitutions de la carte de données expérimentales, nous avons établi un graphe de synthèse (figure n° 8) des différents variogrammes expérimentaux.

## Campagne 2

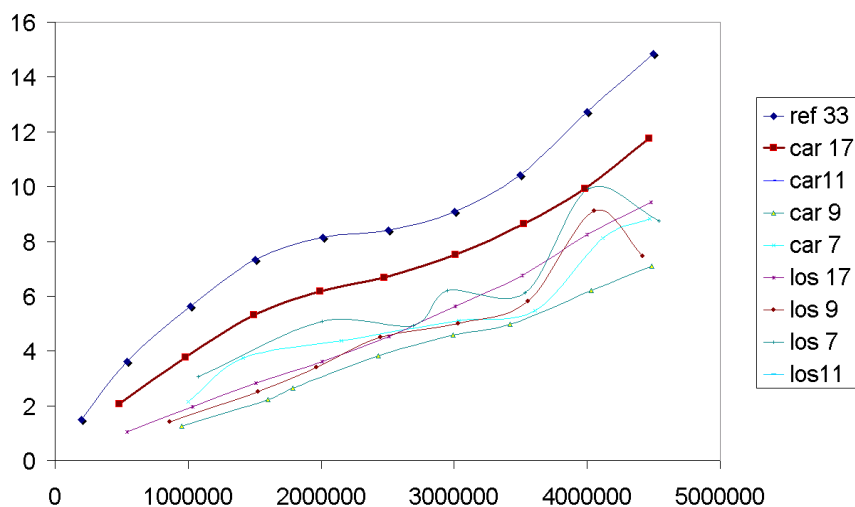
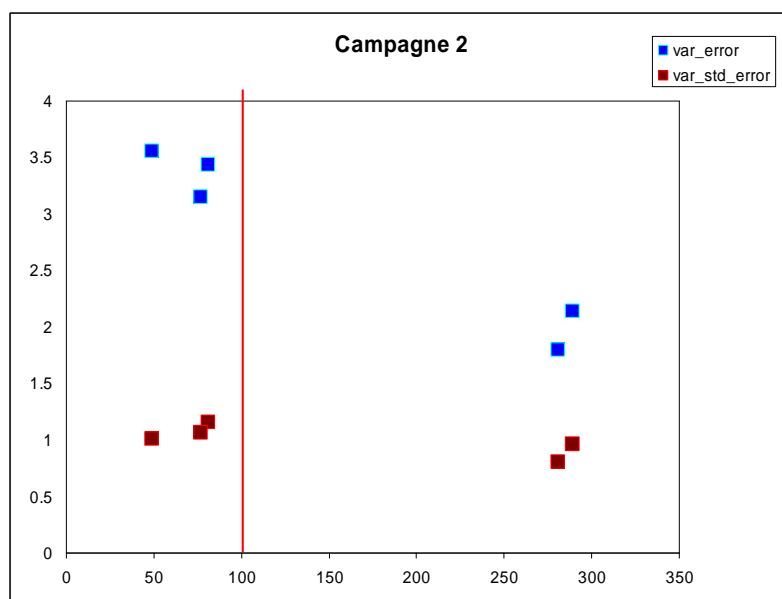


Figure n°38 : Comparaison des différents variogrammes

L'analyse de ce graphique montre un net changement de structure du variogramme avec la diminution du nombre de points, elle apparaît clairement à partir de la grille 9x9 (car 9) pour le maillage carré et (los 9) pour le maillage losange.

Une analyse de la variance des erreurs d'estimation confirme ce résultat.



Figures n°39 : variance de l'erreur de Krigeage en fonction du nombre de point de mesure pour les maillages carré

On remarque une dégradation significative de la variance de l'erreur de krigeage à partir d'un seuil, évalué pour ce jeu de données expérimental aux alentours de 100 points.

## IV DISCUSSION

A partir des différentes expériences sur les maillages numériques concernant les deux périodes de mesures étudiées, il est possible de tirer un certain nombre de conclusions concernant la recherche d'un optimum d'échantillonnage.

1) Tout d'abord la structure des variogrammes expérimentaux se détériore nettement en fonction du nombre de points utilisés. Cette détérioration existe pour les deux géométries de maillage considérées ; elle est particulièrement marquée à partir du schéma de dégradation de 9x9 soit 81 points,.

2) On remarque ensuite une dégradation significative de la variance de l'erreur de krigeage à partir d'un seuil, évalué pour le jeu de données expérimental aux alentours de 100 points. Cela correspond à une densité d'échantillonnage de 2,29 points pour 100 km<sup>2</sup> ou 1 point pour 43,5 km<sup>2</sup>.

**2<sup>EME</sup> PARTIE :**

**SURVEILLANCE DES VILLES**

**DE 10000 A 100000 HABITANTS**

## ETUDE D'UNE VILLE DE MOINS DE 100 000 HABITANTS : CAS DE CHERBOURG-OCTEVILLE.

### I. INTRODUCTION

#### I.1. Présentation du jeu de données initial

Le jeu de données étudié a été élaboré par l'association agréée de surveillance AIRCOM. Il est constitué de deux campagnes de mesure, d'une durée de 14 jours chacune, des concentrations en dioxyde d'azote dans l'air ambiant, par échantillonnage passif, sur la communauté urbaine de Cherbourg-Octeville.

Avec 91 717 habitants, la communauté urbaine de Cherbourg regroupe un peu plus de 18% de la population du département de la Manche (Fig. 40). Au sein de cet ensemble, 48% des habitants résident sur le territoire de Cherbourg-Octeville.

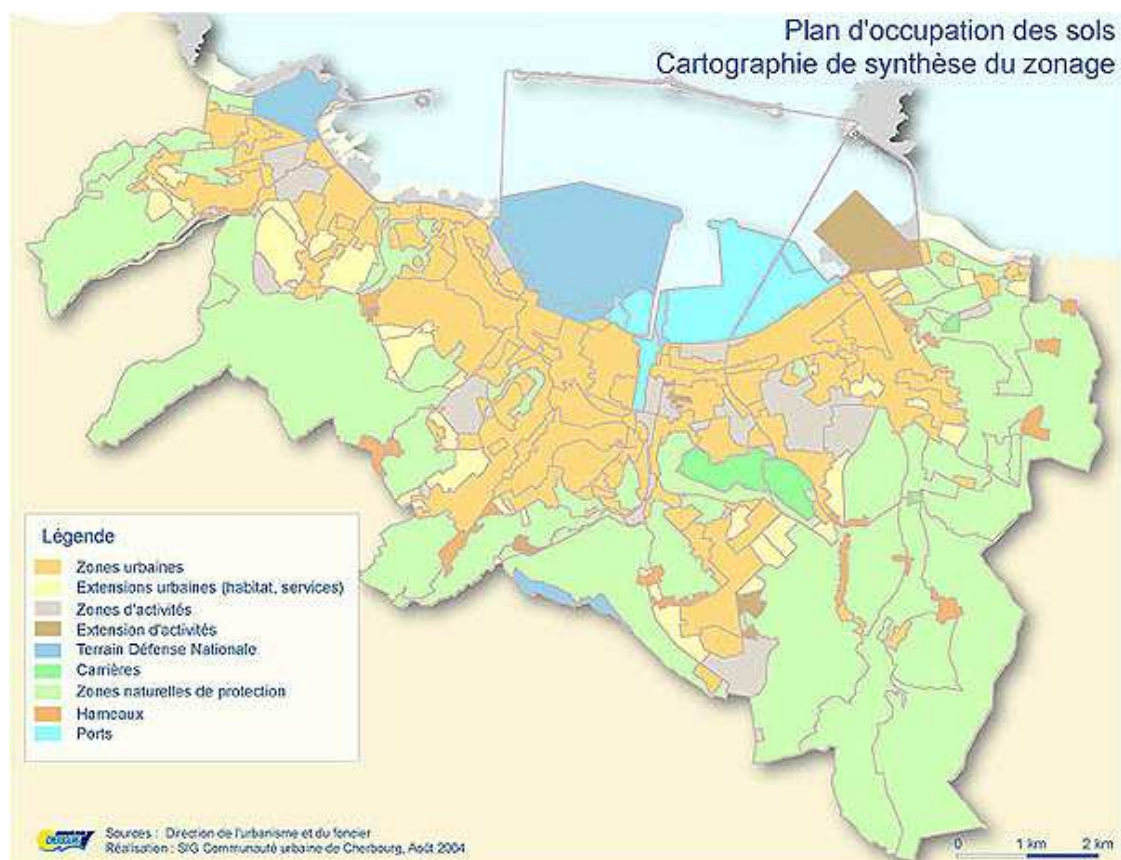


Figure 40 – Communauté Urbaine de Cherbourg-Octeville (plan d'occupation des sols).

88 sites de mesure ont été disposés dans l'agglomération de Cherbourg, sur un domaine d'environ 14 km d'est en ouest et de 10 km du nord au sud (Fig. 41). Les valeurs prises en compte pour cette étude correspondent à la moyenne des deux campagnes ou, lorsque l'une des deux valeurs fait défaut, à la valeur existante.

La stratégie d'échantillonnage suivie pour ces campagnes est celle classiquement adoptée par la plupart des AASQA : un resserrement des points de mesure en centre-ville et une disposition plus espacée en périphérie. Tous les détails sur ces campagnes sont consignés dans le rapport d'AIRCOM [6].

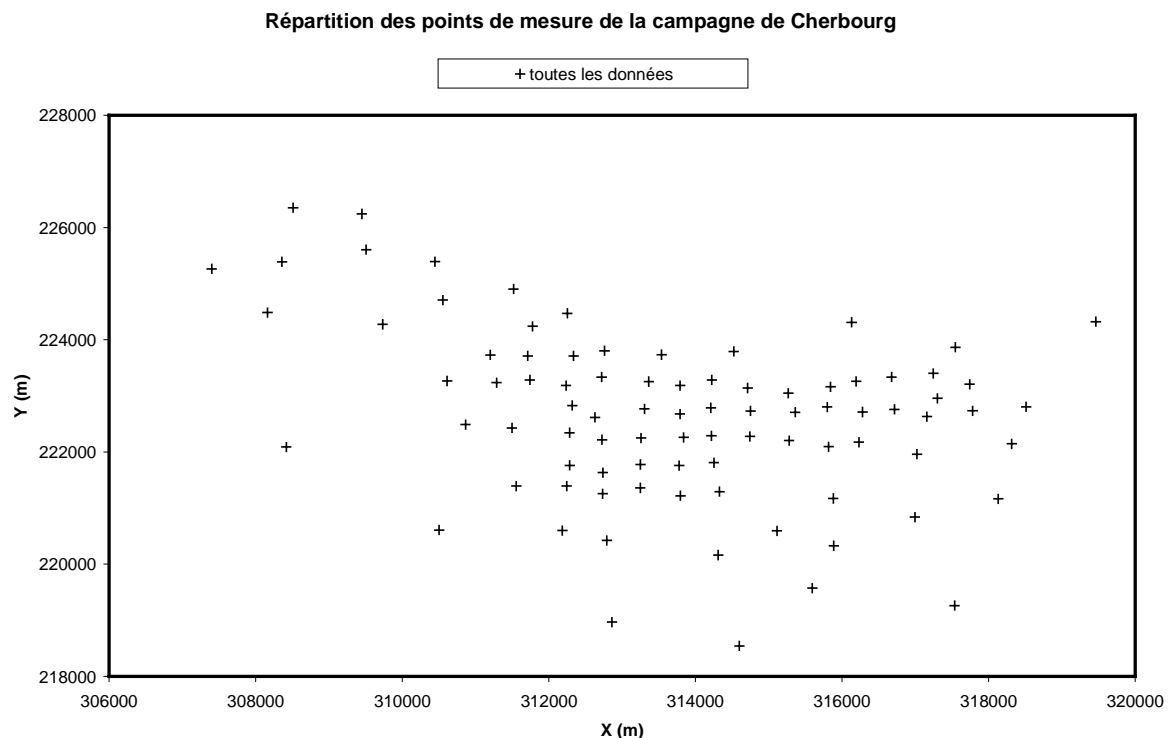


Figure 41 – Carte de répartition des points de mesure de la campagne de Cherbourg.

## I.2. Méthodologie de sélection des données

Nous souhaitons échantillonner les données suivant deux maillages réguliers :

- l'un à mailles carrées de 1 km de côté ;
- l'autre à mailles carrées de 2 km de côté.

Pour cela, nous avons calé sur le domaine d'étude un maillage régulier de mailles 1km ou 2km. Puis nous avons sélectionné, pour chaque maille, le point le plus proche du centre de la maille.

Le jeu de données correspondant à une maille régulière de 1 km est appelé « sélection 1 ». Il comprend 43 données (Fig. 42).

Les données non sélectionnées par ce maillage constituent le jeu de données « complémentaires de la sélection 1 », qui comprend 45 données.

Le jeu de données correspondant à une maille régulière de 2 km est appelé « sélection 2 ». Il comprend 21 données (Fig. 43).

Les données non sélectionnées par ce maillage constituent le jeu de données « complémentaires de la sélection 2 », qui comprend 67 données.

L'ensemble de ces sélections est représenté à la Figure 44. Les caractéristiques statistiques des différents jeux de données ainsi constitués sont données au Tableau 9.

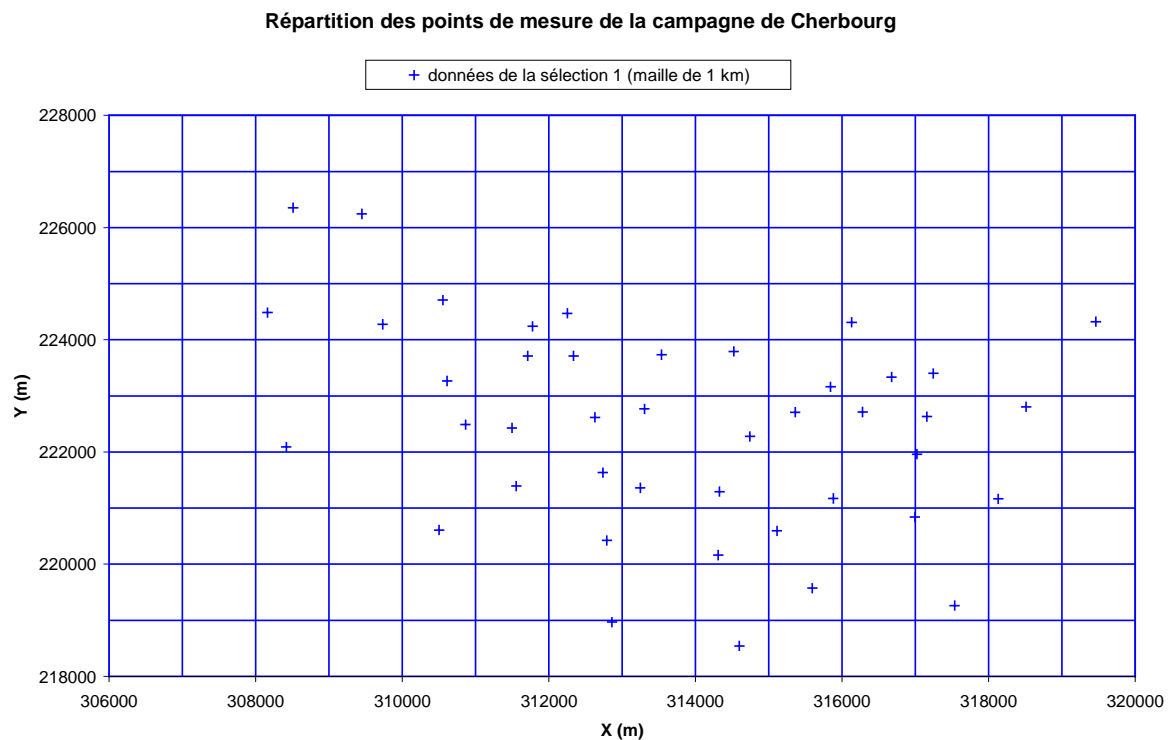


Figure 42 – Carte de répartition des points de la « sélection 1 » (mailles de 1km).

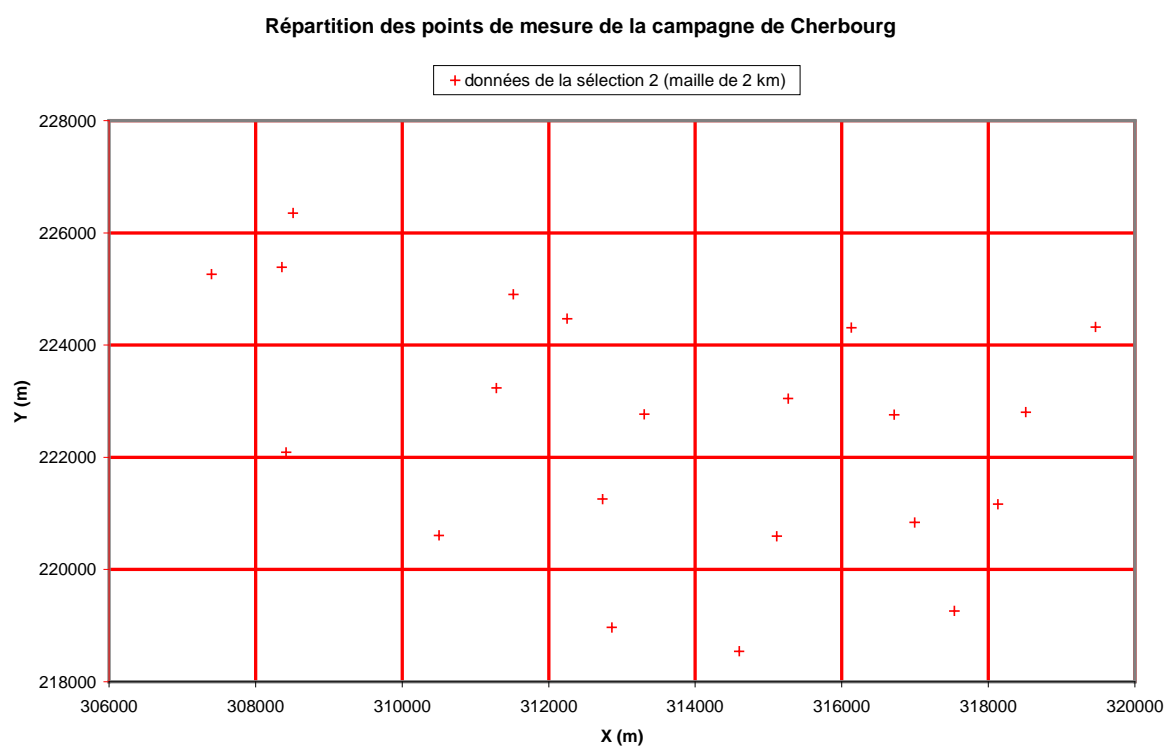


Figure 43 – Carte de répartition des points de mesure de la « sélection 2 » (mailles de 2 km).

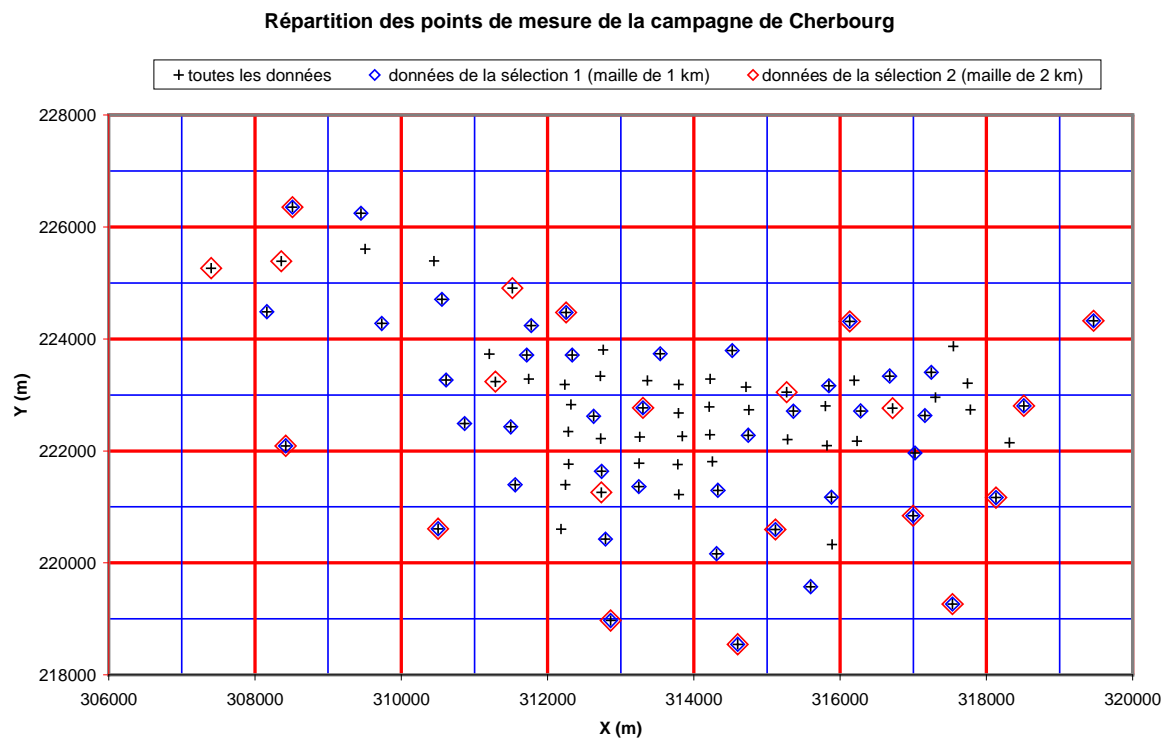


Figure 44 – Carte de répartition des points de mesure échantillonnés selon les différents maillages.

Tableau 9 - Principales caractéristiques des jeux de données étudiés.

	Toutes les données	Sélection 1 (1 donnée par maille de 1km×1km)	Données complémentaires de la sélection 1	Sélection 2 (1 donnée par maille de 2km×2km)	Données complémentaires de la sélection 2
Nombre de données	88	43	45	21	67
Minimum ( $\mu\text{g}/\text{m}^3$ )	6,85	6,85	9,60	7,60	6,85
25 <sup>ème</sup> percentile ( $\mu\text{g}/\text{m}^3$ )	12,80	10,50	13,80	11,35	14,05
Médiane ( $\mu\text{g}/\text{m}^3$ )	17,78	15,75	21,05	13,8	20,55
75 <sup>ème</sup> percentile ( $\mu\text{g}/\text{m}^3$ )	26,25	24,68	27,40	23,80	27,52
Maximum ( $\mu\text{g}/\text{m}^3$ )	40,80	38,00	40,80	30,25	40,80
Moyenne ( $\mu\text{g}/\text{m}^3$ )	19,81	17,86	21,67	16,51	20,97
Ecart-type ( $\mu\text{g}/\text{m}^3$ )	8,31	7,98	8,20	6,90	8,45
Variance ( $\mu\text{g}/\text{m}^3$ )	69,11	63,70	67,24	47,60	71,40
CV%	0,42	0,45	0,39	0,42	0,40



## II. ETUDE DU JEU DE DONNEES INITIAL ET ELABORATION DE LA CARTE DE REFERENCE.

Le but de cette partie est d'établir la carte de référence, c'est-à-dire la carte de répartition des concentrations en dioxyde d'azote établie avec le maximum de points de mesure. Elle servira par la suite à comparer les cartes obtenues avec un nombre de points plus faibles.

### II.1. Etude du variogramme expérimental

Le variogramme expérimental retenu est omnidirectionnel, avec 12 pas de 390 m chacun. La nuée variographique ne présentant pas d'anomalie, aucune donnée n'a été écartée (Fig. 45).

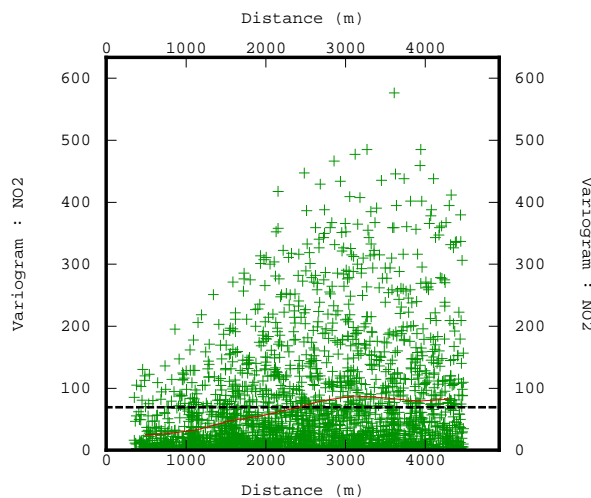


Figure 45 – Variogramme expérimental (pas de 390 m, 12 pas) avec toutes les données.

### II.2. Ajustement du variogramme modélisé et test de validation croisée

Deux modèles différents ont été ajustés aux données expérimentales. Ils sont composés tous deux d'une structure pépitique (de même palier=17) additionnée d'une structure gaussienne ou cubique.

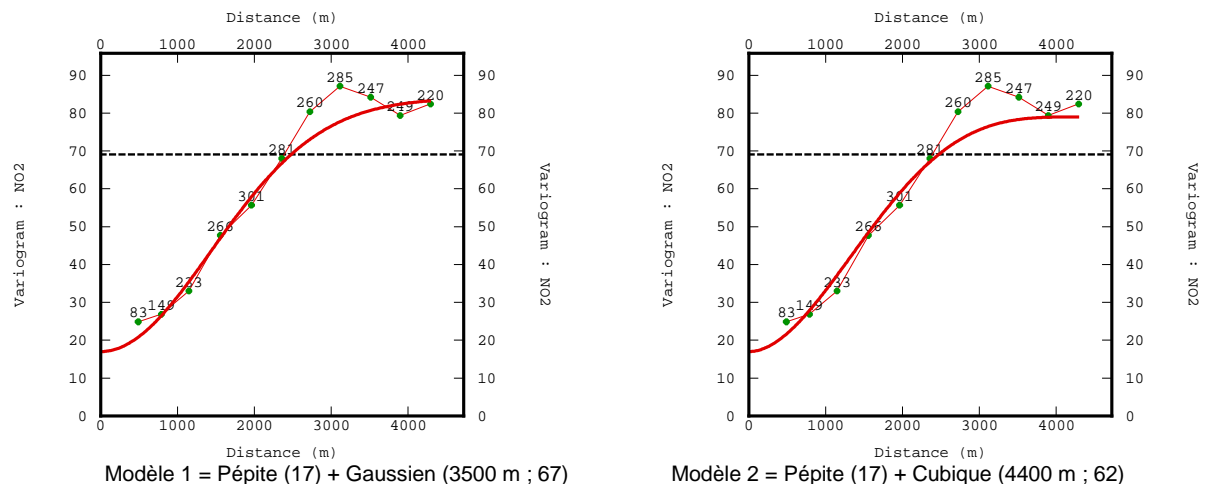


Figure 46 – Variogrammes expérimentaux et modélisés.

Afin d'apprécier leur qualité respective, chacun d'eux est soumis au test de validation croisée. Ce test consiste à calculer, au moyen du modèle, une estimation de la concentration en polluant, pour chaque point de l'espace où celle-ci a été mesurée (pour cela, on ôte momentanément la valeur mesurée du jeu de données). Il faut préalablement définir un voisinage, délimitant de nombre de données mesurées prises en compte pour le calcul d'estimation.

En raison du faible nombre de points de mesure (< 100), c'est le voisinage unique qui a été considéré, c'est-à-dire que, pour chaque estimation, c'est l'ensemble des points de mesure qui est pris en compte. Ce voisinage unique sera celui utilisé dans l'ensemble de l'étude.

Dans le test de validation croisée, les valeurs estimées par le modèle sont comparées aux valeurs mesurées (valeurs vraies), notamment en calculant l'erreur d'estimation. Sous ISATIS, l'erreur d'estimation est définie par la différence entre la valeur estimée moins la valeur vraie.

**Tableau 10 – Statistiques de l'erreur d'estimation par validation croisée sur toutes les données**

(a) avec le modèle 1= Pépîte (17) + Gaussien (3500 m ; 67)

<i>MODELE 1</i>	moyenne	variance
Erreur	0,13066	19,86196
Erreur standardisée	0,00979	0,91117

1 valeur non robuste sur 88 valeurs.

(b) avec le modèle 2 = Pépîte (17) + Cubique (4400 m ; 62).

<i>MODELE 2</i>	moyenne	variance
Erreur	0,14896	19,79324
Erreur standardisée	0,01120	0,84959

0 valeur non robuste sur 88 valeurs.

Les caractéristiques statistiques des erreurs d'estimation (Tab. 10) montrent que les deux modèles conduisent à des résultats semblables et assez satisfaisants : l'erreur moyenne est proche de zéro et la variance de l'erreur standardisée est voisine de 1. Le nombre de valeurs mal estimées (non robustes), pour lesquelles l'écart-type de l'erreur d'estimation standardisée dépasse 2,5 en valeur absolue, est limité (0 ou 1).

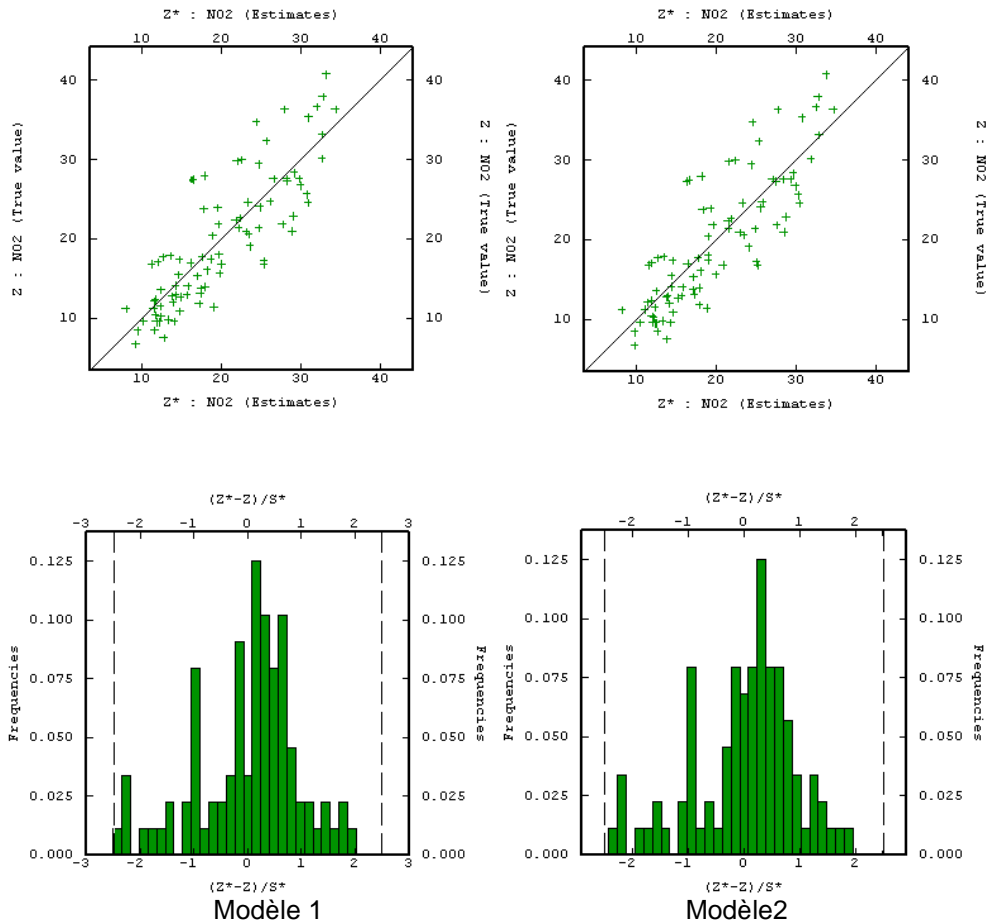
**Tableau 11 – Résultats de la validation croisée sur toutes les données pour le krigeage ordinaire ponctuel avec le modèle 1 = Pépîte (17) + Gaussien (3500 m ; 67).**

		Validation croisée par krigeage ordinaire ponctuel selon le modèle 1	
Toutes les données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation standardisée
Nombre de données	88	88	88
Minimum	6,85	8,02	-2,51
25 <sup>ème</sup> percentile	12,80	13,92	-0,45
Médiane	17,775	18,51	0,19
75 <sup>ème</sup> percentile)	26,25	25,43	0,60
Maximum	40,80	34,44	1,94
Etendue (= Max – Min)	33,95	26,42	4,45
Moyenne)	19,81	19,94	0,01
Ecart-type	8,31	7,03	0,95
Variance	69,11	49,42	0,91
CV %	0,42	0,35	
Corrélation mesures/estimations		r=0,8442	

**Tableau 12 – Résultats de la validation croisée sur toutes les données pour le krigeage ordinaire ponctuel avec le modèle 2 = Pépité (17) + Cubique (4400 m ; 62).**

		Validation croisée par krigeage ordinaire ponctuel selon le modèle 2	
Toutes les données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation standardisée
Nombre de données	88	88	88
Minimum	6,85	8,20	-2,43
25 <sup>ème</sup> percentile	12,80	13,94	-0,41
Médiane	17,775	18,65	0,21
75 <sup>ème</sup> percentile	26,25	25,36	0,60
Maximum	40,80	34,76	1,86
Etendue (= Max – Min)	33,95	26,56	4,29
Moyenne	19,81	19,96	0,01
Ecart-type	8,31	6,94	0,92
Variance	69,11	48,13	0,85
CV %	0,42	0,35	
Corrélation mesures/estimations		r=0,8448	

Comme le montrent les tableaux 11 et 12 et la figure 47 suivante, les valeurs estimées sont bien corrélées aux valeurs mesurées, avec un coefficient de corrélation supérieur à 0,84 pour les deux modèles.



**Figure 47 – Résultats de validation croisée pour les modèles 1 et 2.**

### II.3.Choix du meilleur modèle et établissement de la carte de référence

Bien que les différences entre les deux modèles soient minimales, nous orientons notre choix vers le modèle 2. En effet, celui-ci conduit à des valeurs estimées plus proches des mesures et à une erreur d'estimation plus resserrée (variance de l'écart-type d'erreur plus faible).

Pour dresser la carte de référence de l'étude, c'est-à-dire la carte réalisée à partir du plus grand nombre de mesures disponibles, nous estimons les valeurs de concentration en polluant aux nœuds d'une grille régulière en 2D de 250 m de côté.

Deux techniques de krigeage ordinaire sont utilisées, afin d'être comparées : le krigeage ponctuel et le krigeage de bloc.

Pour le krigeage ponctuel, les caractéristiques de la grille d'estimation sont :

$$\begin{aligned} X_{\min} &= 307000 \text{ m}, & 57 \text{ nœuds}, & & X_{\max} &= 321000, & \Delta X &= 14000 \text{ m}, \\ Y_{\min} &= 218000 \text{ m}, & 41 \text{ nœuds}, & & Y_{\max} &= 228000, & \Delta Y &= 10000 \text{ m}. \end{aligned}$$

Sous ISATIS, lors du krigeage de bloc, les valeurs sont estimées aux nœuds d'une grille correspondant aux centres des blocs. Pour retrouver exactement le même domaine d'estimation que pour le krigeage ponctuel, il faut donc définir une nouvelle grille, décalée d'un demi-bloc (d'une demi-maille).

Pour le krigeage de bloc, les caractéristiques de la grille d'estimation sont ainsi :

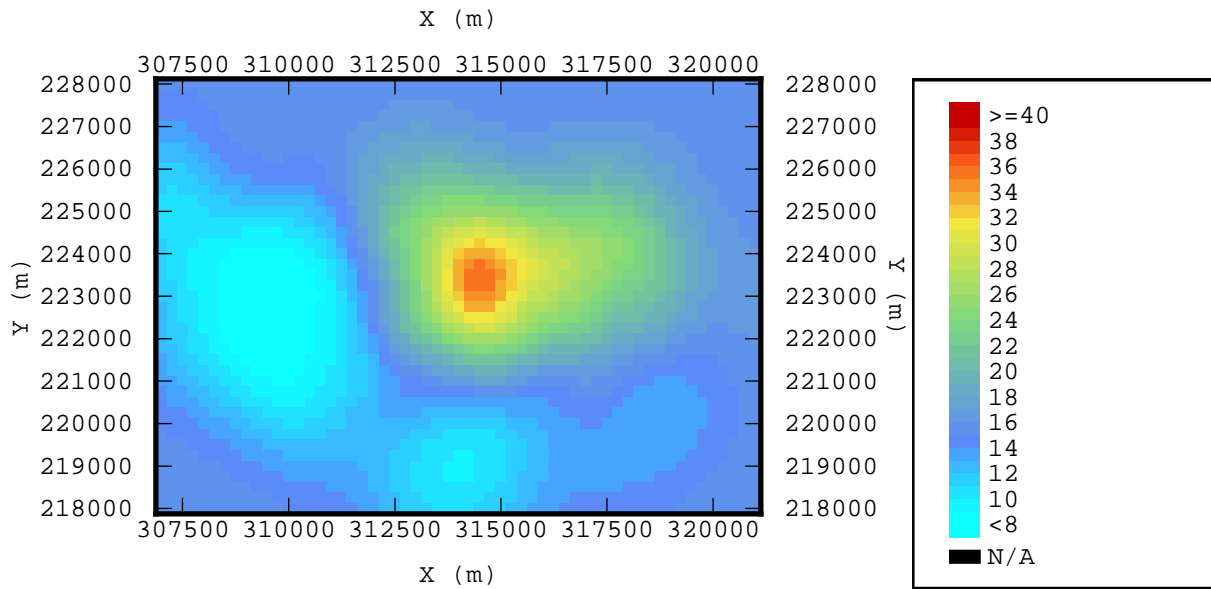
$$\begin{aligned} X_{\min} &= 307125 \text{ m}, & 56 \text{ nœuds}, & & X_{\max} &= 320875, & \Delta X &= 13750 \text{ m}, \\ Y_{\min} &= 218125 \text{ m}, & 40 \text{ nœuds}, & & Y_{\max} &= 227875, & \Delta Y &= 9750 \text{ m}. \end{aligned}$$

**Tableau 13 – Estimation sur une grille régulière de 250 m de côté, par krigeage ordinaire, ponctuel ou de bloc, selon le modèle 2, pour toutes les données.**

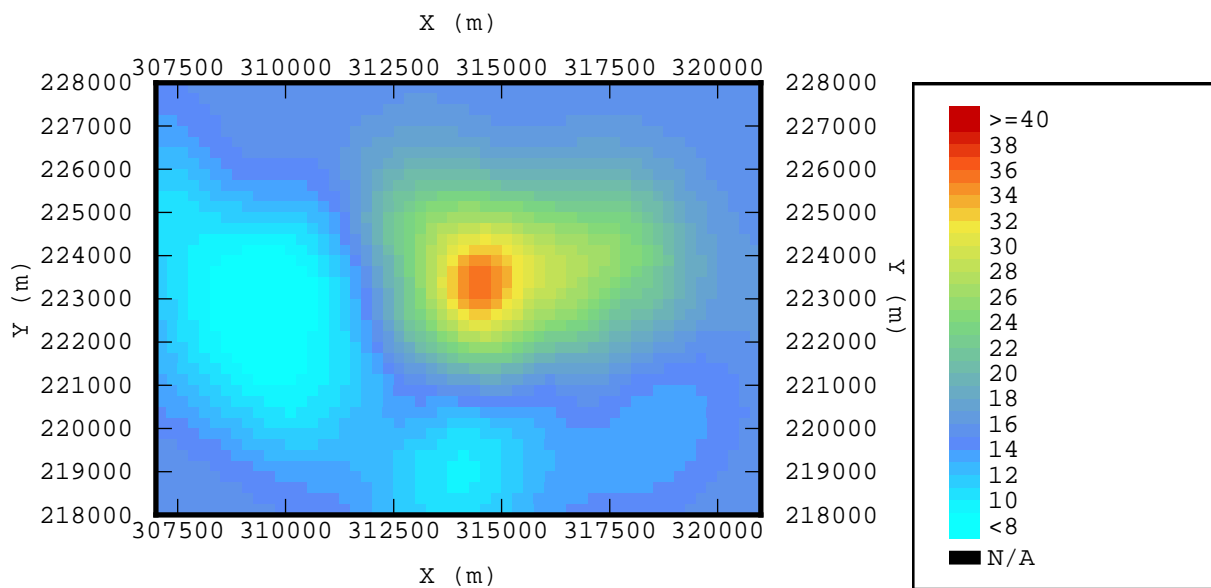
<i>TOUTES LES DONNEES</i>	Mesures ( $\mu\text{g}/\text{m}^3$ )	Estimation par krigeage ordinaire	
		Valeurs estimées modèle 2, krigeage <b>ponctuel</b> ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées modèle 2, krigeage de <b>bloc</b> ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	88	2337	2240
Minimum	6,85	7,26	7,29
25 <sup>ème</sup> percentile	12,80	12,85	12,80
Médiane	17,775	15,34	15,32
75 <sup>ème</sup> percentile	26,25	17,53	17,76
Maximum	40,80	35,73	35,62
Moyenne	19,81	16,02	16,08
Ecart-type	8,31	5,21	5,30
Variance	69,11	27,14	28,08
CV %	0,42	0,33	0,33

**Tableau 14 – Ecart-types de krigeage, sur une grille régulière de 250 m de côté, par krigeage ordinaire, ponctuel ou de bloc, selon le modèle 2, pour toutes les données.**

<i>TOUTES LES DONNEES</i>	Estimation par krigeage ordinaire	
	Ecart-types de krigeage modèle 2, krigeage <b>ponctuel</b> ( $\mu\text{g}/\text{m}^3$ )	Ecart-types de krigeage modèle 2, krigeage de <b>bloc</b> ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	2337	2240
Minimum	4,42	1,57
Médiane	6,51	4,79
Maximum	9,25	8,25
Moyenne	6,77	5,04
Ecart-type	1,70	2,23
Variance	2,88	4,98
CV %	0,25	0,44

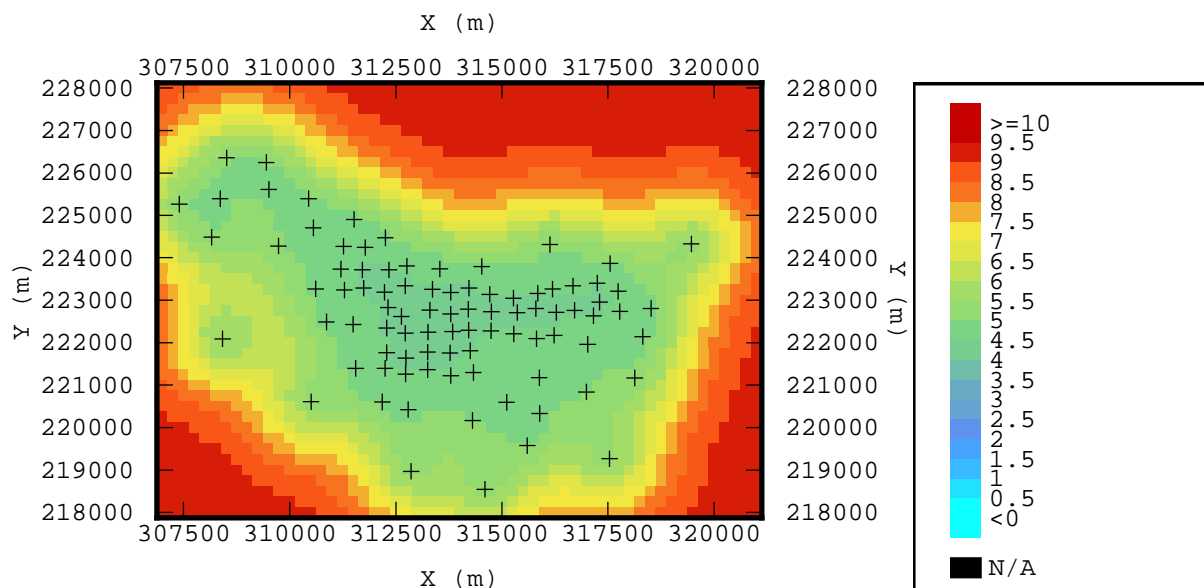


**Figure 48 – Estimation par krigeage ponctuel, à partir de toutes les données, avec le modèle 2, sur une grille régulière de 250 m de côté.**

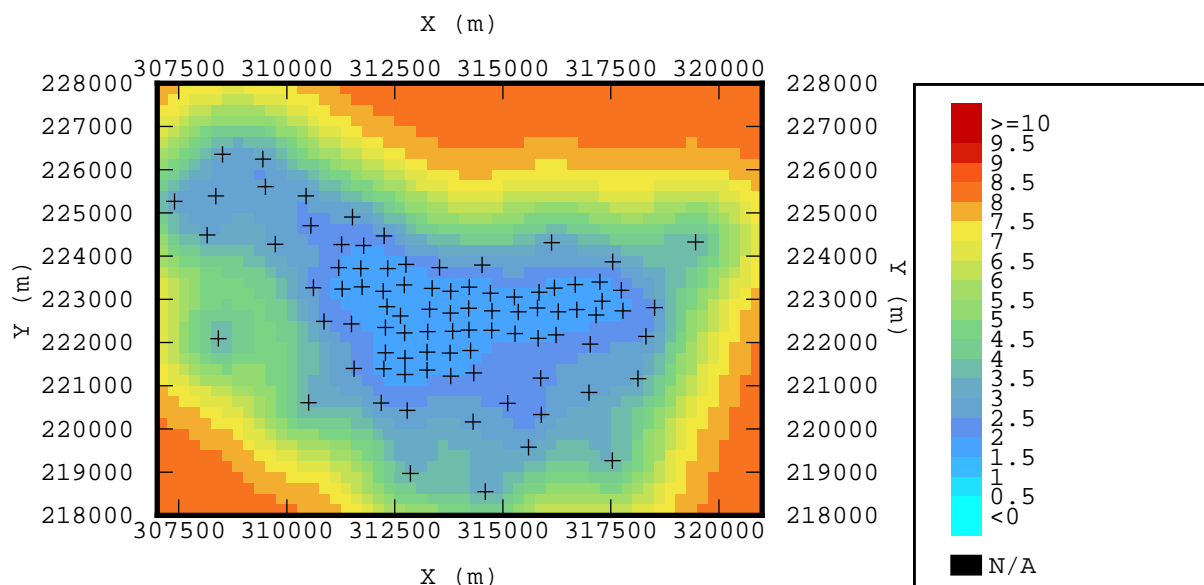


**Figure 49 – Estimation par krigeage de bloc, à partir de toutes les données, avec le modèle 2, sur une grille régulière de 250 m de côté.**

Les cartes des valeurs estimées (Fig. 48 et 49) montrent que la différence entre krigeage ponctuel et krigeage de bloc est, dans ce cas, peu perceptible : les niveaux de concentrations obtenus sont les mêmes et les contours des zones de pollution sont semblables.



**Figure 50 – Ecart-type de krigeage, par krigeage ponctuel, à partir de toutes les données, avec le modèle 2, sur une grille régulière de 250 m de côté.**



**Figure 51 – Ecart-type de krigeage, par krigeage de bloc, à partir de toutes les données, avec le modèle 2, sur une grille régulière de 250 m de côté.**

La différence entre krigeage ponctuel et krigeage de bloc est bien plus marquée lorsque l'on examine les cartes d'écart-type de krigeage (Fig. 50 et 51) : les niveaux des écarts-type de krigeage sont bien plus faibles dans le krigeage de bloc ( $5,0 \mu\text{g}/\text{m}^3$  en moyenne) que dans le krigeage ponctuel ( $6,8 \mu\text{g}/\text{m}^3$  en moyenne). Ceci reflète le fait que la quantité à estimer est moins variable, donc l'estimation plus précise.

### III. ETUDE DE LA SELECTION DE DONNEES N°1 (MAILLES DE 1 KM)

La sélection n°1 comprend 43 données, qui correspondent à 1 point de mesure par maille carrée de 1km<sup>2</sup>.

#### III.1. Etude du variogramme expérimental

Le variogramme expérimental retenu est omnidirectionnel, avec 13 pas de 368 m chacun.

Aucune donnée n'a été écartée.

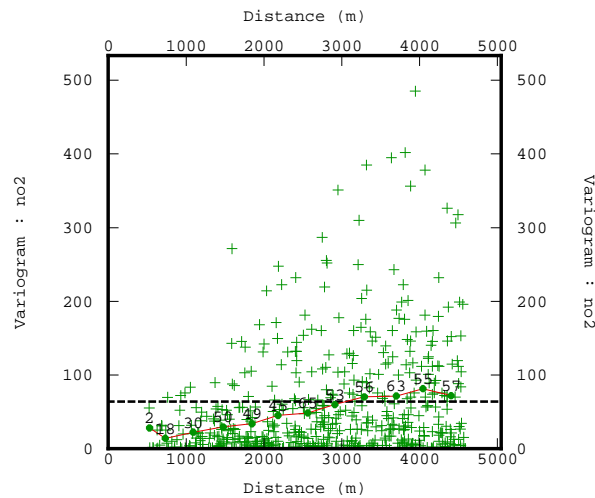
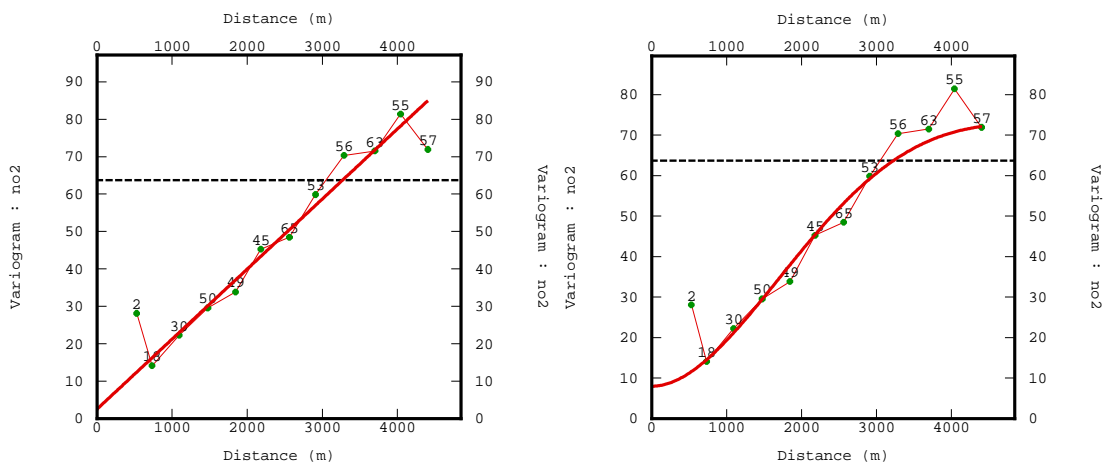


Figure 52 – Variogramme expérimental (pas de 368 m, 13 pas) des données de la sélection 1.

#### III.2. Ajustement du variogramme modélisé et test de validation croisée

Deux modèles différents ont été ajustés aux données expérimentales. Ils sont composés tous deux d'une structure pépitique additionnée d'une structure linéaire ou cubique.



Modèle 1 = pépité(2,5) + linéaire (2270 m ; 42,5)

Modèle 2 = pépité(8) + cubique (5600 m ; 65)

Figure 53 – Variogrammes modélisés pour la sélection 1.

**Tableau 15 – Statistiques sur l'erreur d'estimation obtenues par validation croisée (a) avec le modèle 1.**

<i>MODELE 1</i>	moyenne	variance
Erreur	0,11450	17,70800
Erreur standardisée	0,00910	0,88609

0 valeur non robuste sur 43 données.

**(b) avec le modèle 2.**

<i>MODELE 2</i>	moyenne	variance
Erreur	0,22715	17,73639
Erreur standardisée	0,02043	1,15679

1 valeur non robuste sur 43 données.

**Tableau 16 – Résultats de la validation croisée sur toutes les données pour le krigeage ordinaire ponctuel avec le modèle 1= pépite (2,5) + linéaire (2270 m ; 42,5).**

		Validation croisée par krigeage ordinaire ponctuel selon le modèle 1	
<i>SELECTION 1 (1 KM)</i>	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation standardisée
Nombre de données	43	43	43
Minimum	6,85	7,13	-2,31
25 <sup>ème</sup> percentile	10,50	11,82	-0,47
Médiane	15,75	17,79	0,02
75 <sup>ème</sup> percentile	24,68	24,26	0,67
Maximum	38,00	30,81	2,15
Etendue (= Max – Min)	31,15	23,68	4,46
Moyenne	17,86	17,97	0,01
Ecart-type	7,98	7,16	0,94
Variance	63,70	51,29	0,89
CV %	0,45	0,40	
Corrélation mesures/estimations		r=0,8510	

**Tableau 17 – Résultats de la validation croisée sur toutes les données pour le krigeage ordinaire ponctuel avec le modèle 2 = pépite (8) + Cubique (5600 m ; 65).**

		Validation croisée par krigeage ordinaire ponctuel selon le modèle 2	
<i>Sélection 1 (1 km)</i>	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation standardisée
Nombre de données	43	43	43
Minimum	6,85	7,01	-2,47
25 <sup>ème</sup> percentile	10,50	12,28	-0,61
Médiane	15,75	17,66	0,22
75 <sup>ème</sup> percentile	24,68	23,51	0,60
Maximum	38,00	31,12	2,50
Etendue (= Max – Min)	31,15	24,11	4,97
Moyenne	17,86	18,09	0,02
Ecart-type	7,98	6,95	1,08
Variance	63,70	48,25	1,16
CV %	0,45	0,38	
Corrélation mesures/estimations		r=0,8497	



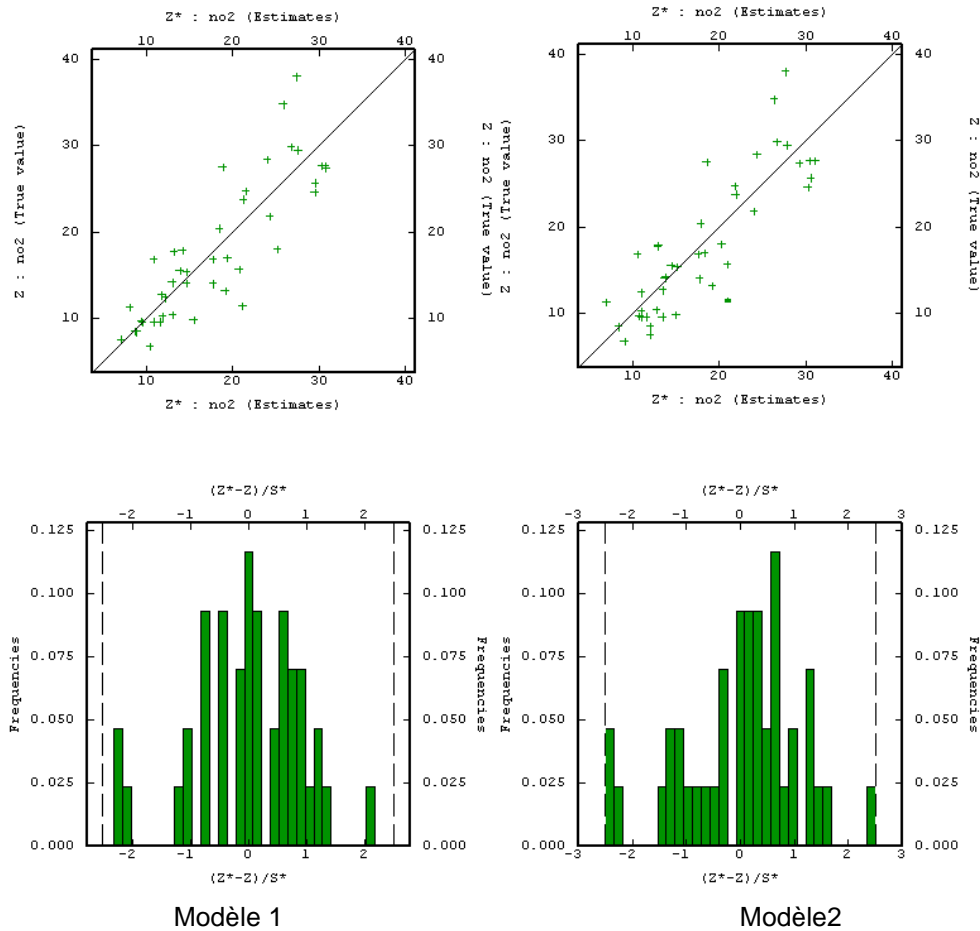


Figure 54 – Résultats de validation croisée pour la sélection 1.

Les résultats de validation croisée montrent que les deux modèles conviennent. Le modèle 1 conduit à des estimations un peu mieux corrélées aux valeurs vraies et à des erreurs d'estimation d'étendue plus petite.

### III.3. Test de validation absolue menée sur les données complémentaires de la sélection 1

Le test de validation absolue consiste à effectuer des estimations en des points où les valeurs vraies sont connues, mais sans que ces valeurs aient été préalablement utilisées dans la procédure d'analyse variographique. Il diffère donc du test de validation croisée, où les mêmes valeurs servent à la fois pour l'étude variographique et pour la validation.

Le test de validation absolue est donc ici effectué en estimant les valeurs de concentration, aux points de mesure du jeu de données complémentaires de la sélection n°1 (cf. I.2.). Ces données complémentaires n'ont, en effet, pas été utilisées pour l'étude variographique.

Les informations fournies pour effectuer la validation absolue sont donc :

- le modèle de variogramme,
- les mesures aux points de la sélection 1,
- les coordonnées (X, Y) des points complémentaires de la sélection 1.

Les estimations sont ensuite comparées aux mesures, en calculant notamment l'erreur d'estimation.

**Tableau 18 – Résultats de la validation absolue pour le krigeage ordinaire ponctuel avec le modèle 1= pépité(2,5) + linéaire (2270 m ; 42,5).**

		Validation absolue par krigeage ordinaire ponctuel selon le modèle 1	
<i>DONNEES COMPLEMENTAIRES DE LA SELECTION 1</i>	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	45	45	45
Minimum	9,60	11,06	-12,0028
25 <sup>ème</sup> percentile	13,80	14,12	-3,9500
Médiane	21,05	20,63	-0,0328
75 <sup>ème</sup> percentile	27,40	27,16	2,1734
Maximum	40,80	31,31	12,2417
Moyenne	21,67	21,00	-0,6710
Ecart-type	8,20	6,55	5,2778
Variance	67,24	42,83	27,8547
CV %	0,39	0,31	
Corrélation mesures/estimations		r=0,7658	

**Tableau 19 – Résultats de la validation absolue pour le krigeage ordinaire ponctuel avec le modèle 2 = pépité(8) + cubique (5600 m ; 65).**

		Validation absolue par krigeage ordinaire ponctuel selon le modèle 2	
<i>DONNEES COMPLEMENTAIRES DE LA SELECTION 1</i>	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	45	45	45
Minimum	9,60	10,80	-11,5603
25 <sup>ème</sup> percentile	13,80	14,09	-3,9272
Médiane	21,05	20,70	-0,1733
75 <sup>ème</sup> percentile	27,40	26,22	2,5367
Maximum	40,80	31,09	11,6167
Moyenne	21,67	21,04	-0,6265
Ecart-type	8,20	6,45	5,2406
Variance	67,24	41,59	27,4642
CV %	0,39	0,31	
Corrélation mesures/estimations		r=0,7691	

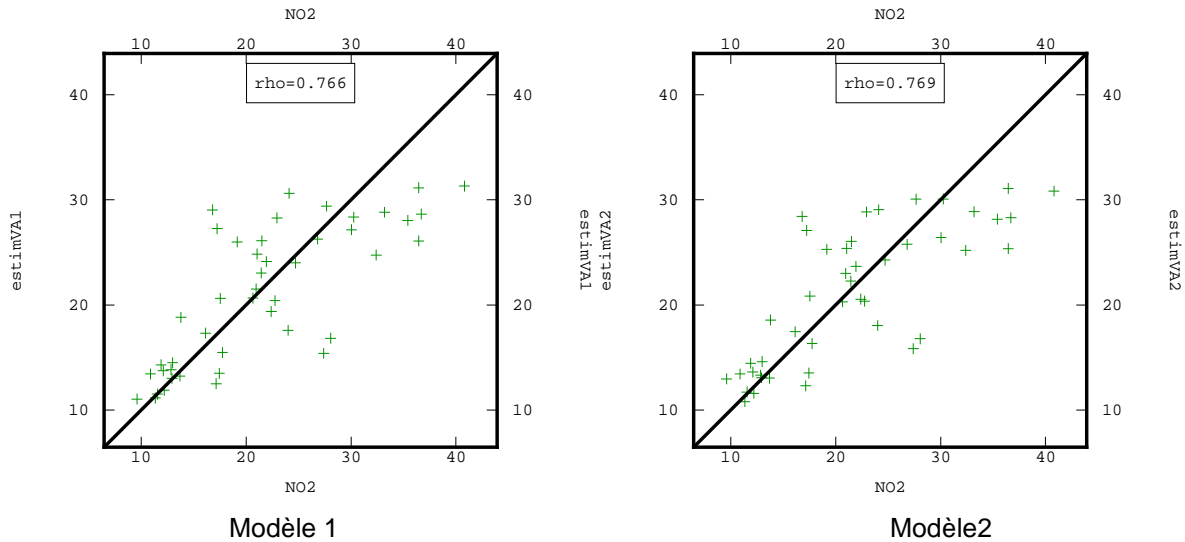


Figure 55 – Test de validation absolue : valeurs estimées en fonction des valeurs vraies.

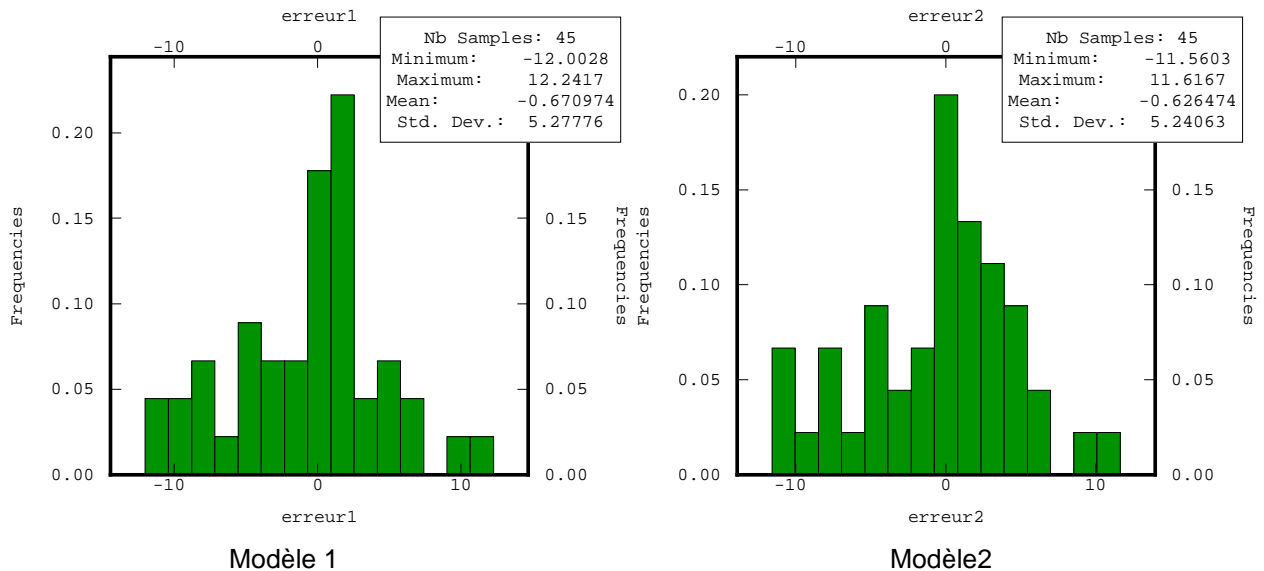


Figure 56 – Test de validation absolue : distribution des erreurs d'estimation.

### III.4.Choix du meilleur modèle et établissement de la carte de la sélection 1

Contrairement à ce que pouvait laisser penser le test de validation croisée, la validation absolue montre que c'est plutôt le modèle 2 (à structures pépétique et cubique) qui estime le mieux les données. En effet, les valeurs estimées par le modèle 2 sont mieux corrélées (Fig. 55) et la distribution des erreurs est plus étroite et mieux centrée autour de zéro (Fig. 56).

Pour dresser la carte de la sélection1, nous estimons les valeurs de concentration en polluant par krigeage, soit ponctuel soit de bloc, en utilisant ce modèle et les mêmes grilles que pour les cartes de référence (cf. II.3.).

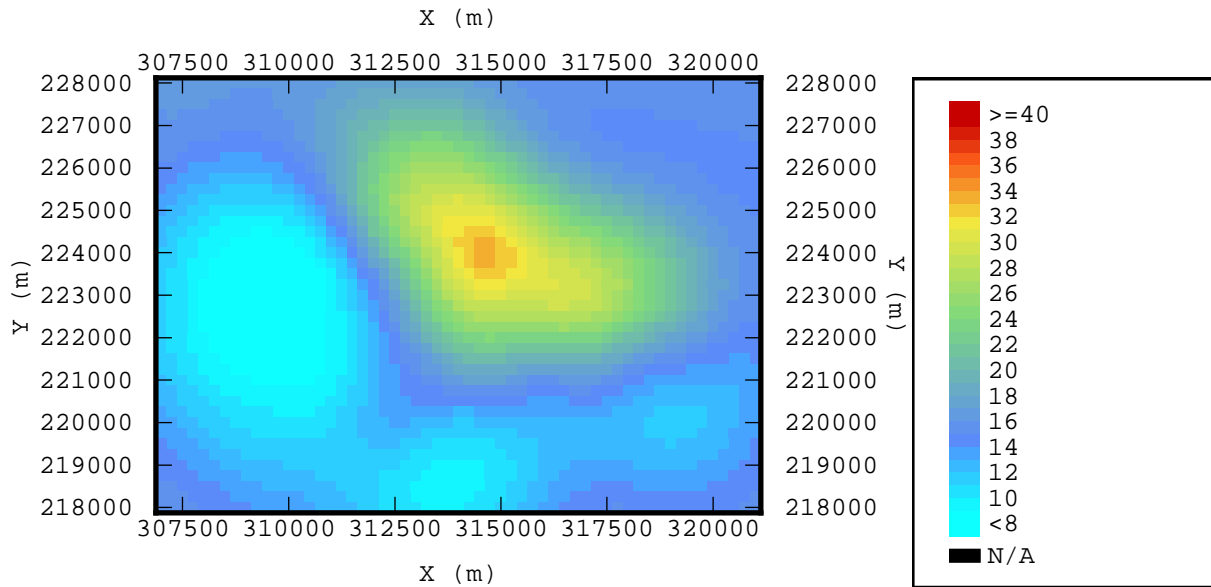
**Tableau 20 – Estimation sur une grille régulière de 250 m de côté, par krigeage ordinaire ponctuel ou de bloc, selon le modèle 2, à partir des données de la sélection 1.**

<i>SELECTION 1</i>	Mesures ( $\mu\text{g}/\text{m}^3$ )	Estimation par krigeage ordinaire	
		Valeurs estimées modèle 2, krigeage <b>ponctuel</b> ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées modèle 2, krigeage de <b>bloc</b> ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	43	2337	2240
Minimum	6,85	6,90	6,92
25 <sup>ème</sup> percentile	10,50	12,05	11,97
Médiane	15,75	14,92	14,90
75 <sup>ème</sup> percentile	24,68	18,66	19,03
Maximum	38,00	33,40	33,38
Moyenne	17,86	16,10	16,17
Ecart-type	7,98	5,83	5,93
Variance	63,70	34,02	35,17
CV %	0,45	0,36	0,37

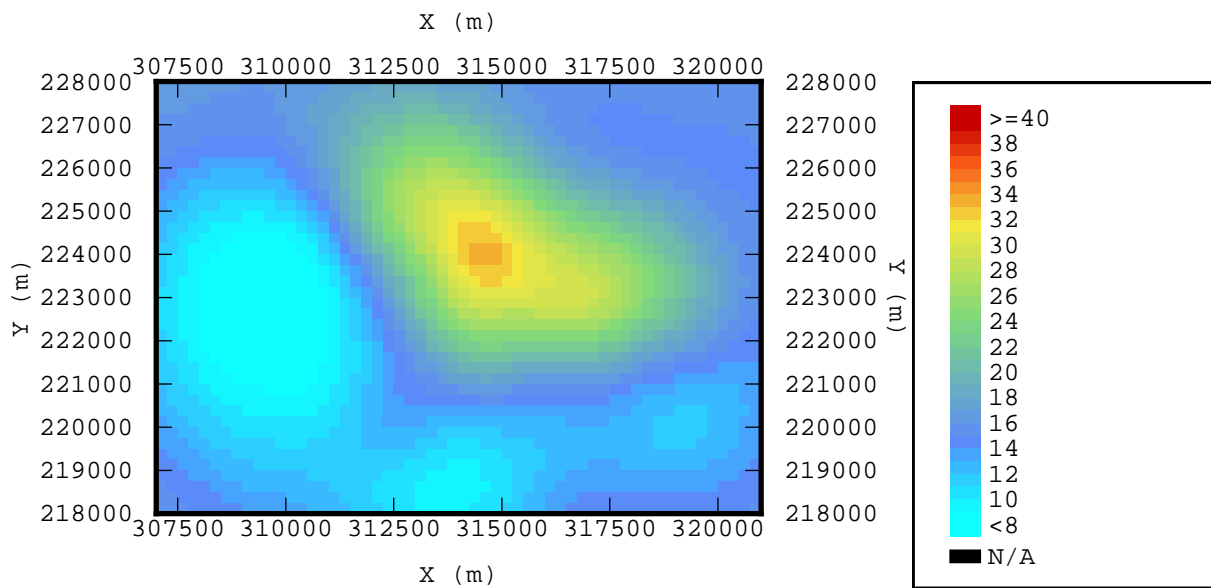
**Tableau 21 – Ecart-types de krigeage, sur une grille régulière de 250 m de côté, par krigeage ordinaire ponctuel ou de bloc, selon le modèle 2, pour les données de la sélection 1.**

<i>SELECTION 1</i>	Estimation par krigeage ordinaire	
	Ecart-types de krigeage modèle 2, krigeage <b>ponctuel</b> ( $\mu\text{g}/\text{m}^3$ )	Ecart-types de krigeage modèle 2, krigeage de <b>bloc</b> ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	2337	2240
Minimum	3,19	1,48
25 <sup>ème</sup> percentile	3,73	2,36
Médiane	4,95	3,85
75 <sup>ème</sup> percentile	7,26	6,45
Maximum	9,04	8,56
Moyenne	5,52	4,44
Ecart-type	1,93	2,26
Variance	3,72	5,10
CV %	0,35	0,51

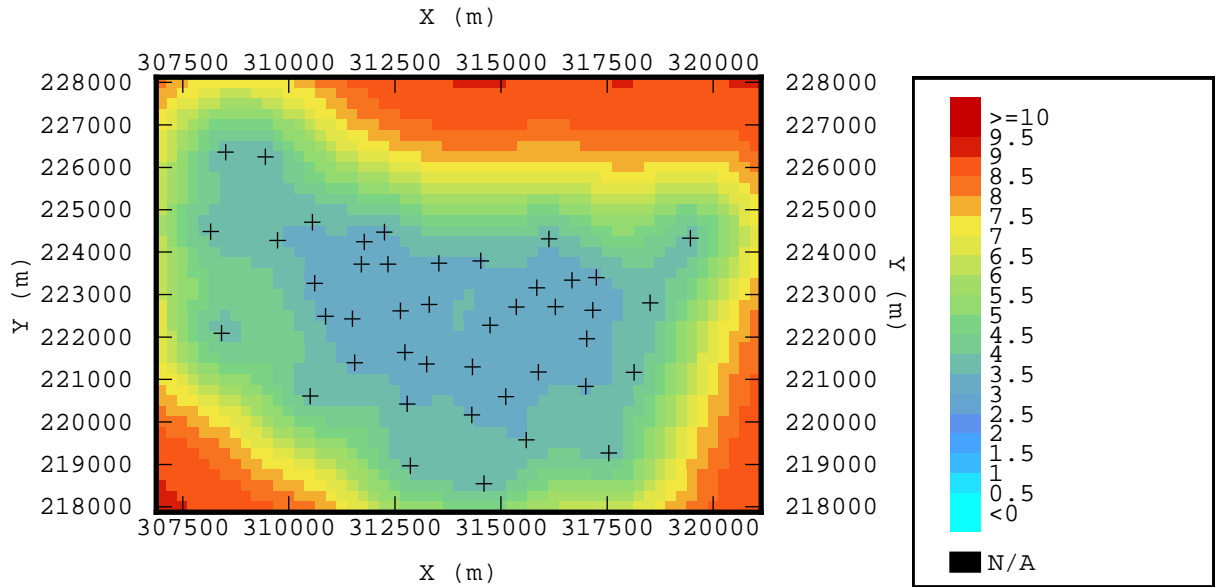
Comme précédemment, la différence entre krigeage ponctuel et krigeage de bloc est peu perceptible en ce qui concerne les valeurs estimées (Tab. 20 et Fig. 57 et 58), mais bien plus nette en ce qui concerne les écarts-type de krigeage (Tab. 21 et Fig. 59 et 60) : le krigeage de bloc conduit à des écarts-type de krigeage plus faibles et constitue donc la meilleure technique des deux.



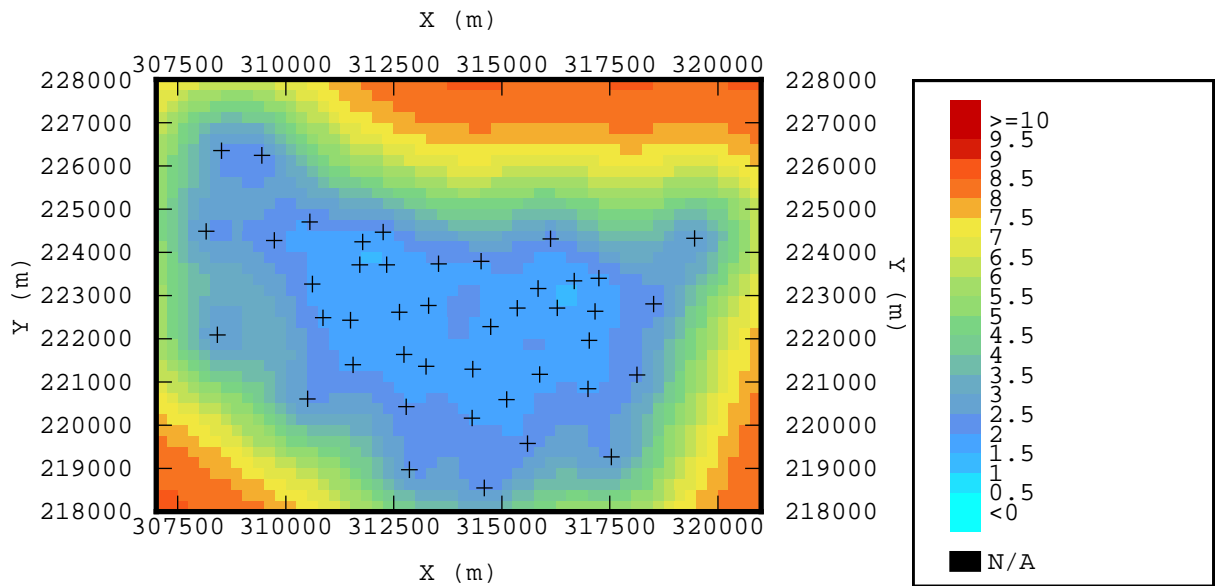
**Figure 57 – Estimation par krigeage ponctuel, à partir des données de la sélection 1, avec le modèle 2, sur une grille régulière de 250 m de côté**



**Figure 58 – Estimation par krigeage de bloc, à partir des données de la sélection 1, avec le modèle 2, sur une grille régulière de 250 m de côté**



**Figure 59 – Ecarts-type de krigeage ponctuel, à partir des données de la sélection 1, avec le modèle 2, sur une grille régulière de 250 m de côté**



**Figure 60 – Ecarts-type de krigeage de bloc, à partir des données de la sélection 1, avec le modèle 2, sur une grille régulière de 250 m de côté**

## IV. ETUDE DE LA SELECTION DE DONNEES N°2 (MAILLES DE 2 KM)

Sélection des données n°2 : 1 donnée par maille carrée de 2km×2km. Cette sélection comprend 21 données.

### IV.1. Etude du variogramme expérimental

Le variogramme expérimental retenu est omnidirectionnel, avec 9 pas de 800 m chacun.

Aucune donnée n'a été écartée.

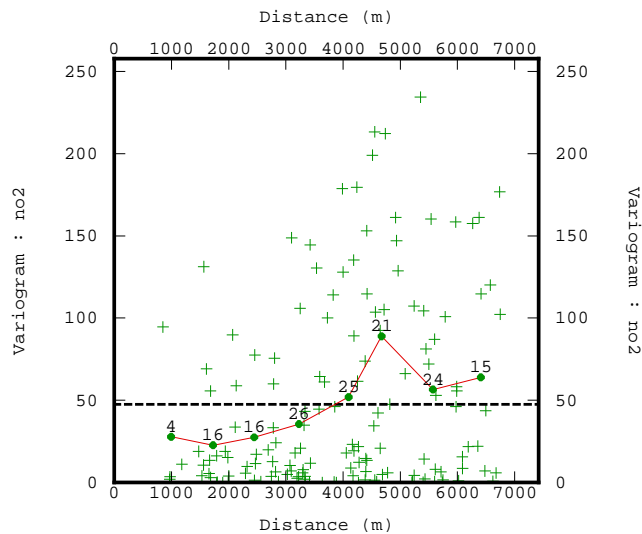


Figure 61 – Variogramme expérimental (pas de 800 m, 9 pas) des données de la sélection 2.

### IV.2. Ajustement du variogramme modélisé et test de validation croisée

Le modèle de variogramme qui semble le mieux convenir pour la sélection 2 est :  
Modèle = pépite (12) + cubique (10000 m ; 55)

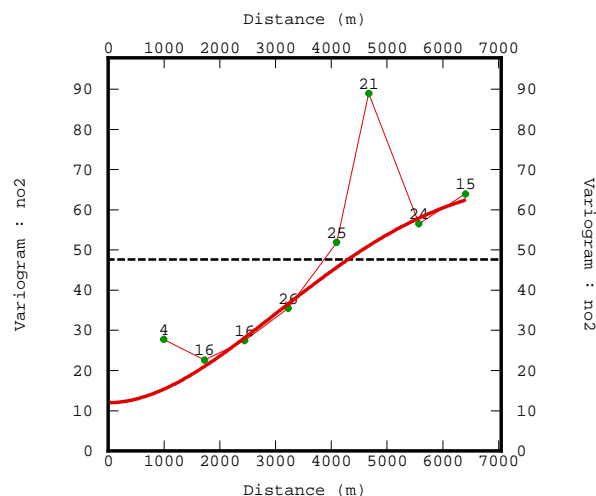
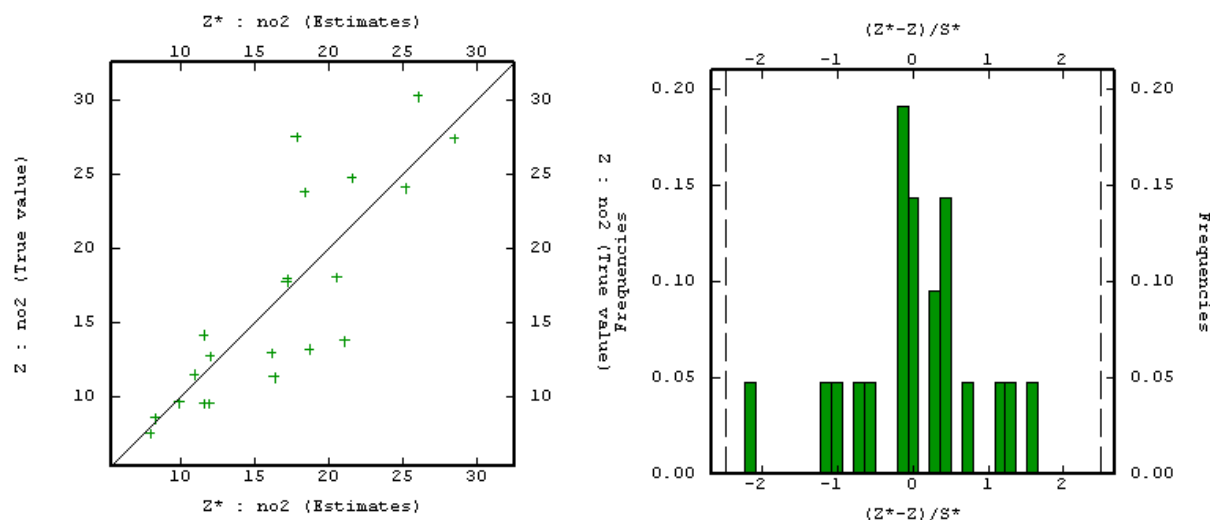


Figure 62 – Variogramme modélisé pour la sélection 2.

**Tableau 22 – Statistiques sur l'erreur d'estimation obtenues par validation croisée pour la sélection 2.**

	moyenne	variance
Erreur	0,13388	14,08863
Erreur standardisée	0,01407	0,71423

0 valeur non robuste sur 21 données.



**Fig. 63 – Test de validation croisée pour la sélection 2.**

### IV.3. Test de validation absolue mené sur les données complémentaires de la sélection 2

Comme précédemment, le test de validation absolue est effectué par rapport aux données complémentaires de la sélection n°2, non utilisées dans l'analyse variographique.

**Tableau 23 – Résultats de la validation absolue pour le krigeage ordinaire ponctuel avec le modèle 2 = pépite(12) + cubique (10000 m ; 55).**

Données complémentaires de la sélection 2	Mesures ( $\mu\text{g}/\text{m}^3$ )	Validation absolue par krigeage ordinaire ponctuel	
		Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	66	66	66
Minimum	6,85	9,65	-13,84
25 <sup>ème</sup> percentile	14,05	16,22	-4,35
Médiane	20,55	20,33	0,15
75 <sup>ème</sup> percentile	27,53	24,42	2,52
Maximum	40,80	28,33	8,54
Moyenne	20,97	20,03	-0,88
Ecart-type	8,45	5,09	5,08
Variance	71,37	25,93	25,83
CV %	0,40	0,25	
Corrélation mesures/estimations		r=0,8300	



Le test de validation absolue montre que le modèle choisi permet de bien estimer les données, puisque le coefficient de corrélation entre estimations et valeurs vraies est de 0,83.

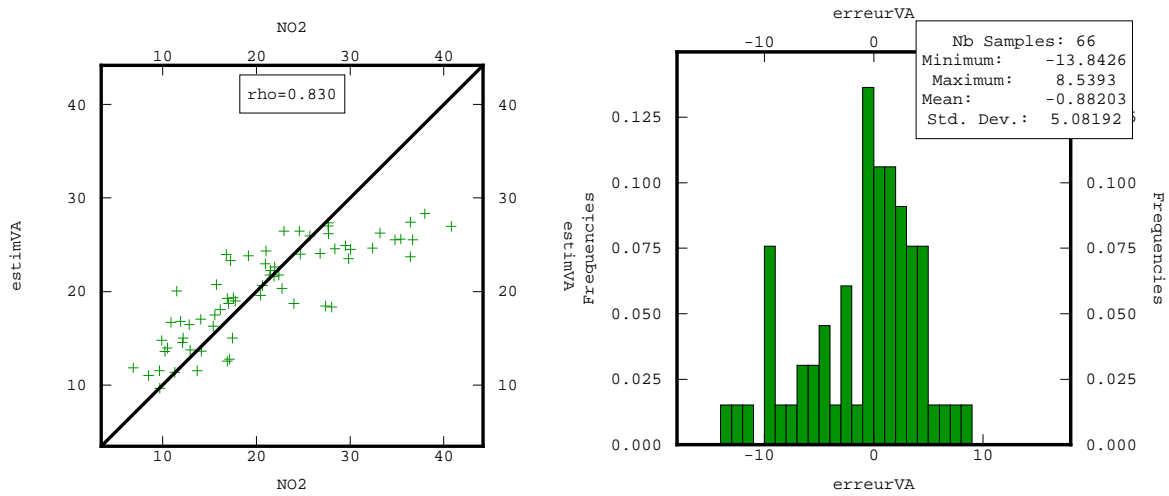


Figure 64 – Résultats de la validation absolue : estimations en fonction des valeurs vraies (à gauche) et distribution des erreurs d'estimation (à droite).

#### IV.4. Etablissement de la carte pour la sélection 2

La carte de répartition des concentrations en NO<sub>2</sub> est établie à partir du modèle testé précédemment et de la sélection 2 (21 données), sur les mêmes grilles de 250 m que pour les cartes de référence, soit par krigeage ponctuel, soit par krigeage de bloc.

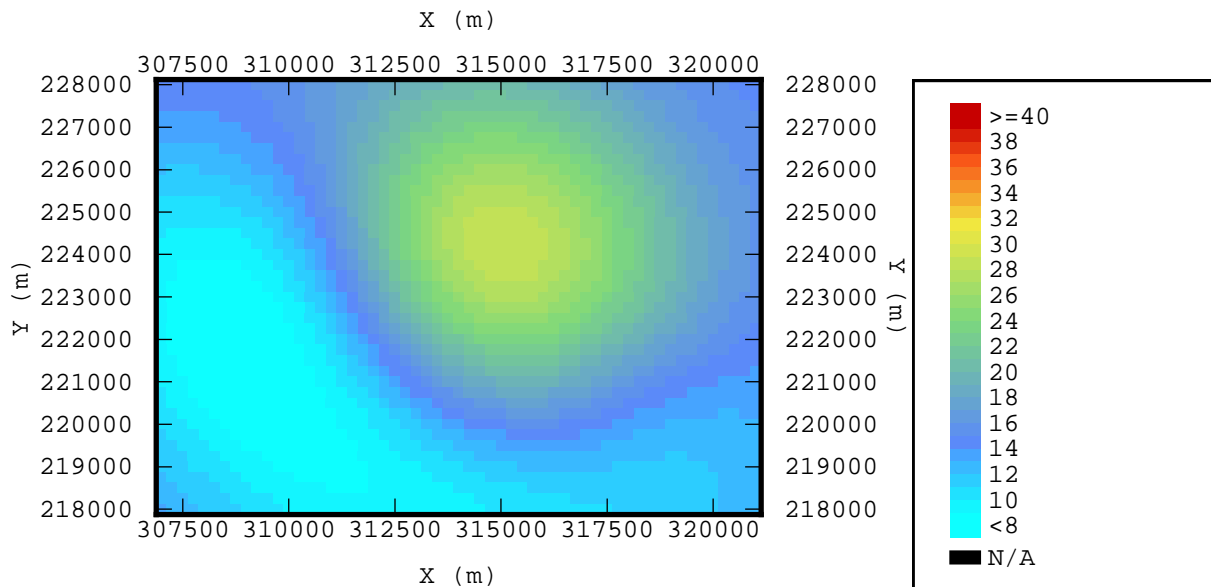
Tableau 24 – Estimation sur une grille régulière de 250 m de côté, par krigeage ordinaire ponctuel ou de bloc, à partir des données de la sélection 2.

Sélection 2	Mesures (µg/m <sup>3</sup> )	Estimation par krigeage ordinaire	
		Valeurs estimées krigeage <b>ponctuel</b> (µg/m <sup>3</sup> )	Valeurs estimées krigeage de <b>bloc</b> (µg/m <sup>3</sup> )
Nombre de données	21	2337	2240
Minimum	7,60	7,65	7,65
25 <sup>ème</sup> percentile	11,35	11,22	11,27
Médiane	13,8	15,19	15,37
75 <sup>ème</sup> percentile	23,80	20,36	20,64
Maximum	30,25	28,90	28,89
Moyenne	16,51	16,04	16,16
Ecart-type	6,90	5,76	5,82
Variance	47,60	33,18	33,84
CV %	0,42	0,36	0,36

**Tableau 25 – Ecart-types de krigeage, sur une grille régulière de 250 m de côté, par krigeage ordinaire ponctuel ou de bloc, pour les données de la sélection 2.**

<i>SELECTION 2</i>	Estimation par krigeage ordinaire	
	Ecart-types de krigeage, ponctuel ( $\mu\text{g}/\text{m}^3$ )	Ecart-types de krigeage de bloc ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	2337	2240
Minimum	3,95	1,90
Médiane	4,63	2,98
Maximum	8,13	7,21
Moyenne	5,08	3,50
Ecart-type	1,08	1,40
Variance	1,17	1,95
CV %	0,21	0,40

Le krigeage de bloc conduit à des valeurs estimées plus proches des valeurs vraies (Tab. 24 et Fig. 66), et à des écarts-type de krigeage plus faibles (Tab. 25 et Fig. 68) et constitue donc, comme précédemment, la meilleure technique des deux.



**Figure 65 – Carte des estimations par krigeage ponctuel, à partir de la sélection 2.**

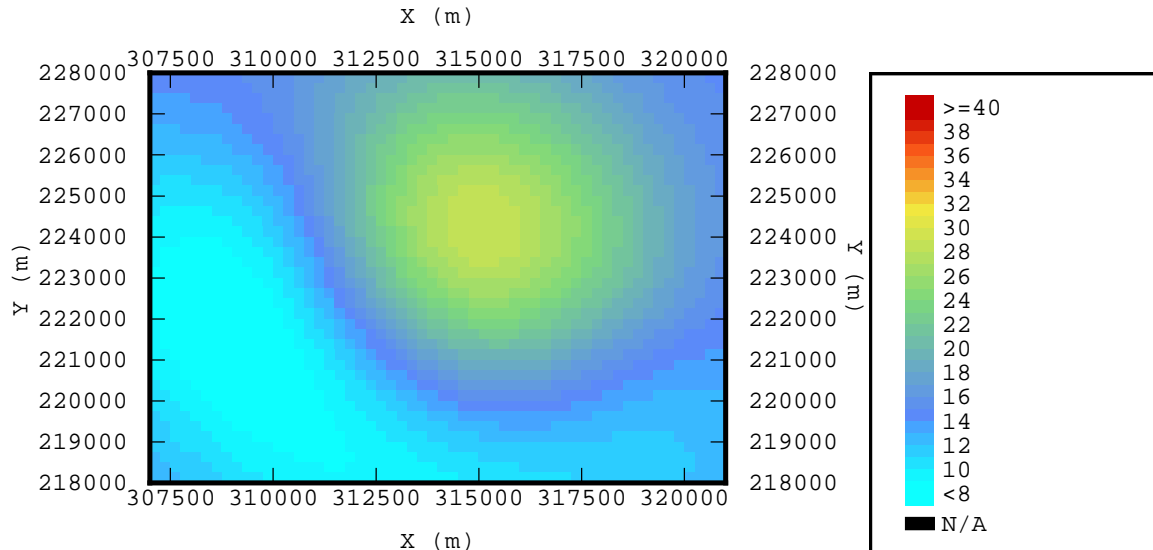


Figure 66 – Carte des estimations par krigeage de bloc, à partir de la sélection 2.

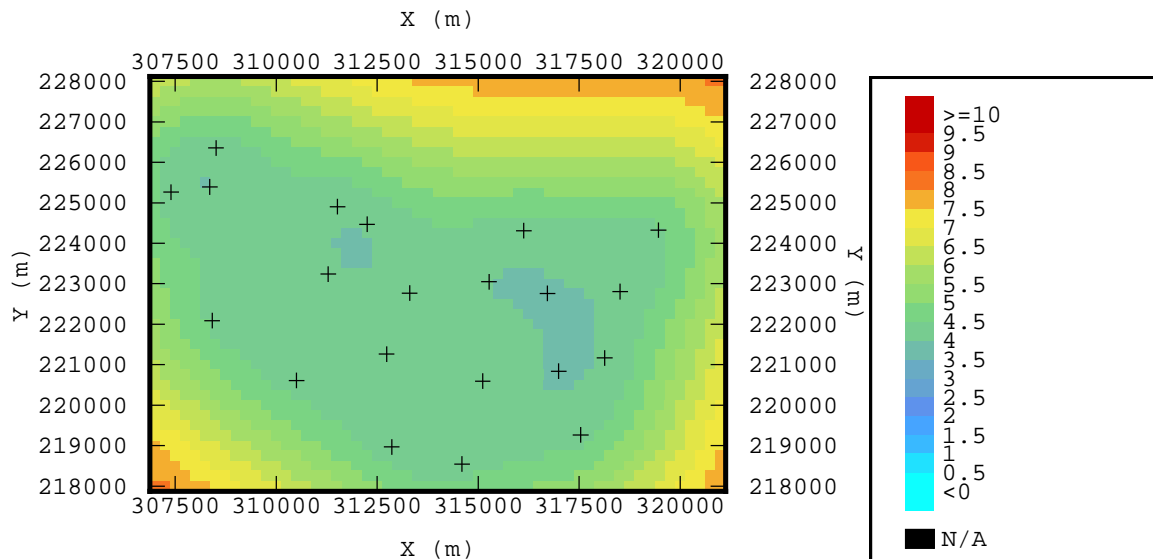


Figure 67 – Carte des écarts-type par krigeage de ponctuel, à partir de la sélection 2.

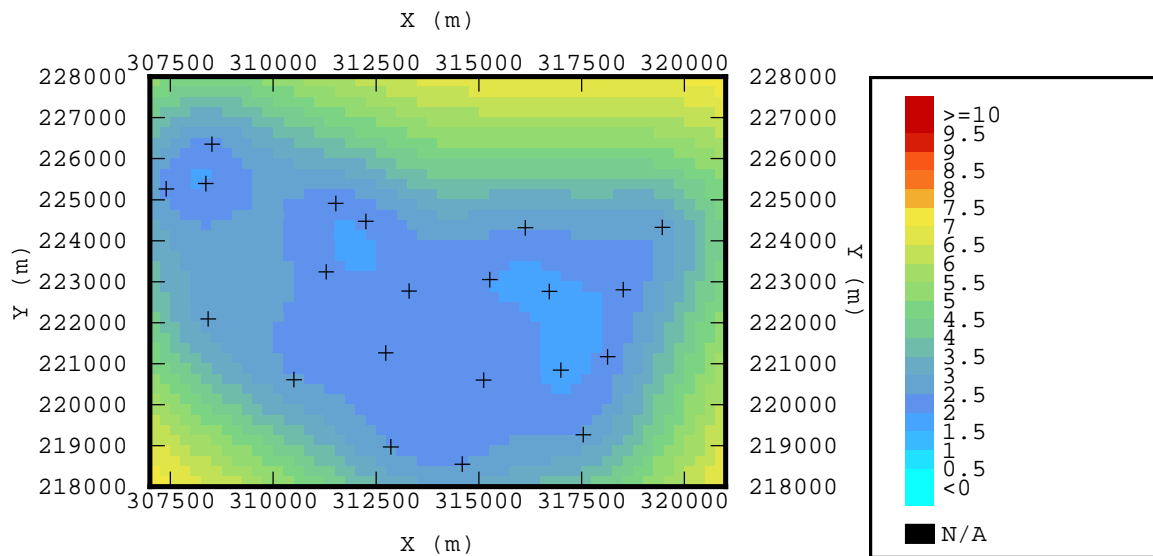


Figure 68 – Carte des écarts-type par krigeage de bloc, à partir de la sélection 2.

## V COMPARAISON DES CARTES ET DISCUSSION

A partir des différentes sélections un tableau de synthèse des résultats de validation peut être établi :

Distribution des valeurs		Toutes les données	Sélection 1	Sélection 2
nombre de points		88	43	21
Min ( $\mu\text{g}/\text{m}^3$ )		6,85	6,85	7,60
Max ( $\mu\text{g}/\text{m}^3$ )		40,80	38,00	30,25
Moyenne ( $\mu\text{g}/\text{m}^3$ )		19,81	17,86	16,51
ecart-type ( $\mu\text{g}/\text{m}^3$ )		8,31	7,98	6,90
Variance ( $\mu\text{g}/\text{m}^3$ )		69,11	63,70	47,60
<b>Validation croisée</b>				
erreur				
	mean	0,14896	0,22715	0,13388
	var	19,79324	17,73639	14,08863
std. Error				
	mean	0,01120	0,02043	0,01407
	var	0,84959	1,15679	0,71423
<b>Validation absolue</b>				
Error				
	mean	-	- 0,6265	- 0,8820
	var	-	27,4642	25,8259

Tableau 26 : synthèse des résultats de validation des différents sélections.

Mise à part la variance de l'erreur d'estimation obtenue par validation croisée (Fig. 69), les autres paramètres statistiques de validation ne présentent pas de variation monotone avec le nombre de points.

La variance de l'erreur d'estimation obtenue par validation croisée (Fig. 69) augmente lorsque le nombre de points augmente, ce qui pourrait s'expliquer par la difficulté des modèles à estimer une variable dont la distribution des valeurs s'élargit.

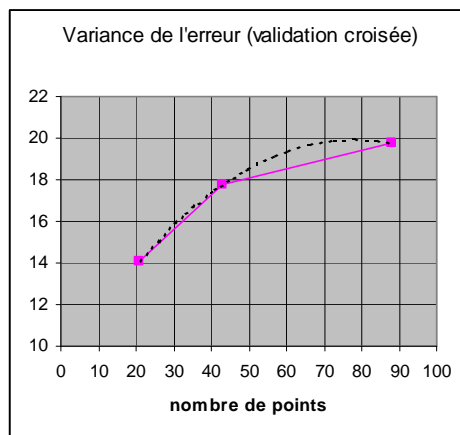
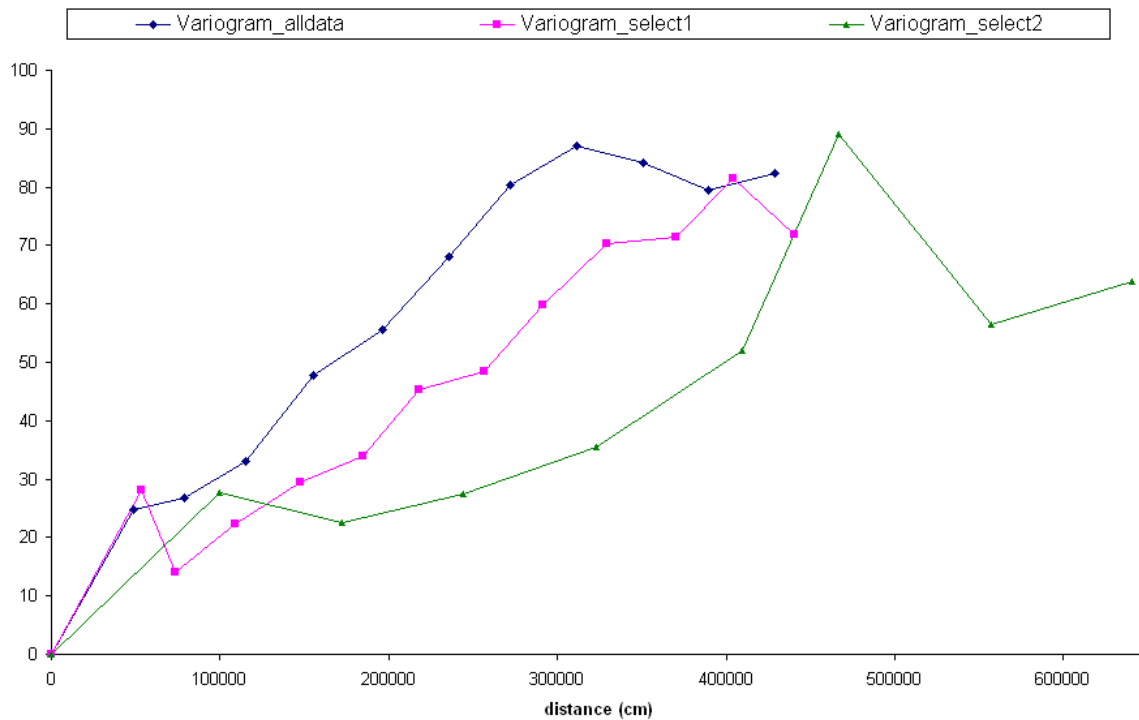


Figure 69 – Evolution de la variance de l'erreur, calculée par validation croisée, en fonction du nombre de points sélectionnés.

Nous avons établi le graphe de synthèse des différents variogrammes expérimentaux (Fig. 70).



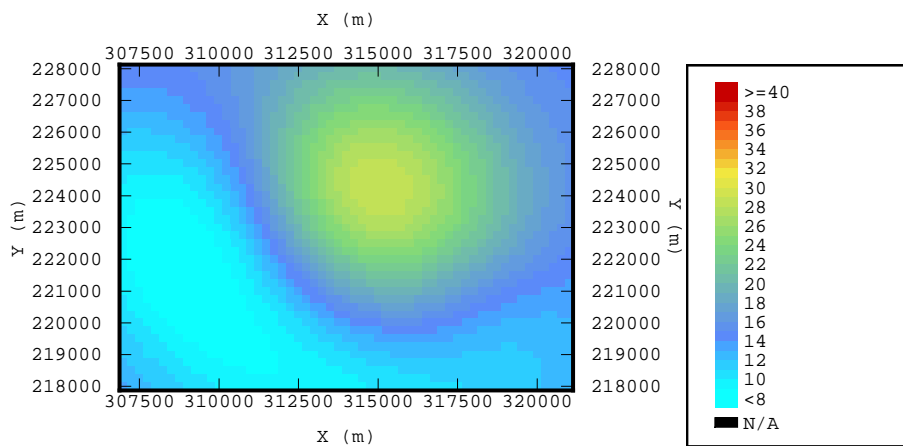
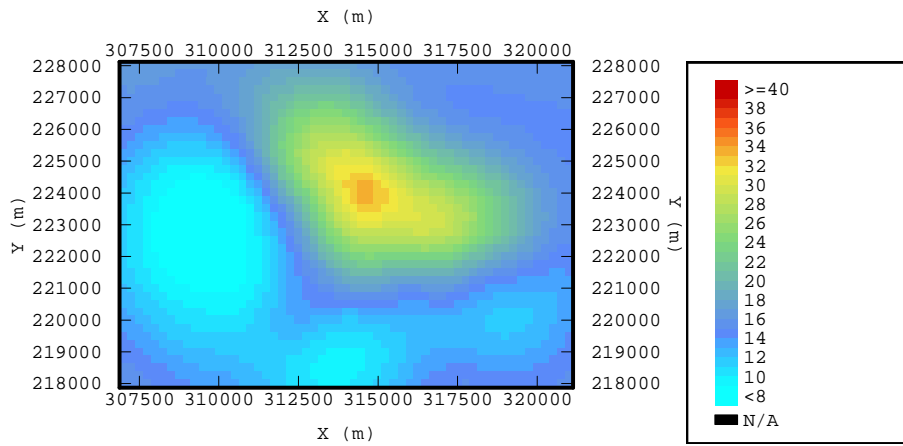
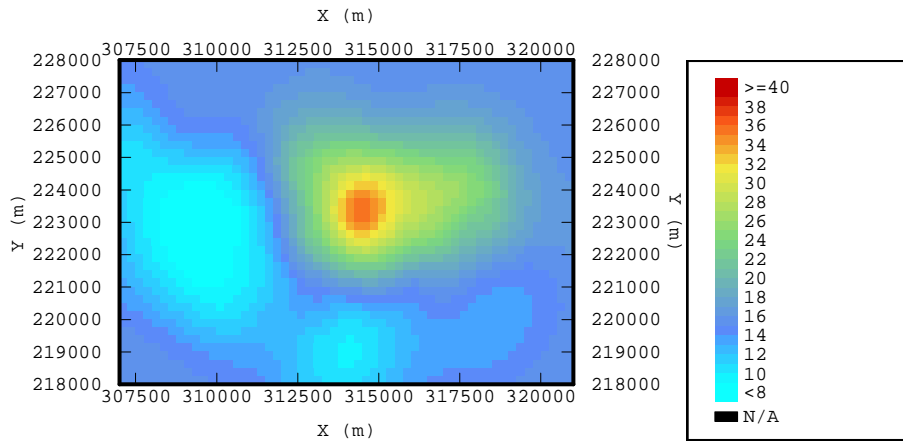
**Figure 70 : Comparaison des différents variogrammes**

L'analyse de ce graphique montre un net changement de structure du variogramme avec la diminution du nombre de points ; elle apparaît clairement à partir de la sélection 2 (21 points). Cependant, dès la sélection 1 (43 points), on constate que la portée du variogramme devient plus importante.

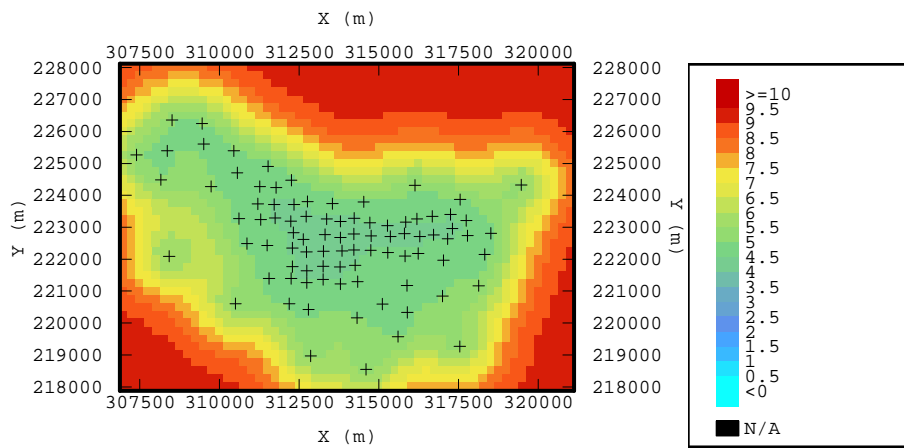
En effet, la diminution homogène (isotrope) du nombre de points, sans privilégier aucune zone particulière, conduit à la diminution des points situés en centre-ville (peu nombreux et localisés) au profit des points situés en périphérie (plus nombreux et bien répartis). Ceci est visible en examinant les paramètres statistiques (moyenne des valeurs ; étendue) qui diminuent avec la diminution du nombre de points.

Or les mesures localisées en périphérie présentent une continuité spatiale (une homogénéité) plus importante que les mesures de centre-ville, ce qui peut se quantifier par la portée spatiale du variogramme qui augmente.

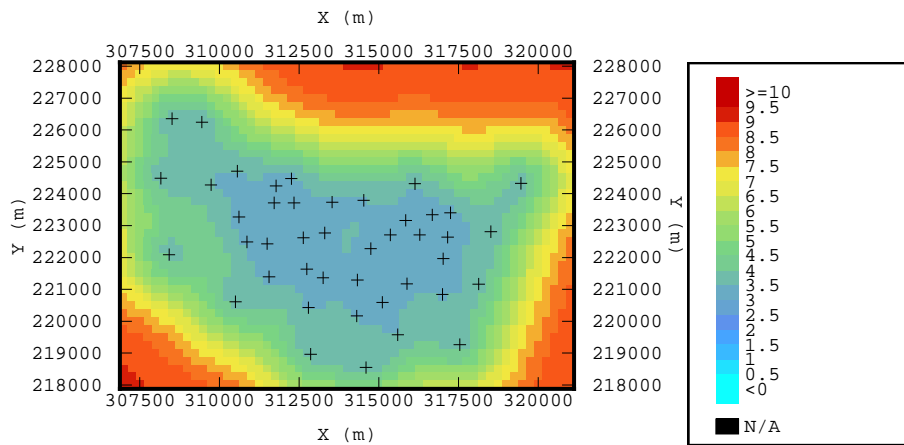
Cela est clairement visible sur les cartes de la figure 71, où l'on voit progressivement disparaître les fortes concentrations (« tâche rouge ») du centre-ville.



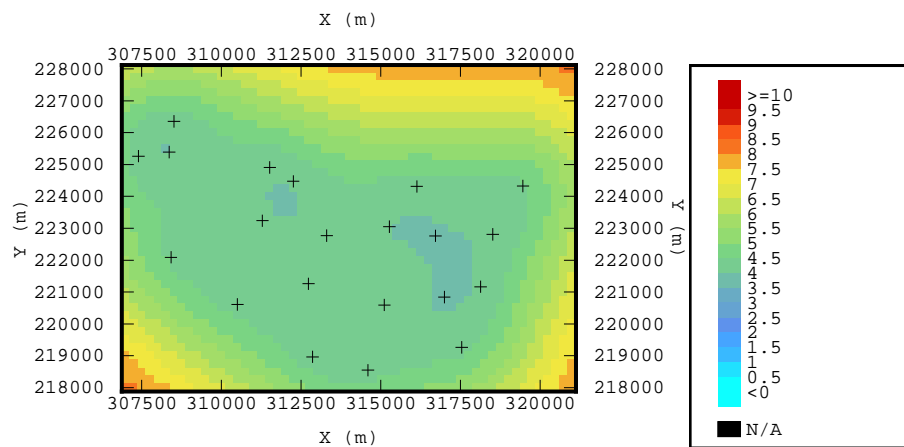
**Figure 71 – Comparaison des cartes d'estimation établies par krigeage ponctuel sur la grille de référence.**



(a) Toutes les données (88 points)



(b) Sélection 1 (43 points)



(c) Sélection 2 (21 points)

**Figure 72 – Comparaison des cartes d'écart-type de krigeage établies par krigeage ponctuel sur la grille de référence.**

La comparaison des cartes d'écart-types de krigeage est, elle aussi, riche d'enseignements. En écartant de notre examen les bords du domaine (peu couverts par les mesures et donc présentant des valeurs d'écart-types très élevées), on peut noter que la carte correspondant à la sélection 1 présente des écart-types de krigeage plus faibles que pour les deux autres cartes.

Ceci est d'autant plus remarquable que les trois modèles de variogrammes utilisés comportent tous les mêmes structures :

Toutes les données	: pépité(17) + cubique( 4400 m, 62)
Sélection 1	: pépité( 8) + cubique( 5600 m, 65)
Sélection 2	: pépité(12) + cubique(10000 m, 55)

En effet, on sait qu'un changement de structure peut induire des variations des écart-type de krigeage.

Ces résultats de krigeage semblent donc accorder une confiance plus importante à la carte établie à partir d'un nombre assez limité de données (43 points) mais bien réparties dans l'espace, plutôt qu'à un nombre plus que double de données (88 points) mais réparties de façon inhomogène dans l'espace (selon un échantillonnage stratifié). Ces résultats, intéressants, méritent d'être confirmés par d'autres études.

Toutefois, lorsque le nombre de données est trop faible (21 points, sélection 2), la confiance accordée aux résultats de krigeage diminue.

Cette étude semble donc montrer qu'un optimum est obtenu pour environ 43 points de mesure pour un domaine de  $125 \text{ km}^2$  ( $12,5 \text{ km} \times 10 \text{ km}$ ), soit 34,4 points pour  $100 \text{ km}^2$  ou 1 point par maille de  $2,90 \text{ km}^2$ .



# Etude d'une ville de moins de 100 000 habitants :

## CAS DE LA ROCHELLE.

### B.I. DESCRIPTION DU CAS

#### B.I.1. Présentation du jeu de données initial

Le jeu de données étudié a été fourni par l'association agréée de surveillance ATMO Poitou-Charentes. Il concerne la mesure du dioxyde d'azote à La Rochelle (76594 et 147000 habitants pour la ville et la communauté d'agglomération, superficie : 29,2 km<sup>2</sup> et 206 km<sup>2</sup> pour la ville et la communauté d'agglomération). Il est constitué de huit campagnes de mesure isolées, d'une durée de 14 jours chacune, conduites par échantillonnage passif au cours de l'année 2004.

65 sites de mesure, dont un site sous influence industrielle, ont été disposés sur un domaine d'environ 17 km d'est en ouest et de 20 km du nord au sud. Leur répartition correspond à la stratégie d'échantillonnage couramment adoptée par les AASQA : un resserrement des points de mesure en centre-ville et une disposition plus espacée en périphérie (Figure 73).

Les valeurs prises en compte pour cette étude correspondent à la moyenne des 8 campagnes. Leurs statistiques sont fournies dans le Tableau 26. Le point de type industriel, qui présente une concentration moyenne très supérieure à celles des autres sites, a été éliminé.

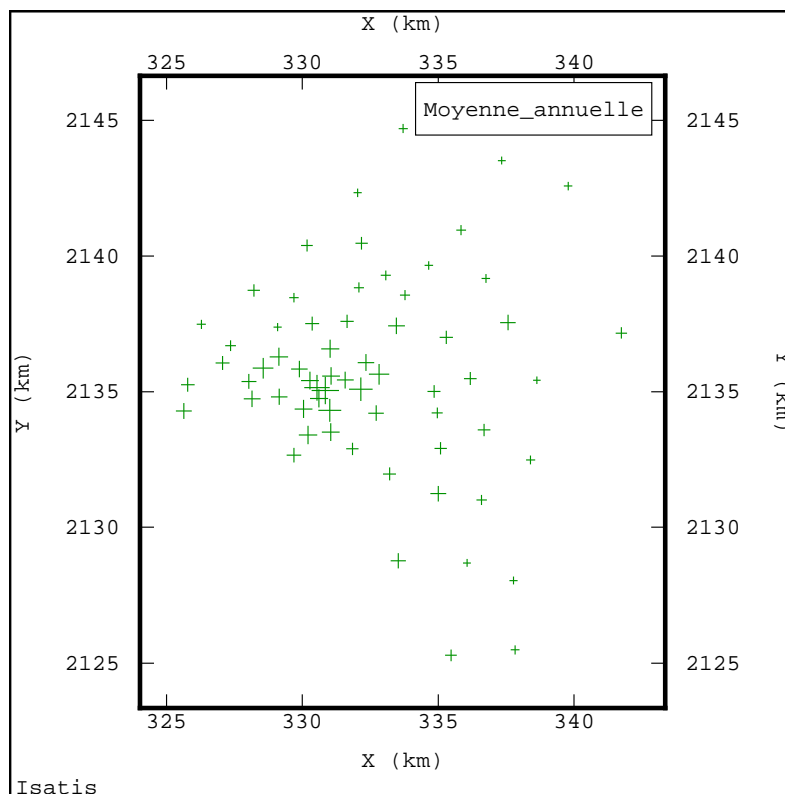


Figure 73 - Carte de répartition des points de mesure

#### B.I.2. Méthodologie de sélection des données

Les données sont sélectionnées selon trois maillages réguliers :

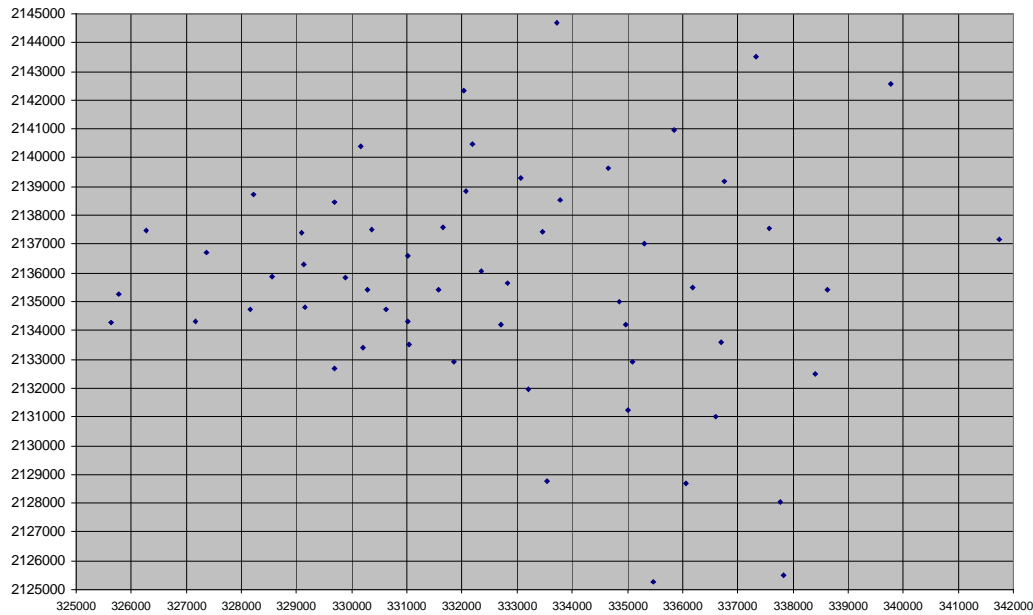
- mailles carrées de 1 km de côté (sélection 1 notée S1) ;

- mailles carrées de 1,5 km de côté (sélection 2 notée S2)
- mailles carrées de 2 km de côté (sélection 3 notée S3).

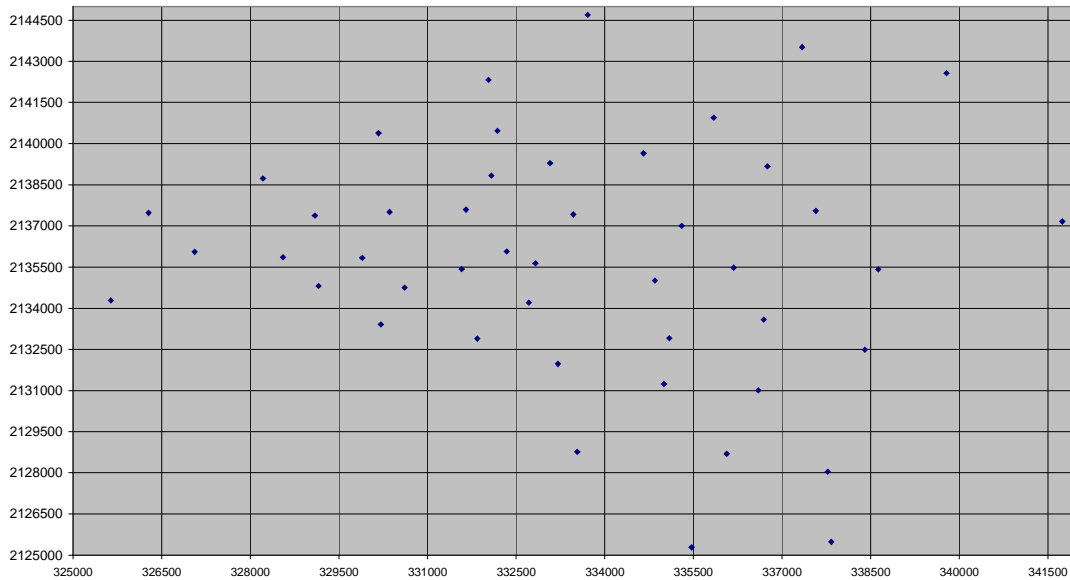
La procédure est la suivante : un maillage régulier de résolution 1km, 1.5km ou 2km calé est calé sur le domaine d'étude puis, pour chaque maille, on sélectionne le point de mesure le plus proche du centre de celle-ci.

Les jeux de données complémentaires de ces trois sélections (notés  $C_{S1}$ ,  $C_{S2}$ ,  $C_{S3}$ ) sont également constitués.

Le Tableau 27 présente les statistiques élémentaires de ces différents jeux.



**Figure 74 – Carte de répartition des points de la sélection 1 (mailles de 1km).**



**Figure 75 – Carte de répartition des points de la sélection 2 (mailles de 1.5km)**

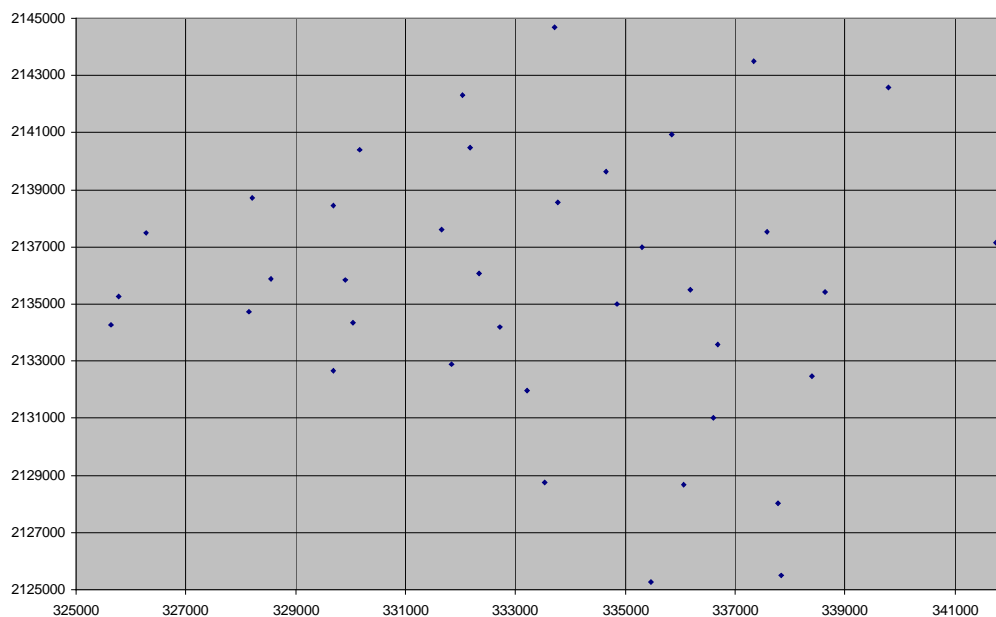


Figure 76 – Carte de répartition des points de la sélection 3 (mailles de 2 km).

Tableau 27 - Principales caractéristiques des jeux de données étudiés.

	Toutes les données	S1	C <sub>S1</sub>	S2	C <sub>S2</sub>	S3	C <sub>S3</sub>
Nombre de données	64	57	7	46	18	38	26
Minimum	7.34	7.34	11.69	7.34	8.58	7.34	7.65
25 <sup>ème</sup> percentile	8.89	8.73	12.59	8.23	11.46	8.23	10.52
médiane	10.96	10.59	14.18	10.57	14.14	10.53	13.41
75 <sup>ème</sup> percentile	13.39	13.20	19.20	12.74	14.45	11.96	14.44
Maximum	20.66	17.70	20.66	16.28	20.66	16.28	20.66
Moyenne	11.50	10.98	15.77	10.64	13.72	10.34	13.20
Ecart-type	3.07	2.59	3.37	2.41	3.46	2.20	3.37
Variance	9.45	6.71	11.33	5.79	11.95	4.85	11.34
CV%	27	24	21	23	25	21	26

## B.II. ETUDE DU JEU DE DONNEES INITIAL ET ELABORATION DE LA CARTE DE REFERENCE

Le but de cette partie est d'élaborer la carte de référence, c'est-à-dire la carte de répartition des concentrations de dioxyde d'azote établie avec le maximum de points de mesure. On comparera par la suite à cette carte les cartes obtenues à partir d'un nombre réduit de points.

### Remarque :

***Dans une première étape, cette étude est réalisée sans prise en compte de variables auxiliaires. Or des données de population et d'émission, susceptibles d'expliquer les concentrations de NO<sub>2</sub> et par suite, d'améliorer la précision des cartographies, sont disponibles. Dans une seconde étape, il conviendra d'examiner si l'utilisation de telles données peut modifier les conclusions sur l'échantillonnage. Nous reviendrons sur ce point dans la discussion.***

### B.II.1. Etude du variogramme expérimental

Deux sites de mesure implantés au centre-ville accroissent de façon notable la variabilité spatiale des concentrations. Sans ces deux sites, le variogramme est plus structuré. Quelques points ressortent néanmoins de la nuée variographique (points marqués en bleus dans la figure): ils correspondent à des couples de sites positionnés selon l'axe  $N160 \pm 30$ . Cette observation est cohérente avec la carte variographique qui met en évidence une anisotropie de directions  $N70$  (direction de plus grande continuité) et  $N160$ .

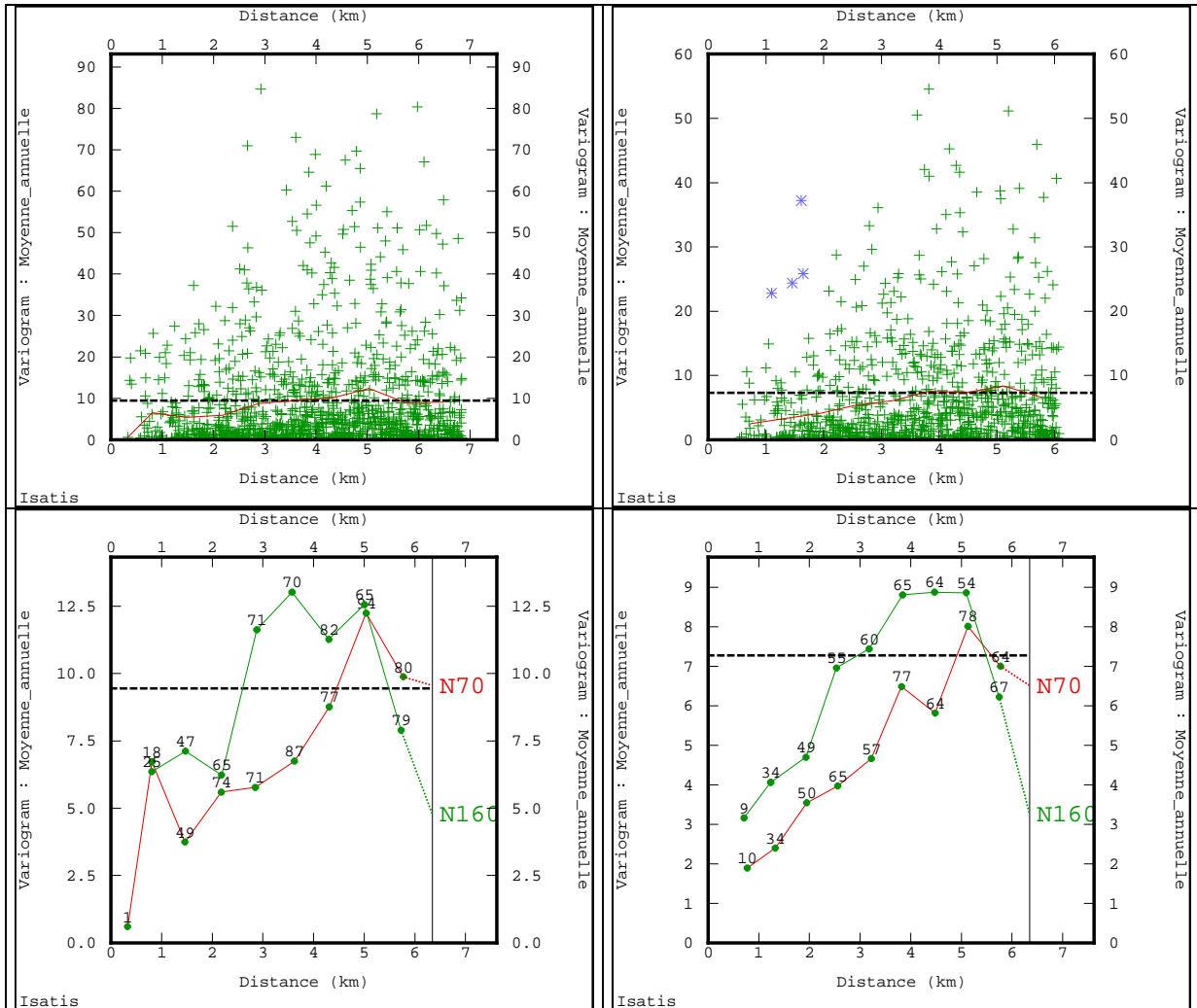


Figure 77 – Ligne du haut : nuée variographique avec (fig. de gauche, pas = 720 m, 10 pas), et sans (fig. de droite, pas = 640 m, 10 pas) les points du centre-ville. Ligne du bas : Variogramme bidirectionnel avec et sans ces deux points.

### B.II.2. Modélisation du variogramme et validation croisée

Des modèles ont été ajustés sur les variogrammes expérimentaux omni- et bidirectionnels construits avec et sans les deux points de centre-ville. Ils se composent tous d'une structure pépitiqque additionnée d'une structure sphérique ou cubique. Avec les deux points du centre, la structure plus erratique du variogramme rend la modélisation moins aisée.

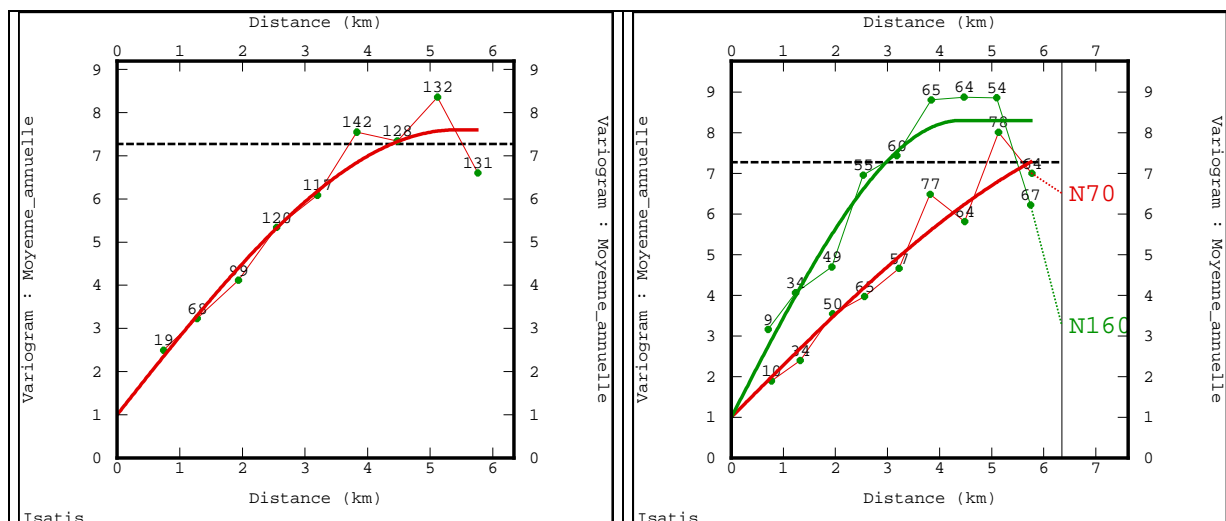


Figure 78– Variogrammes expérimentaux et modélisés sur 62 sites de mesure.

Comme dans l'étude de Cherbourg, la qualité respective des différents modèles est appréciée par validation croisée. Un voisinage unique est considéré pour l'estimation.

Pour tous les modèles, les statistiques d'erreur sont calculées sur les 64 sites d'une part, sur les 62 sites d'autre part.

Le **modèle bidirectionnel sphérique**, ajusté sur le variogramme expérimental calculé avec 62 données, fournit les résultats les plus satisfaisants. Il est pris pour référence.

Tableau 28 – Statistiques sur l'erreur d'estimation obtenues par validation croisée avec le modèle = Pépite (1) + Sphérique (N70- 8500 m, N160-4400m; 7.3)

64 données	moyenne	variance
Erreur	0.07626	3.16898
Erreur standardisée	0.02072	1.09469

2 valeurs non robustes sur 64 valeurs (les deux points du centre).

62 données	moyenne	variance
Erreur	0.06164	2.40735
Erreur standardisée	0.01534	0.67710

0 valeur non robuste sur 62 valeurs.

Tableau 29 – Résultats de la validation croisée sur toutes les données pour le krigeage ordinaire ponctuel, modèle de référence

Toutes les données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Validation croisée par krigeage ordinaire ponctuel sur 64 données			
		Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	64	64	64	64	64
Minimum	7.34	8.01	-4.59	-3.37	0.23
25 <sup>ème</sup> percentile	8.89	9.62	-1.30	-0.62	6.72
Médiane	10.96	10.69	0.43	0.2	11.43
75 <sup>ème</sup> percentile	13.39	13.66	1.54	0.74	19.11
Maximum	20.66	17.62	3.24	2.3	39.81
Etendue (= Max – Min)	13.32	9.61	-	-	-
Moyenne	11.50	11.58	0.08	0.02	13.12
Ecart-type	3.07	2.53	1.78	1.05	8.52
Variance	9.45	6.40	3.17	1.09	72.55
CV %	27	22			
Corrélation mesures/estimations		r=0,815			

Les valeurs estimées sont bien corrélées aux valeurs mesurées, avec un coefficient de corrélation supérieur à 0,80 (Tableau 29, Figure 79). Les deux points de plus grande erreur standardisée (données non robustes) sont les deux points du centre-ville, où les concentrations sont sous-estimées. (Pour ces deux sites, il ne s'agit pas d'une validation croisée au sens strict puisqu'ils n'ont pas servi à calculer le variogramme. L'estimation en ces points ne s'améliore cependant pas lorsqu'on utilise le modèle de variogramme établi sur 64 données.)

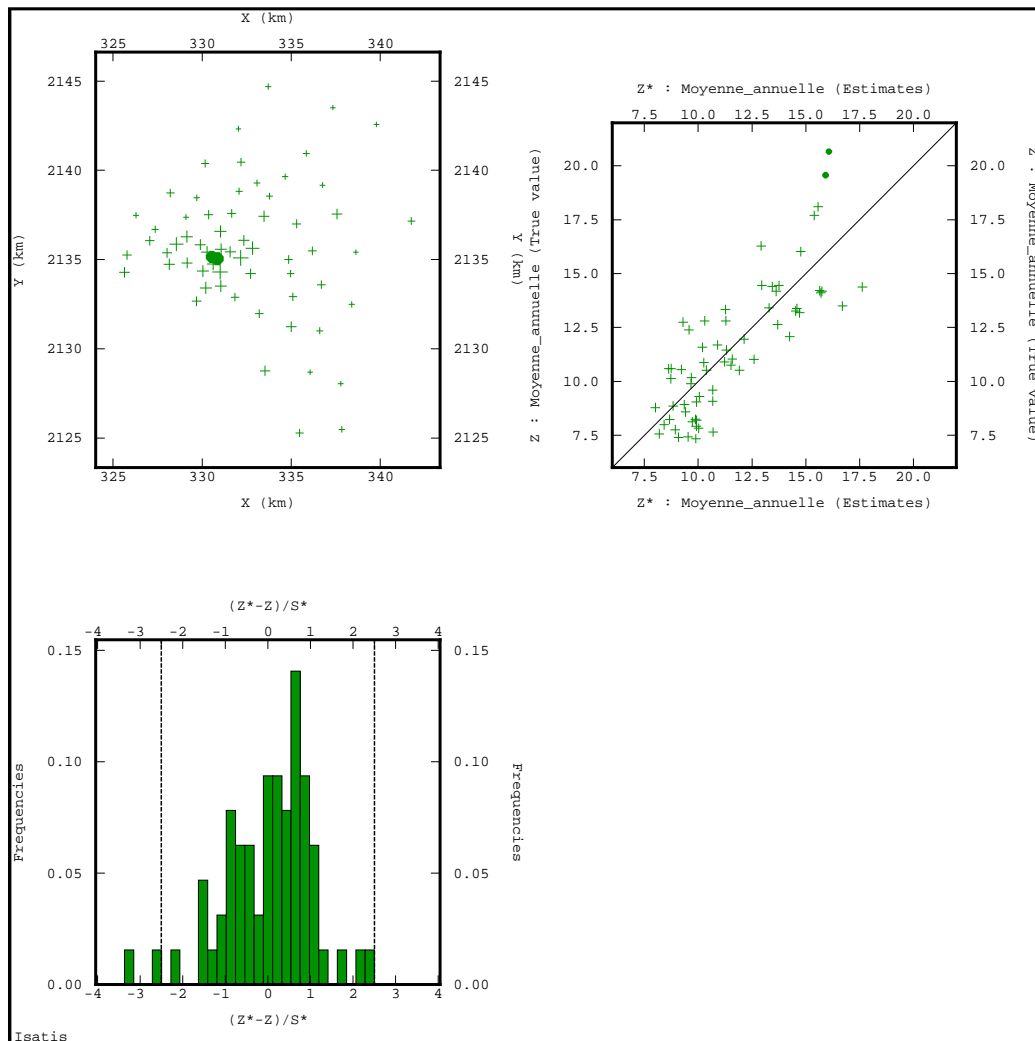


Figure 79 – Résultats de validation croisée. Les deux points qui s'écartent de la bissectrice correspondent aux deux points du centre-ville.

### B.II.3. Etablissement de la carte de référence

Afin de dresser la carte de référence de l'étude, c'est-à-dire la carte réalisée à partir du plus grand nombre de mesures disponibles, nous estimons les valeurs de concentration de  $\text{NO}_2$  aux nœuds d'une grille 2D régulière. La distance entre sites étant d'au moins 320m dans la zone d'échantillonnage la plus dense, une maille d'estimation de 500m de côté est choisie. Cette maille sera conservée dans toute l'étude.

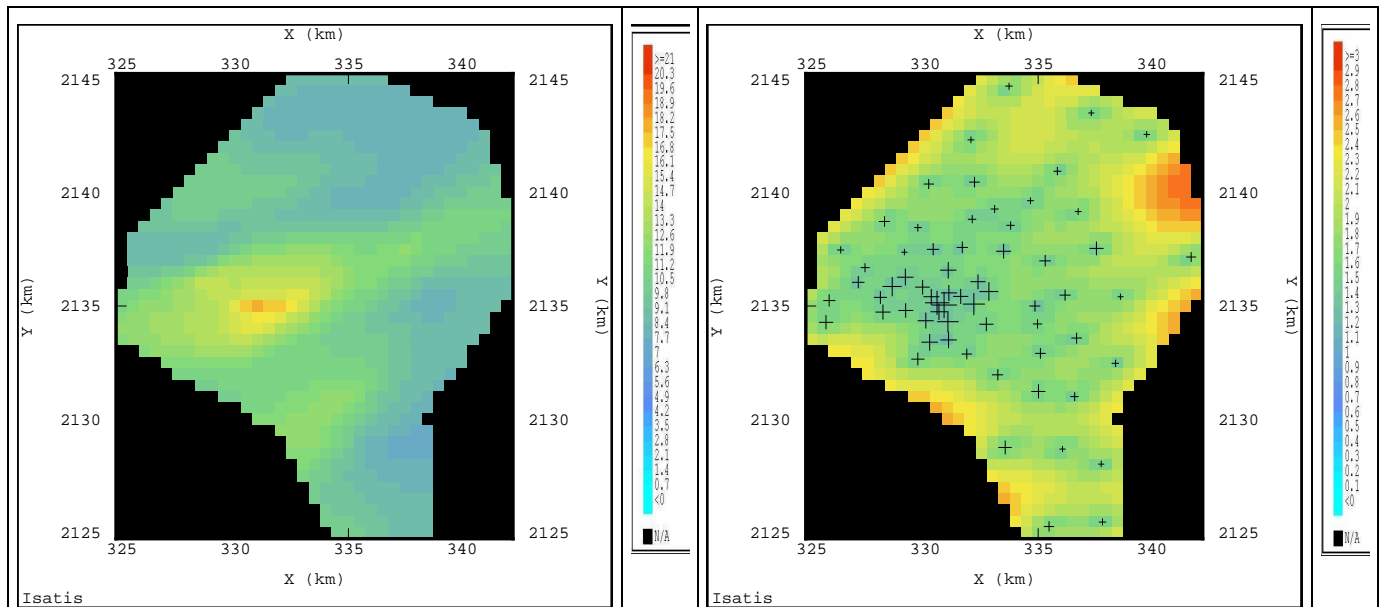
Les caractéristiques de la grille d'estimation sont :

Xmin=325000 m, 35 nœuds, Xmax=342000,  $\Delta X=17000$  m,  
 Ymin=2125000 m, 41 nœuds, Ymax=2145000,  $\Delta X=20000$  m.  
 Compte tenu du nombre limité de données (moins d'une centaine), un voisinage d'estimation **unique** est adopté.

**Tableau 30 – Estimation sur une grille régulière de 500 m de côté, par krigeage ordinaire ponctuel**

Toutes les données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées par krigeage ponctuel ( $\mu\text{g}/\text{m}^3$ )	écarts-types de krigeage ( $\mu\text{g}/\text{m}^3$ )
Nombre de données	64	934	934
Minimum	7.34	7.56	1.08
25 <sup>ème</sup> percentile	8.89	8.61	1.66
Médiane	10.96	9.72	1.85
75 <sup>ème</sup> percentile	13.39	11.12	2.10
Maximum	20.66	17.53	2.77
Moyenne	11.50	10.11	1.89
Ecart-type	3.07	1.86	0.30
Variance	9.45	3.45	0.09
CV %	27	18	16

Statistiques des valeurs estimées à l'intérieur du polygone définissant les contours de l'agglomération



**Figure 80 – Estimation par krigeage ponctuel, à partir de toutes les données, avec le modèle de référence (fig. de gauche) et écart-type de krigeage (fig. de droite)**

### B.III. ETUDE DE LA SELECTION DE DONNEES N°1 (MAILLE DE 1 KM)

La sélection n°1 comprend 57 données, qui correspondent à 1 point de mesure par maille carrée de  $1\text{km}^2$ .

### B.III.1. Etude du variogramme expérimental

Le variogramme expérimental comprend 10 pas de 640 m chacun. Il montre là encore une anisotropie (N70, N160). Aucun point n'est éliminé, les sites 43 et 45 précédemment écartés n'étant pas inclus dans cette sélection de points.

Les mêmes couples de points que précédemment ressortent de la nuée variographique.

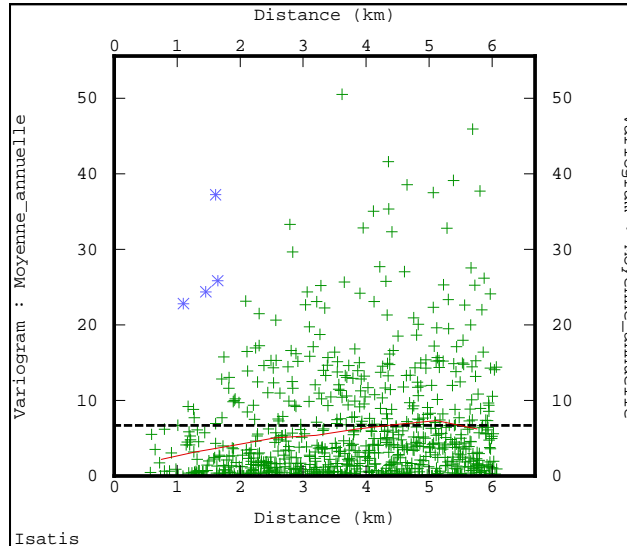


Figure 81– Nuée variographique (pas de 640 m, 10 pas) des données de la sélection 1.

### B.III.2. Modélisation du variogramme et test de validation croisée

Comme précédemment, un modèle omnidirectionnel et un modèle bidirectionnel sont ajustés et soumis au test de validation croisée.

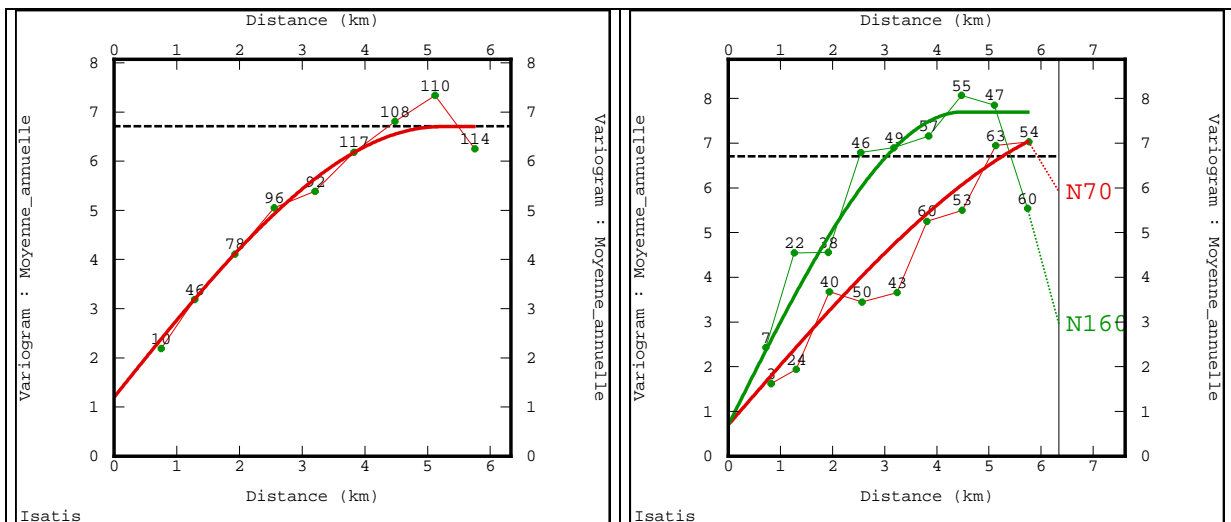


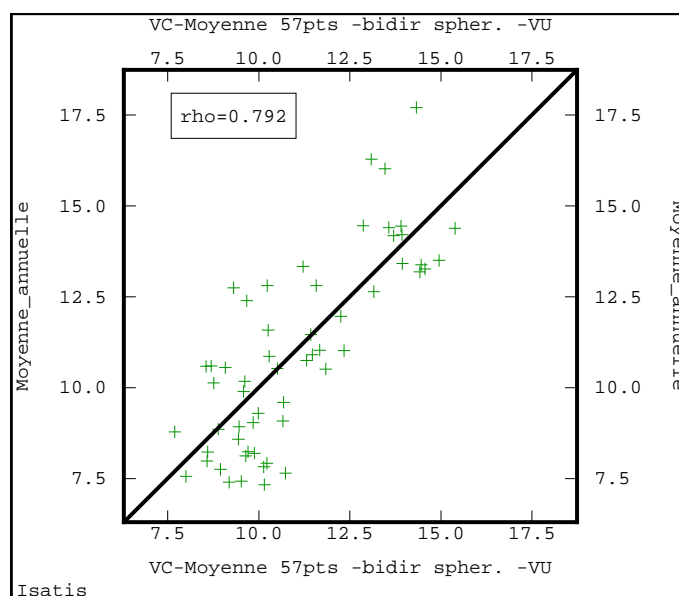
Figure 82 – Variogrammes expérimentaux et modélisés sur 57 sites de mesure

Les statistiques de validation croisée sont favorables au modèle bidirectionnel. Toutes les données sont robustes.



**Tableau 31 – Résultats de la validation croisée sur les données de la sélection 1 pour le krigeage ordinaire ponctuel, modèle bidirectionnel de la sélection 1 : modèle = Pépite (0.7) + Sphérique (N70- 7800 m, N160-4500m; 7.3)**

Validation croisée par krigeage ordinaire ponctuel sur 57 données					
57 données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	57	57	57	57	57
Minimum	7.34	7.70	-3.43	-2.23	0.09
25 <sup>ème</sup> percentile	8.73	9.50	-1.12	-0.56	5.43
Médiane	10.59	10.29	0.53	0.28	10.77
75 <sup>ème</sup> percentile	13.20	12.92	1.26	0.70	18.14
Maximum	17.70	15.39	3.09	1.90	40.37
Etendue (= Max – Min)	10.36	7.69	-	-	-
Moyenne	10.98	11.04	0.06	0.02	12.55
Ecart-type	2.59	2.04	1.58	0.86	9.15
Variance	6.71	4.14	2.5	0.74	83.74
CV %	24	18			
Corrélation mesures/estimations		r=0.792			



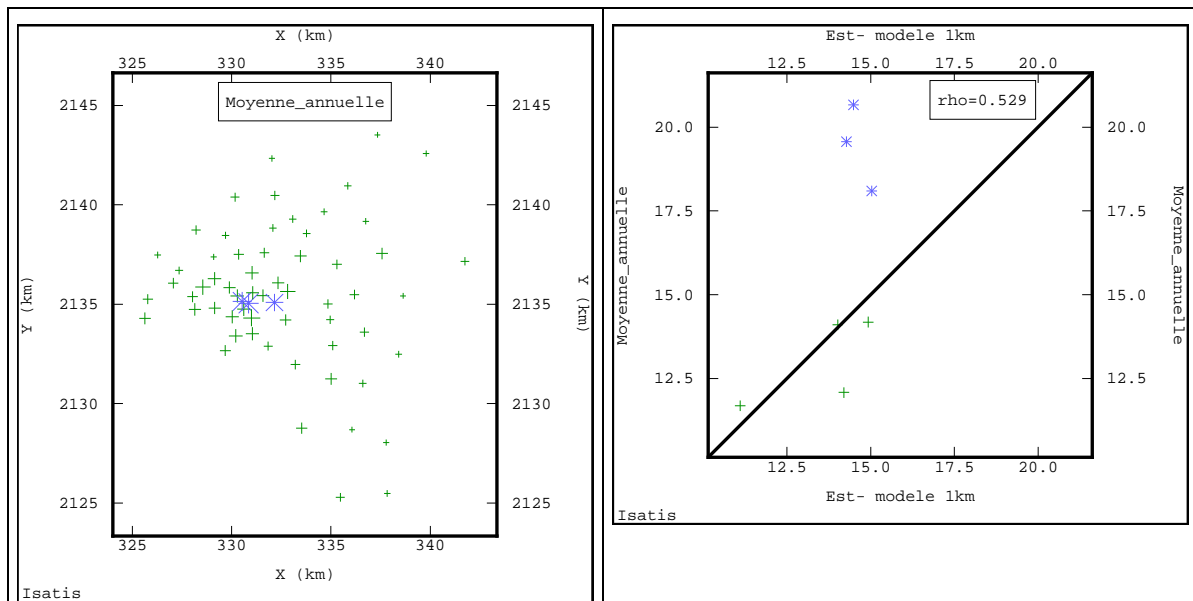
**Figure 83 – Validation croisée : nuage de corrélation entre les concentrations estimées et mesurées aux points de la sélection 1**

### B.III.3. Test de validation absolue sur le complémentaire de la sélection N°1

Les concentrations de  $\text{NO}_2$  sont estimées aux points qui n'ont pas servi à la construction du variogramme, puis sont comparées aux valeurs réelles. Les valeurs supérieures à  $15 \mu\text{g}/\text{m}^3$  sont sous-estimées, en particulier les mesures aux points 43 et 45 (NB : il s'agit des deux points de centre-ville retirés de l'analyse variographique de référence).

**Tableau 32 – Résultats de la validation absolue sur le complémentaire de la sélection 1 pour le krigeage ordinaire ponctuel, modèle de la sélection 1**

		Validation absolue par krigeage ordinaire ponctuel sur 7 données			
7 données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	7	7	7	7	7
Minimum	11.69	11.11	-6.18	-4.74	0.64
Maximum	20.66	15.04	2.12	1.43	29.89
Etendue (= Max – Min)	8.97	4.93	-	-	-
Moyenne	15.77	14.01	-1.76	-1.33	14.62
Ecart-type	3.37	1.23	2.91	2.17	10.53
Variance	11.33	1.52	8.46	4.70	110.96
CV %	21	9			
Corrélation mesures/estimations		r=0.529			



**Figure 84 – Nuage de corrélation entre les concentrations estimées et mesurées au points complémentaires de la sélection 1. En bleu : points où les concentrations sont sous-estimées.**

### B.III.4. Etablissement de la carte pour la sélection 1

La carte d'estimation est proche de la carte initiale ; on note une atténuation des concentrations dans le centre-ville.

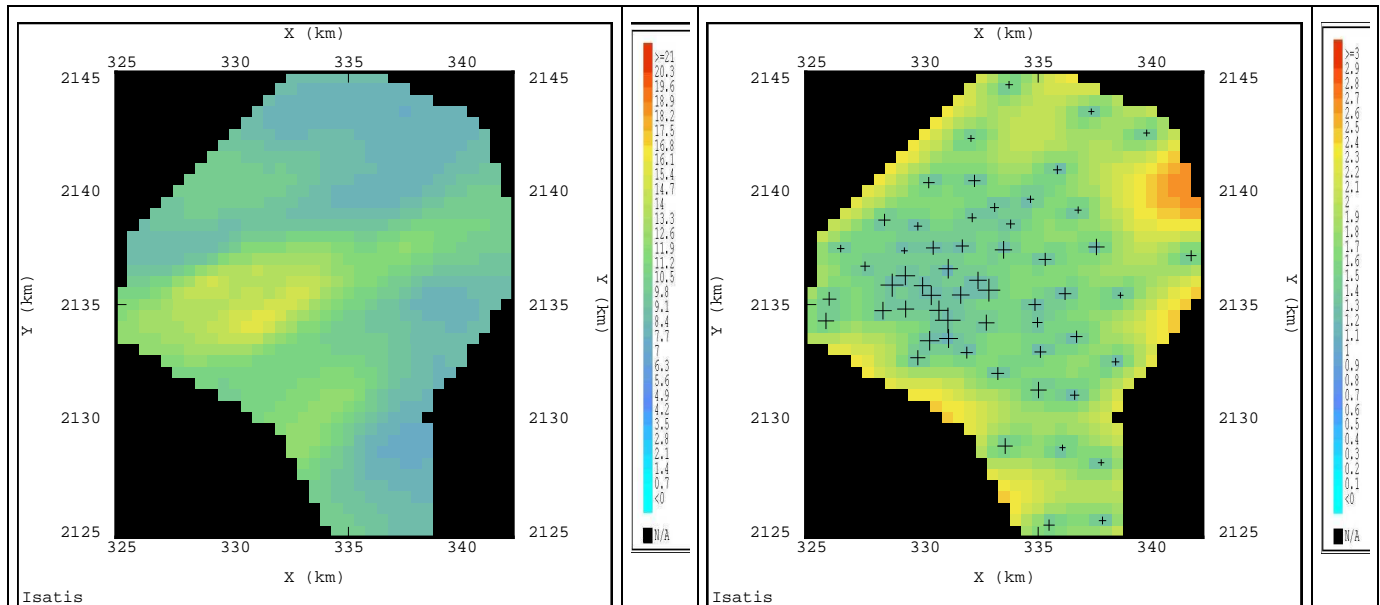


Figure 85 – Estimation par krigeage ponctuel, à partir des données de la sélection 1, avec le modèle de la sélection1 (fig. de gauche) et écart-type de krigeage (fig. de droite)

## B.IV. ETUDE DE LA SELECTION DE DONNEES N°2 (MAILLE DE 1.5 KM)

La sélection n°2 comprend 46 données, qui correspondent à 1 point de mesure par maille carrée de 1,5km<sup>2</sup>.

### B.IV.1. Etude du variogramme expérimental

Le variogramme expérimental comprend 13 pas de 640 m chacun. Aux distances considérées, il présente une non stationnarité dans la direction N70 tandis que dans la direction perpendiculaire (N160), il atteint rapidement un palier. Trois couples de valeurs contrastées, deux desquels sont alignés dans la direction N160, ressortent de la nuée variographique. Aucun point n'est éliminé (notons que les sites 43 et 45 ne sont pas inclus dans cette sélection de points).

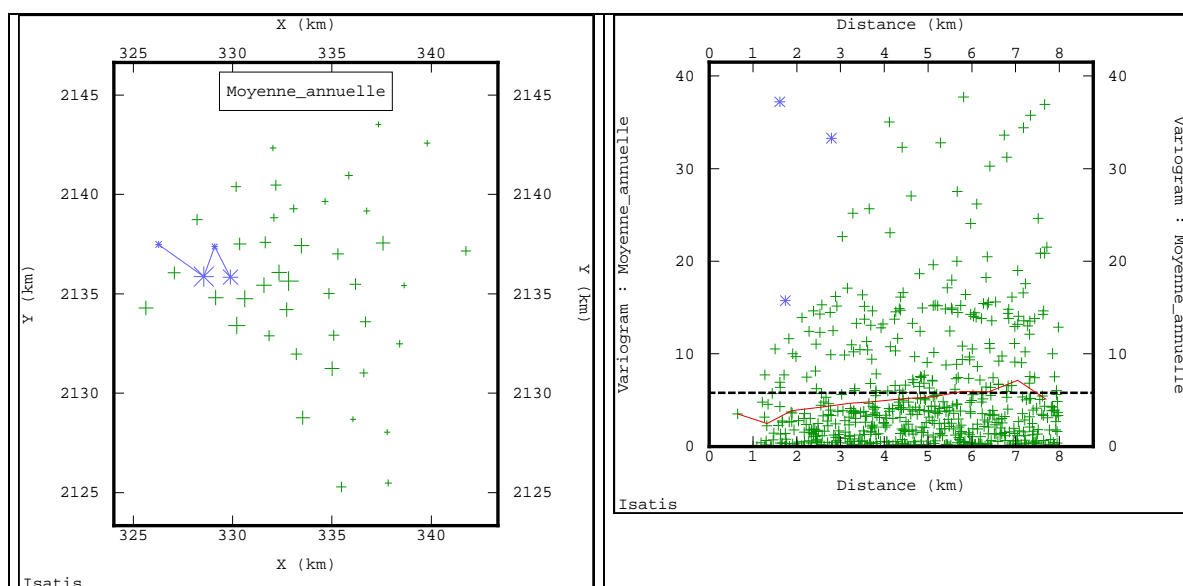


Figure 86– Nuée variographique (pas de 640 m, 13 pas) des données de la sélection 2.

### B.IV.2. Modélisation du variogramme et test de validation croisée

Différents modèles omnidirectionnels et bidirectionnels sont ajustés. Le modèle bidirectionnel présenté en Figure 86 fournit les meilleurs résultats de validation croisée. Bien que la qualité d'ajustement se dégrade au-delà de 5km, la variance d'erreur et la corrélation sont respectivement plus faible et plus élevée en voisinage unique.

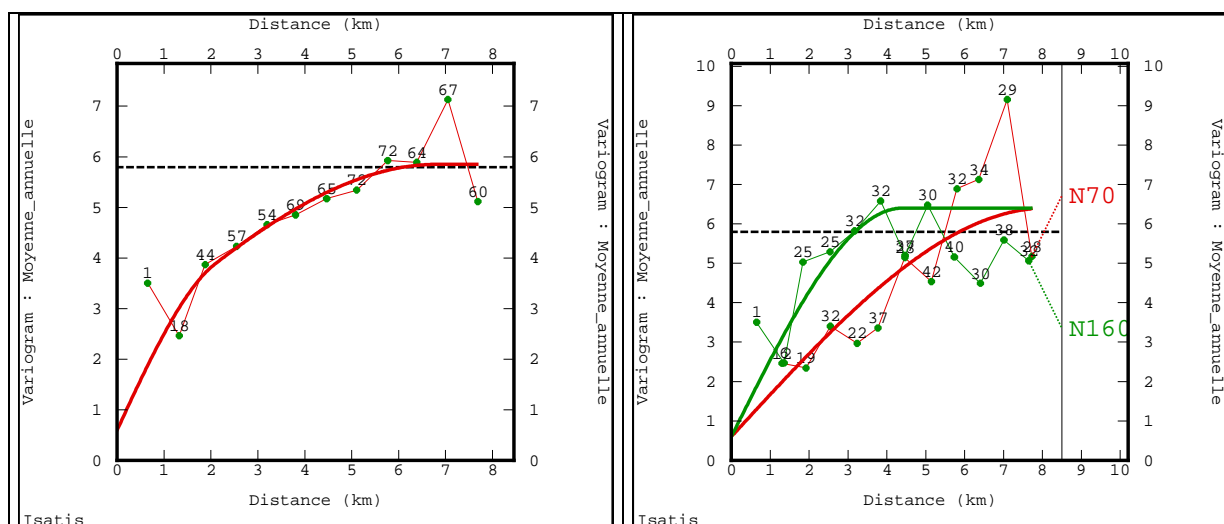
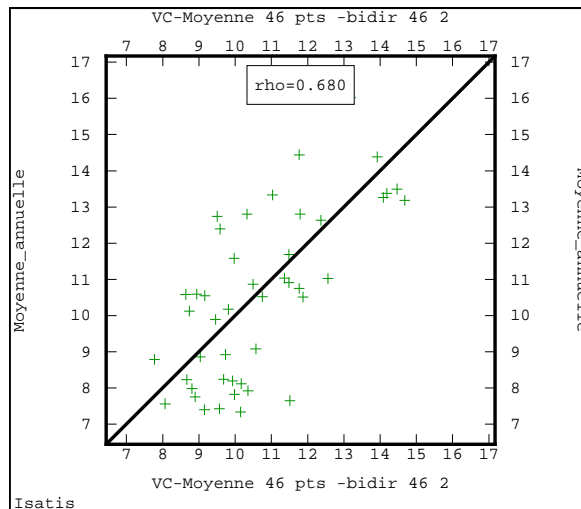


Figure 87 – Variogrammes expérimentaux et modélisés sur 46 sites de mesure

**Tableau 33 – Résultats de la validation croisée sur les données de la sélection 2 pour le krigeage ordinaire ponctuel, modèle bidirectionnel de la sélection 2. modèle = Pépite (0.6) + Sphérique (N70- 8100 m, N160-4400m; 7.3)**

57 données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Validation croisée par krigeage ordinaire ponctuel sur 57 données			
		Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	46	46	46	46	46
Minimum	7.34	7.78	-4.16	-2.67	2.16
25 <sup>ème</sup> percentile	8.23	9.46	-1.01	-0.63	6.13
Médiane	10.57	10.25	0.54	0.20	13.44
75 <sup>ème</sup> percentile	12.74	11.77	1.46	0.76	21.87
Maximum	16.28	14.68	2.81	2.42	38.31
Etendue (= Max – Min)	8.94	6.90	-	-	-
Moyenne	10.64	10.69	0.15	0.01	14.38
Ecart-type	2.41	1.76	1.67	1.00	9.06
Variance	5.79	3.08	2.79	1.00	82.09
CV %	23	16			
Corrélation mesures/estimations		r=0.680			



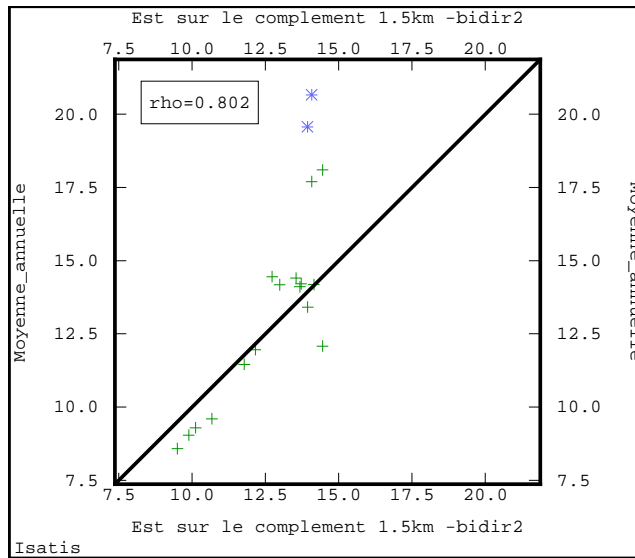
**Figure 88 – Validation croisée : nuage de corrélation entre les concentrations estimées et mesurées aux points de la sélection 1**

### B.IV.3. Test de validation absolue sur le complémentaire de la sélection 2

L'estimation des valeurs inférieures à  $15 \mu\text{g}/\text{m}^3$  est tout à fait satisfaisante. En revanche, comme dans la sélection précédente, les concentrations aux points 43 et 45 sont sous-estimées.

**Tableau 34 – Résultats de la validation absolue sur le complémentaire de la sélection 2 pour le krigeage ordinaire ponctuel, modèle de la sélection 2**

		Validation absolue par krigeage ordinaire ponctuel sur 28 données			
7 données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	18				
Minimum	8.58	9.50	-6.59	-5.38	0.21
P25	11.46	11.78	-1.72	-1.16	3.45
P50	14.14	13.61	-0.23	-0.17	9.09
P75	14.45	14.08	0.82	0.55	19.59
Maximum	20.66	14.45	2.37	1.66	31.88
Etendue (= Max – Min)	12.08	4.95			
Moyenne	13.72	12.77	-0.95	-0.76	11.24
Ecart-type	3.46	1.63	2.36	1.80	9.07
Variance	11.95	2.66	5.56	3.23	82.21
CV %	25	13			
Corrélation mesures/estimations		r=0.802			



**Figure 89 – Nuage de corrélation entre les concentrations estimées et mesurées au points complémentaires de la sélection 2. En bleu : points où les concentrations sont sous-estimées.**

#### B.IV.4. Etablissement de la carte pour la sélection 2

Visuellement, la carte d'estimation ressemble à la carte obtenue pour la sélection 1 ; elle est légèrement moins contrastée.

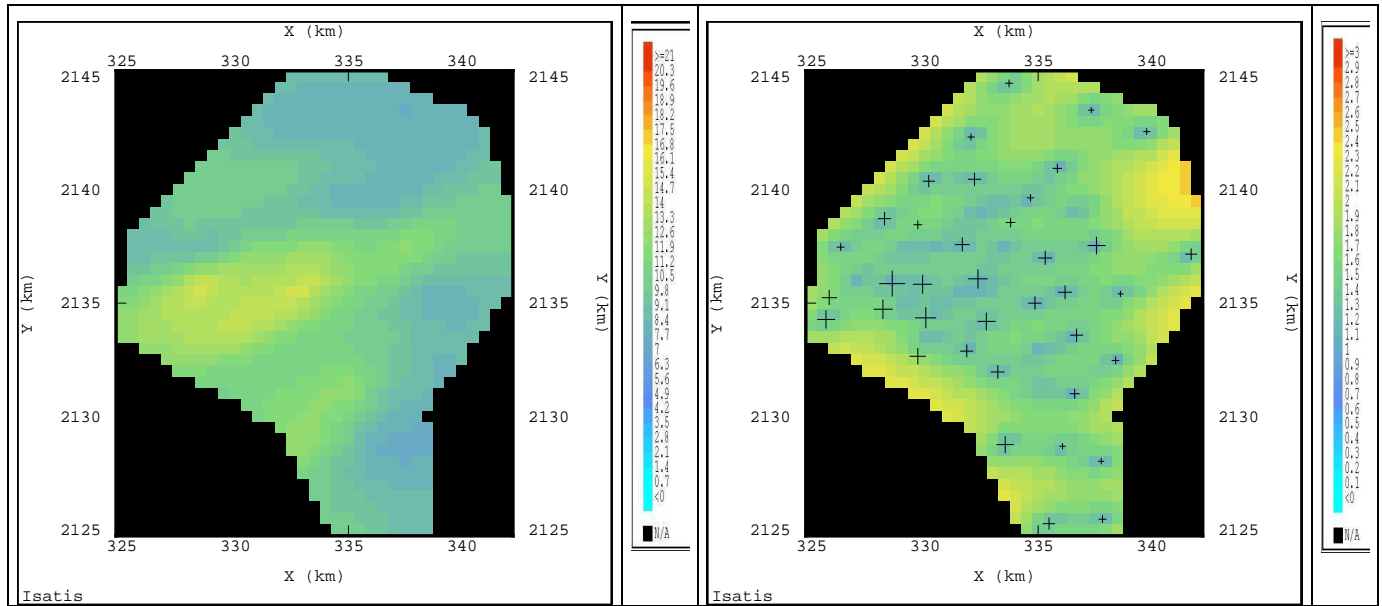


Figure 90 – Estimation par krigeage ponctuel, à partir des données de la sélection 2, avec le modèle bidirectionnel établi sur 46 points (fig. de gauche) et écart-type de krigeage (fig. de droite)

## B.V. ETUDE DE LA SELECTION DE DONNEES N°3 (MAILLE DE 2 KM)

### B.V.1. Etude du variogramme expérimental

Le variogramme expérimental de la figure 91 comprend 12 pas de 710 m. Il présente un pic à 3km, dû principalement au site 63 : ce point de concentration supérieure à  $15 \mu\text{g}/\text{m}^3$  est situé sur la bordure sud d'une zone moins polluée ( $<9 \mu\text{g}/\text{m}^3$ ). Une anisotropie (N70, N160) est encore mise en évidence.

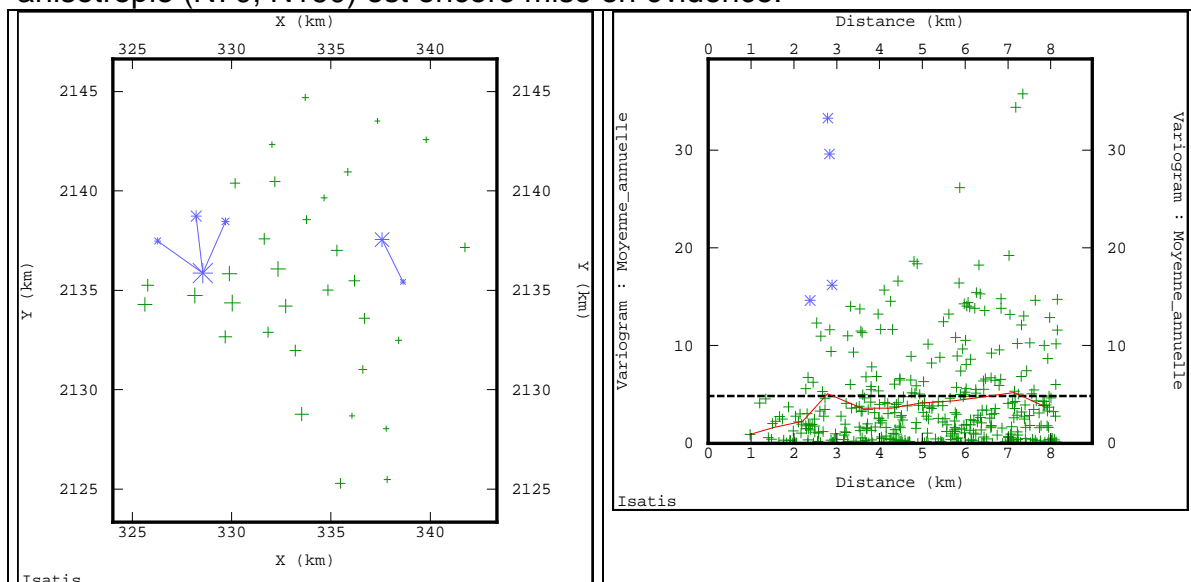


Figure 91– Nuée variographique (pas de 710 m, 12 pas) des données de la sélection 3.

## B.V.2. Modélisation du variogramme et test de validation croisée

L'ajustement est effectué avec et sans le point 63, sur les variogrammes omnidirectionnel et bidirectionnel (Figure 92).

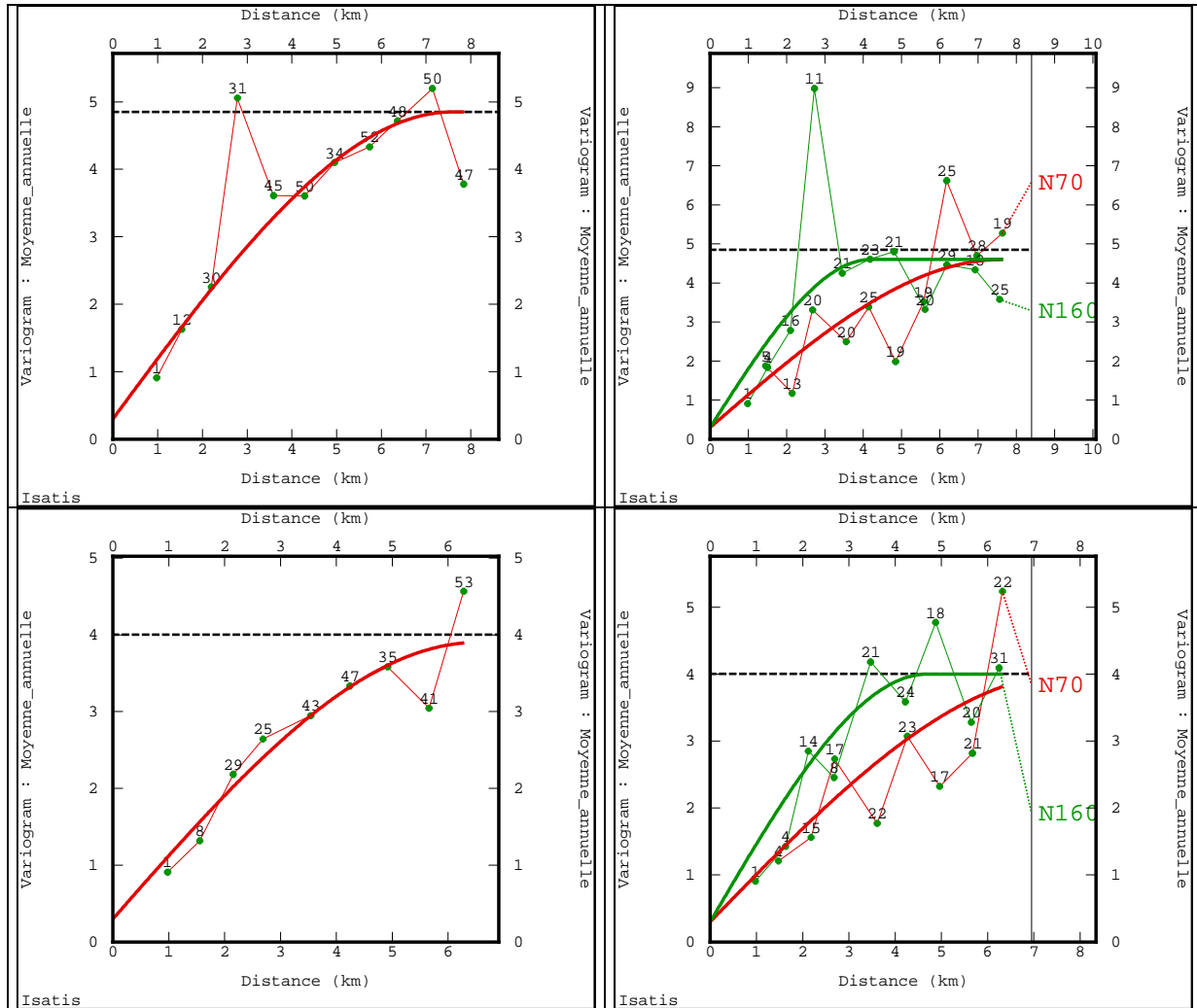


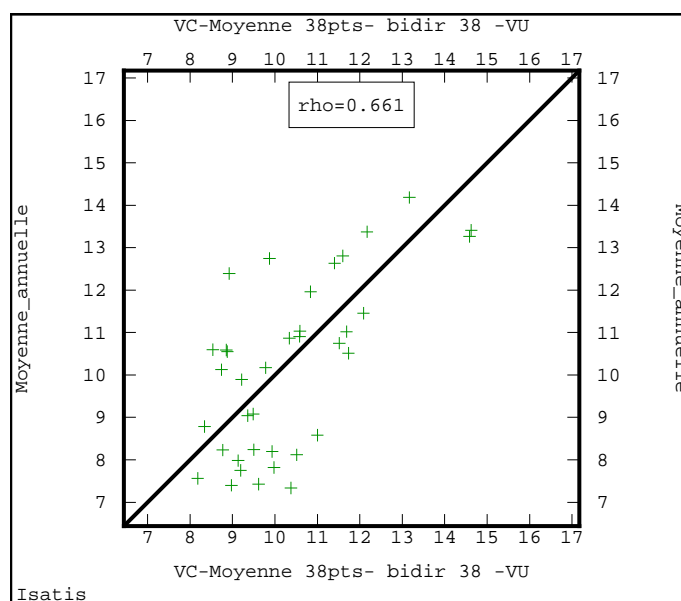
Figure 92 – Variogrammes expérimentaux et modélisés sur 38 sites de mesure (ligne du haut), sur 37 sites de mesure (ligne du bas)

Ces différents modèles sont soumis au test de validation croisée (réestimation sur la totalité des 38 sites à chaque fois). Le modèle bidirectionnel calé sur 38 points est finalement retenu.



**Tableau 35 – Résultats de la validation croisée sur les données de la sélection 3 pour le krigeage ordinaire ponctuel, modèle bidirectionnel de la sélection 3. modèle = Pépite (0.3) + Sphérique (N70- 7600 m, N160-4200m; 4.3)**

		Validation croisée par krigeage ordinaire ponctuel sur 38 données			
38 données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	38	38	38	38	38
Minimum	7.34	8.18	-4.02	-2.93	3.0
25 <sup>ème</sup> percentile	8.23	9.14	-1.21	-0.74	6.53
Médiane	10.53	9.96	0.36	0.22	9.85
75 <sup>ème</sup> percentile	11.96	11.51	1.26	0.80	21.20
Maximum	16.28	14.63	3.04	1.79	41.45
Etendue (= Max – Min)	8.94	6.45	-	-	-
Moyenne	10.34	10.38	0.03	0.01	13.99
Ecart-type	2.20	1.59	1.66	1.02	9.40
Variance	4.85	2.53	2.75	1.04	88.43
CV %	21	15			
Corrélation mesures/estimations		r=0.661			



**Figure 93 – Validation croisée : nuage de corrélation entre les concentrations estimées et mesurées aux points de la sélection 3**

### B.V.3. Test de validation absolue sur le complémentaire de la sélection 3

Tableau 36 – Résultats de la validation absolue sur les données de la sélection 3 pour le krigeage ordinaire ponctuel, modèle de la sélection 3

38 données	Mesures ( $\mu\text{g}/\text{m}^3$ )	Validation croisée par krigeage ordinaire ponctuel sur 38 données			
		Valeurs estimées ( $\mu\text{g}/\text{m}^3$ )	Erreur d'estimation	Erreur d'estimation standardisée	Erreur d'estimation relative (%)
Nombre de données	26	26	26	26	26
Minimum	7.65	9	-6.85	-5.49	0.31
25 <sup>ème</sup> percentile	10.52	10.84	-2.39	-1.63	3.63
Médiane	13.41	12.49	-0.45	-0.35	13.54
75 <sup>ème</sup> percentile	14.44	13.81	0.56	0.45	20.55
Maximum	20.66	14.56	3.39	2.41	44.30
Etendue (= Max – Min)	13.01	5.56	-	-	-
Moyenne	13.20	12.24	-0.96	-0.77	14.34
Ecart-type	3.67	1.59	2.50	1.94	11.54
Variance	11.34	2.53	6.24	3.77	133.13
CV %	23	13			
Corrélation mesures/estimations		r=0.712			

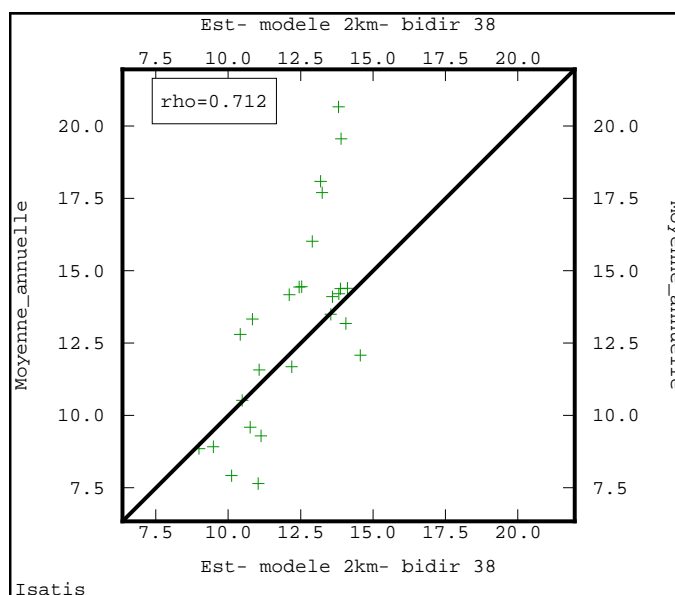


Figure 94 - Nuage de corrélation entre les concentrations estimées et mesurées au points complémentaires de la sélection 32. En bleu : points où les concentrations sont sous-estimées.

### B.V.4. Etablissement de la carte pour la sélection 3

La carte d'estimation est plus uniforme que les précédentes.

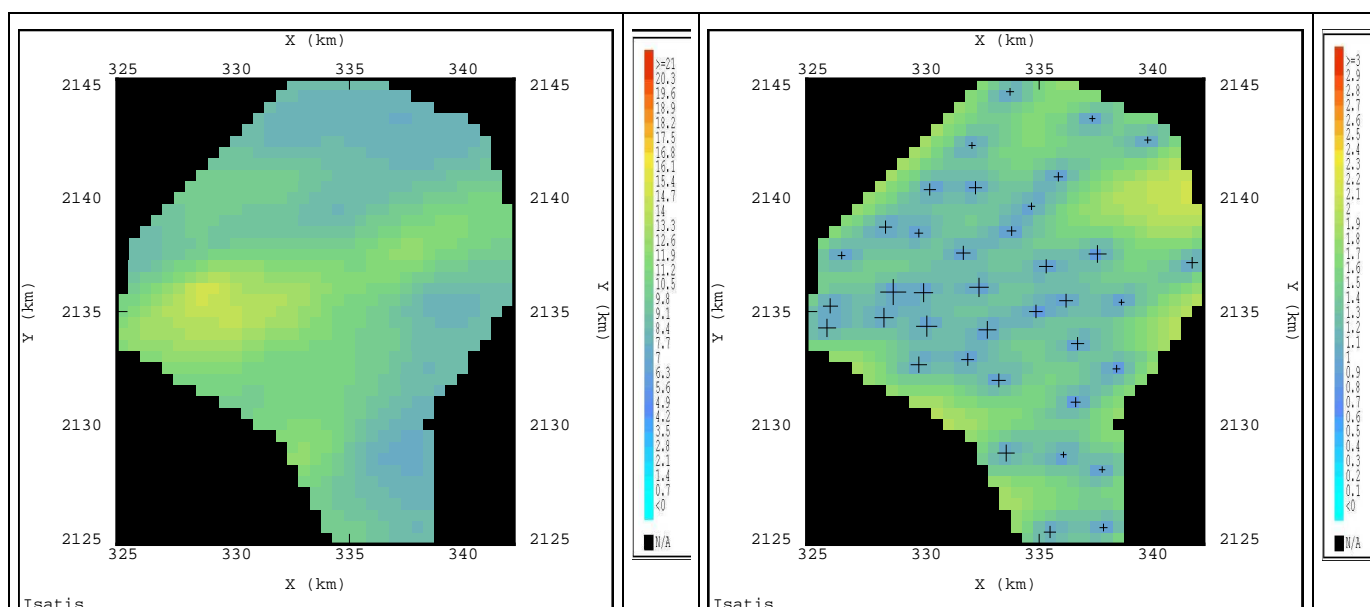


Figure 95 – Estimation par krigeage ponctuel, à partir des données de la sélection 3, avec le modèle bidirectionnel établi sur 38 points (fig. de gauche) et écart-type de krigeage (fig. de droite)

### B.V.5. Test de validation absolue sur l'intersection des complémentaires

Trois jeux de données indépendants sont définis :

- l'intersection des complémentaires des sélections 1 et 2 (6 données) ;
- l'intersection des complémentaires des sélections 2 et 3 (12 données) ;
- l'intersection des complémentaires des sélections 1, 2 et 3 (5 données).

Les concentrations sont estimées sur ces intersections à l'aide des données des sélections S1, S2 ou S3 puis elles sont confrontées aux valeurs expérimentales. Dans la quasi totalité des cas, les valeurs estimées s'écartent de plus ou plus de la bissectrice au fur et à mesure que le nombre de points d'échantillonnage diminue. La variance de l'erreur augmente.

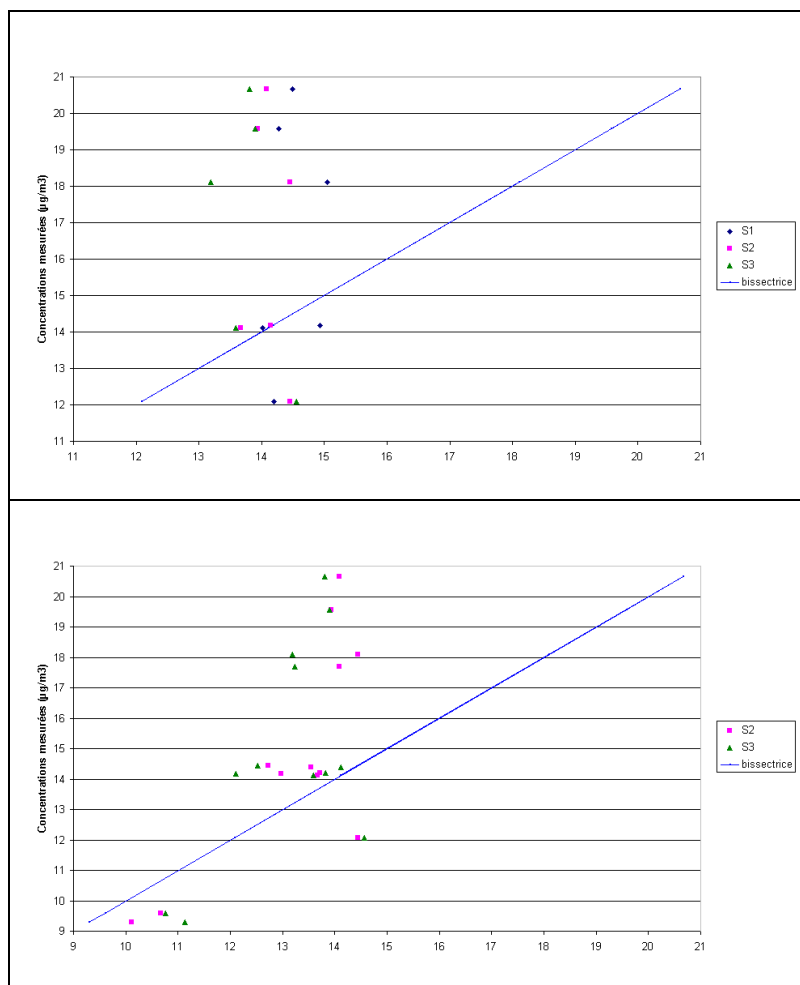


Figure 96- Estimations sur les intersections  $C_{S1} \cap C_{S2}$  et  $C_{S1} \cap C_{S2} \cap C_{S3}$  (figure du haut) et sur l'intersection  $C_{S2} \cap C_{S3}$  (figure du bas)

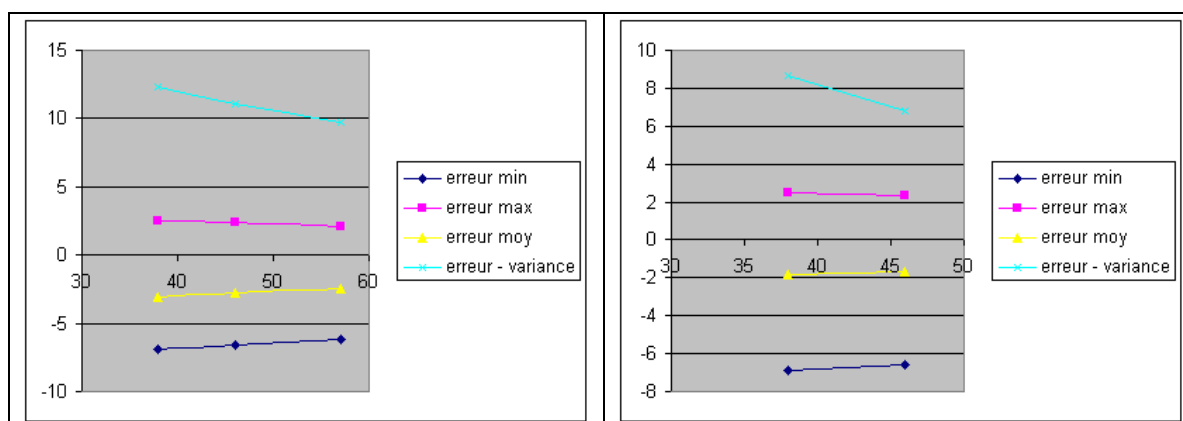


Figure 97- Statistiques d'erreur en fonction du nombre de données – Estimations effectuées sur les intersections  $C_{S1} \cap C_{S2} \cap C_{S3}$  (figure de gauche) et sur l'intersection  $C_{S2} \cap C_{S3}$  (figure de droite)

## B.VI. COMPARAISON DES ESTIMATIONS ET DISCUSSION

### B.IV.1. Influence du nombre de données sur le variogramme expérimental

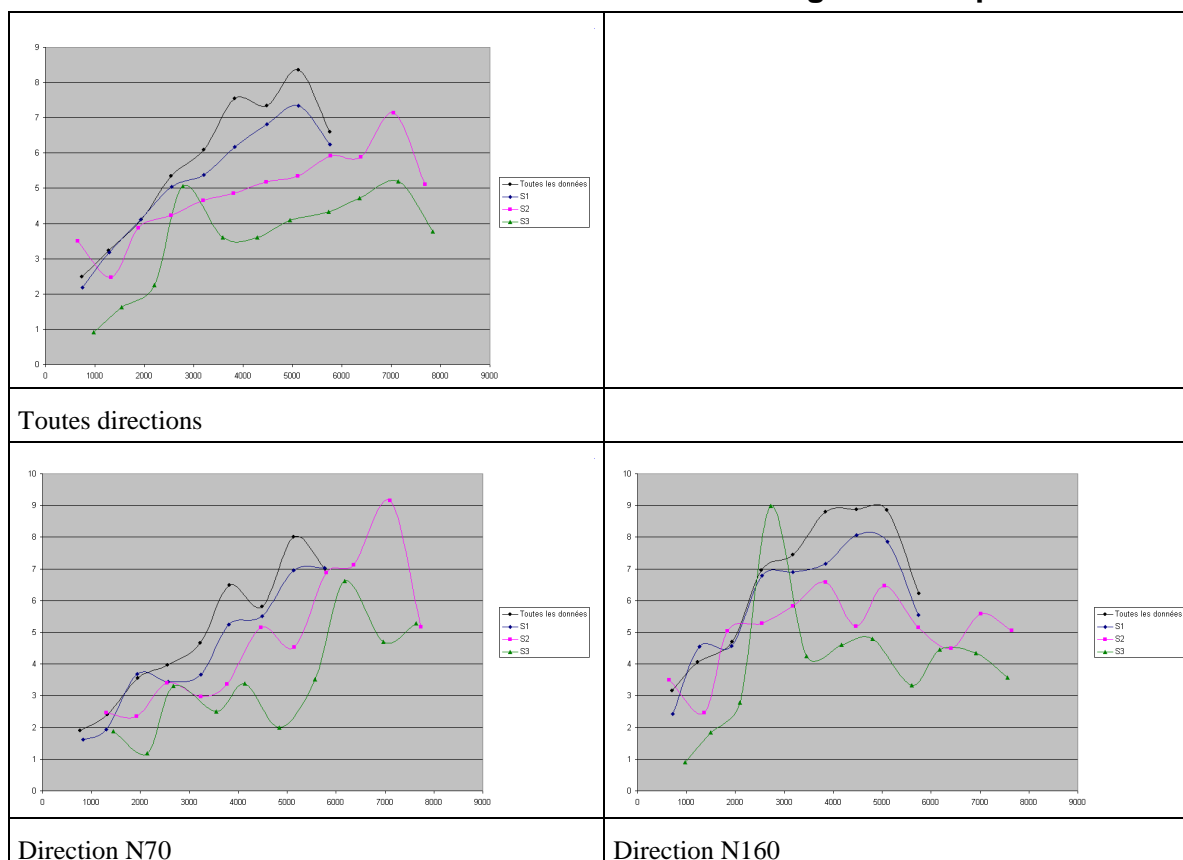


Figure 98 – Variogrammes expérimentaux omnidirectionnels (figure du haut) et bidirectionnels (figures du bas) obtenus pour les trois sélections

Quand la maille d'échantillonnage se relâche, le variogramme expérimental s'écarte du variogramme de référence. Ce changement de structure est plus particulièrement marqué entre la sélection 2 et la sélection 3.

### B.VI.2. Influence sur l'estimation du nombre de données de krigeage pour un modèle établi de variogramme

#### B.VI.2.1. ESTIMATION SUR UN JEU DE DONNEES INDEPENDANT

Le but de ces calculs est d'évaluer isolément l'effet du nombre de données expérimentales sur l'estimation, en supposant la structure spatiale connue.

Le modèle de variogramme utilisé est le modèle de référence.

Les données d'entrée du krigeage sont successivement les sélections S1, S2 et S3.

Les concentrations sont estimées sur les intersections des complémentaires.

**Tableau 37 – Estimation sur les intersections des complémentaires avec le variogramme de référence – effet du nombre de données**

Jeu de validation	$C_{S1} \cap C_{S2} \cap C_{S3}$		$C_{S1} \cap C_{S2}$		$C_{S2} \cap C_{S3}$	
Nb de données	5		6		12	
Statistique	$\rho$	$\sigma_{K_{moy}}$	$\rho$	$\sigma_{K_{moy}}$	$\rho$	$\sigma_{K_{moy}}$
S1 (57 données)	0.490	1.52	0.255	1.53		
S2 (46 données)	-0.057	1.58	-0.049	1.59	0.752	1.65
S3 (38 données)	-0.445	1.75			0.615	1.78

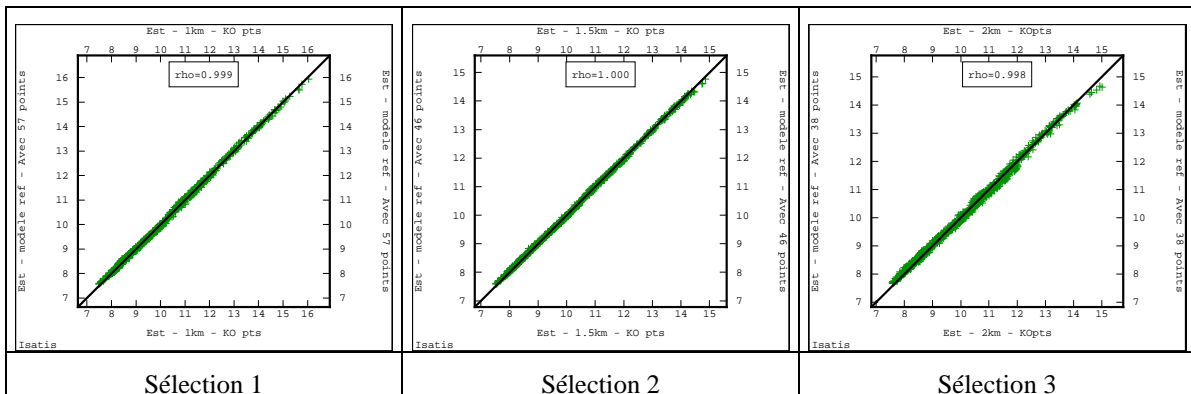
$\rho$  : corrélation mesures-estimations

$\sigma_{K_{moy}}$  : écart-type moyen de l'erreur de krigeage

**A modèle constant, la diminution du nombre de données utilisées dans le krigeage détériore systématiquement la corrélation entre concentrations estimées et mesurées et accroît l'écart-type de l'erreur de krigeage.**

#### B.VI.2.2. ESTIMATION SUR LA GRILLE

Si aux modèles propres à chaque sélection, on substitue le modèle de référence, les estimations sont peu modifiées. Ce résultat traduit la robustesse du krigeage vis-à-vis du modèle variographique. Seule l'estimation de la sélection 3, i.e. de la sélection la plus pauvre en mesures, montre une légère sensibilité au variogramme (Figure 89).



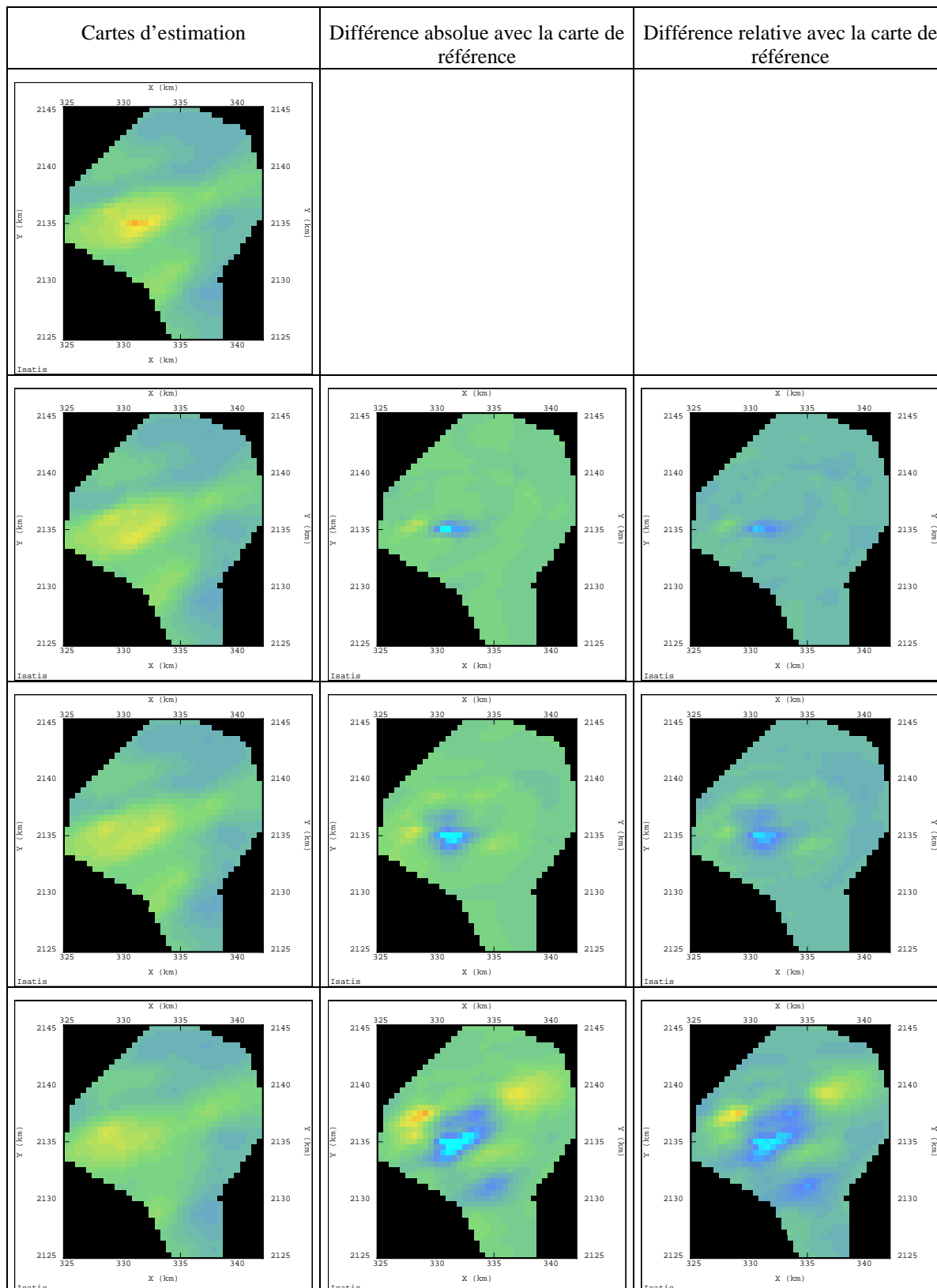
**Figure 99 – Comparaison entre les estimations obtenues avec le modèle de référence (Y) et le modèle propre à la sélection (X).**

Ainsi la dégradation des cartes est moins liée à celle du variogramme (perte d'information sur la structure spatiale du phénomène) qu'à la diminution du nombre de données. L'évolution à modèle constant de l'écart-type moyen de krigeage l'atteste (on rappelle que cet écart-type ne dépend que du modèle de variogramme et de la configuration des points.)

**Tableau 38 – Statistiques de l'écart-type de l'erreur de krigeage. Le modèle de variogramme est le modèle de référence.**

	Référence	S1	S2	S3
Min	1.28	1.29	1.38	1.41
Max	2.77	2.77	2.78	2.82
Moy	1.89	1.90	1.93	1.97
Variance	0.089	0.086	0.079	0.073

### B.VI.3. Effet du nombre de données expérimentales sur la qualité des cartes



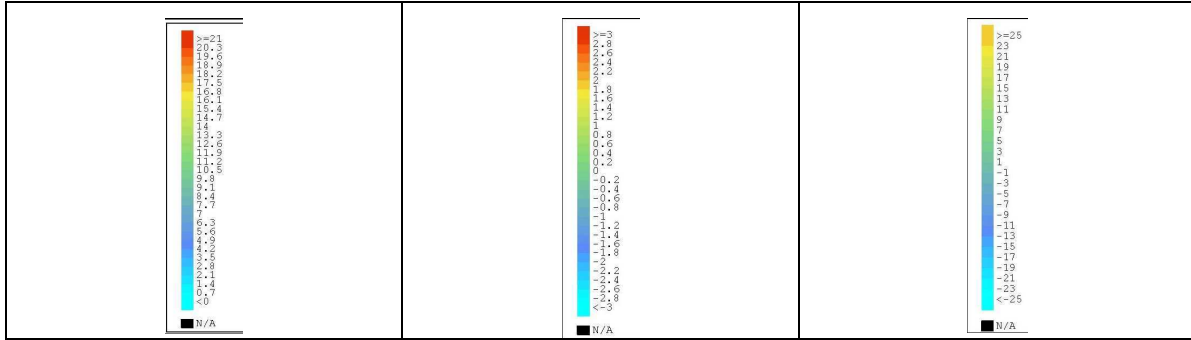


Figure 100 – Comparaison entre les cartes d'estimation obtenues avec un jeu réduit de points (de haut en bas : S1, S2, S3) et la carte de référence

En valeur absolue, l'écart avec la carte de référence reste limité à  $3 \mu\text{g}/\text{m}^3$ . En valeur relative, il peut atteindre 25 % sur une surface d'autant plus étendue que le nombre de points est plus faible. Pour les sélections 1 et 2, la dégradation affecte la zone la plus polluée ; pour la sélection 3, elle touche également les zones périphériques, avec un net effet de lissage: diminution des estimations dans le centre-ville, augmentation dans la partie nord du domaine.

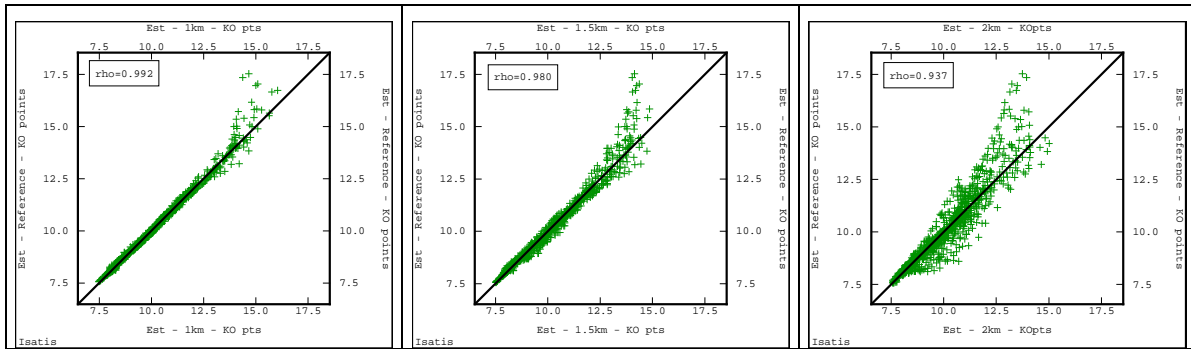


Figure 101 – Comparaison entre l'estimation de référence et les estimations obtenues à partir des trois sélections étudiées (intérieur du polygone de La Rochelle)

Tableau 39 – Statistiques de l'estimation de référence et des estimations obtenues à partir des trois sélections étudiées (intérieur du polygone de La Rochelle)

	Mesures ( $\mu\text{g}/\text{m}^3$ )	Ref	Sélection 1km	Sélection 1.5km	Sélection 2km
Nombre de données	64	934	934	934	934
Minimum	7.34	7.56	7.48	7.52	7.57
25 <sup>ème</sup> percentile	8.89	8.61			
Médiane	10.96	9.72	9.75	9.81	9.91
75 <sup>ème</sup> percentile	13.39	11.12	11.13	11.19	11.02
Maximum	20.66	17.53	16.04	14.86	15.01
Moyenne	11.50	10.11	10.07	10.08	10.05
Ecart-type	3.07	1.86	1.81	1.75	1.59
Variance	9.45	3.45	3.28	3.06	2.54
CV %	27	18	18	17	16



## CONCLUSIONS ET PERSPECTIVES

### **Objectifs d'échantillonnage et jeux de données étudiés**

Cette étude, prévue sur deux ans (2005-2006), a pour but de fournir des recommandations sur les plans d'échantillonnage spatial adaptés aux polluants mesurés et aux objectifs des campagnes. Complétant les orientations proposées par le *guide des aides de l'ADEME en matière d'équipements pour la surveillance* [1], elle pourra être considérée comme une annexe technique de ce document.

Les présents travaux portent plus particulièrement sur :

- la surveillance en zone rurale non couverte par le réseau de mesure fixe, et par extension, sur la surveillance à l'échelle régionale ;
- la surveillance dans les villes de 10000 à 100000 habitants.

Du fait que les concentrations se répartissent de façon plus ou moins homogène selon les espèces chimiques, il convient de conduire l'analyse polluant par polluant. Les travaux réalisés en 2005 ont été consacrés au dioxyde d'azote et aux stratégies d'échantillonnage régulier. Ils ont pour objet de tester l'influence de la maille d'échantillonnage sur les cartes de concentration établies par traitement géostatistique. Le cas de l'ozone et du benzène sera examiné en 2006. Par manque de données exploitables pour des tests, celui du dioxyde de soufre et des particules ne sera pas traité comme il avait été initialement envisagé.

Les jeux de données étudiés en 2005 sont :

- des résultats de simulations en zone rurale (modélisation réalisée par l'EMD sur une partie de la région Centre, à l'aide du code Transchim). Dans cette étude théorique, ils sont utilisés comme substituts de données d'échantillonnage, ayant l'avantage de représenter une grande densité de points.
- des données de campagnes de mesure à l'échelle urbaine, pour des villes de moins de 100000 habitants (Cherbourg, La Rochelle).

Nous présentons les résultats relatifs à la zone rurale de la région Centre et aux zones urbaines de Cherbourg et La Rochelle. Les résultats obtenus sur la région Champagne-Ardenne ne permettent pas encore de tirer de conclusion car seules les données estivales étaient disponibles au moment de l'étude. Les données hivernales seront analysées en 2006. Les résultats de l'étude de la ville de Le Robert (Martinique) seront présentés en 2006 de façon à fournir une meilleure interprétation des données avec le concours de MADININAIR.

### **Méthodologie**

La méthodologie appliquée consiste à définir, à partir d'un jeu de données initial :

- un jeu de référence, utilisé comme point de comparaison ;
- des jeux de données dégradés correspondant à des mailles d'échantillonnage de plus en plus lâches.

Pour les données simulées, deux géométries de maille sont considérées : maille carrée et maille losange ; pour les données de mesure, on se limite à une maille carrée.

L'étude de chaque maillage (caractérisé par la forme et la taille de la maille) nous conduit à comparer :

- les paramètres statistiques du jeu de données sélectionné par ce maillage avec les paramètres statistiques du jeu de référence ;
- et graphiquement, le variogramme expérimental associé à cette sélection avec le variogramme expérimental de référence.

Des modèles de variogramme sont ajustés et leur qualité est contrôlée :

- par un test de validation croisée (sur les points sélectionnés par le maillage testé) ;
- par un test de validation absolue (sur des points décalés d'1/2 maille de simulation par rapport au jeu de référence, dans le cas des données simulées ; sur les points non sélectionnés par le maillage dans le cas des données de mesure).

Enfin, pour chaque sélection de points, une carte d'estimation est élaborée par krigeage et comparée à la carte krigée de référence.

### **Résultats**

A partir d'un nombre seuil de points d'échantillonnage, le variogramme expérimental se dégrade nettement, traduisant une perte d'information sur la structure spatiale du phénomène. Ce seuil est aux environs de 100 points en zone rurale, en deçà de 46 points pour les deux agglomérations étudiées.

Si les estimations sont robustes vis à vis du modèle de variogramme, en revanche, le nombre de points fourni en entrée du krigeage a une influence déterminante sur la qualité des estimations. Une diminution de l'effectif de données provoque un lissage de plus en plus marqué des estimations, et cela, même lorsque la structure spatiale du phénomène est supposée connue (utilisation du variogramme de référence).

L'homogénéité de l'écart-type de krigeage, encore vraie avec un nombre réduit de points, atteste l'efficacité d'un échantillonnage régulièrement réparti dans l'espace.

Les statistiques d'erreurs ne s'accordent pas pour montrer la supériorité de la maille losange ou de la maille carrée. En validation absolue, la dégradation des résultats avec la réduction du nombre de données est plus sensible pour la maille carrée.

Les tests de validation et la comparaison des cartes avec la carte de référence suggèrent qu'une estimation correcte des concentrations de fond peut être obtenue :

- avec 100 points, soit 1 point par maille de 43,5 km<sup>2</sup> dans la zone rurale étudiée;
- avec 43 à 46 points, soit 1 point par maille de 1 à 2,25 km<sup>2</sup> dans les villes étudiées. Pour reproduire la variabilité accrue des concentrations en centre-ville, des points de mesure supplémentaires y sont nécessaires.

### **Poursuite des travaux sur le dioxyde d'azote**

En s'aidant de données de population ou d'émission, il convient d'examiner de plus près le rôle des variables auxiliaires :

- en zone urbaine, l'élargissement de la maille d'échantillonnage fait perdre en variabilité dans le centre-ville. L'existence de variables auxiliaires corrélées aux concentrations peut-elle compenser cet effet ? Afin d'établir une relation entre concentration et variable auxiliaire, on ne peut néanmoins se passer de mesures. Quelle densité d'échantillonnage demeure nécessaire ?
- de quelle manière les variables auxiliaires peuvent-elles guider l'élaboration d'un plan d'échantillonnage ? une réflexion sur la stratification est à mener.

Afin de confirmer et d'affiner les conclusions du travail conduit sur Cherbourg et La Rochelle, l'agglomération de Le Robert sera considérée en 2006. Comme il a été mentionné, l'étude de l'échantillonnage à l'échelle régionale sera poursuivie grâce aux données mises prochainement à disposition par ATMO Champagne-Ardenne (campagne hivernale de mesure ; simulations ESMERALDA).

La méthodologie développée pour le dioxyde d'azote sera mise en œuvre pour l'ozone et le benzène.

## REFERENCES

- [1] ADEME et MEDD - Guide en matière d'équipements pour la surveillance de la qualité de l'air ambiant dans les associations agréées de surveillance de la qualité de l'air, Décembre 2004, 34 pages.
- [2] Groupe de travail « échantillonneurs passifs », Oct. 2001. Echantillonneurs passifs pour le dioxyde d'azote (guide), 73 pages.
- [3] LCSQA-EMD, Déc. 2001. Application d'un modèle de dispersion de polluants atmosphériques en zone rurale influencée. *P. Olivier et A. Wroblewski*.
- [4] LCSQA-EMD, Déc. 2003. Cartographie de la pollution atmosphérique en zone rurale influencée au moyen d'un modèle déterministe de dispersion des polluants. *A. Wroblewski, E. Perdrix et P. Olivier*.
- [5] INERIS ; Déc. 2003. Méthodes de représentation de la qualité de l'air : guide d'utilisation des méthodes de la géostatistique linéaire. *L. Malherbe et L. Rouil*.
- [6] AIRCOM, 2000. Etude préalable à l'implantation d'une station fixe de mesure de la pollution atmosphérique dans l'agglomération cherbourgeoise. *J. Ozdowski*.

**LISTE DES ANNEXES**

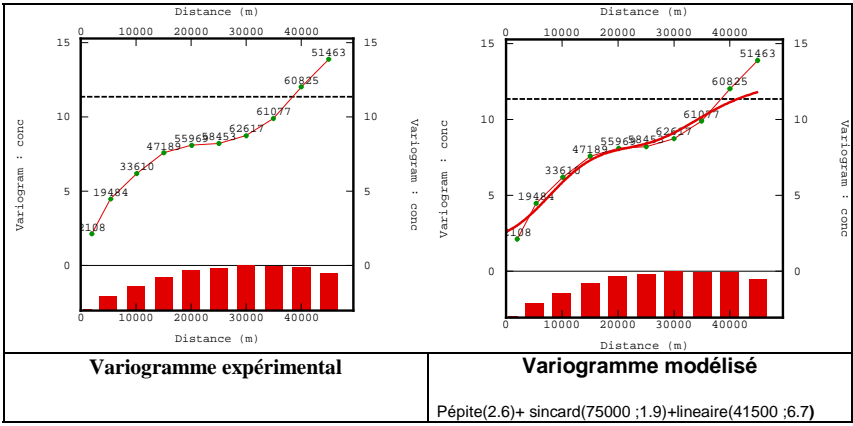
Repère	Désignation	Nombre de pages
Annexe 1	<b><u>1<sup>ERE</sup> PARTIE : COUVERTURE HORS ZONE</u></b> <b><u>AGGLOMERATION</u></b> Fiches résultantes concernant les travaux de la campagnes 1 (10 et 25 juillet 2002)	9
Annexe 2	<b><u>1<sup>ERE</sup> PARTIE : COUVERTURE HORS ZONE</u></b> <b><u>AGGLOMERATION</u></b> Fiches résultantes concernant les travaux de la campagnes 2 (3 au 17 avril 2003)	7

## **Annexe 1**

### **1<sup>ERE</sup> PARTIE : COUVERTURE HORS ZONE AGGLOMERATION**

**Fiches résultantes concernant les travaux de la campagnes 1  
(10 et 25 juillet 2002)**

**Etude de l'échantillonnage sur la grille de référence au pas de 2km**



count	1089
min	5.91
max	40.56
mean	12.43
std.dev	3.47
var	12.06

cross validation

error

mean	-0.00906
var	3.23

std. error

mean	-0.0064
var	1.02

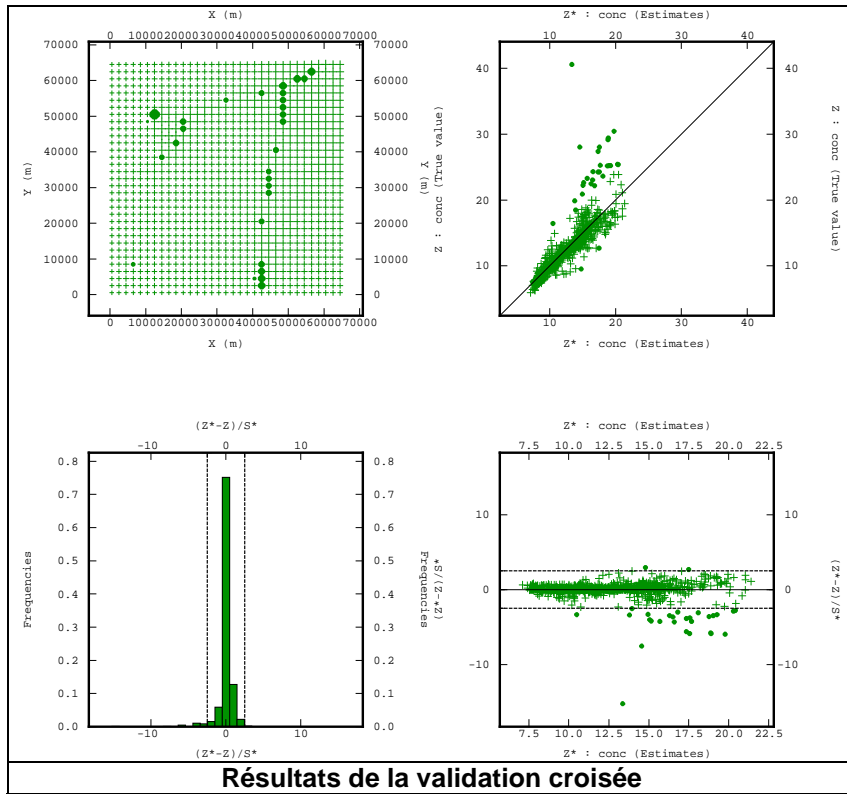
validation absolue

error

mean	0.044
var	1.83

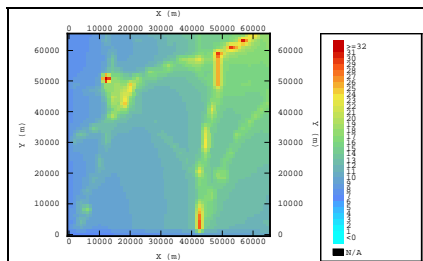
std. error

mean	0.025
var	0.58

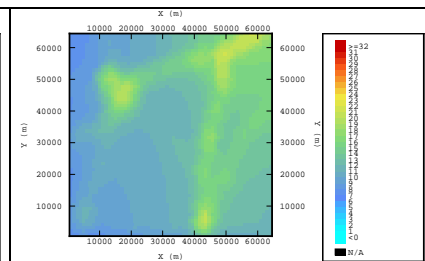


Résultats de la validation croisée

Statistiques

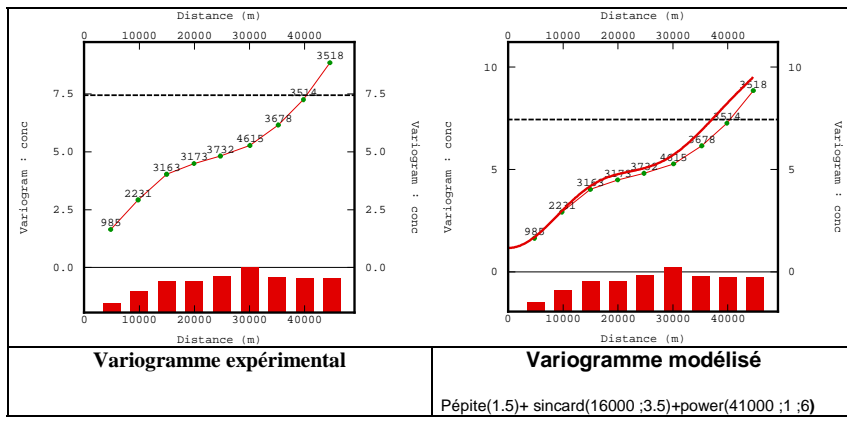


Carte de krigeage avec variogramme modélisé (domaine 33x33)



Carte de krigeage pour la validation absolue

**Etude de l'échantillonnage sur une grille de 17x17 carre au pas de 4km**



Variogramme expérimental

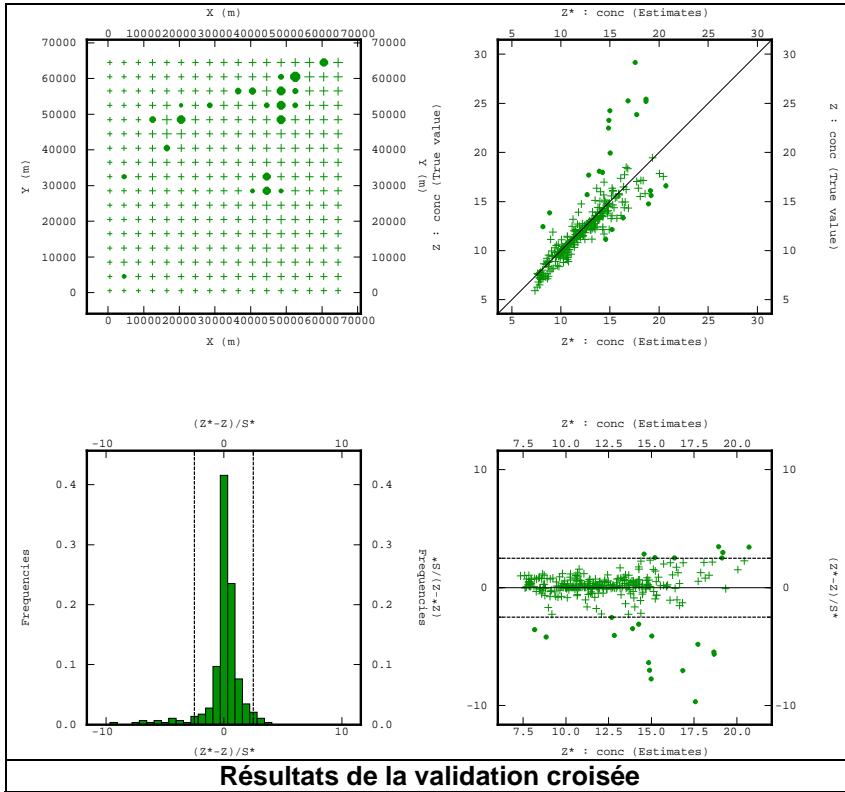
Variogramme modélisé

Pépite(1.5)+ sincard(16000 ;3.5)+power(41000 ;1 ;6)

count	289
min	5.91
max	29.16
mean	12.25
std.dev	3.5
var	12.24

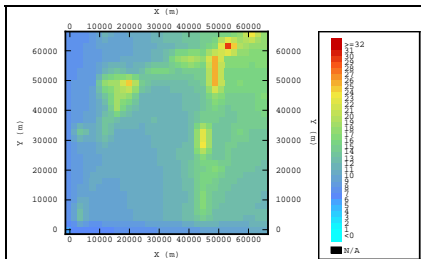
cross validation



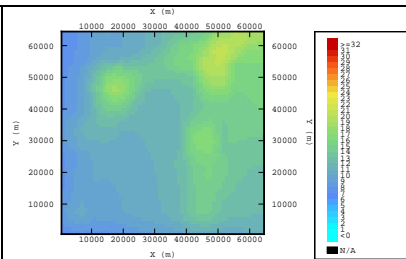


Résultats de la validation croisée

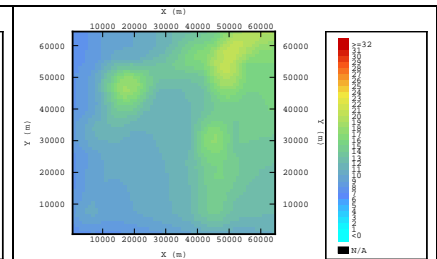
Statistiques



Carte de krigeage avec variogramme modélisé (domaine 17x17)

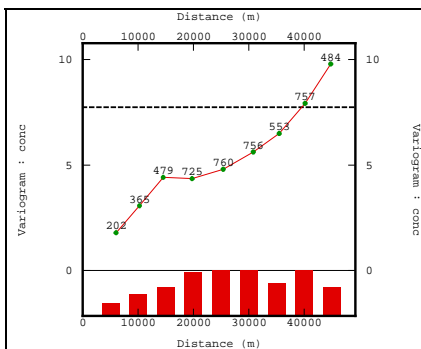


Carte de krigeage pour la validation absolue

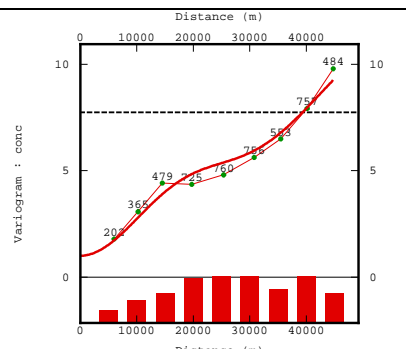


Carte de krigeage pour la validation absolue avec variogramme de ref

**Etude de l'échantillonnage sur une grille de 11x11 carre au pas de 6km**



Variogramme expérimental

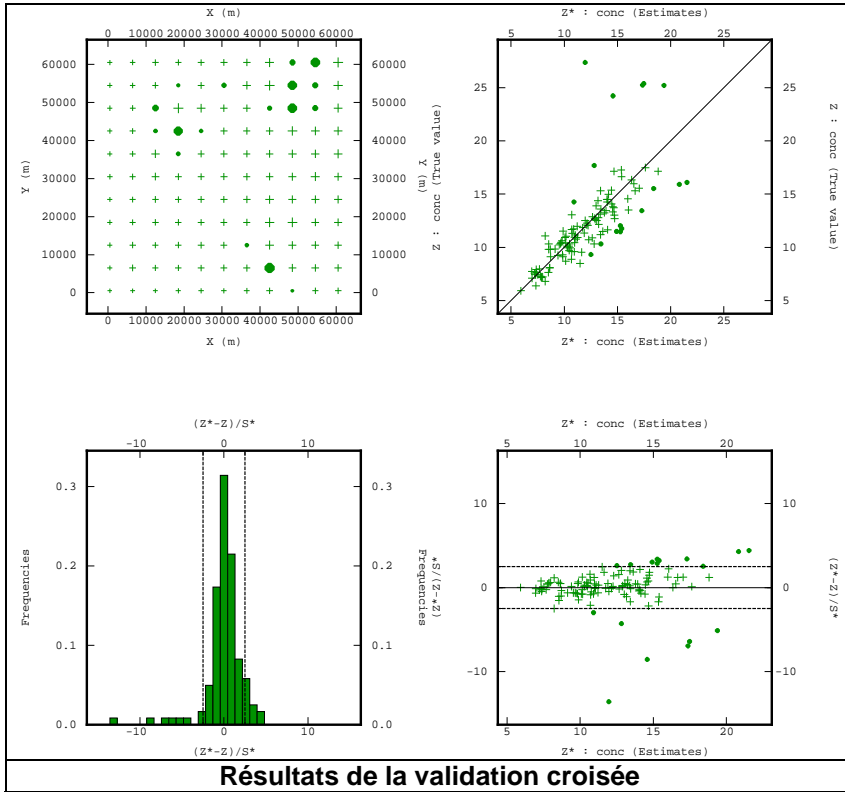


Variogramme modélisé

Pépité(1)+ sincard(90000 ;2)+power(1.8 ;40000 ;5)

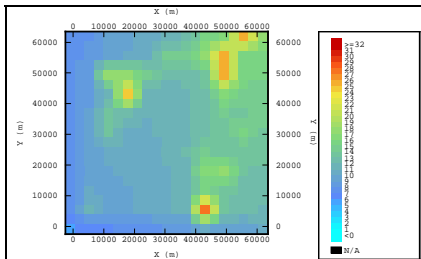
count	121
min	5.91
max	27.37
mean	12.1
std.dev	3.9
var	15.19

cross validation

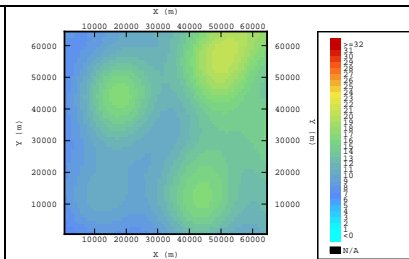


Résultats de la validation croisée

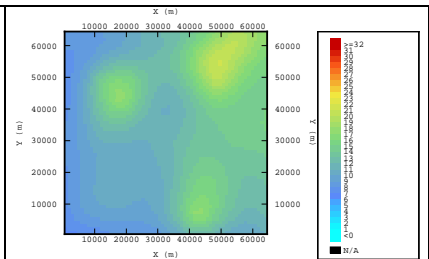
Statistiques



Carte de krigeage avec variogramme modélisé (domaine 11x11)

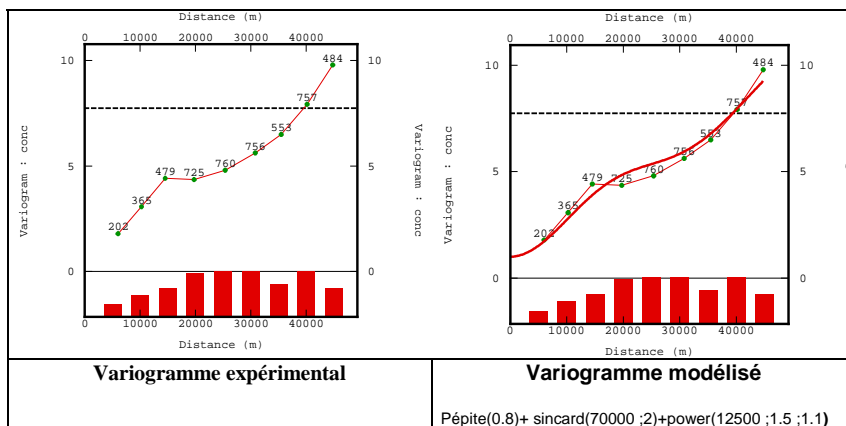


Carte de krigeage pour la validation absolue

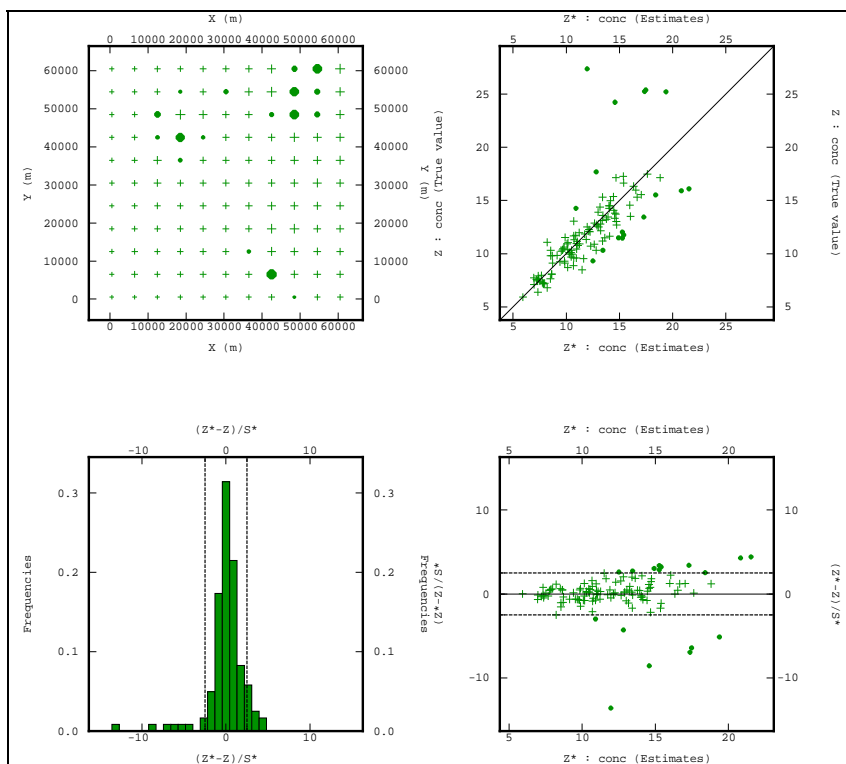


Carte de krigeage pour la validation absolue avec variogramme de ref

# Etude de l'échantillonnage sur une grille de 9x9 carre au pas de 8km



count	81
min	5.91
max	25.42
mean	11.93
std.dev	3.73
var	13.9

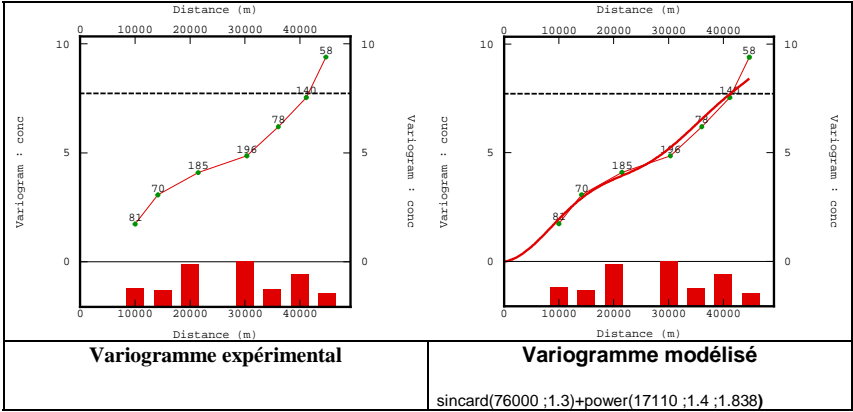


cross validation	
error	
mean	0.0166
var	3.019
std. error	
mean	0.0042
var	1.335

validation absolue	
error	
mean	-0.1807
var	4.712
std. error	
mean	-0.145
var	2.984

utilisation	
variogramme de ref	
error	

**Etude de l'échantillonnage sur une grille de 7x7 carre au pas de 10km**



count	49
min	5.91
max	19.43
mean	11.3
std.dev	2.99
var	8.94

cross validation

error

mean	0.03745
var	0.987

std. error

mean	0.019
var	0.718

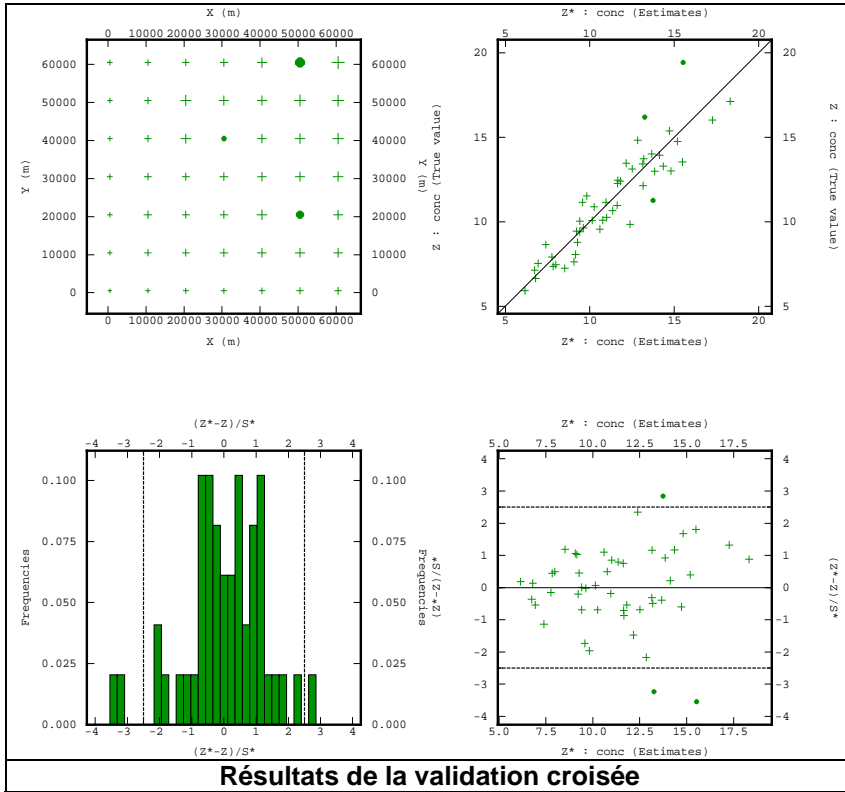
validation absolue

error

mean	-0.6581
var	4.49

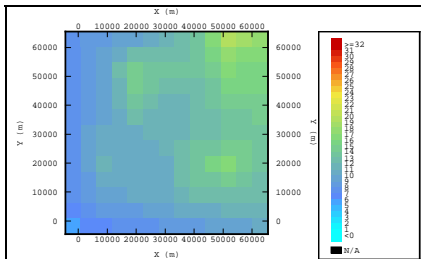
std. error

mean	-0.8641
var	

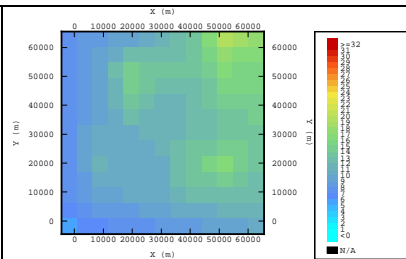


Résultats de la validation croisée

Statistiques

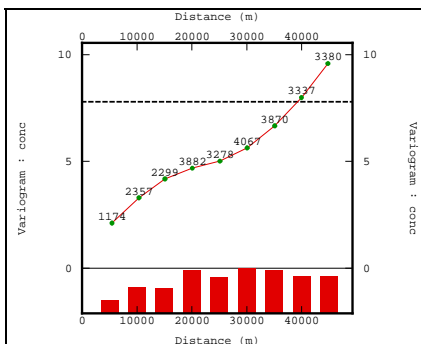


Carte de krigeage avec variogramme modélisé (domaine 7x7)

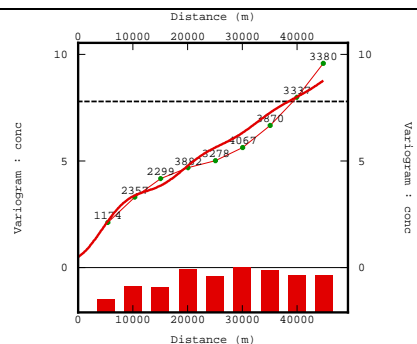


Carte de krigeage pour la validation absolue

**Etude de l'échantillonnage sur une grille de 17x17 losange au pas de 4km**



Variogramme expérimental

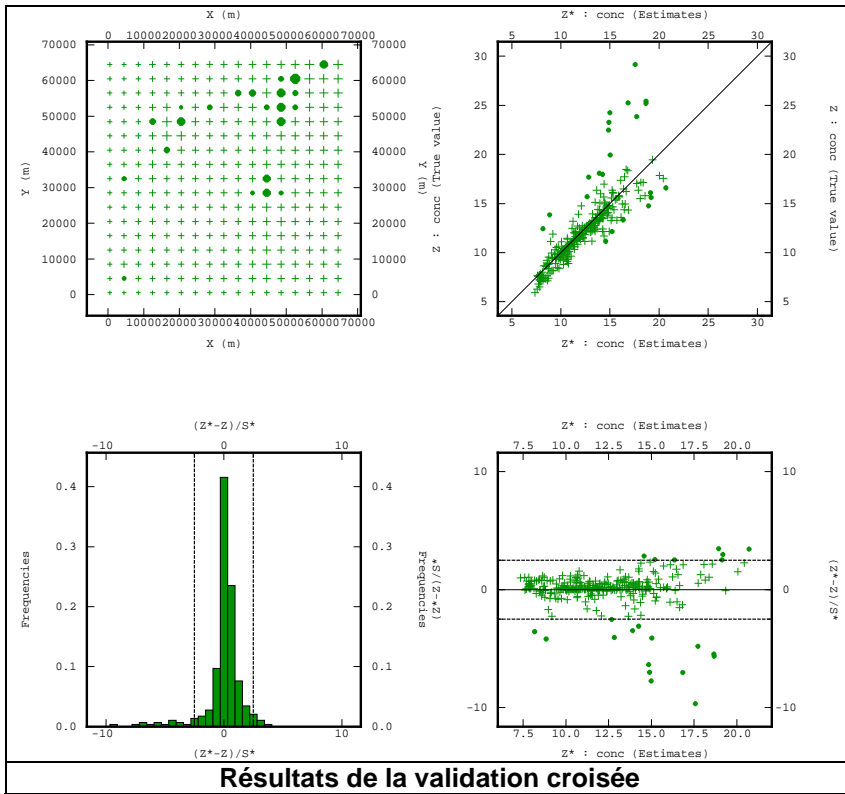


Variogramme modélisé

Pépité(0.5)+ sincard(42800 ;1)+power(10000 ;1 ;1.6)

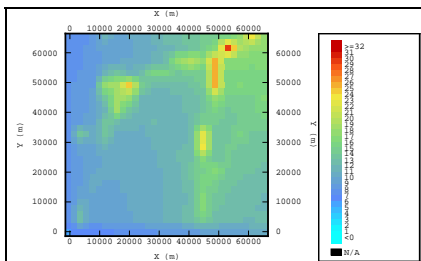
count	281
min	5.91
max	28.01
mean	12.34
std.dev	3.53
var	12.46

cross validation

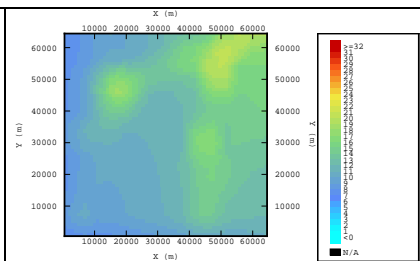


Résultats de la validation croisée

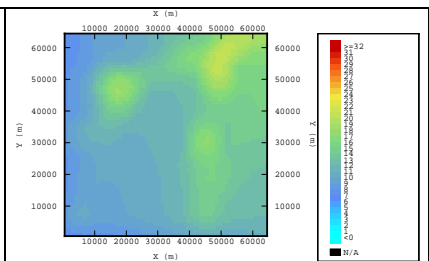
Statistiques



Carte de krigeage avec variogramme modélisé (domaine 17x17)

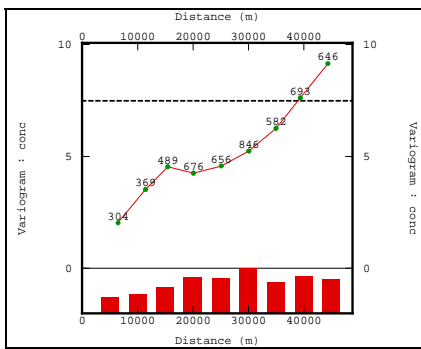


Carte de krigeage pour la validation absolue

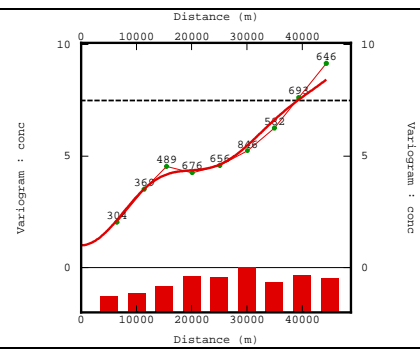


Carte de krigeage pour la validation absolue avec variogramme de ref

**Etude de l'échantillonnage sur une grille de 11x11 losange au pas de 6km**



Variogramme expérimental

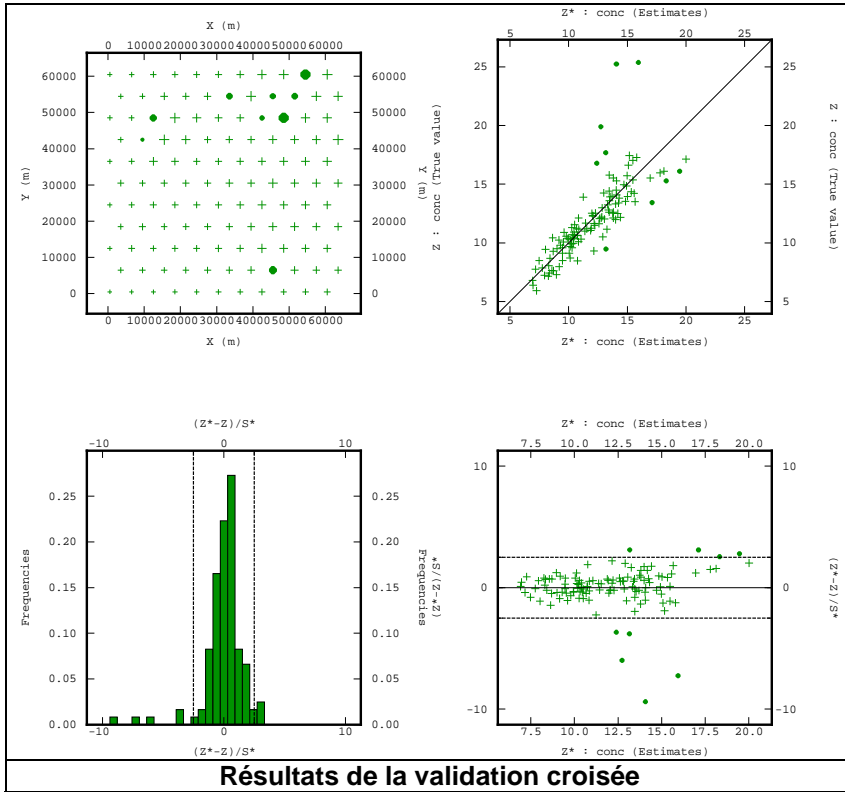


Variogramme modélisé

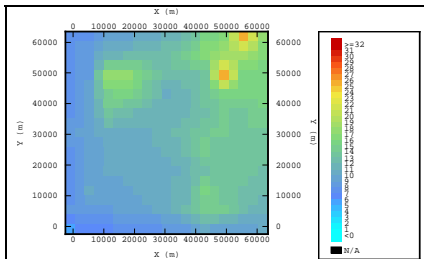
Pépité(1)+ sincard(70000 ;2)+power(40000 ;1.9 ;4.5)

count	121
min	5.91
max	25.38
mean	12.04
std.dev	3.29
var	10.82

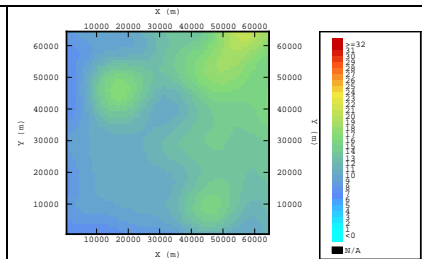
cross validation



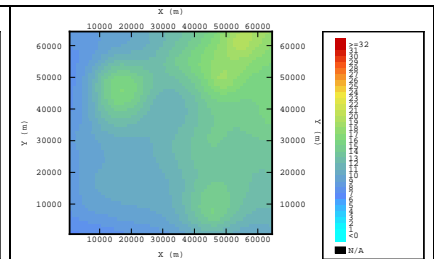
**Statistiques**



Carte de krigeage avec variogramme modélisé (domaine 11x11)

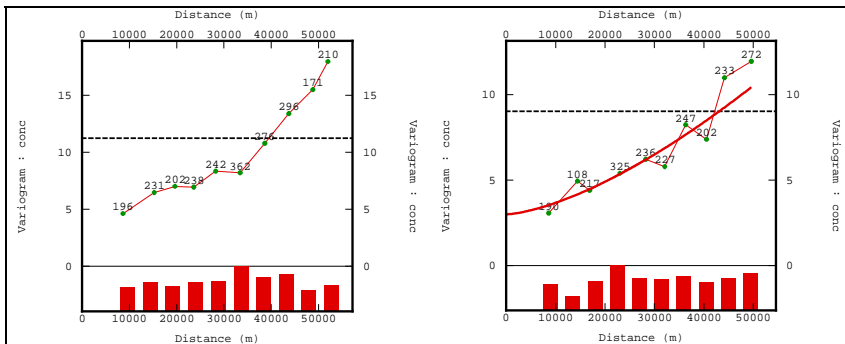


Carte de krigeage pour la validation absolue



Carte de krigeage pour la validation absolue avec variogramme de ref

**Etude de l'échantillonnage sur une grille de 9x9 losange au pas de 8km**

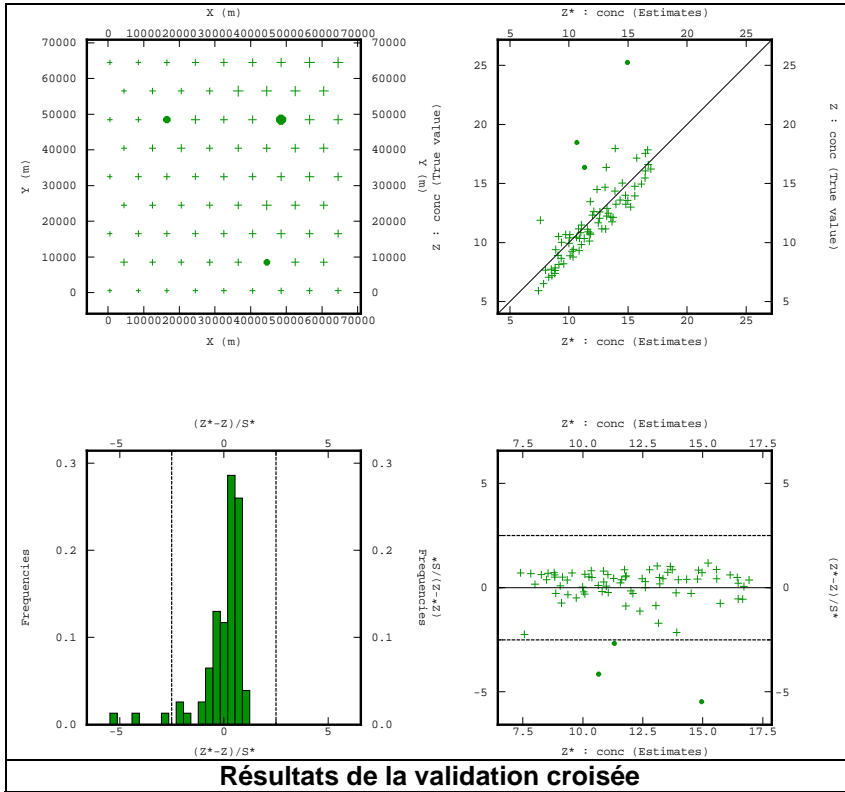


**Variogramme expérimental**

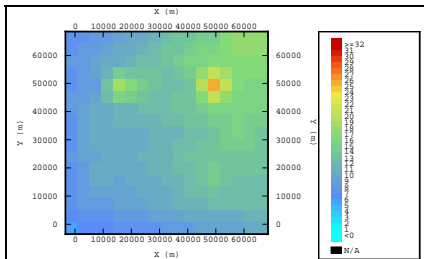
**Variogramme modélisé**  
Pépite(3)+ power(30000 ;1.5 ;3.5)

count	77
min	5.91
max	25.25
mean	11.96
std.dev	3.35
var	11.23

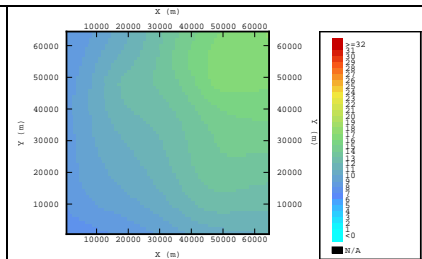
cross validation



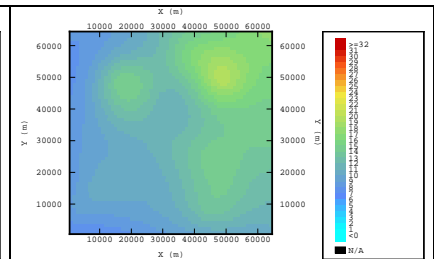
**Statistiques**



Carte de krigeage avec variogramme modélisé (domaine 9x9)

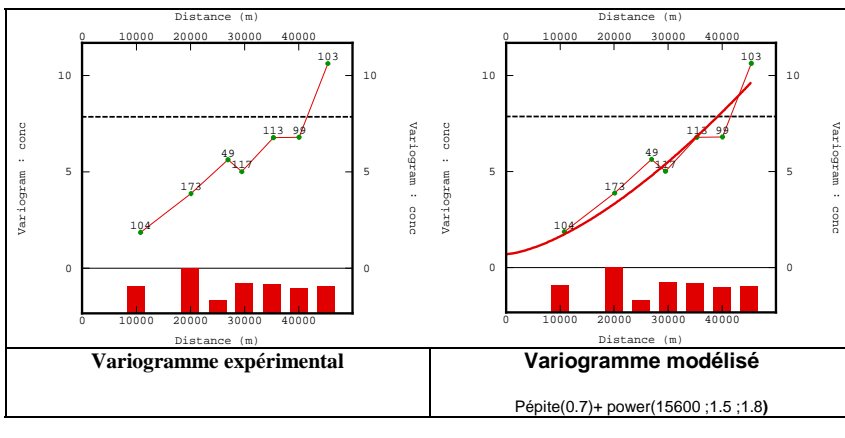


Carte de krigeage pour la validation absolue



Carte de krigeage pour la validation absolue avec variogramme de ref

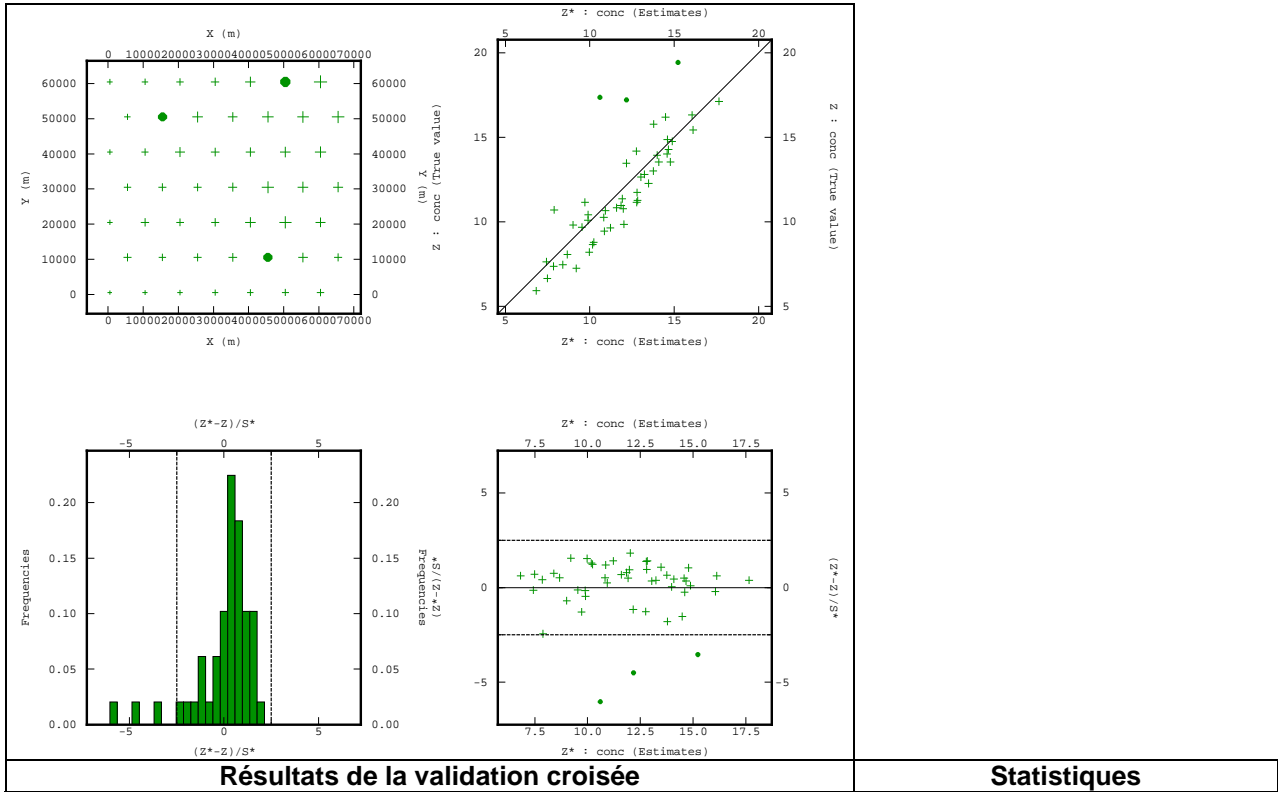
**Etude de l'échantillonnage sur une grille de 7x7 losange au pas de 10km**



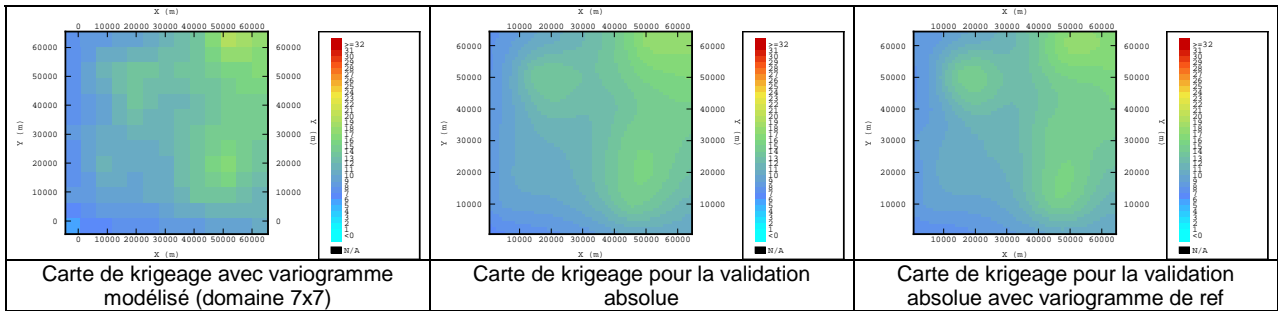
count	49
min	5.91
max	16.43
mean	11.79
std.dev	3.15
var	9.95

cross validation



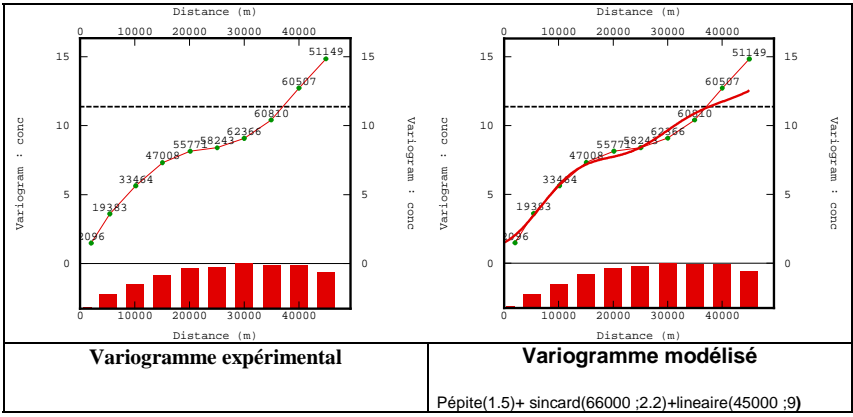


**Statistiques**



**Fiches résultantes concernant les travaux de la campagne 2  
(3 au 17 avril 2003)**

**Etude de l'échantillonnage sur la grille de référence au pas de 2km**



count	1089
min	10.76
max	32.31
mean	18.56
std.dev	3.65
var	11.38

cross validation

error

mean	0.00182
var	2.7167

std. error

mean	0.00063
var	1.3536

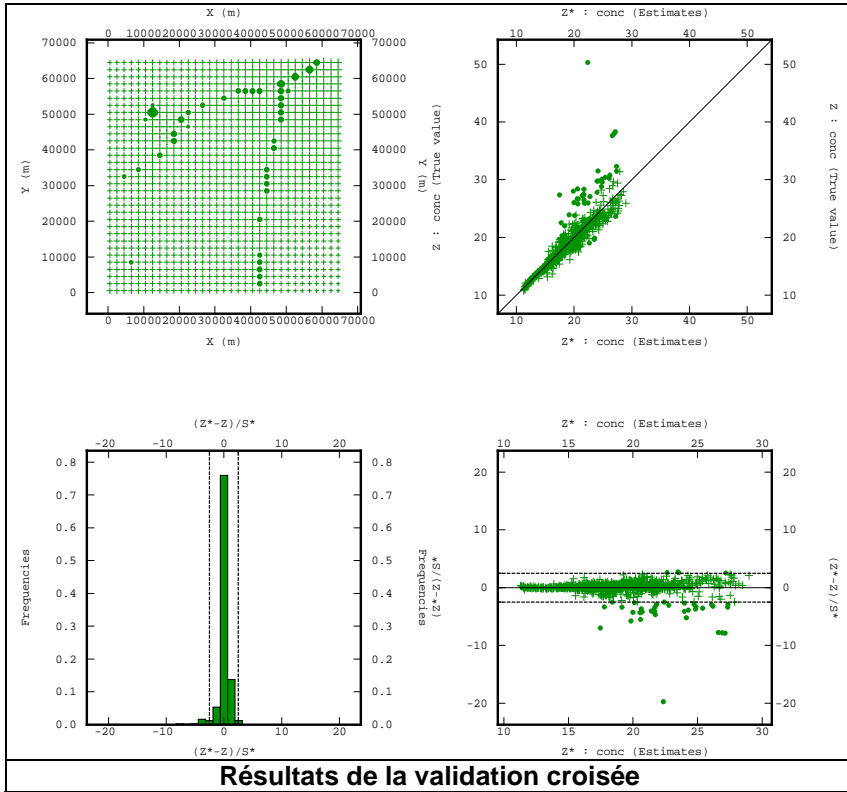
validation absolue

error

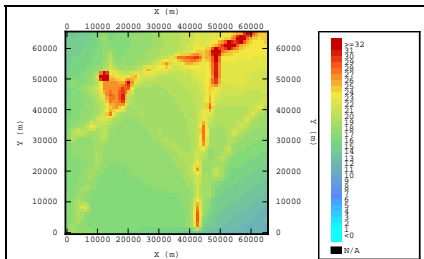
mean	0.0422
var	1.3658

std. error

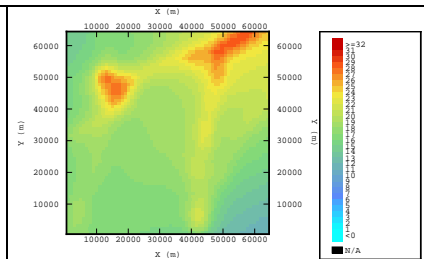
mean	0.0304
var	0.706



**Statistiques**

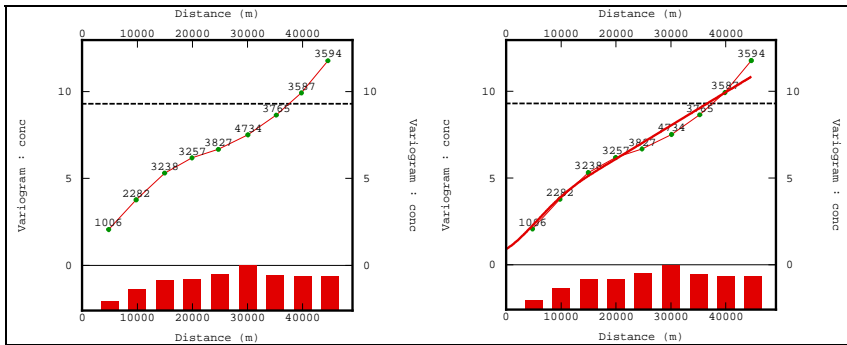


Carte de krigeage avec variogramme modélisé (domaine 33x33)



Carte de krigeage pour la validation absolue

**Etude de l'échantillonnage sur une grille de 17x17 carre au pas de 4km**



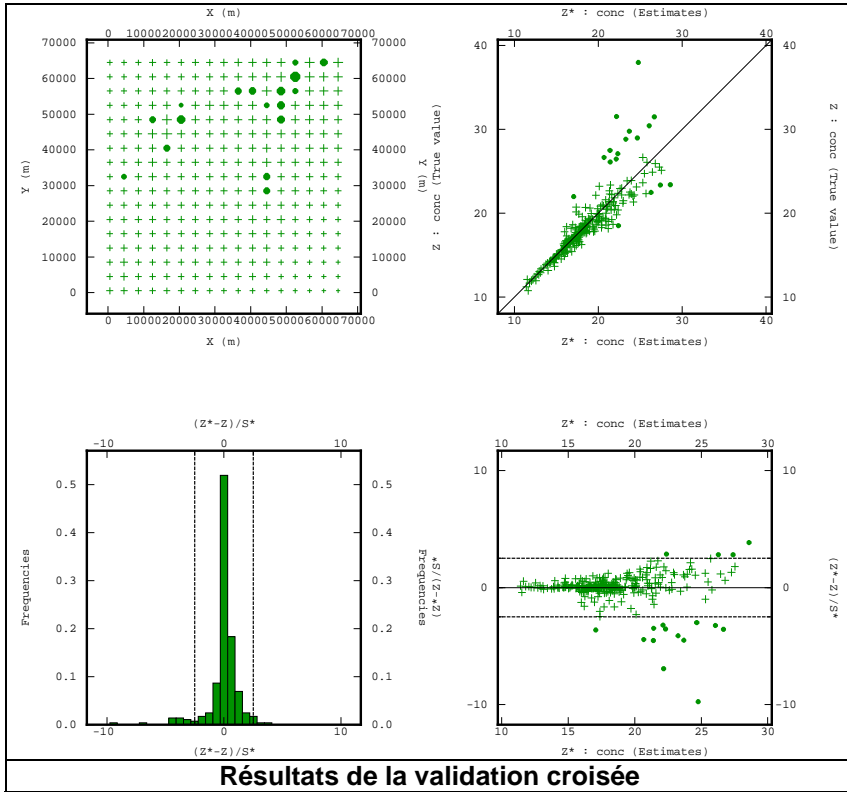
**Variogramme expérimental**

**Variogramme modélisé**

Pépité(0.9)+ sincard(18600 ;1.35)+power(13000 ;1 ;2.5)

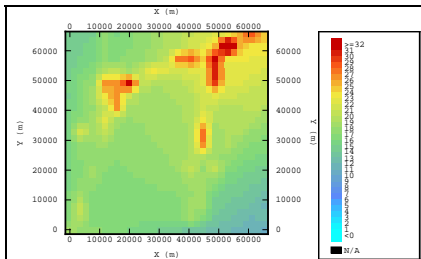
count	289
min	10.76
max	27.5
mean	18.14
std.dev	3.66
var	9.3

cross validation

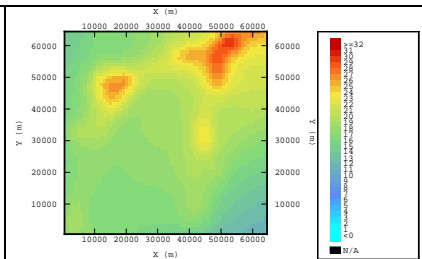


Résultats de la validation croisée

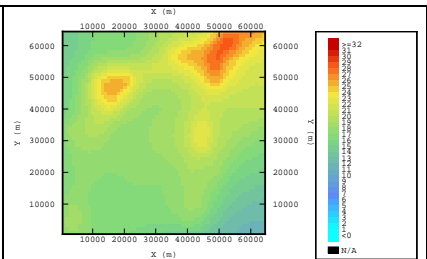
Statistiques



Carte de krigage avec variogramme modélisé (domaine 17x17)

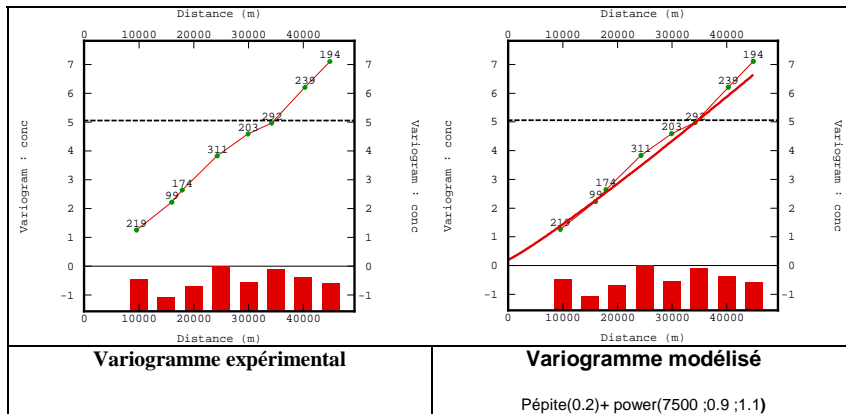


Carte de krigage pour la validation absolue

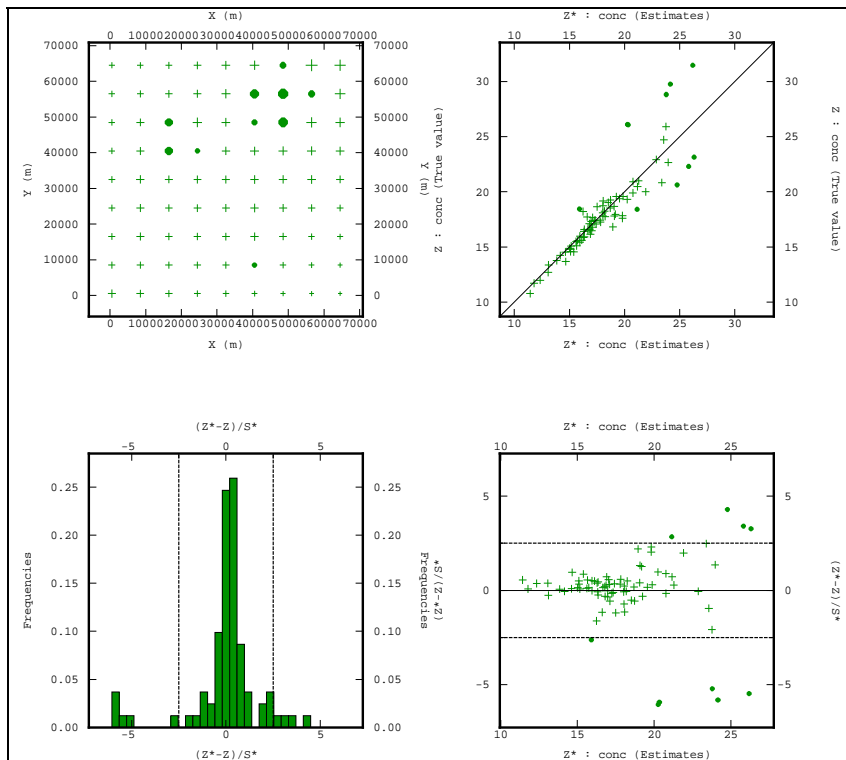


Carte de krigage pour la validation absolue avec variogramme de ref

# Etude de l'échantillonnage sur une grille de 9x9 carre au pas de 8km



count	81
min	10.76
max	20.99
mean	16.99
std.dev	3.83
var	5.05



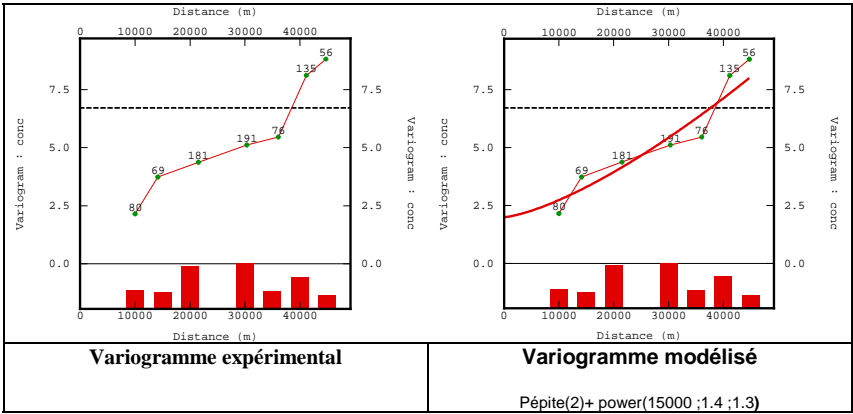
cross validation	
error	mean 0.00345
	var 3.187
std. error	mean 0.0022
	var 3.33

validation absolue	
error	mean -0.1059
	var 3.30
std. error	mean -0.1286
	var 5.050

utilisation  
variogramme de ref

error

**Etude de l'échantillonnage sur une grille de 7x7 carre au pas de 10km**



count	49
min	11.24
max	29.4
mean	17.92
std.dev	3.15
var	9.93

cross validation

error

mean  
0.02481

var  
3.27754

std. error

mean  
0.00729

var  
1.13414

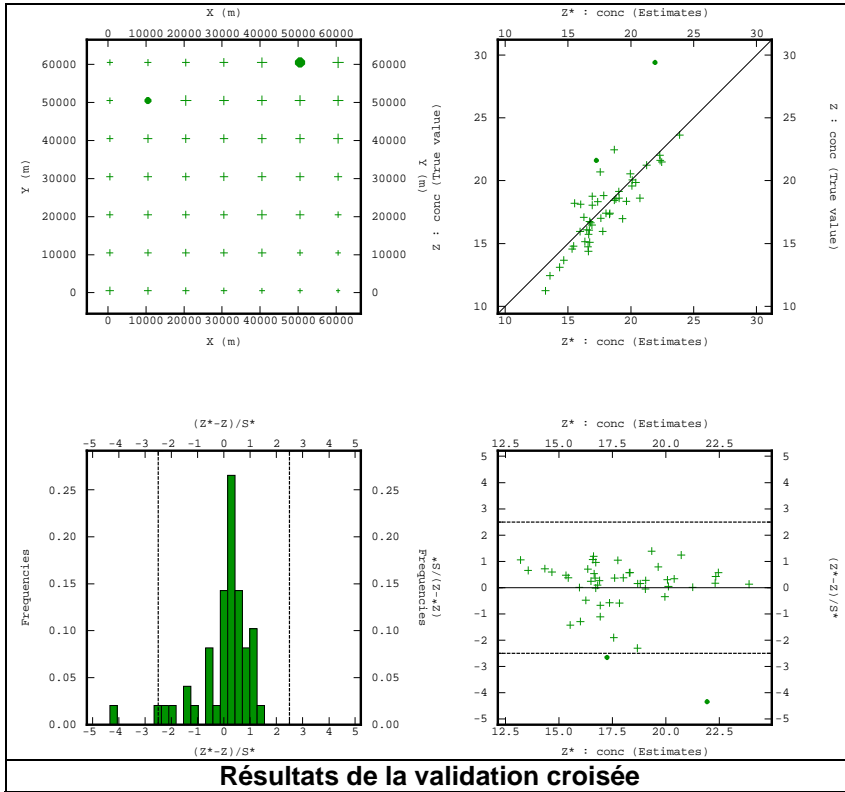
validation absolue

error

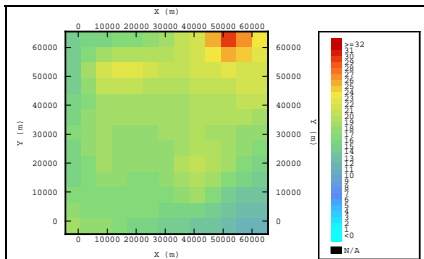
mean  
-0.41

var  
4.3761

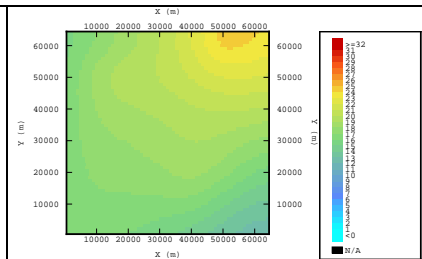
std. error



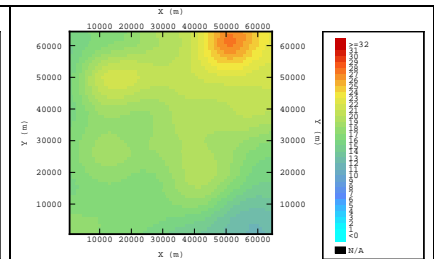
**Statistiques**



Carte de krigeage avec variogramme modélisé (domaine 7x7)

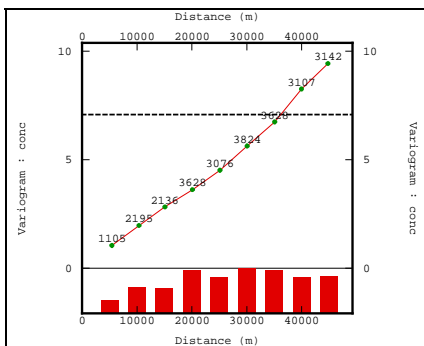


Carte de krigeage pour la validation absolue

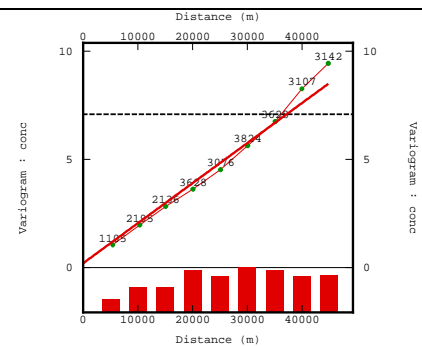


Carte de krigeage pour la validation absolue avec variogramme de ref

**Etude de l'échantillonnage sur une grille de 17x17 losange au pas de 4km**



Variogramme expérimental

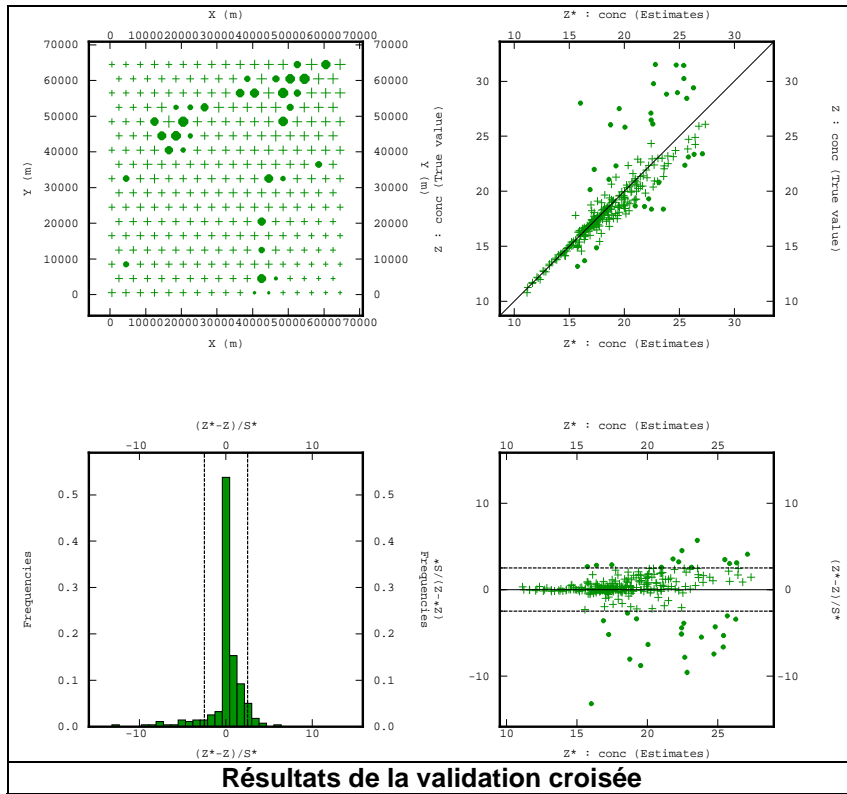


Variogramme modélisé

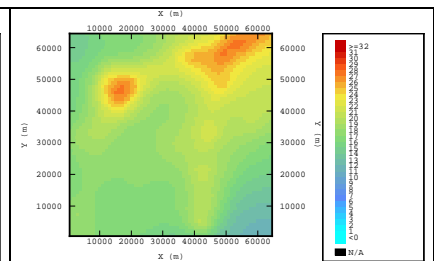
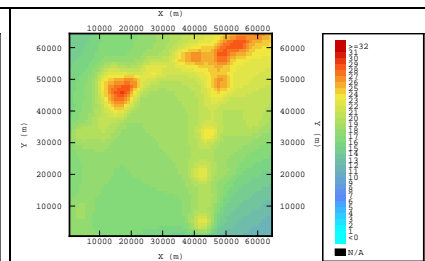
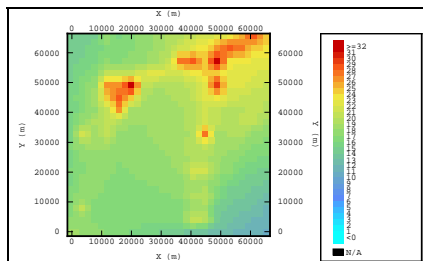
Pépite(0.2)+lineaire(10000 ;1.85)

count	281
min	10.76
max	31.52
mean	18.58
std.dev	3.65
var	13.33

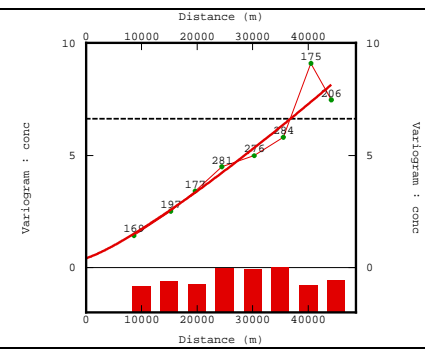
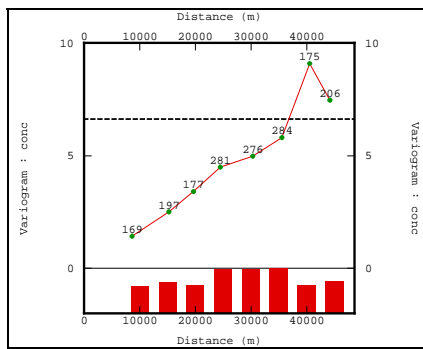
cross validation



**Statistiques**



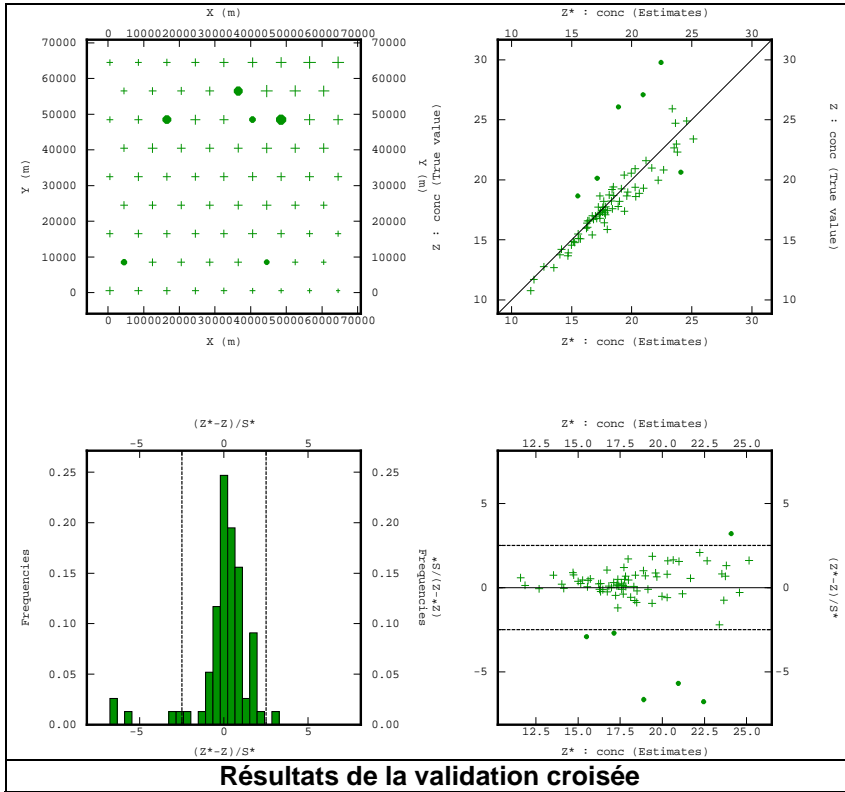
**Etude de l'échantillonnage sur une grille de 9x9 losange au pas de 8km**



count	77
min	10.76
max	29.77
mean	18.25
std.dev	3.46
var	11.96

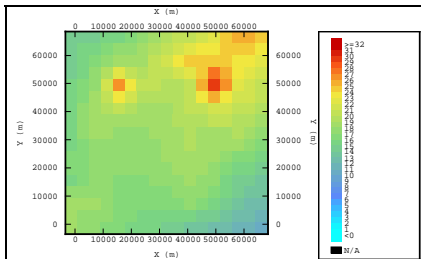
cross validation



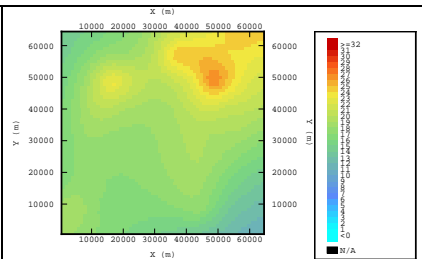


Résultats de la validation croisée

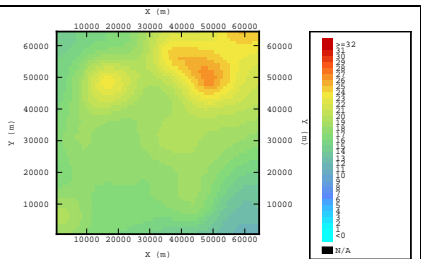
Statistiques



Carte de krigeage avec variogramme modélisé (domaine 9x9)

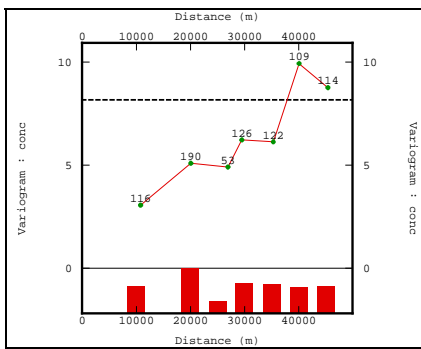


Carte de krigeage pour la validation absolue

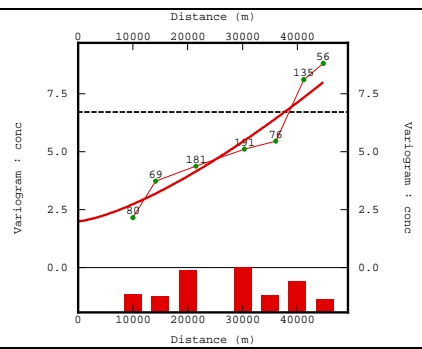


Carte de krigeage pour la validation absolue avec variogramme de ref

**Etude de l'échantillonnage sur une grille de 7x7 losange au pas de 10km**



Variogramme expérimental

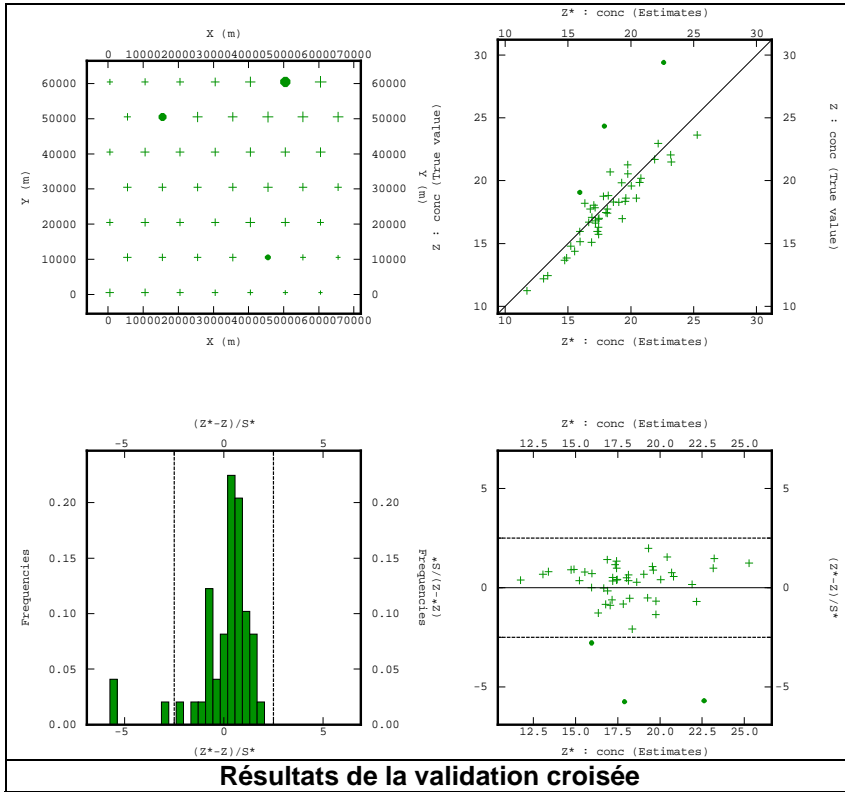


Variogramme modélisé

Pépité(2)+ power(10000 ; 1.25 ; 1)

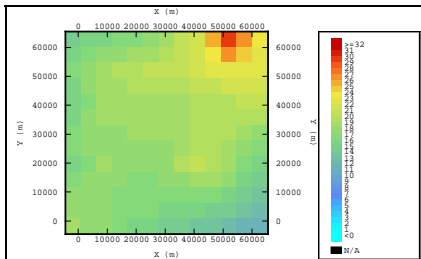
count	49
min	11.24
max	29.4
mean	18.08
std.dev	3.27
var	10.67

cross validation

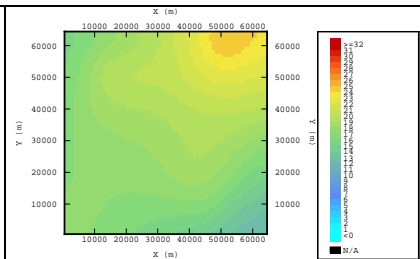


Résultats de la validation croisée

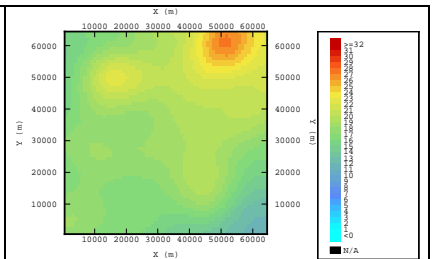
Statistiques



Carte de krigeage avec variogramme modélisé (domaine 7x7)



Carte de krigeage pour la validation absolue



Carte de krigeage pour la validation absolue avec variogramme de ref

## **THEME : TRAITEMENTS NUMERIQUES**

### ***Adaptation des plans d'échantillonnage spatial aux objectifs de campagne***

#### **OBJECTIF**

Proposer des recommandations de plans d'échantillonnage spatial, adaptés aux différents objectifs possibles des campagnes de mesure (validation de points fixes, surveillance en zones rurales et dans les agglomérations de moins de 100 000 habitants, analyse objective).

#### **CONTEXTE**

La définition et la mise en place des PSQA (Plans de Surveillance de la Qualité de l'Air), ainsi que la disponibilité de méthodes d'évaluation de la qualité de l'air complémentaires à la mesure (analyse objective, géostatistiques, modélisation déterministe...) ont engendré un véritable besoin relatif à la définition de méthodes de « référence » pour l'élaboration de plans d'échantillonnage.

Les questions posées par les AASQA concernent aussi bien les critères de dimensionnement de campagnes de mesure que les bonnes pratiques de traitement de données. Il est nécessaire d'aborder ces points en fonction de l'objectif de la campagne.

#### **TRAVAUX PROPOSES POUR 2005**

Le programme proposé en 2005 a pour objectif de tester, et d'évaluer des plans d'échantillonnage spatial, pour différents objectifs de surveillance, et de déterminer les stratégies les mieux adaptées. Ces travaux peuvent être considérés comme une « annexe technique » du guide des aides de l'ADEME en matière d'équipements pour la surveillance qui propose de premières orientations.

Il est proposé de s'intéresser plus particulièrement aux objectifs de surveillance suivants :

##### Dans un premier temps (2005) :

- surveillance des zones rurales non couvertes par les stations de mesures
- surveillance des villes de 10 000 à 100 000 habitants
- implantation d'un nouveau site de mesure

##### Dans un second temps (2006) :

- alimentation des modèles (assimilation de données) et validation de modèles

Il est évident que cette analyse devra être menée polluant par polluant.

Pour chacun de ces objectifs et en fonction des polluants étudiés, le plan d'étude est le suivant :

- présentation et analyse de différents plans d'échantillonnage spatial : recensement des données et études disponibles au sein des AASQA, analyse des techniques mises en œuvre,
- tests effectués sur ces plans d'échantillonnage, notamment à partir de données simulées à une échelle fine, et ce pour différentes situations géographiques et météorologiques,

- réalisation de tests complémentaires à partir de campagnes de tubes passifs disponibles, pour les polluants pertinents (NO<sub>2</sub>, en particulier),
- sélection des stratégies les plus adéquates.

Ces travaux seront menés en concertation et en suivant les recommandations du GT « échantillonnage » qui devrait être ré-activé en 2005. Ils seront également précédés d'une analyse bibliographique des méthodes mises en œuvre en France et en Europe, qui servira de point de départ aux études menées par la suite.

## **DUREE DES TRAVAUX**

Les travaux proposés s'étendront sur deux ans. Ils seront menés conjointement par l'EMD et l'INERIS.

La première année pourra être consacrée au traitement du NO<sub>2</sub>. Pour le NO<sub>2</sub>, les travaux et données disponibles sur la région Nord et le Centre (voir travaux de l'EMD 2002-2004) serviront de base de travail.

Le cas de l'ozone, des BTX, du SO<sub>2</sub> et des PM sera traité durant la deuxième année.

A l'issue des deux années d'étude, un guide de recommandations sera proposé aux AASQA.

## **COLLABORATIONS**

AASQA , ADEME