

Représentation de la qualité de l'air dans les zones peu/pas couvertes par les stations de mesure fixes : partie II, application à la problématique d'une association

Décembre 2002

Giovanni CARDENAS
Laure MALHERBE

Ce document comporte 59 pages (hors couverture et annexes).

	Rédaction	Vérification	Approbation
NOM	Giovanni CARDENAS Laure MALHERBE	Claude MEUNIER Laurence ROUÏL	Michel NOMINE
Qualité	Ingénieurs Etudes et Recherches Direction des Risques Chroniques	Responsable de l'unité 2IEN Ingénieur Etudes et Recherches Direction des Risques Chroniques	Délégué scientifique Direction des Risques Chroniques
Visa			

<u>1.</u>	<u>INTRODUCTION</u>	4
<u>2.</u>	<u>LE DOMAINE D'ÉTUDE - LES DONNÉES</u>	5
2.1	<u>DONNÉES DE STATIONS FIXES EN MILIEU URBAIN</u>	5
2.2	<u>DONNÉES DE MOYENS MOBILES</u>	5
2.3	<u>DONNÉES D'ÉCHANTILLONNEURS PASSIFS</u>	5
2.4	<u>DONNÉES AUXILIAIRES</u>	5
<u>3.</u>	<u>ÉTUDE MONOVARIABLE : ANALYSE EXPLORATOIRE DES DONNÉES</u>	7
3.1	<u>EXAMEN PRÉLIMINAIRE DES DONNÉES</u>	7
3.2	<u>COMPARAISON AVEC LES DONNÉES DE STATIONS FIXES – ÉTUDE DE L'INCERTITUDE DE MESURE</u>	10
3.3	<u>STATISTIQUES ÉLÉMENTAIRES ET ANALYSE STRUCTURALE DES SIX SERIES DES MESURES SELECTIONNEES ET DE LA MOYENNE DES QUATORZE SEMAINES</u>	12
3.3.1	<i>Statistiques élémentaires des six semaines de mesure sélectionnées</i>	12
3.3.2	<i>Dérives spatiales</i>	14
3.3.3	<i>Analyse variographique : variogrammes expérimentaux et modélisation du variogramme</i>	17
3.3.4	<i>Validation croisée</i>	25
<u>4.</u>	<u>ÉTUDE MONOVARIABLE : ESTIMATION PAR KRIGEAGE DE BLOCS</u>	29
4.1	<u>ESTIMATION AVEC MODÈLES ISOTROPES</u>	29
4.2	<u>ESTIMATION AVEC MODELES ANISOTROPES</u>	33
<u>5.</u>	<u>ESTIMATION AVEC VARIABLES AUXILIAIRES</u>	39
5.1	<u>CHOIX DES VARIABLES</u>	39
5.2	<u>ANALYSE DE LA VARIABLE DENSITÉ DE POPULATION</u>	40
5.3	<u>ANALYSE DES SURFACES BÂTIES DE L'ALLIER</u>	42
5.4	<u>ANALYSE DE LA VARIABLE DENSITÉ DE SURFACE BOISÉE</u>	44
5.4.1	<i>Analyse de la zone 1</i>	45
5.4.2	<i>Analyse de la zone 2</i>	45
5.4.3	<i>Analyse de la zone 3</i>	45
5.4.4	<i>Analyse de la zone 4</i>	45
5.5	<u>ANALYSE DU RELIEF</u>	47
5.6	<u>MODÉLISATION MULTIVARIABLE DE L'OZONE</u>	51
<u>6.</u>	<u>CONCLUSION</u>	58
<u>7.</u>	<u>RÉFÉRENCES</u>	59

1. RESUME

Cette étude propose un exemple d'utilisation des méthodes de la géostatistique pour la cartographie de la pollution dans de vastes territoires peu équipés en stations de mesure fixes.

Les données exploitées ont été fournies par l'association ATMO Auvergne et sont issues d'une campagne de mesure de l'ozone réalisée dans l'Allier au cours de l'été 2001. Elles sont enrichies par des données de variables auxiliaires qui caractérisent l'altitude, la population et les zones bâties et boisées de ce département.

L'étude géostatistique a porté principalement sur les données des échantillonneurs passifs. Un échantillonnage selon un maillage régulier (chaque maille ayant 20 km de côté) a permis en effet de mesurer l'ozone dans tout le département pendant quatorze semaines consécutives (de juin à septembre 2001). Six semaines de mesure ont été plus particulièrement analysées afin de mettre en évidence l'évolution des caractéristiques spatiales de l'ozone durant la saison estivale. La moyenne des quatorze semaines de mesure, représentative de la pollution moyenne estivale d'ozone, a été également étudiée.

Une analyse structurale a permis de modéliser la structure de corrélation spatiale de l'ozone pour les différentes périodes considérées, à la suite de quoi les concentrations d'ozone ont été estimées par krigeage ordinaire sur des blocs de 5 km de côté. Deux types de modèles ont été élaborés : des modèles omnidirectionnels qui supposent que la variabilité du phénomène est identique quelle que soit la direction de l'espace, et des modèles bidirectionnels qui tiennent compte des différences entre les directions. Les résultats obtenus dans les deux cas sont proches. Une étude plus approfondie des anisotropies pourrait être conduite par comparaison avec des données météorologiques. Il importe de noter que les modèles évoluent dans le temps et qu'une même structure n'a pu être modélisée pour toutes les semaines.

Dans une deuxième étape, les variables auxiliaires disponibles ont été étudiées dans le but d'améliorer la précision des cartographies. Deux variables explicatives, l'altitude et la densité de population, ont été sélectionnées et combinées entre elles. La nouvelle variable auxiliaire ainsi créée, corrélée linéairement à la variable de concentration, a été prise pour dérive externe et introduite dans le krigeage. Les nouvelles cartes d'estimation sont comparables aux cartes précédemment réalisées, mais elles sont plus détaillées et font mieux ressortir les variations des concentrations en fonction du relief.

Cette étude confirme l'importance de l'échantillonnage. Le maillage adopté est approprié et permet de caractériser la structure spatiale de l'ozone dans le département. Un échantillonnage plus resserré en certains endroits et la pose de tubes multiples serait utile pour affiner le modèle ajusté aux courtes distances et prendre éventuellement en compte l'erreur de mesure. La structure spatiale de l'ozone varie au cours des quatorze semaines de mesure, de là l'intérêt d'étendre l'échantillonnage à plusieurs semaines afin de caractériser la pollution moyenne saisonnière.

La question de l'incertitude des cartes, partiellement abordée dans ce travail, sera plus particulièrement développée en 2003.

2. INTRODUCTION

Cette étude a pour objet d'appliquer les méthodes de cartographie décrites dans la partie théorique¹ au cas particulier d'une Association, afin d'en apprécier la pertinence et d'illustrer la démarche par laquelle on peut extraire des mesures et des données auxiliaires une information spatiale et temporelle sur la qualité de l'air.

Le département de l'Allier, zone principalement rurale et équipée d'un petit nombre de stations fixes, entre complètement dans le cadre de cette étude et a été pris pour champ d'application. Des campagnes de mesure de l'ozone y ont été conduites pendant l'année 2001, ce qui permet de disposer de jeux de données sur ce département. Ceux-ci ont été fournis à l'INERIS par ATMO Auvergne. La région Midi-Pyrénées, par son étendue, la présence de petites et moyennes agglomérations et l'importance des zones rurales devait initialement servir de domaine d'étude pour la cartographie du NO₂, mais des difficultés techniques n'ont pas permis à l'ORAMIP de délivrer les données dans les délais requis.

Dans cette étude, nous nous attachons essentiellement à **la mise en œuvre des techniques de la géostatistique**. Le but est de fournir, à l'aide de ces outils, une représentation cartographique de l'ozone pendant la saison estivale dans le département de l'Allier. La dynamique de l'ozone à très court terme n'est pas concernée par le présent travail.

Ce rapport aborde les différentes étapes de l'étude géostatistique :

1. analyse exploratoire des données,
2. modélisation de la structure spatiale de l'ozone et estimation des concentrations,
3. analyse des corrélations avec les variables auxiliaires et utilisation de ces variables dans la cartographie.

Il met en évidence pour chacune de ces étapes les points les plus importants ou les plus délicats. Il a été conçu dans un but pédagogique, afin de permettre aux AASQA n'ayant que peu d'expérience des méthodes de géostatistique d'appréhender cette problématique.

¹ Rapport INERIS/LCSQA 2002. Estimation de la qualité de l'air dans les zones peu/pas couvertes. Partie 1 : stratégies et préconisations

3. LE DOMAINE D'ETUDE - LES DONNEES

Le champ de l'étude couvre le département de l'Allier ; il s'étend sur une longueur de 125 km et une largeur voisine de 100 km (Figure 1). Le jeu de données fourni à l'INERIS par ATMO Auvergne résulte des mesures **d'ozone** effectuées dans ce département au cours de l'année 2001 et il est complété par des données de variables auxiliaires. Il se compose de :

3.1 DONNEES DE STATIONS FIXES EN MILIEU URBAIN

Situées dans l'agglomération de Montluçon, une station urbaine (Montluçon Château) et une station périurbaine (Montluçon Hippodrome) ont mesuré continûment l'ozone pendant toute l'année 2001. Les données de la station fixe rurale de Busset sont disponibles à partir du 12 juin 2001, date à laquelle cette station a été mise en service.

3.2 DONNEES DE MOYENS MOBILES

Une remorque équipée d'analyseurs a stationné pendant dix à quinze jours en différents points des environs de Moulins et de Vichy. Ces points sont au nombre de six, trois pendant la saison hivernale (décembre 2000 à février 2001) et trois pendant la saison estivale (mai - juin 2001). Un camion laboratoire a été installé pendant trois semaines à Montluçon (du 26 juin au 17 juillet 2001).

3.3 DONNEES D'ECHANTILLONNEURS PASSIFS

Une campagne de mesure de l'ozone au moyen de tubes à échantillonnage passif a été conduite dans tout le département du 18 juin 2001 au 24 septembre 2001. Vingt-neuf sites ont fait l'objet d'un échantillonnage durant quatorze semaines consécutives. Des mesures ont été effectuées en trois autres sites pendant quatre semaines. Les données de concentration d'ozone issues de cette campagne sont disponibles semaine par semaine (valeurs hebdomadaires) et en moyenne sur la période étudiée.

La disposition des points d'échantillonnage suit un maillage régulier, de pas égal à 20 km. Cet échantillonnage est resserré dans les zones urbanisées (Montluçon, Moulins, Vichy, Figure 1).

3.4 DONNEES AUXILIAIRES

Ces données proviennent de l'INSEE. Elles apportent de nombreuses informations sur les communes qui composent l'Allier (superficie, population, altitude), la situation et la superficie des zones bâties et boisées, le réseau routier.

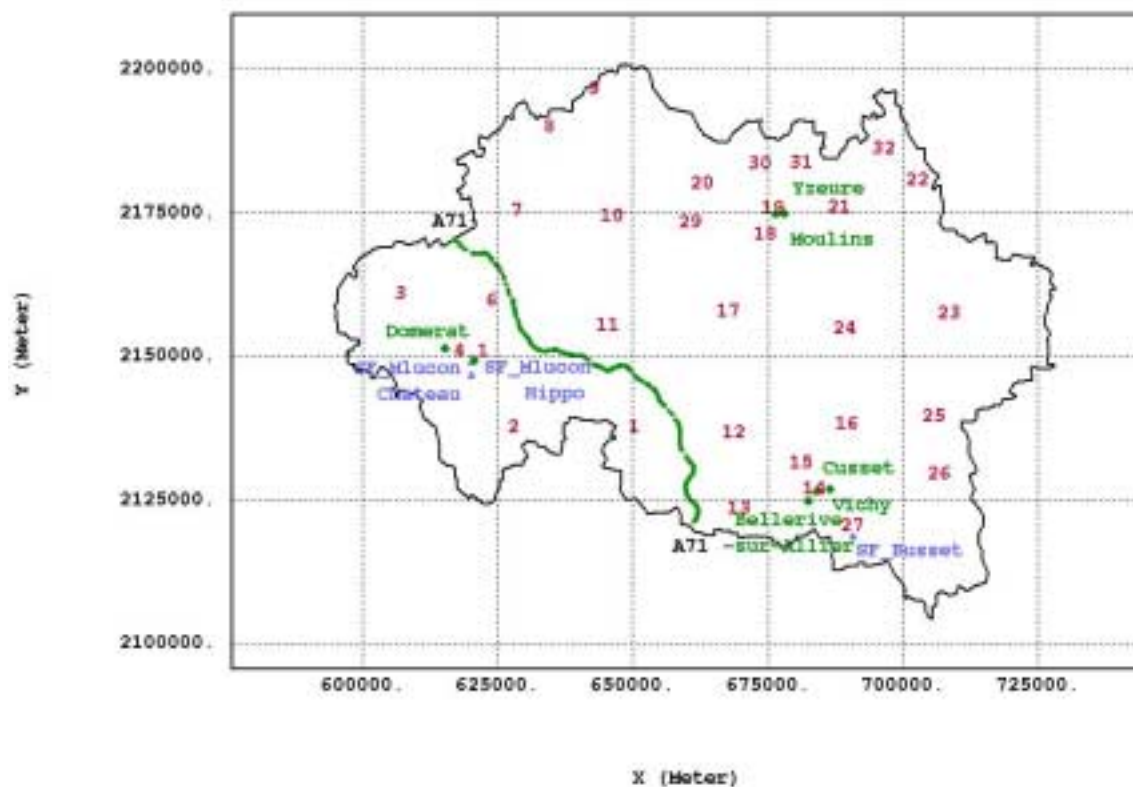


Figure 1 : Département de l'Allier. Localisation des points de mesure. Les coordonnées sont en Lambert II étendu.

4. ETUDE MONOVARIABLE : ANALYSE EXPLORATOIRE DES DONNEES

Cette étape vise à dégager les caractéristiques spatiales de la pollution d'ozone dans l'Allier. Les résultats de cette analyse guideront le choix des méthodes d'estimation à employer. Les données traitées sont les données des échantillonneurs passifs (campagne de mesure du 18 juin au 24 septembre 2001).

Les données de moyens mobiles ne sont pas employées. Elles couvrent une partie de la saison hivernale ainsi que quelques semaines estivales antérieures à la campagne d'échantillonnage passif. On pourrait envisager d'en faire usage et d'exploiter aussi les données de stations fixes afin d'extrapoler à d'autres périodes les informations issues des mesures de tubes. Cela nécessite la mise en œuvre d'une méthodologie adaptée qui pourra faire l'objet de développements futurs.

4.1 EXAMEN PRELIMINAIRE DES DONNEES

L'évolution des concentrations d'ozone au cours de la saison estivale (Tableau 1 et Figure 2) est liée à celle des températures, comme cela a été mis en évidence dans le rapport de campagne d'ATMO Auvergne [1]. D'un point de vue spatial, les concentrations dans les villes sont toujours plus faibles que les concentrations en milieu rural.

Numéro de la semaine	DATES	Nombre de points	Minimum ($\mu\text{g}/\text{m}^3$)	Médiane ($\mu\text{g}/\text{m}^3$)	Maximum ($\mu\text{g}/\text{m}^3$)	Moyenne ($\mu\text{g}/\text{m}^3$)	Ecart-type ($\mu\text{g}/\text{m}^3$)
1	18-25 Juin	25	48.30	90.70	125.80	88.94	15.89
2	25-2 Juillet	32	60.00	82.70	113.40	85.12	13.09
3	2-9 Juillet	32	68.40	85.95	114.30	89.25	12.17
4	9-16 Juillet	32	40.60	64.95	87.80	65.61	10.75
5	16-23 Juillet	31	60.00	71.20	95.70	73.98	9.08
6	23-30 Juillet	28	73.60	91.75	123.00	95.82	14.21
7	30 -6 Août	29	72.90	93.60	139.30	98.45	16.55
8	6-13 Août	29	53.50	76.70	101.30	76.26	12.46
9	13-20 Août	29	52.70	74.50	100.10	74.32	12.17
10	20-27 Août	29	54.40	77.40	111.30	76.93	17.06
11	27-03 Sept	29	62.80	74.00	99.30	76.68	11.39
12	03-10 Sept	29	46.20	59.90	87.20	61.70	10.83
13	10-17 Sept	29	36.40	50.60	88.40	56.13	12.91
14	17-24 Sept	29	42.00	61.00	85.20	62.34	10.62
	Moyenne	32	60.36	75.39	101.01	76.76	10.06

Tableau 1 : Statistiques par série de mesure. Les concentrations maximales sont atteintes pendant la semaine 7, les concentrations minimales, pendant la semaine 13.

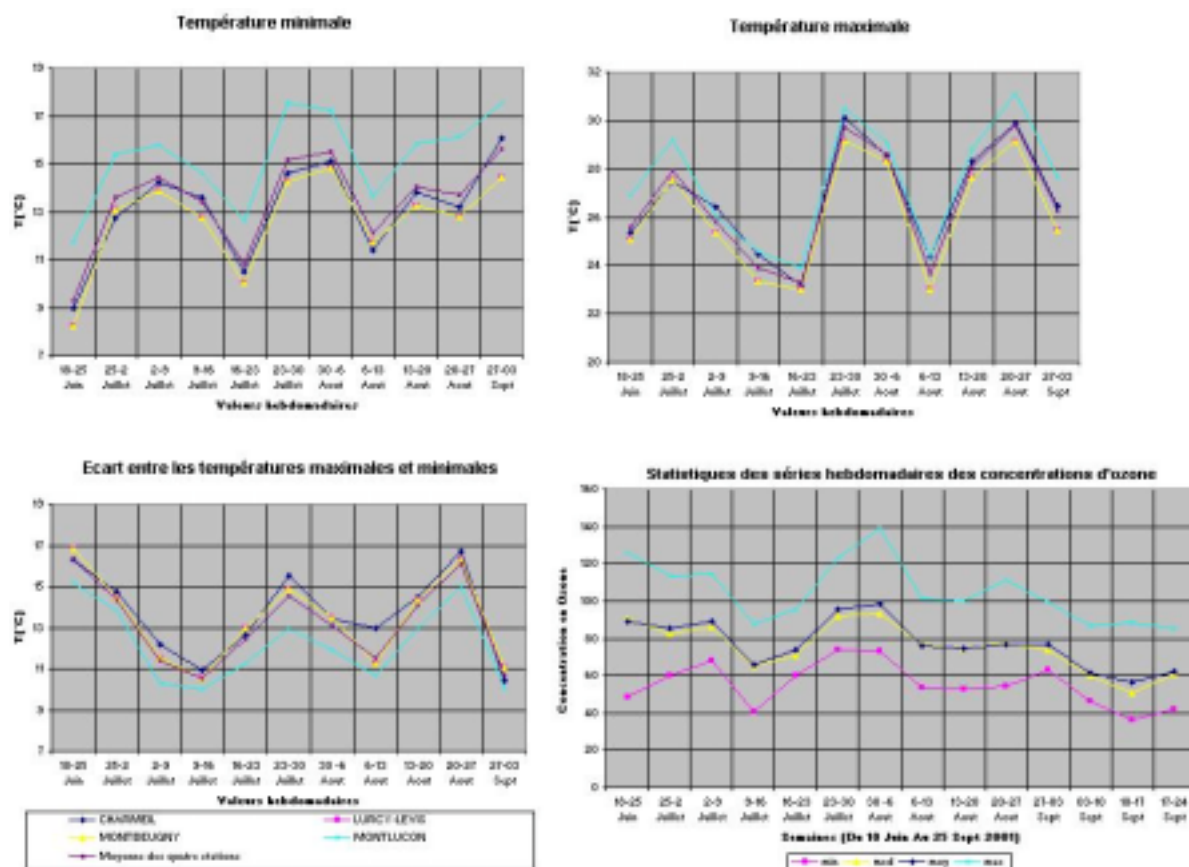


Figure 2 : Evolution des températures et des concentrations d'ozone pendant la campagne de mesure.

Ces observations sont complétées par une **analyse en composantes principales (ACP)**, destinée à mettre en évidence des similarités ou des oppositions entre les semaines de mesure et à repérer les variables les plus corrélées à la pollution d'ozone (Figure 3). Cette méthode consiste à remplacer une famille de variables par de nouvelles variables de variance maximale, non corrélées deux à deux et qui sont des combinaisons linéaires des variables d'origine. Ces nouvelles variables, appelées *composantes principales*, définissent des plans factoriels qui servent de base à une représentation graphique plane des variables initiales.

Les variables introduites dans l'analyse en composantes principales sont la concentration moyenne sur toute la période de mesure, les quatorze concentrations hebdomadaires, les coordonnées de l'espace x et y, l'altitude et la densité de population. Les valeurs de ces variables sont connues en au moins 25 des 32 points d'échantillonnage. En raison de leur hétérogénéité, ces variables sont préalablement centrées et réduites.

Le diagramme en haut à droite de la Figure 3 indique le pourcentage de l'inertie totale du nuage des variables porté par chaque axe factoriel. Ce pourcentage est respectivement de 66%, 10% et 5% pour les trois premiers axes qui sont retenus pour la représentation graphique.

En ACP normée, les variables projetées sur chaque plan factoriel se trouvent à l'intérieur d'un cercle de rayon unité. Plus une variable est projetée vers le bord du cercle, mieux elle est représentée. Par ailleurs, deux variables bien représentées et proches l'une de

l'autre sont corrélées positivement tandis que deux variables qui s'opposent sont corrélées négativement. Une orthogonalité entre deux variables traduit une absence de corrélation.

La première composante principale sépare les variables de concentration (concentration moyenne, minimale et maximale et concentrations des quatorze semaines) de la densité de population. Compte tenu des différences de concentration entre ville et campagne, cette opposition était prévisible. On observe de plus une corrélation positive entre la concentration d'ozone et l'altitude. Aucune corrélation n'apparaît avec les coordonnées de l'espace.

Les autres composantes sont essentiellement liées à la position géographique des tubes passifs, représentée par les coordonnées X et Y, et à la densité de population. Cela se vérifie aisément en examinant le plan F2-F3 dans lequel toutes les variables de concentration sont projetées vers le centre du cercle.

Il faut noter que le nuage des variables de concentration associées aux quatorze semaines de mesure est dispersé autour de la concentration moyenne. Les points du nuage sont néanmoins assez proches dans les différentes projections, ce qui laisse supposer que le comportement spatial de l'ozone varie peu d'une semaine à l'autre. Cette hypothèse sera vérifiée lors de l'analyse variographique. Certaines semaines, telles la 3 et la 8, se distinguent cependant. La semaine 3 correspond au premier pic de concentration observé, la semaine 8 à la retombée du deuxième pic liée à une diminution brutale de température.

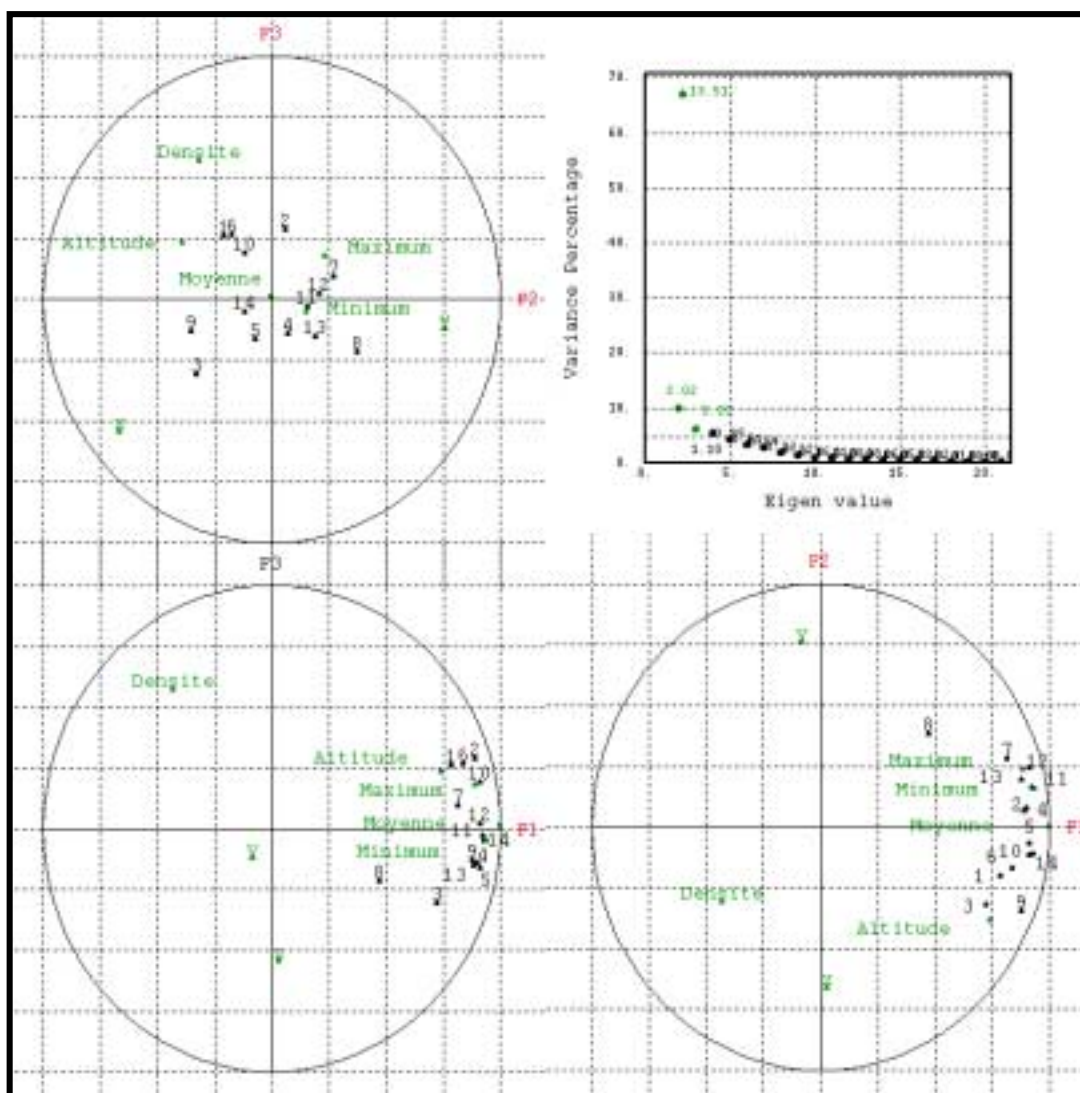


Figure 3 : ACP des séries des mesures. Projection du nuage des variables sur les trois premiers plans factoriels

L'examen préliminaire des données et les résultats de cette ACP nous conduisent à sélectionner pour cette étude six semaines de mesure : les semaines 1, 3, 6, 7 et 8 et 13. Comme il a été mentionné, les périodes 3 et 8 se démarquent du nuage des variables, les semaines 1 et 6 semblent en revanche très similaires, les périodes 6, 7 et 8 marquent le sommet et la fin du deuxième pic de concentration. La semaine 13, en fin de saison estivale est celle pendant laquelle ont été mesurées les concentrations les plus faibles de la campagne. L'intérêt de cette sélection est d'avoir un échantillon de semaines représentatif de toutes les situations possibles.

4.2 COMPARAISON AVEC LES DONNEES DE STATIONS FIXES – ETUDE DE L'INCERTITUDE DE MESURE

Des échantillonneurs passifs ont été installés à proximité de trois analyseurs (stations fixes). Bien qu'elles varient dans le temps de façon similaire (Figure 4), les concentrations d'ozone obtenues par les tubes sont presque systématiquement supérieures à celles des

stations fixes. Si l'on considère les analyseurs comme plus fiables, cette constatation montre la tendance des échantillonneurs passifs à surestimer les concentrations d'ozone. L'importance de cette surestimation varie néanmoins selon les sites. Elle n'est de plus observée qu'en trois points. Aucune correction ne sera appliquée aux valeurs des tubes.

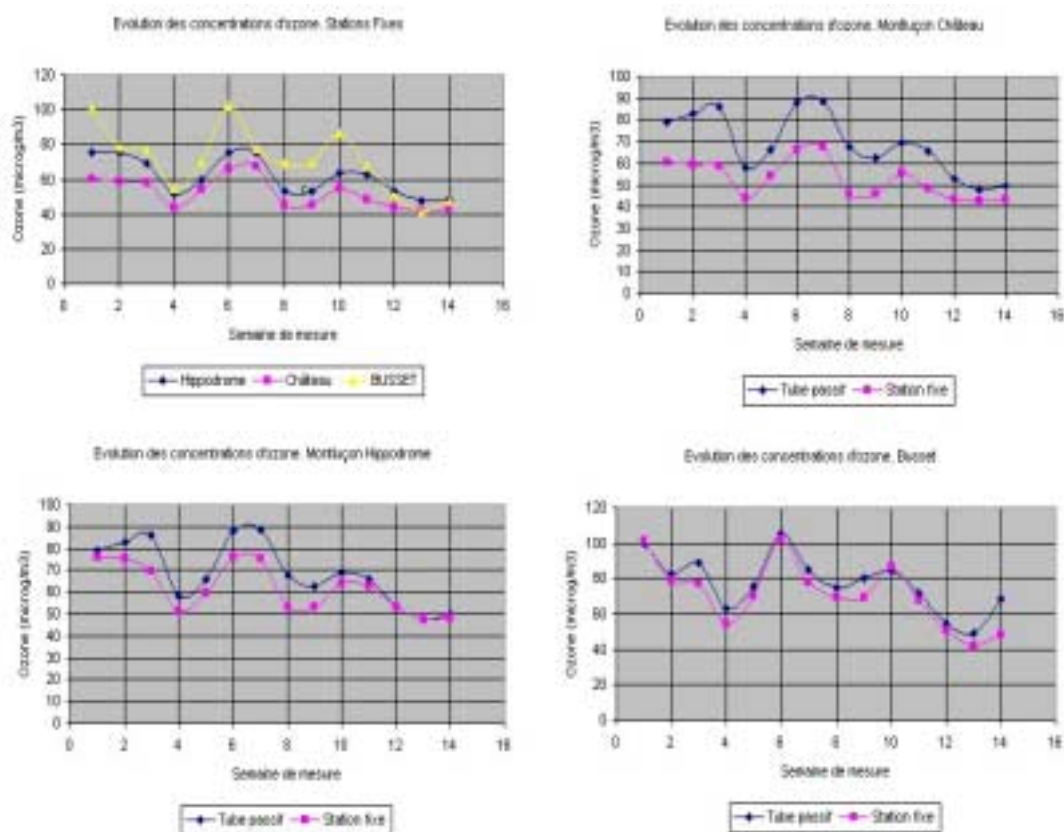


Figure 4 : Comparaison entre les données de tubes passifs et de stations fixes à Montluçon Château (site urbain), Montluçon Hippodrome (site périurbain) et Busset (site rural). Les données de stations fixes sont la moyenne sur chacune des quatorze périodes d'échantillonnage des concentrations horaires mesurées par les analyseurs. La différence entre les deux techniques de mesure est moindre pour le site rural.

Notons que la variable mesurée par les tubes est en réalité l'absorbance, c'est-à-dire la quantité absorbée sur les tubes pendant une période déterminée. Les données communiquées par ATMO Auvergne sont les concentrations d'ozone déduites des valeurs d'absorbance par la loi de Fick. Ces données s'accompagnent d'une incertitude liée d'une part à la mesure et d'autre part au choix du coefficient de diffusion utilisé dans la loi de Fick.

Une autre méthode d'estimation des concentrations d'ozone, étudiée en particulier par AIR NORMAND, consiste à effectuer une régression orthogonale entre les concentrations mesurées par les analyseurs et les absorbances des tubes disposés à proximité de ces derniers [2][3]. La relation ainsi établie est appliquée à l'ensemble des tubes afin d'estimer les concentrations en tout point d'échantillonnage. L'incertitude sur les concentrations est estimée selon les principes décrits dans la norme NF ENV 13005 et peut être introduite comme erreur de mesure dans l'algorithme géostatistique. Cette option n'a pas été ici retenue :

- du fait que nous travaillons sur les données de concentration déjà transformées par ATMO Auvergne ;

- du fait que le calcul de l'incertitude
 - suppose que l'on a pu identifier une relation entre l'écart-type de l'absorbance et l'absorbance moyenne, grâce à un ou plusieurs sites pourvus de multiples tubes ;
 - s'applique d'autant mieux que plusieurs tubes ont été exposés en chaque point d'échantillonnage.

Dans cette étude seuls deux sites ruraux (Agonges et Busset) et un site urbain (Montluçon Château) ont été équipés respectivement de trois, trois et deux tubes.

A l'erreur de mesure s'ajoute par ailleurs la variabilité des concentrations d'ozone à petite échelle, variabilité considérée d'ordinaire comme faible mais sur laquelle nous ne disposons pas d'information détaillée compte tenu de la maille d'échantillonnage.

4.3 STATISTIQUES ELEMENTAIRES ET ANALYSE STRUCTURALE DES SIX SERIES DE MESURES SELECTIONNEES ET DE LA MOYENNE DES QUATORZE SEMAINES

Nous présentons dans les paragraphes qui suivent :

- l'analyse exploratoire des données de tubes conduite séparément pour les semaines de mesure sélectionnées à la suite de l'ACP. Cette analyse détaillée dans le temps n'a pas pour fin d'étudier exhaustivement les quatorze semaines de mesure mais d'apprécier l'évolution structurale de la pollution d'ozone au cours de la saison estivale.
- l'analyse exploratoire des données de tubes réalisée sur la moyenne des 14 semaines de mesure (18 juin 2001- 25 septembre 2001). On s'intéresse ici à la structure spatiale moyenne de la pollution estivale d'ozone (année 2001).

Cette étude est conduite entièrement avec le logiciel de géostatistique Isatis (® Géovariances)

4.3.1 Statistiques élémentaires des six semaines de mesure sélectionnées

L'histogramme des concentrations moyennes d'ozone a été établi à partir des données des trente² tubes passifs (Figure 5 et Figure 6). En raison de ce nombre limité de valeurs, il serait audacieux d'en proposer une interprétation approfondie. Quelques caractéristiques se dégagent cependant. Les concentrations se répartissent dans un intervalle de valeurs assez resserré (le coefficient de variation, c'est-à-dire le rapport écart-type/moyenne, est compris dans tous les cas entre 13 et 23%) et aucune valeur extrême, qui pourrait sembler anormale, ne s'écarte de la distribution.

Les histogrammes des six semaines de mesure sont logiquement plus dispersés que l'histogramme de la moyenne des quatorze semaines, en particulier ceux des semaines 1 et 7. Toutefois, il n'apparaît pas de point aberrant.

² Comme cela est justifié dans l'analyse variographique, les trois sites urbains et périurbains de Montluçon sont considérés comme un seul point.

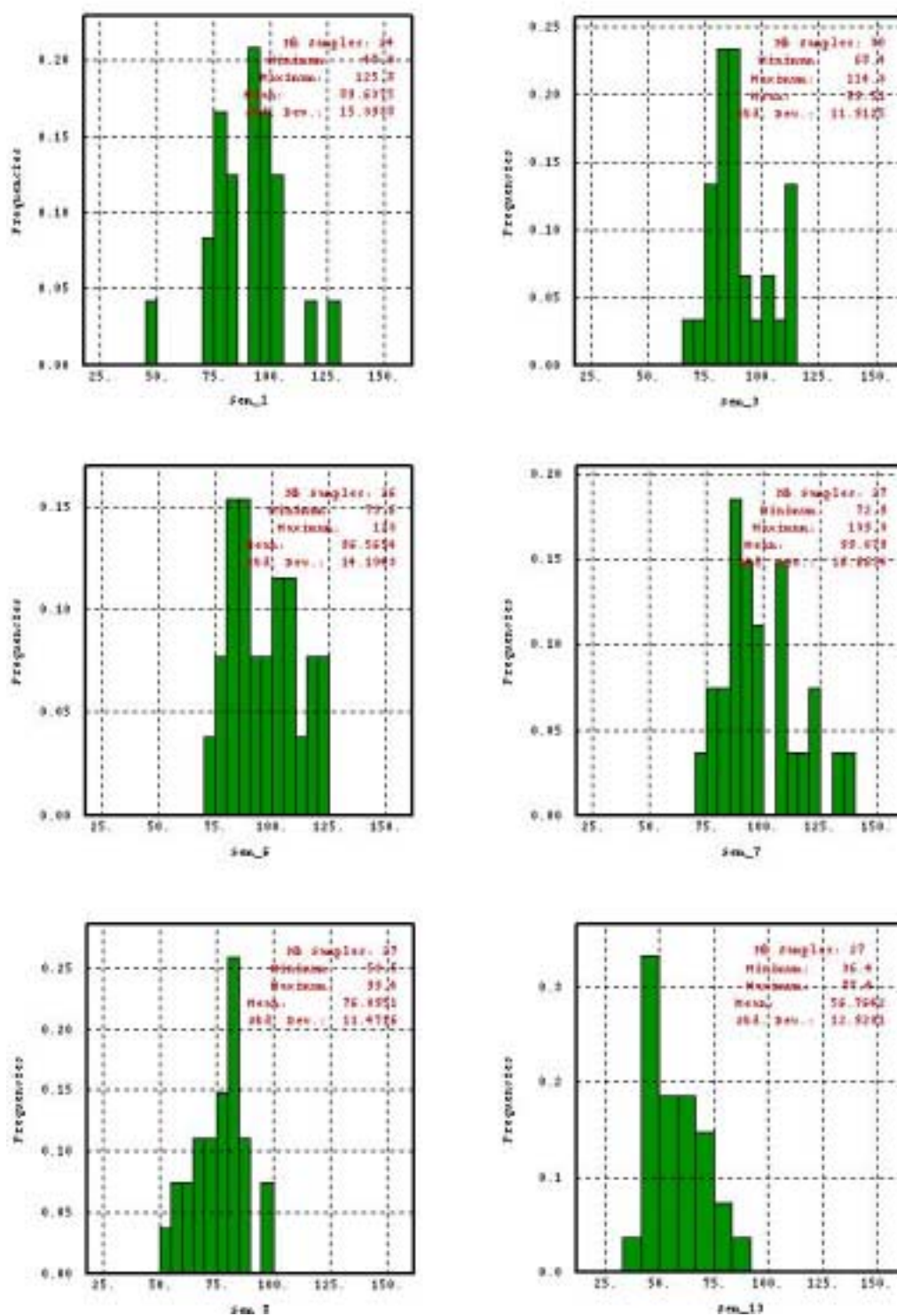


Figure 5 : Histogramme et principales statistiques des semaines 1, 3, 6, 7, 8 et 13

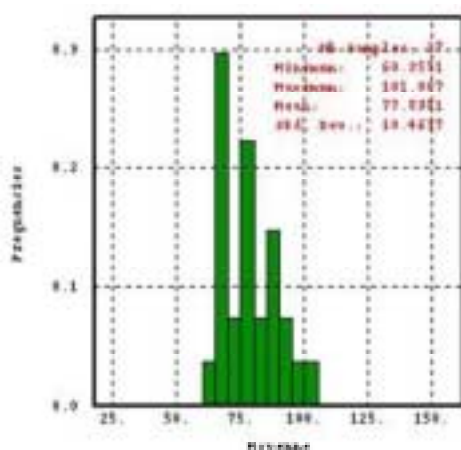


Figure 6 : Histogramme et principales statistiques de la moyenne des quatorze semaines

4.3.2 Dérives spatiales

Etant donné les grandes dimensions du domaine d'étude, on peut soupçonner l'existence d'une dérive spatiale, c'est-à-dire d'un comportement tendanciel de l'ozone en fonction des coordonnées de l'espace. Le tracé de l'évolution des concentrations en fonction de X (axe EO) puis de Y (axe NS) aide à détecter la présence d'une telle dérive qui, si elle existe, devrait se traduire graphiquement par une croissance ou une décroissance du nuage de point dans tout le domaine d'étude.

Les points de la courbe noire (Figure 7 respectivement Figure 8) représentent la moyenne des concentrations par classe de valeurs de X (respectivement de Y) pour les six semaines étudiées –Largeur d'une classe : 25 km-. La même analyse est conduite pour la moyenne des quatorze semaines (Figure 9).

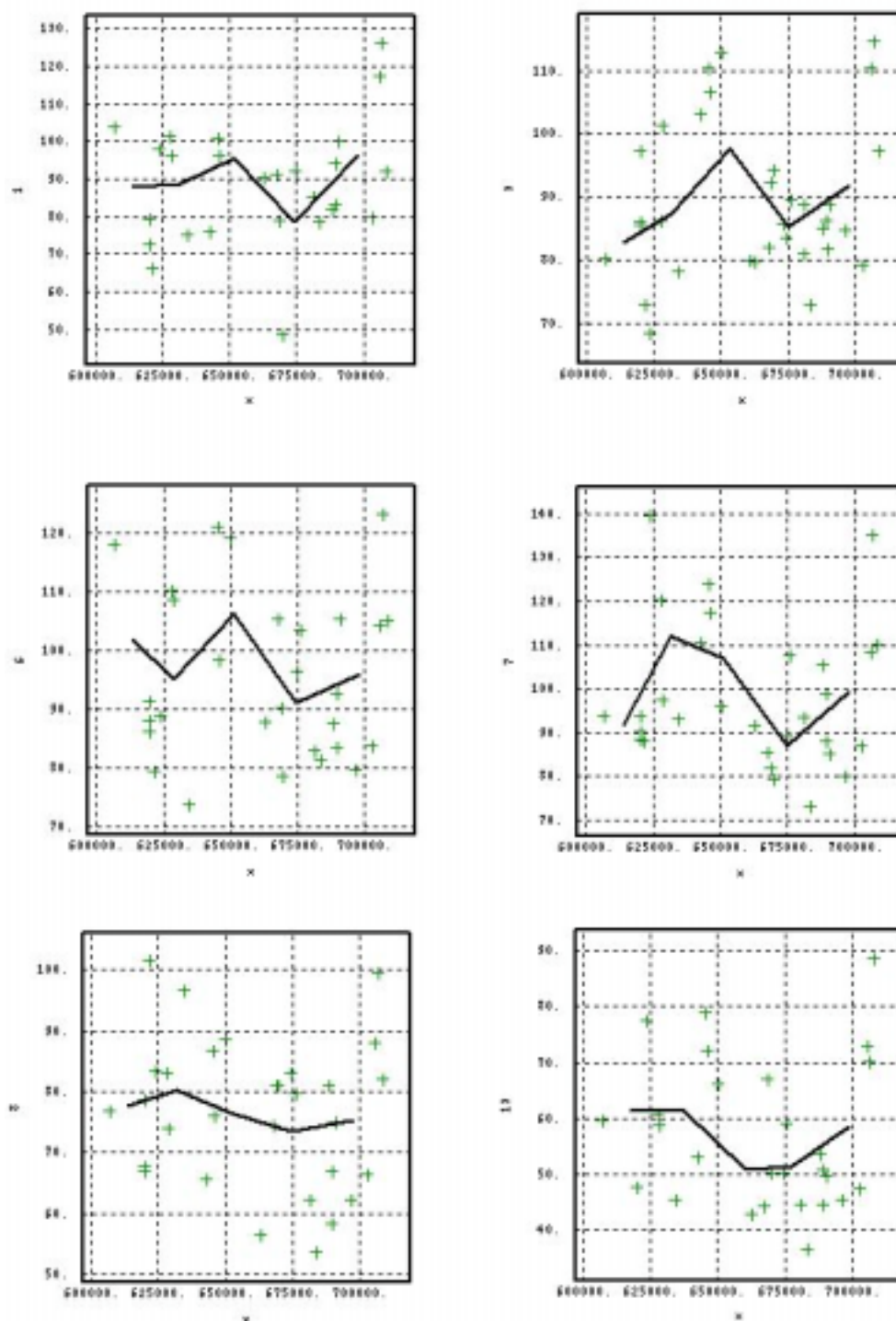


Figure 7 : Evolution des concentrations d’ozone en fonction de X.

Aucune dérive à l’échelle du domaine d’étude ne ressort de ces graphiques (Figure 7). Le comportement de l’ozone selon x est en moyenne stationnaire. L’allure des courbes peut s’expliquer en partie par les modulations du relief et par la présence des villes où les concentrations sont localement plus faibles.

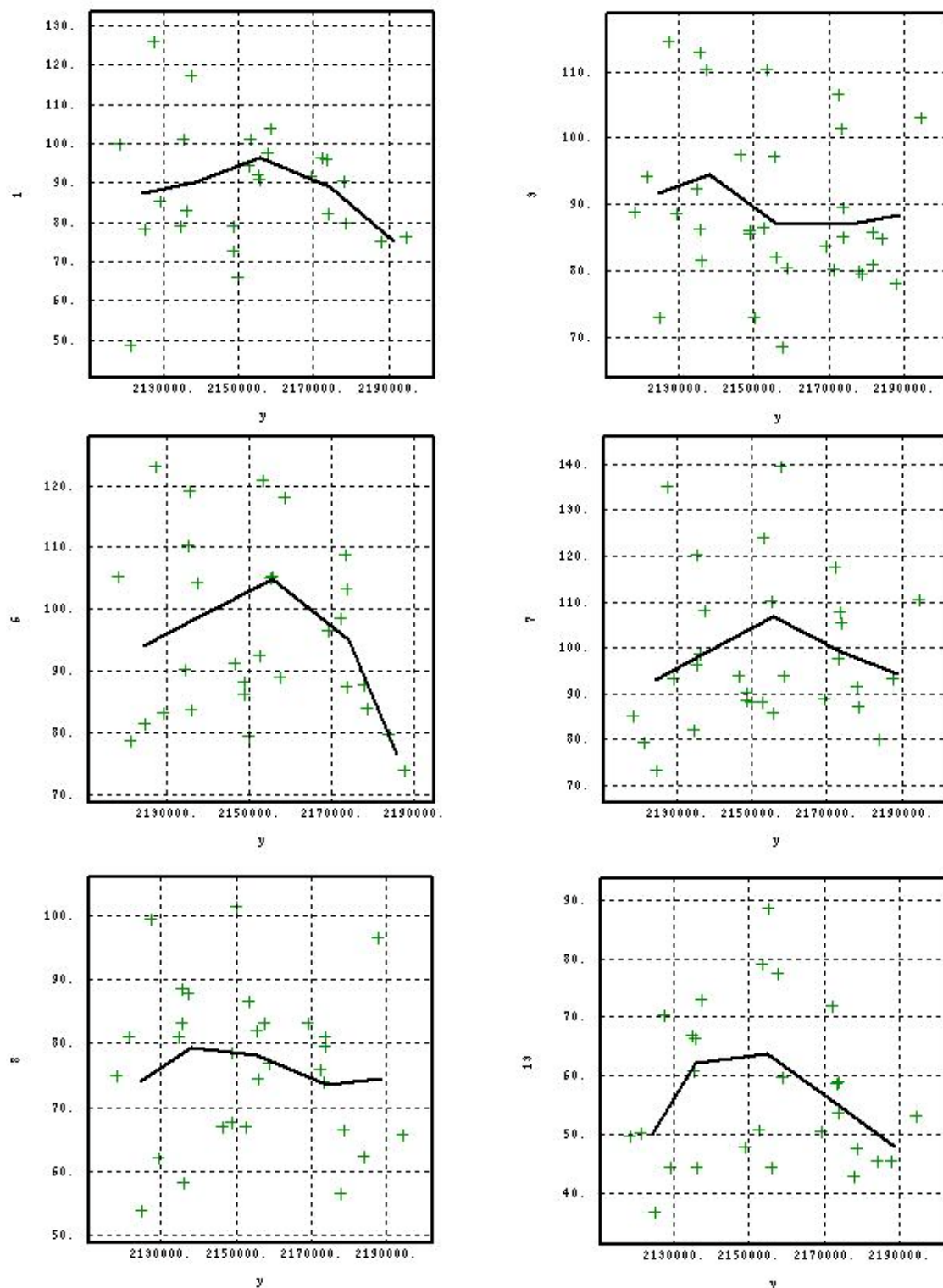


Figure 8 : Evolution des concentrations d'ozone en fonction de Y.

Les concentrations d'ozone ont tendance à diminuer avec la latitude à partir de $Y = 2150$ km (Figure 8), c'est-à-dire de la limite Nord du Massif Central. Ces courbes ne permettent pas de conclure à la présence d'une dérive mais confirment le lien qui existe entre concentration et altitude.

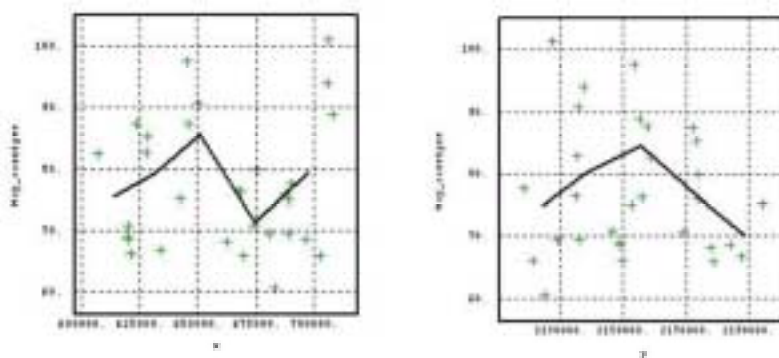


Figure 9 : Concentrations d’ozone en fonction des coordonnées de l’espace. Moyenne des quatorze semaines de mesure. Les conclusions tirées pour les six semaines valent aussi pour la moyenne. Aucune dérive fonction de X ou de Y n’est clairement mise en évidence,

4.3.3 Analyse variographique : variogrammes expérimentaux et modélisation du variogramme

L’analyse variographique a pour but d’étudier la structure de corrélation spatiale de la pollution d’ozone. Elle est effectuée sur les données de concentration moyenne et sur celles des six périodes de mesure. Cette analyse exige un examen préalable de l’échantillonnage dont sont issues les données. Comme cela a été mentionné (3.3), l’échantillonnage suit un schéma régulier (maille de 20 km) et présente un resserrement dans les zones plus urbanisées (points 18 et 19 autour de Moulins, points 14 et 15 autour de Vichy, distants d’environ 5 km).

La ville de Montluçon et ses alentours immédiats comptent trois points de mesure particulièrement proches à l’échelle du domaine d’étude (Montluçon Hippodrome, Montluçon Château, Désertines). Ces trois points, relativement peu touchés par la pollution d’ozone, sont entourés par des zones rurales de plus forte concentration. Aussi risquent-ils d’influencer l’analyse spatiale, en introduisant artificiellement une variabilité qui ne représente pas le comportement réel de l’ozone. Il a donc été décidé d’en faire la moyenne. Le point moyen résultant n’a plus la même nature que les autres données. Néanmoins, compte tenu du comportement similaire des trois points, exception faite de quelques écarts, et étant donné l’échelle du domaine d’étude, il peut être considéré comme représentatif de la concentration d’ozone à Montluçon (**Figure 10**). Le nombre total de sites de mesure est désormais égal à 30.

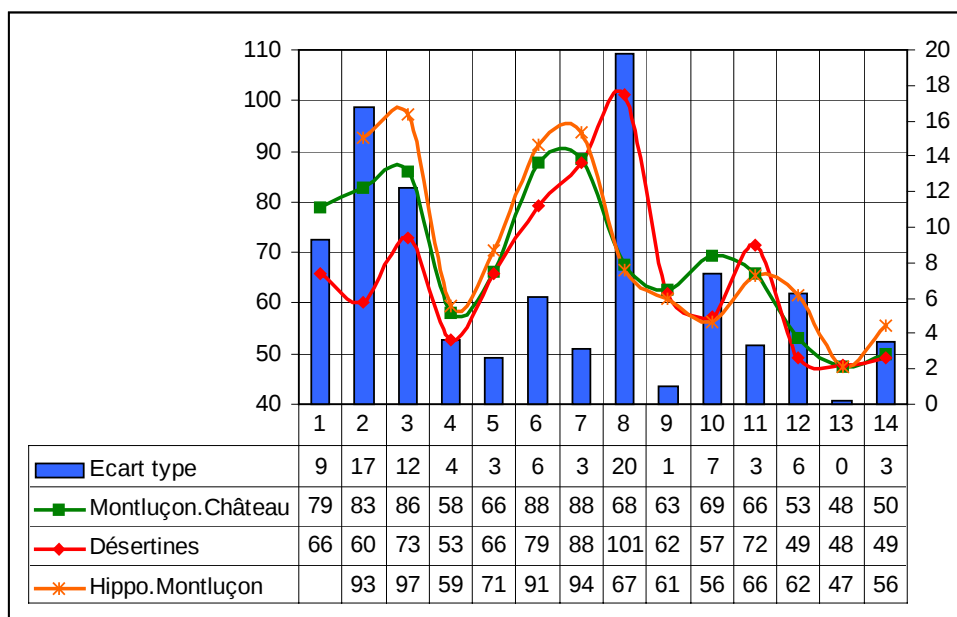


Figure 10 : Concentrations d'ozone mesurées à Montluçon et dans ses alentours. L'écart-type des concentrations, faible en général, est supérieur à 10 µg/m³ pour les semaines 2, 3 et 8. La concentration du point moyen résultant est donc plus incertaine pour ces semaines.

a) Variogrammes omnidirectionnels

Variogrammes expérimentaux

Le variogramme expérimental est déduit du nuage variographique qui représente la variance entre chaque paire de points de mesure. Il s'écrit :

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{N(h)} [z(s_i) - z(s_j)]^2$$

N(h) : nombre de couples de points distants de h

z(s_i) : concentration au point de mesure s_i.

Il s'obtient en pratique en faisant la moyenne des points du nuage par classe de distance. Ces classes sont définies par le pas de calcul du variogramme et par la tolérance sur la distance qui est généralement d'un demi-pas.

Les variogrammes expérimentaux des six semaines de mesure sélectionnées et de la moyenne des quatorze semaines sont calculés dans toutes les directions (Figure 11). Les paramètres de calcul de ces variogrammes (largeur et nombre de classes de distance) sont ajustés de façon que ceux-ci apparaissent le plus structurés possible, c'est-à-dire qu'ils présentent une croissance jusqu'à un éventuel palier. Le calcul des variogrammes omnidirectionnels permet d'avoir une première connaissance de la structure spatiale des données et de l'effet de pépité, connaissance qui pourra être affinée au fur et à mesure que l'on avance dans l'analyse.

Le terme *effet de pépité* désigne la limite du variogramme lorsque h tend vers 0. Celle-ci représente la variabilité entre deux points quand la distance qui les sépare devient très faible. Elle est liée aux erreurs de mesure d'une part, et à la variabilité des concentrations à petite distance d'autre part. L'effet de pépité est théoriquement identique quelle que soit la direction considérée.

Remarque : à cause du nombre limité de points de mesure, les sites urbains et périurbains sont employés au même titre que les sites ruraux dans le calcul des variogrammes expérimentaux.

- Dans la gamme des distances entre points inférieures à 6 km, ils interviennent seuls dans le calcul du variogramme du fait de l'échantillonnage. A petite échelle, le variogramme expérimental est donc représentatif du comportement de l'ozone en milieu urbain et périurbain. Un resserrement des points d'échantillonnage en zone rurale permettrait de mieux caractériser la structure spatiale de l'ozone aux très courtes distances.
- Dans la gamme des distances entre points supérieures à 6 km, on s'est assuré qu'ils ne perturbaient pas la structure d'ensemble du variogramme expérimental, bien que celui-ci prît en compte leur interaction avec les sites ruraux.

Modélisation des variogrammes omnidirectionnels

La modélisation consiste à ajuster un modèle mathématique au variogramme expérimental. Ce modèle est la somme d'un effet de pépité si celui-ci est non nul, et d'une ou plusieurs fonctions choisies parmi celles qu'autorise la géostatistique. Dans les sept cas étudiés (semaines 1, 3, 6, 7, 8, 13 et moyenne des quatorze semaines) différents modèles ont été testés (modèles sphériques, cubiques et gaussiens) afin de sélectionner ceux qui s'accordaient le mieux avec les données expérimentales. A cause du petit nombre de couples de points distants de moins de 6 km, l'effet de pépité s'est révélé peu aisé à déterminer. La validation croisée a été retenue comme critère de sélection afin de trancher entre des modèles qui, visuellement, semblaient conduire à une même qualité d'ajustement.

L'effet de pépité, les fonctions mathématiques utilisées et les paramètres du variogramme (portée, palier) sont différents pour chaque période (Tableau 2, Figure 11 et Figure 12). Le rapport entre effet de pépité et palier est le plus important pour la semaine 7, semaine durant laquelle ont été enregistrées les concentrations les plus élevées. Cela correspond à un phénomène moins structuré. Dans ce cas, l'estimation en un point sera plus proche de la moyenne des données expérimentales. Chaque période étudiée a donc sa propre structure de corrélation spatiale. Les semaines 1 et 6, pourtant proches dans l'ACP, conduisent ainsi à des modèles de variogramme différents.

Semaine	Pépité	Modèle	Palier	Portée
1	105	Sphérique 1	8.4	22km
		Sphérique 2	156	44km
3	30	Cubique	150	40km
6	40	Sphérique	200	40km
7	78	Sphérique 1	95	24km
		Sphérique 2	122	40km
8	60	Sphérique	85	22km
13	20	Cubique	202	38km
Moyenne	30	Cubique	100	41.5km

Tableau 2 : Modèles omnidirectionnels ajustés aux six séries hebdomadaires

Le *palier* est la valeur de stabilisation du variogramme expérimental. La *portée* est la distance à laquelle le palier est atteint.

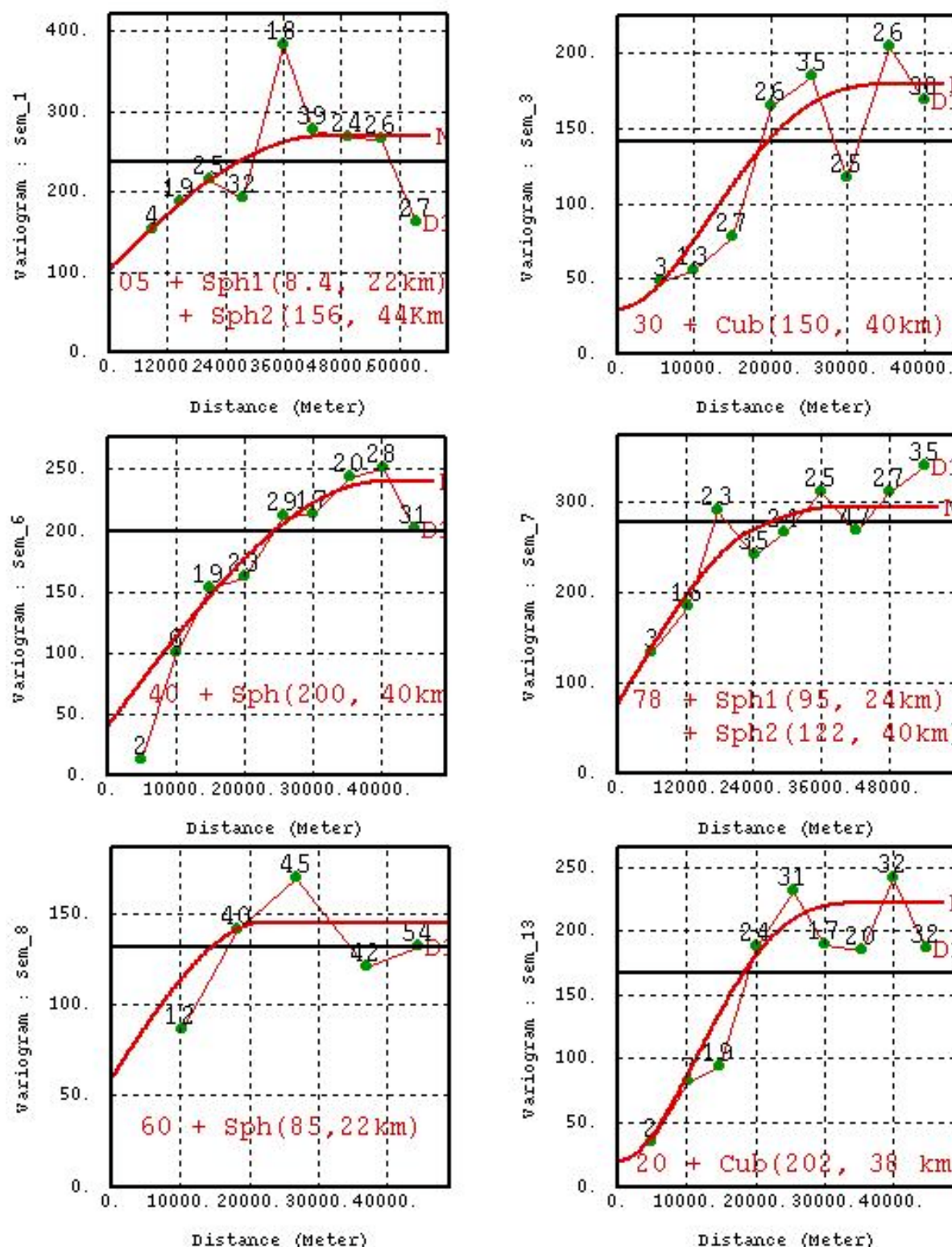


Figure 11 : Variogrammes omnidirectionnels. Le variogramme de la semaine 13 a été calculé avec un pas supérieur à celui des autres variogrammes parce qu'aucune structure n'était visible avec un pas plus petit.

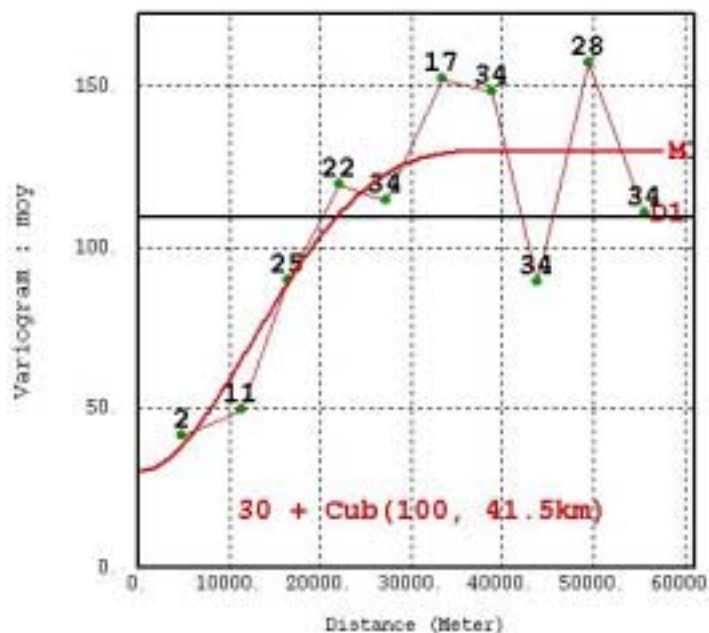


Figure 12 : Variogramme omnidirectionnel. Moyenne des quatorze semaines de mesure.

b) Variogrammes bidirectionnels

Variogrammes expérimentaux

Cette seconde étape de l'analyse variographique a pour but de déceler d'éventuelles directions d'anisotropie. A cause du relief ou des conditions météorologiques, notamment des vents dominants, il est possible que l'ozone se comporte différemment selon les directions. Un examen visuel de la configuration des points d'échantillonnage et des valeurs de concentration en ces points peut donner une première idée des directions suivies préférentiellement par la pollution d'ozone. La carte variographique (Figure 13), représentation plane du variogramme expérimental, constitue aussi un outil efficace pour détecter les anisotropies. La recherche des anisotropies est ensuite affinée en calculant le variogramme expérimental dans plusieurs directions et en étudiant si une direction se différencie d'une autre.

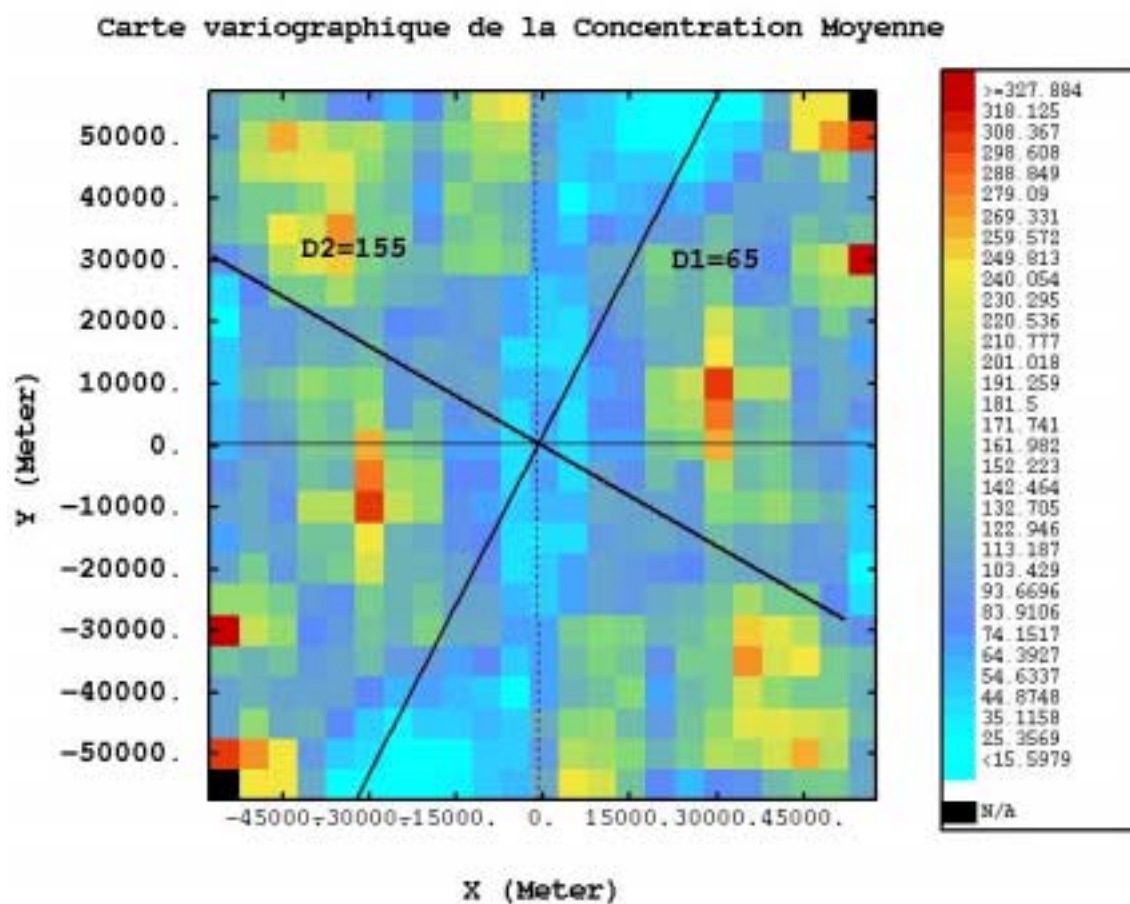


Figure 13 : Carte variographique de la moyenne des quatorze semaines de mesure (carte des isovalues du variogramme expérimental dans toutes les directions de l'espace). Cette carte fait apparaître une anisotropie de 65° environ par rapport à l'axe est-ouest.

Dans notre cas, les anisotropies mises en évidence sont de nature géométrique : le palier est le même quelle que soit la direction, mais la portée change. La direction de plus grande continuité fait le plus souvent un angle de 60° avec l'horizontale (semaines 1, 6, 8 et moyenne). Cet angle est de 120°, 135° et 130° pour les semaines 3, 7 et 13 ce qui pourrait s'expliquer par des conditions météorologiques différentes.

Modélisation des variogrammes bidirectionnels

Cette modélisation s'appuie sur les résultats de l'analyse omnidirectionnelle. Pour tous les variogrammes (Figure 14 et Figure 15), l'effet de pépité est pareil à celui de l'étude omnidirectionnelle (4.3.3 et Tableau 3) et les fonctions variographiques employées dans chaque modèle sont identiques à celles des modèles isotropes, exception faite de la semaine 1. Les portées varient entre 40 et 70 km dans la direction de plus grande continuité et entre 15 et 35 km dans la direction perpendiculaire. Les portées des deux directions encadrent la portée du variogramme omnidirectionnel.

Certaines semaines, telles la 7 et la 13, sont plus délicates à modéliser mais il semble que le variogramme soit plus structuré dans la direction 130/135° et présente dans cette direction une portée plus longue.

<i>Semaine</i>	<i>Pépite</i>	<i>Modèle</i>	<i>Palier</i>	<i>Portée D1</i>	<i>Portée D2</i>
1	105	Gaussienne	200	60°, 55km	150°, 30km
3	30	Cubique	160	120°, 70km	30°, 35km
6	40	Sphérique	195	60°, 70km	150°, 26km
7	78	Sphérique	206	45°, 15km	135°, 40km
8	65	Sphérique	96	60°, 55km	150°, 20km
13	20	Cubique	185	40°, 30km	130°, 48km
Moyenne	30	Cubique	100	65°, 60.5km	155°, 28km

Tableau 3 : Modèles bidirectionnels ajustés aux variogrammes expérimentaux des six séries hebdomadaires

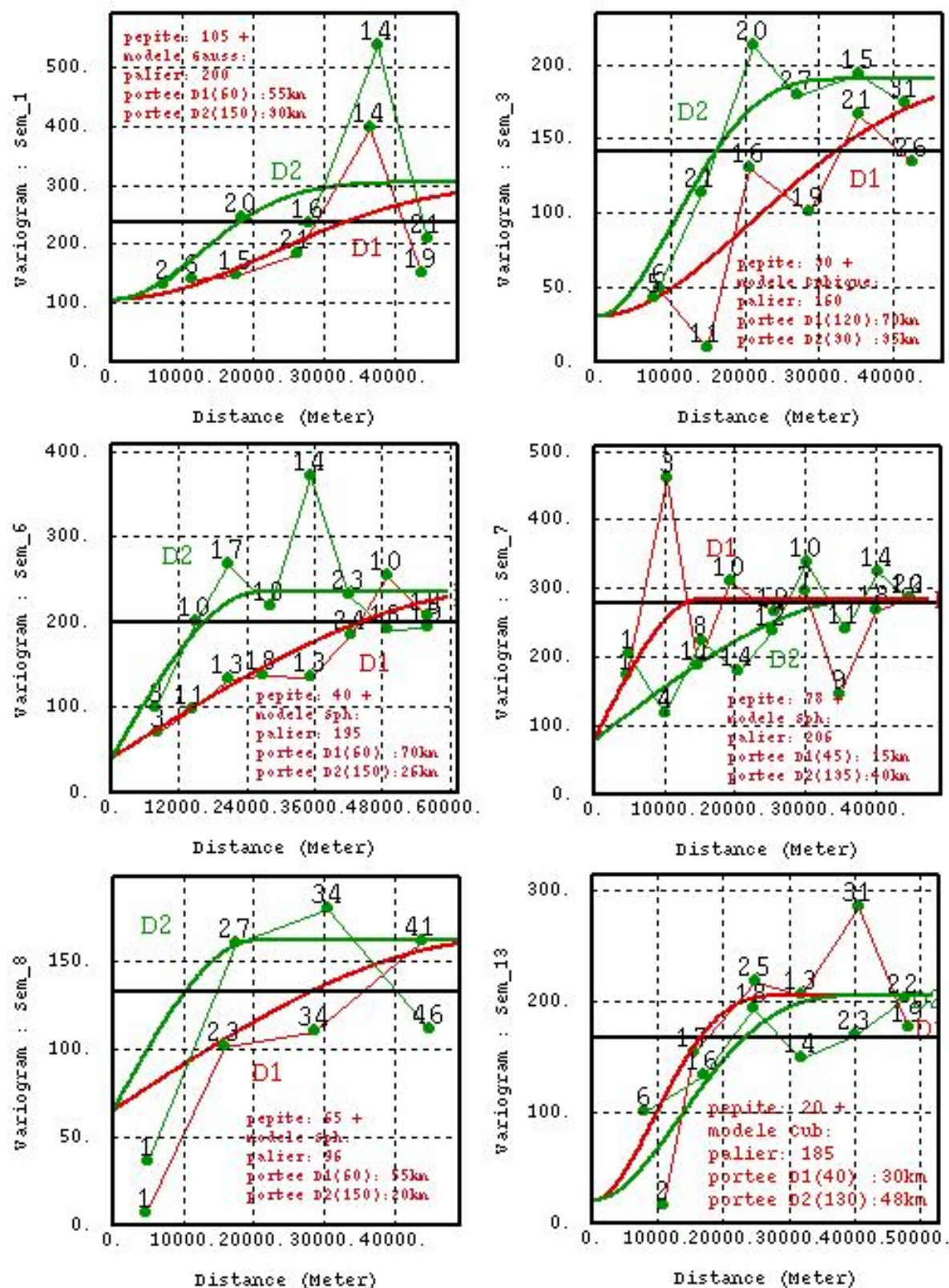


Figure 14 : Variogrammes bidirectionnels des six semaines de mesure sélectionnées.

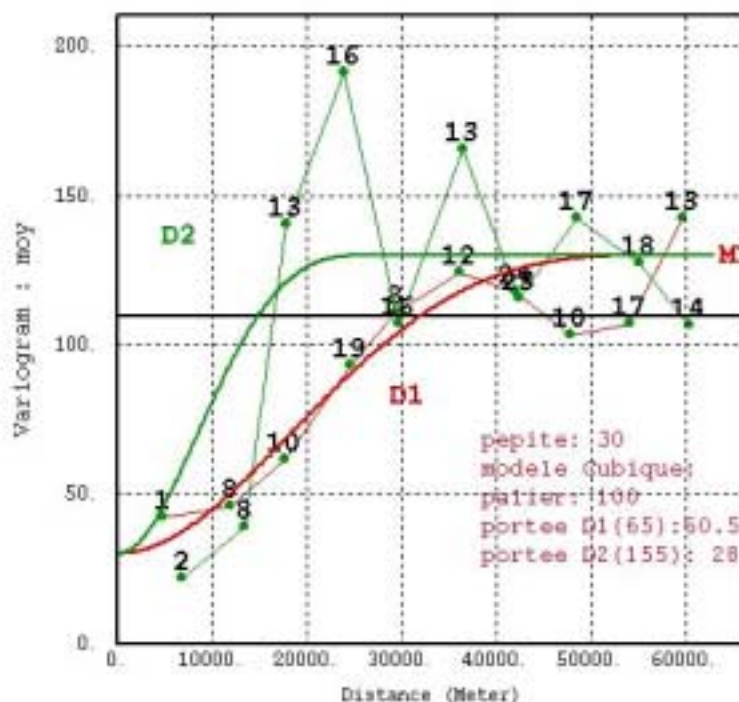


Figure 15 : Variogramme bidirectionnel. Moyenne des quatorze semaines de mesure.

c) Bilan

L'analyse variographique nous confirme qu'à l'échelle de travail l'hypothèse de **stationnarité** est vérifiée. Tous les variogrammes atteignent un palier. Ils se stabilisent à une distance comprise entre 15 et 70 km selon la semaine et selon la direction d'anisotropie.

Des anisotropies existent et sont modélisées. Elles ne semblent pas dues uniquement à la présence des villes et des montagnes puisqu'elles varient dans le temps mais elles pourraient être liées à la météorologie. Aussi faudrait-il comparer les directions de ces anisotropies à des données météorologiques. En particulier les changements de direction d'anisotropie observés durant les semaines consécutives 6, 7 et 8 mériteraient d'être comparés à des mesures de vent.

Certains variogrammes ne sont pas très structurés (ex : semaine 7), de là l'intérêt que pourra présenter l'usage de variables auxiliaires afin d'améliorer l'estimation. Il n'a pas été possible de définir de modèle variographique commun à toutes les périodes. Chacune présente une structure spatiale qui lui est propre. Cela montre la pertinence d'étendre l'échantillonnage à plusieurs semaines quand le but est d'évaluer la structure spatiale moyenne de la pollution estivale d'ozone.

4.3.4 Validation croisée

La validation croisée consiste à éliminer temporairement un point de l'ensemble des données puis à kriger sa valeur à l'aide des données restantes et du modèle de variogramme qui a été ajusté. Cette opération est répétée pour tous les points.

Ainsi l'on obtient pour chaque point s_i de concentration mesurée $Z(s_i)$ une valeur estimée $Z^*(s_i)$ et un écart-type de krigeage $\sigma_K(s_i)$. L'erreur d'estimation correspondante est alors $[Z^*(s_i) - Z(s_i)]$. Si le modèle correspond bien aux données, on espère donc que :

- la moyenne des erreurs d'estimation est égale à zéro : $E[Z^*(s_i) - Z(s_i)] = 0$
- la moyenne des erreurs réduites d'estimation est égale à zéro : $E\left[\frac{Z^*(s_i) - Z(s_i)}{\sigma_K(s_i)}\right] = 0$
- la variance des erreurs réduites d'estimation est égale à 1,0 : $Var\left[\frac{Z^*(s_i) - Z(s_i)}{\sigma_K(s_i)}\right] = 1$

Ces statistiques sont calculées pour chacun des modèles omnidirectionnels et bidirectionnels (Tableau 4). Tous donnent en général des moyennes proches de zéro et des variances de l'erreur réduite proches de 1. Ces modèles sont jugés acceptables.

Série	Nb	Erreur		Erreur réduite	
		Moyenne (=0)	Variance	Moyenne (=0)	Variance (1)
Sem 1	24	-0,09	148,99	0,0035	0,75
Sem 3	30	-0,45	117,76	0,0087	1,29
Sem 6	26	-0,66	147,13	-0,0286	0,85
Sem 7	27	-0,45	286,68	-0,0086	1,00
Sem 8	27	-0,55	125,33	-0,04351	0,83
Sem13	27	-0,89	142,07	-0,0333	1,06
Moyenne	27	-0,10	65,19	0,0036	0,80

Tableau 4 : Tableau de validation croisée (modèles anisotropes)

Choix du voisinage

L'estimation en un point peut se faire à l'aide de toutes les autres données, dans ce cas il est dit *unique*, ou seulement à l'aide des données contenues dans un voisinage de ce point. Le voisinage employé pour la validation croisée des modèles omnidirectionnels est un voisinage unique. Dans le cas des modèles bidirectionnels, il a une forme d'ellipse qui respecte l'orientation et le rapport d'anisotropie (qui est en général de 2 :1). Différents voisinages elliptiques, définis d'après la géométrie du domaine, ont été testés par validation croisée. De cette façon nous avons estimé qu'un voisinage de 40 km par 20 km pour toutes les semaines était le plus adéquat. Le seul paramètre qui change est l'orientation de l'ellipse.

Le voisinage elliptique utilisé dans la validation croisée du modèle qui a été choisi pour l'estimation de la moyenne des 14 séries hebdomadaires est orienté dans la direction 65° (Figure 16). La valeur estimée au point numéro 17 (centre de l'ellipse) (76,12 $\mu\text{g}/\text{m}^3$) est très proche de la valeur mesurée (76,14 $\mu\text{g}/\text{m}^3$). L'écart-type d'estimation en ce point est donc très faible (-0.0017).

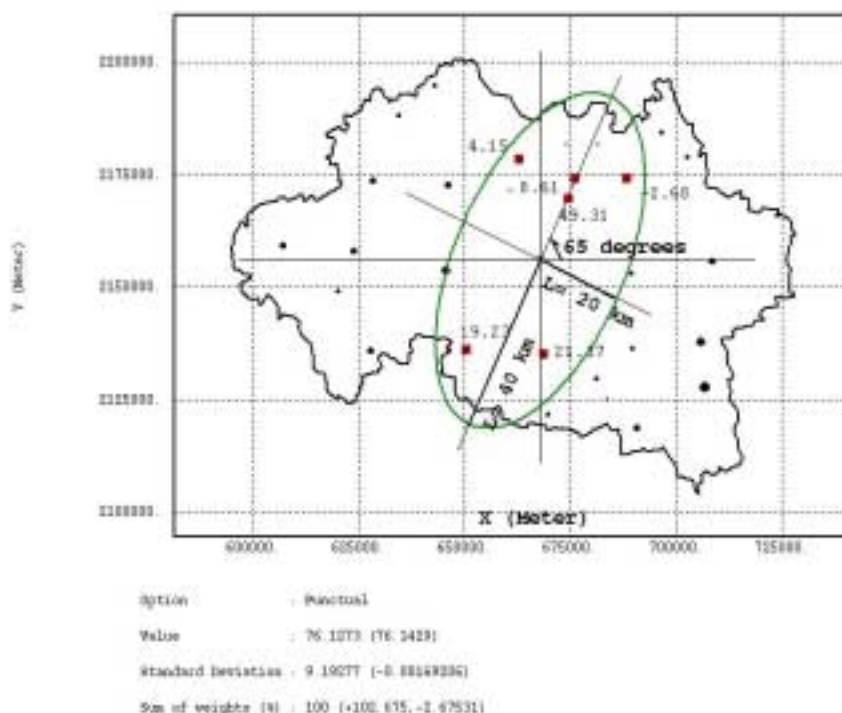


Figure 16 : Voisinage utilisé pour l'estimation. Six points interviennent dans l'estimation de la concentration au point 17 (centre de l'ellipse). Les poids affectés à ces points varient entre -2,68 et 49,31.

Le krigeage a un petit effet de lissage. Les valeurs estimées par validation croisée sont comprises dans une plage plus resserrée que les valeurs vraies (65 à 95 $\mu\text{g}/\text{m}^3$ au lieu de 62 à 102 $\mu\text{g}/\text{m}^3$). Les graphiques de la Figure 17 permettent de distinguer les points où la concentration a été sous-estimée (7 points verts), surestimée (7 points bleus) et bien estimée (12 points rouges). Les tubes passifs correspondant à chaque groupe sont localisés aléatoirement.

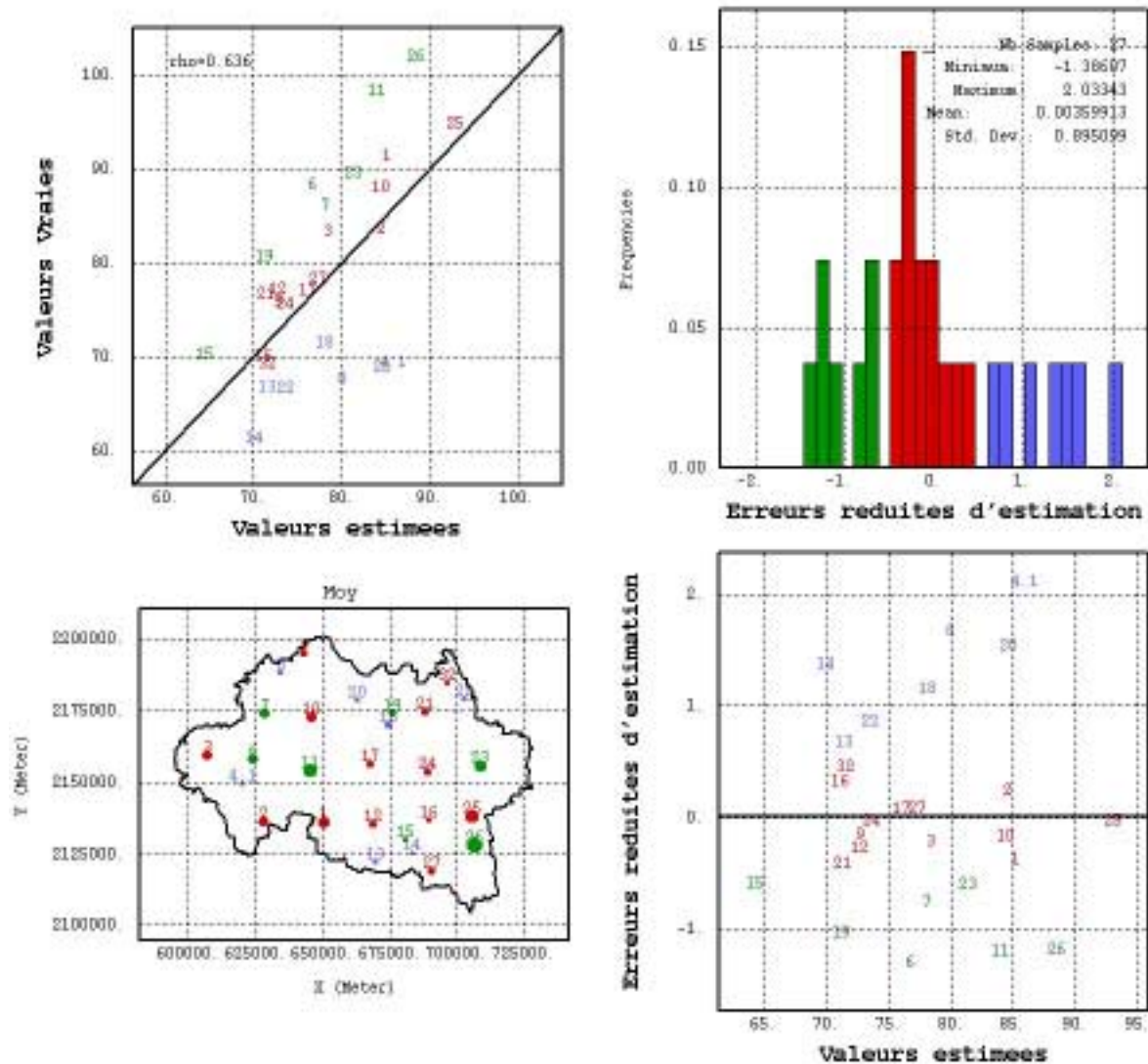


Figure 17 : Validation croisée du modèle bidirectionnel de la moyenne des 14 séries hebdomadaires

5. ETUDE MONOVARIABLE : ESTIMATION PAR KRIGEAGE DE BLOCS

5.1 ESTIMATION AVEC MODELES ISOTROPE

Les concentrations d’ozone sont estimées par krigeage ordinaire sur des blocs carrés de 5 km de côté qui couvrent le domaine d’étude (Figure 20 et Figure 22). Ce choix de la maille d’estimation permet d’obtenir une description suffisamment fine de la pollution par l’ozone à l’échelle du domaine d’étude. Pour chaque période, le modèle variographique utilisé dans les équations de krigeage est le modèle omnidirectionnel défini au chapitre 4.3.3. Le voisinage de krigeage est un voisinage unique.

Les moyennes des concentrations estimées sur toute la grille d’estimation sont proches de celles des données ponctuelles expérimentales (Figure 18). Les écarts-types de ces estimations sont en revanche inférieurs aux écarts-types expérimentaux, ce qui s’explique par le changement de support (passage de données ponctuelles à des valeurs estimées sur des blocs) et par l’effet de lissage du krigeage.

L’écart-type de krigeage représente l’incertitude associée à la valeur interpolée. Plus il est faible, plus l’estimation de la concentration moyenne du bloc est précise. L’écart-type de krigeage ne dépend pas des valeurs numériques des données mais uniquement du variogramme et de l’implantation relative du bloc à estimer et des points d’échantillonnage. Il est sensible au modèle de variogramme, en particulier au palier et à l’effet de pépité et doit être interprété avec prudence. Il présente l’intérêt de mettre en évidence les zones où par suite d’un échantillonnage insuffisant, l’estimation est plus imprécise.

Comme les points d’échantillonnage sont régulièrement répartis, l’écart-type de krigeage est assez homogène dans le domaine d’étude (Figure 19 et Figure 21). Il augmente localement là où une donnée d’échantillonnage est manquante et sur les pourtours du domaine. Si l’on compare entre elles les périodes de mesure, par exemple la semaine 1 et la semaine 3 ou la semaine 6 et la semaine 7, on constate qu’à niveaux de concentration équivalents, les écarts-types de krigeage sont plus grands lorsque le modèle variographique est moins structuré (modèle où la part relative de l’effet de pépité est importante). Cette observation est conforme à la théorie.

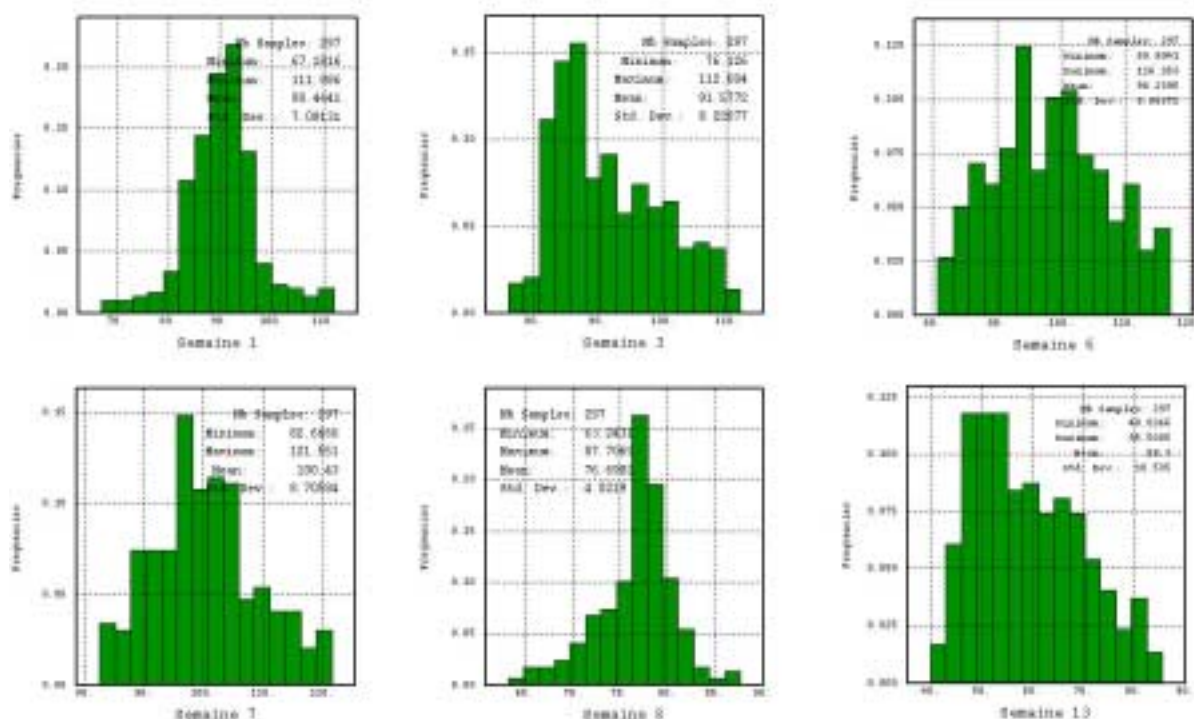


Figure 18 : Histogrammes des valeurs estimées par krigeage avec modèles isotropes.

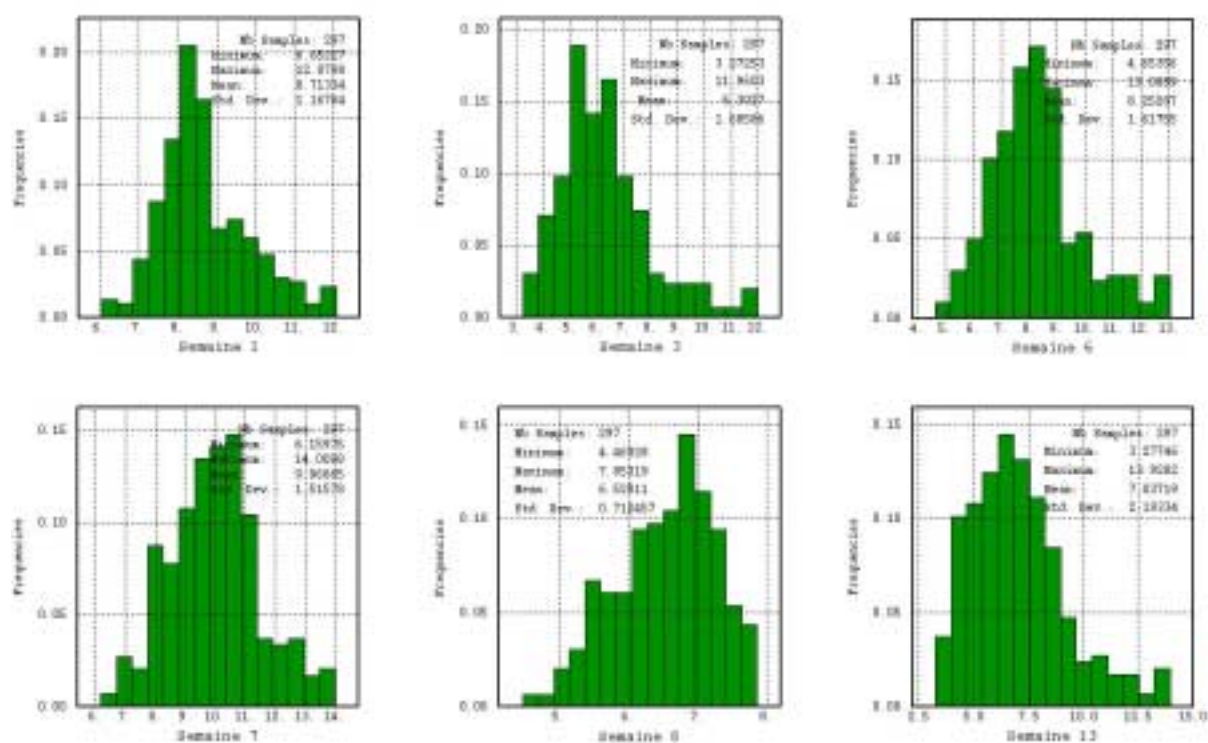


Figure 19 : Histogrammes des écarts-types de krigeage avec modèles isotropes

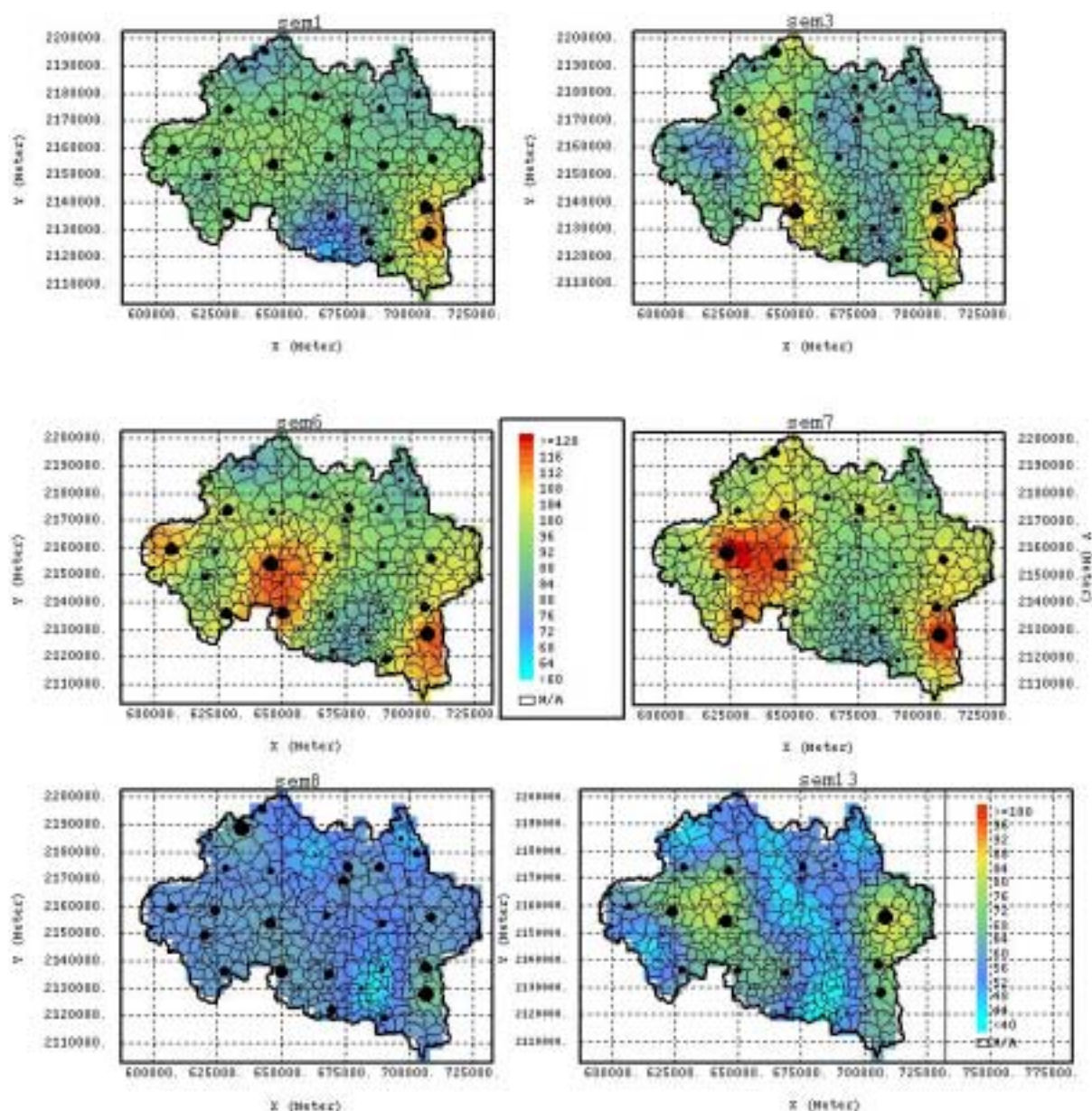


Figure 20 : Cartes des concentrations d’ozone estimées par krigeage à partir des modèles isotropes.

Les points noirs correspondent aux points d’échantillonnage. Leur taille est proportionnelle à la valeur de concentration. L’échelle de couleurs de la semaine 13 est différente de celle des autres semaines parce que les concentrations mesurées pendant cette période sont plus faibles.

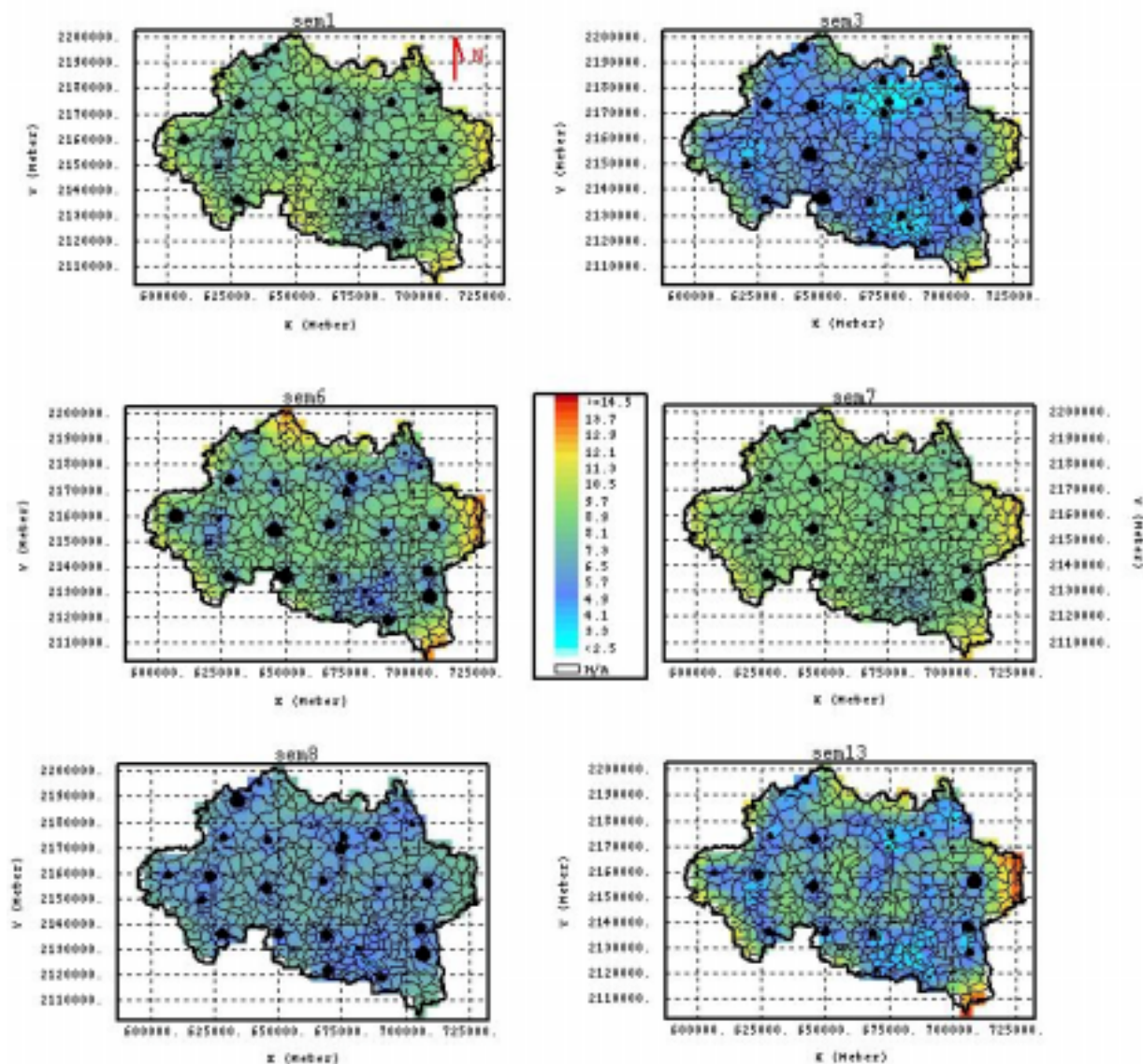


Figure 21 : Cartes de l'écart-type de krigeage avec modèles isotropes.

L'écart-type de krigeage, qui mesure la précision de l'estimation, est plus faible autour des points expérimentaux et dans les zones où l'échantillonnage est plus resserré.

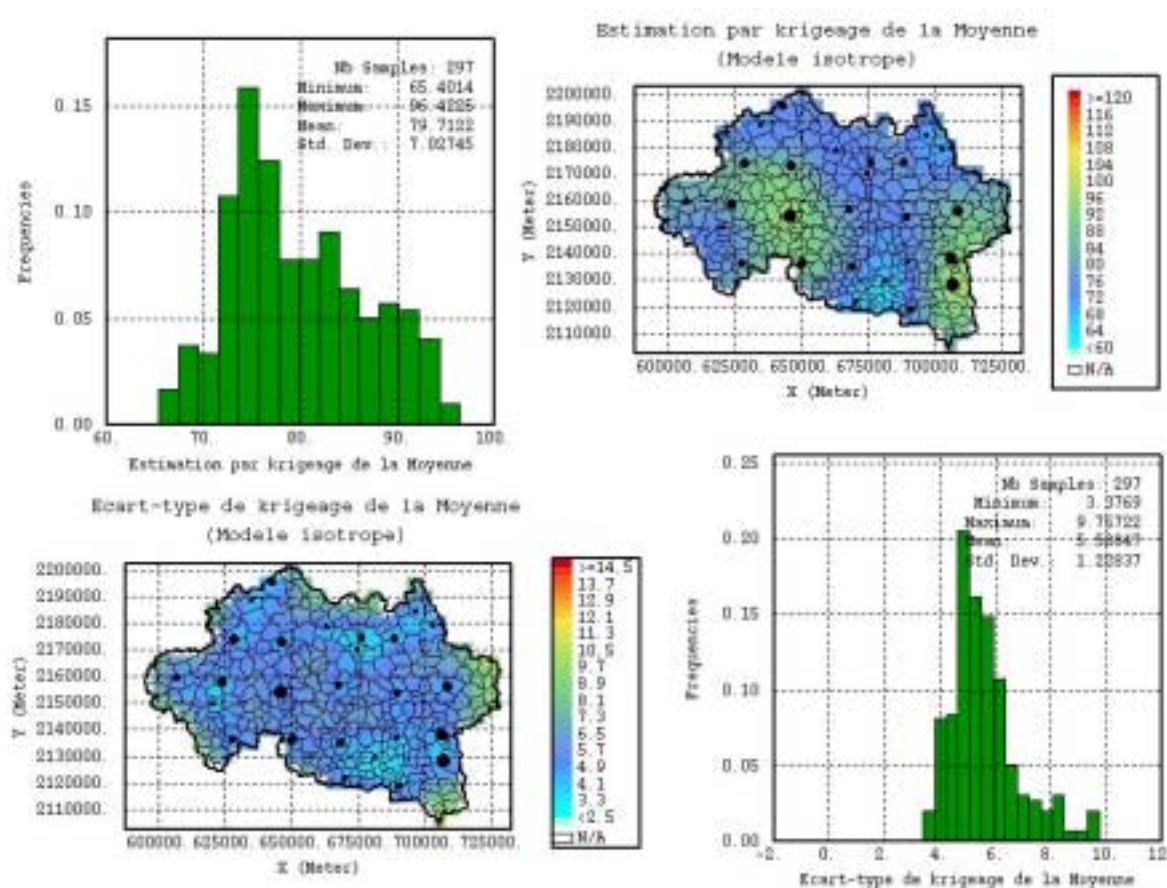


Figure 22 : Modèle isotrope : résultats de l'estimation pour la moyenne de 14 semaines.

La carte de concentration moyenne fait ressortir deux zones de plus forte concentration au centre ouest et au sud-est du département. Les écarts-types de krigeage sont relativement faibles.

5.2 ESTIMATION AVEC MODELES ANISOTROPES

Suivant le même procédé que dans l'étude omnidirectionnelle, les concentrations d'ozone sont estimées par krigeage ordinaire sur des blocs carrés de 5 km de côté (Figure 23 à Figure 27). Pour chaque période, le modèle variographique utilisé dans les équations de krigeage est le modèle bidirectionnel défini au chapitre 4.3.3. Le voisinage de krigeage est le voisinage glissant elliptique qui a été employé pour la validation croisée de ce modèle.

Les anisotropies modélisées sont visibles sur les cartes de concentration qui par ailleurs ressemblent aux représentations isotropes précédemment obtenues. L'anisotropie est une information contenue dans les données expérimentales ; elle fait partie des propriétés structurales du phénomène étudié. Les cartes issues de l'analyse et de la modélisation anisotropes sont donc théoriquement plus réalistes. Dans notre cas, nous ne pouvons être aussi affirmatifs, dans la mesure où les variogrammes anisotropes ont été établis à partir d'un nombre restreint de données et qu'ils n'ont pas toujours été aisés à modéliser. Une étude plus poussée sur l'origine de ces anisotropies, qui sort du cadre de ce rapport, serait en outre nécessaire. Remarquons que les écarts-type de krigeage sont en moyenne plus faibles mais qu'ils sont plus dispersés. L'effet de lissage est moindre dans le cas anisotrope. La

comparaison avec les cartes obtenues à l'aide de variables auxiliaires fournira des éléments d'appréciation supplémentaires.

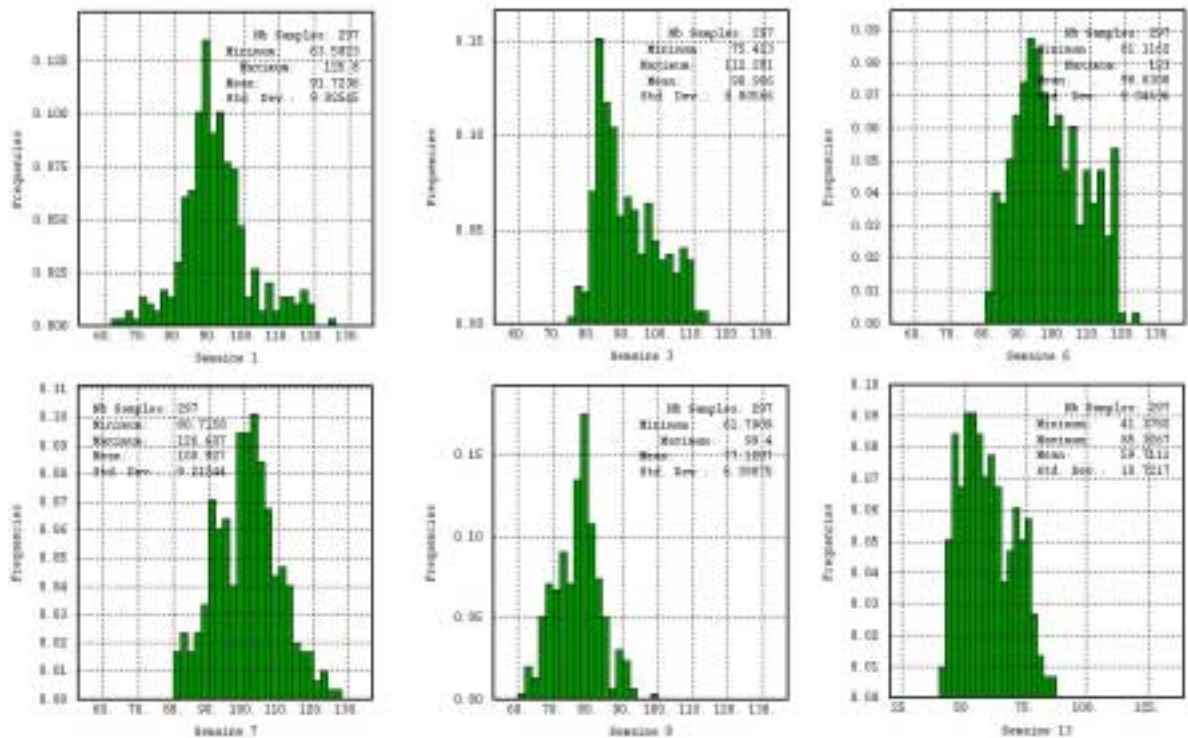


Figure 23 : Histogrammes des valeurs estimées par krigeage avec modèles anisotropes.

Comme dans l'étude isotrope, la moyenne des concentrations estimées est voisine de la moyenne des données mesurées. L'écart-type est en revanche plus faible.

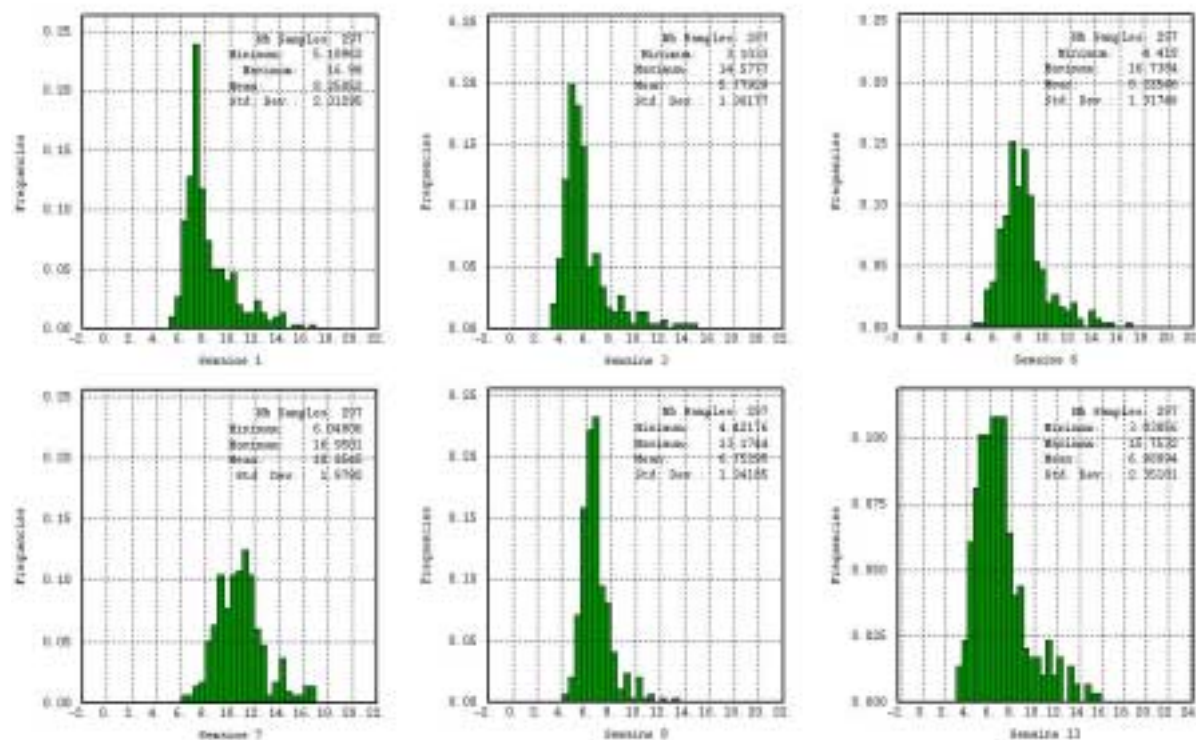


Figure 24 : Histogrammes de l'écart-type de krigeage avec modèles anisotropes

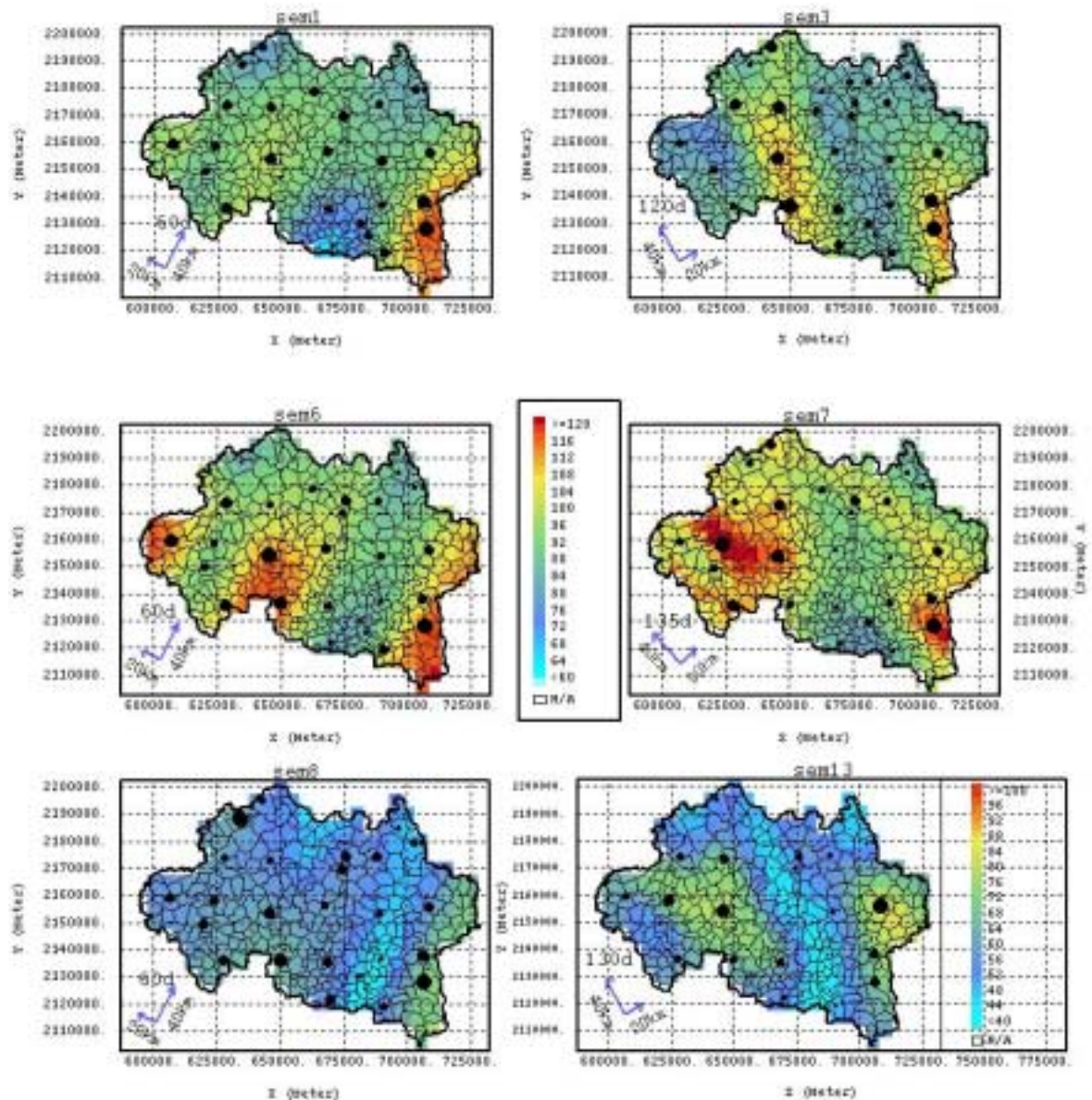


Figure 25 : Cartes des concentrations d’ozone estimées par krigeage de blocs à partir des modèles anisotropes. Les flèches au bas des cartes indiquent les directions principales d’anisotropie. Les distances marquées à côté de ces flèches précisent la taille du voisinage de krigeage dans ces directions.

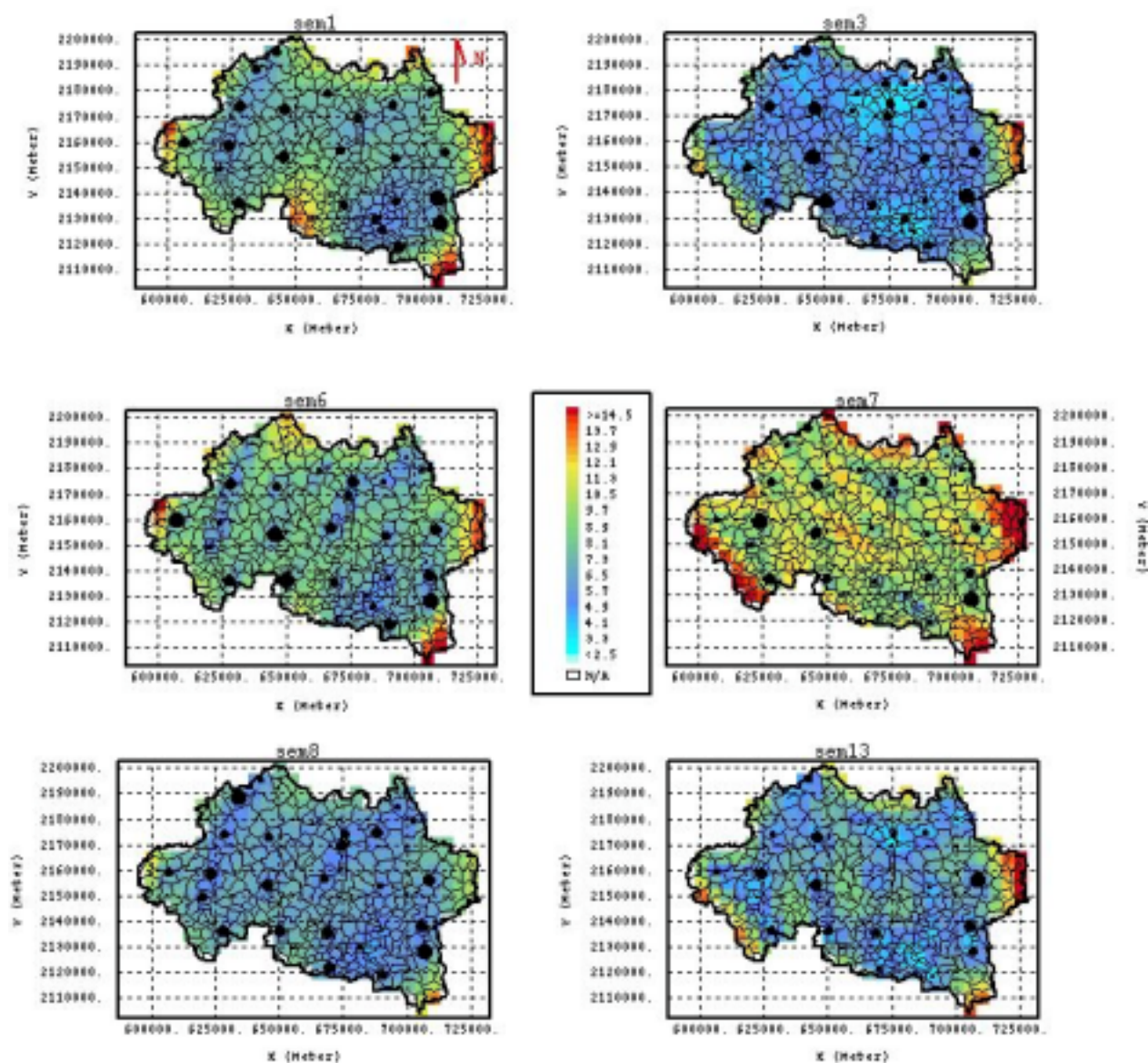


Figure 26 : Cartes de l'écart-type de l'estimation par krigeage avec modèles anisotropes

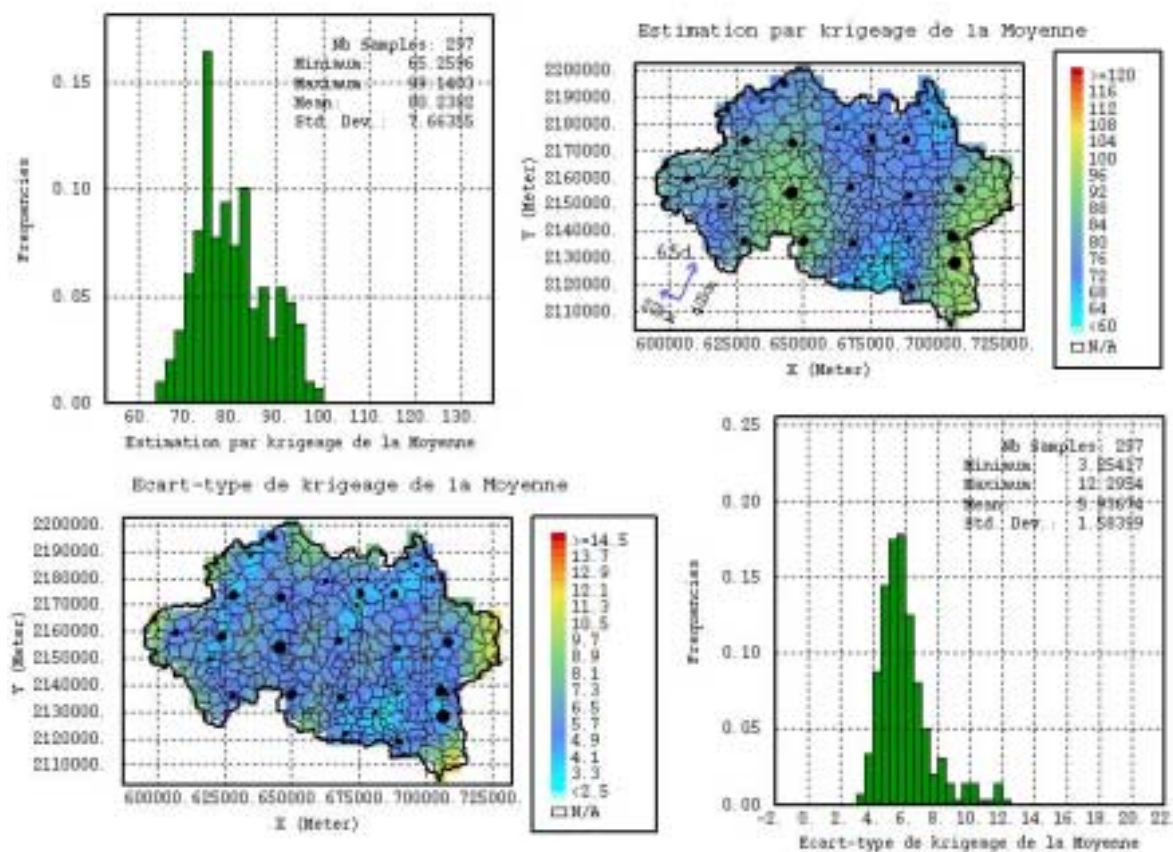


Figure 27 : Modèle anisotrope : résultats de l'estimation pour la moyenne de 14 semaines. La zone de plus forte concentration à l'est du département s'étend dans la direction de plus grande continuité.

6. ESTIMATION AVEC VARIABLES AUXILIAIRES

Cette phase de l'étude a pour fin d'améliorer les estimations précédemment obtenues en exploitant les relations entre les concentrations d'ozone et des variables connues sur un plus grand nombre de points.

Une étude géostatistique multivariable, qui mettrait en jeu les différentes variables chimiques et météorologiques qui interviennent dans la formation et le transport de l'ozone, permettrait de caractériser plus finement la distribution spatiale de ce polluant pendant la période étudiée. A défaut de telles données, nous pouvons prendre en compte d'autres facteurs liés plus indirectement à la pollution par l'ozone. Les méthodes de la géostatistique permettent en effet d'intégrer des variables dites *auxiliaires* dans la matrice de calcul des estimations [4]. Nous avons jugé que la technique du **krigeage avec dérive externe** était la plus appropriée compte tenu des données disponibles. Cette technique d'estimation repose sur la connaissance d'une variable auxiliaire censée mesurer indirectement le même phénomène que celui qui est étudié mais connue sur un plus grand nombre de points. Pour la mettre en œuvre il convient de rechercher une variable **corrélée linéairement** à la variable de concentration et dont les valeurs soient disponibles en tout point de la grille d'estimation.

Dans le krigeage avec dérive externe, la variable de concentration se décompose de la façon suivante :

$$Z(x)=(a_0+a_1s(x))+R(x)$$

$s(x)$: dérive externe déterministe

$R(x)$: résidu aléatoire, qui peut être modélisé numériquement par Isatis

6.1 CHOIX DES VARIABLES

Afin de vérifier si les conditions de vent sont ou non responsables des anisotropies identifiées dans l'étude monovariante, il serait intéressant de pouvoir faire usage de données auxiliaires météorologiques. Ces données n'étant pas disponibles, une autre source possible de variabilité a été étudiée : le relief.

La densité de population a été également considérée ainsi que l'influence des zones bâties et des zones boisées :

- Les valeurs de densité de population sont fournies pour l'ensemble des communes qui composent le département de l'Allier. Celles-ci sont représentées dans un système d'information géographique (SIG) sous la forme de polygones adjacents.
- Les zones boisées et les zones bâties sont également représentées sous la forme de polygones dont on peut calculer la superficie.

Or dans un krigeage avec dérive externe, les valeurs des variables auxiliaires doivent être disponibles sur un maillage régulier qui est alors pris pour grille d'estimation. Afin de satisfaire à ces exigences, une transformation a été préalablement effectuée par SIG à partir des variables auxiliaires initiales. De cette façon trois nouvelles variables ont été créées sur une grille de pas égal à 5 km : les variables *densité de surface bâtie* et *densité de surface boisée*, exprimées en km² de polygone par km² de maille (km²/km²) et la variable *densité de population* en nombre d'habitants par km² de maille (hab./km²).

6.2 ANALYSE DE LA VARIABLE DENSITE DE POPULATION

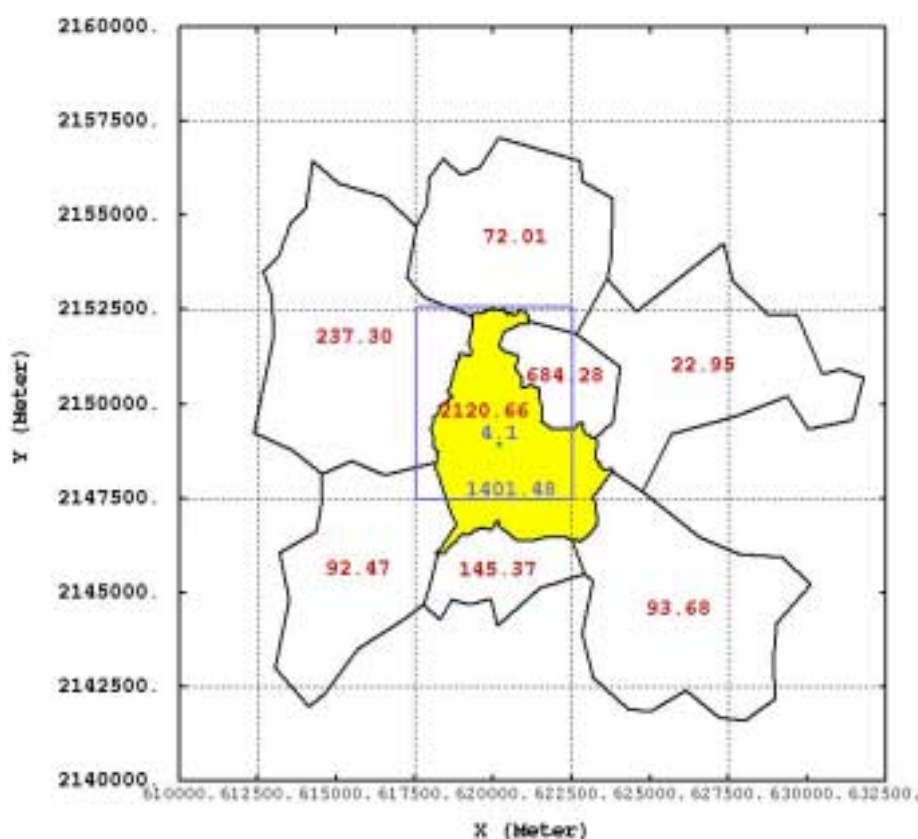


Figure 28: Calcul de la densité de population de Montluçon par SIG.

La densité de population de la maille entourée de bleu est égale à la moyenne pondérée des densités de population des cinq communes de cette maille. Les pondérateurs sont proportionnels aux surfaces des polygones incluses dans la maille. La densité de population ainsi calculée est de 1401 hab/km² alors que celle de la commune de Montluçon est de 2 121 hab/Km². Cette valeur est plus faible à cause des communes voisines moins densément peuplées.

Une grille de densité de population a été créée selon le principe décrit à la Figure 28. Mais afin d'étudier les corrélations entre la densité de population et les concentrations d'ozone, il est nécessaire que les valeurs numériques de ces deux variables soient connues aux mêmes points. Deux procédés sont envisagés pour attribuer une valeur de densité de population au point d'implantation d'un tube : soit la valeur affectée au point est celle de la commune dans laquelle celui-ci se trouve, soit elle est égale à la valeur de la maille dont le centre est le plus proche du tube. La Figure 29 présente le nuage de corrélation entre les données de densité qui résultent de ces deux modes d'affectation. On observe quelques différences importantes qui concernent plus particulièrement les points d'échantillonnage situés dans les villes ou à proximité de ces dernières. La seconde façon de faire a été préférée ; de cette manière, la densité de population attribuée à un point (*migrée vers un point*) est toujours représentative d'une même superficie (5 km x 5 km).

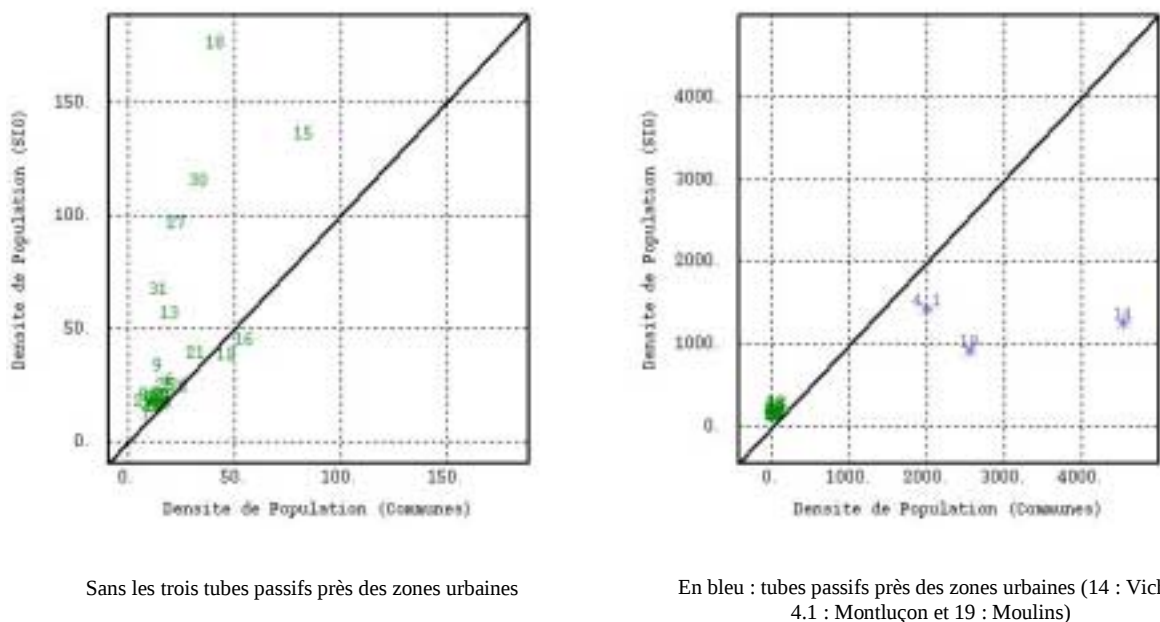
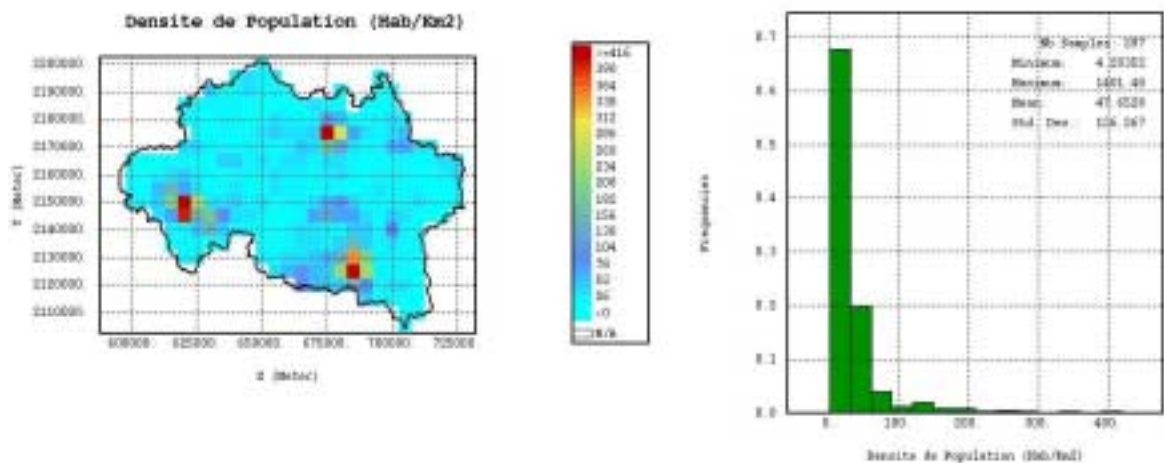


Figure 29 : Nuage de corrélation entre les densités de population affectées de deux manières différentes aux points d’implantation des tubes passifs (hab/km²), ligne droite : bissectrice

La distribution spatiale de la variable de densité de population est influencée par la présence des agglomérations. L’Allier compte trois zones urbaines importantes qui correspondent aux communes de Domérat - Montluçon, Moulins - Yzeure et Bellerive -Vichy – Cusset (Figure 30).

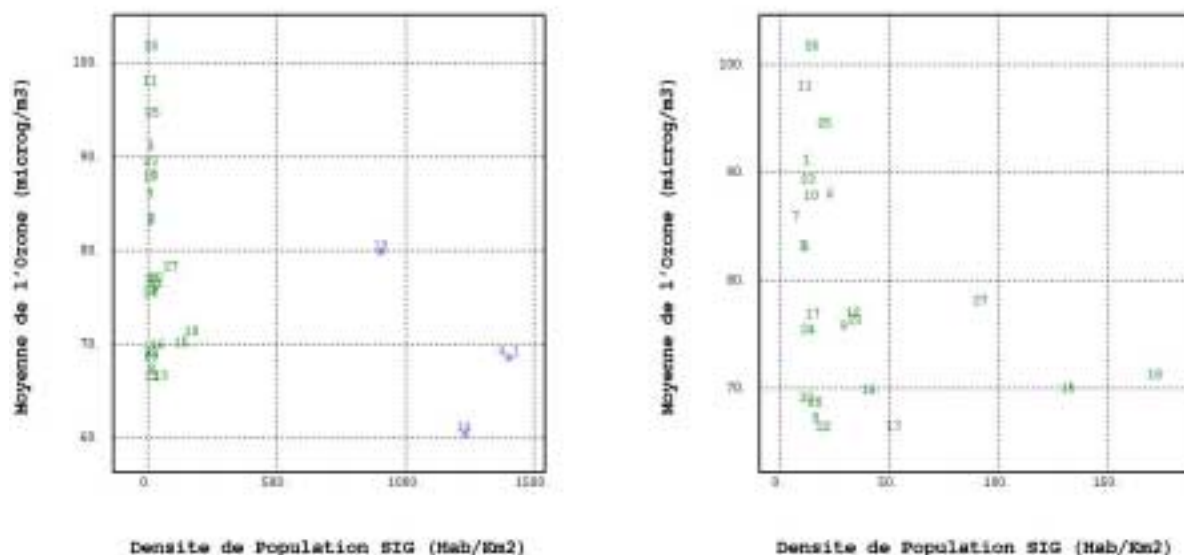
Dans le reste du département les valeurs de densité de population sont très faibles, puisqu’il s’agit de zones rurales.



Histogramme sans les trois grilles près des zones urbaines

Figure 30 : Carte et histogramme de la densité de population (SIG) de l’Allier (Hab/Km2)

L'examen du nuage de corrélation entre la densité de population et la concentration moyenne d'ozone des quatorze séries hebdomadaires (Figure 31) montre que les niveaux de concentration dans les villes (points n°4.1, 14 et 19) et à proximité de ces dernières (18, 15 et 27) sont en général plus faibles que dans le reste du département (points n°4.1, 14 et 19). La courbe que l'on peut ajuster à ces deux graphiques est de la forme : $Y = -a * \ln(X) + b$.



En bleu : tubes passifs près des zones urbaines (14 : Vichy, 4.1 : Montluçon et 19 : Moulins)

Sans les trois tubes passifs près des zones urbaines

La relation entre ozone et population s'écrit : $[O_3] = -3,4 \ln(\text{pop}) + 90$

Figure 31 : Nuage de corrélation entre la densité de population et la concentration moyenne de l'ozone

6.3 ANALYSE DES SURFACES BATIES DE L'ALLIER

Les polygones qui représentent les surfaces bâties du département de l'Allier ont été transformés en données de densité de bâti (km^2/km^2) régulièrement réparties sur un maillage (Figure 32).

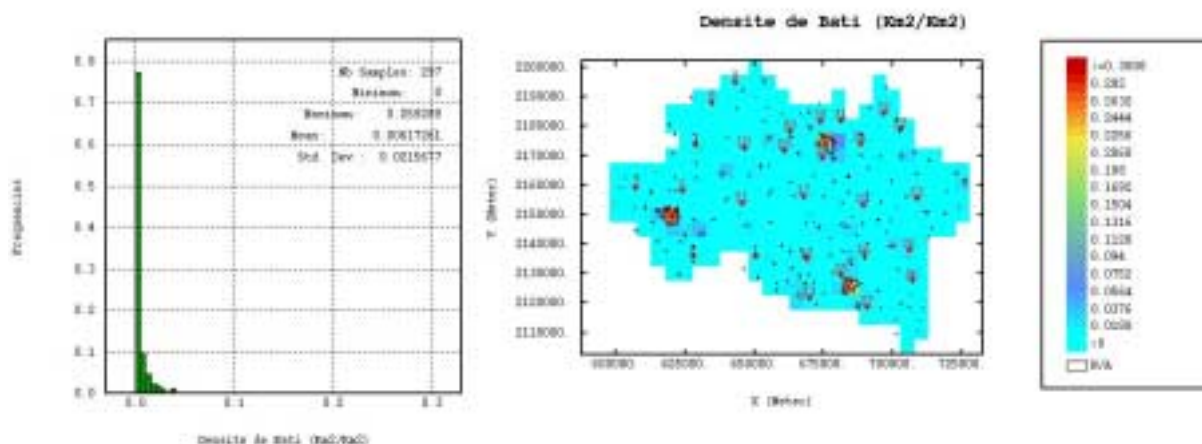
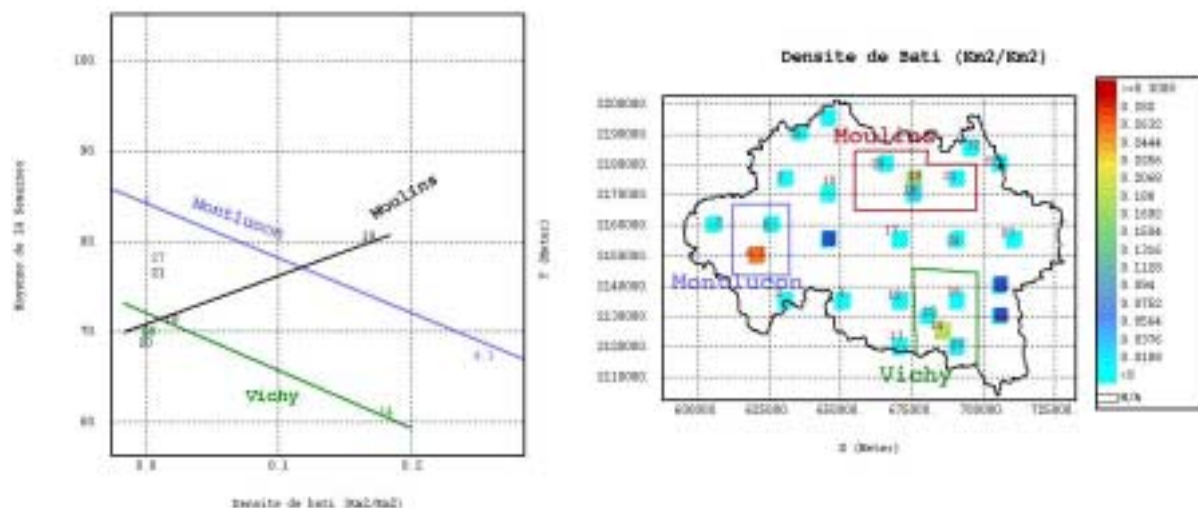


Figure 32 : Histogramme et carte de densité de bâti avec implantation des zones bâties

L'analyse de cette variable est conduite de la même façon que l'étude de la densité de population. A chaque point de mesure est affectée la valeur de densité de bâti de la maille dont le centre est le plus proche du point (Figure 33). Cette variable atteint environ $0.2 \text{ km}^2/\text{km}^2$ aux points 4.1, 14 et 19 (centres urbains) et prend des valeurs quasi nulles partout ailleurs.



Lignes droites colorées : régression linéaire entre les valeurs de concentration mesurées dans les villes et à proximité de celles-ci et les valeurs de densité de bâti

Figure 33 : Analyse de la variable *densité de bâti*. Les carrés colorés représentent les densités de bâti affectée aux différents points de mesure.

La valeur de la concentration moyenne d'ozone dans les villes est comparée aux valeurs des tubes situés dans les alentours. Les concentrations d'ozone sont plus faibles dans les centres urbains, là où le bâti est le plus dense. Un point d'échantillonnage situé à Moulins fait néanmoins exception.

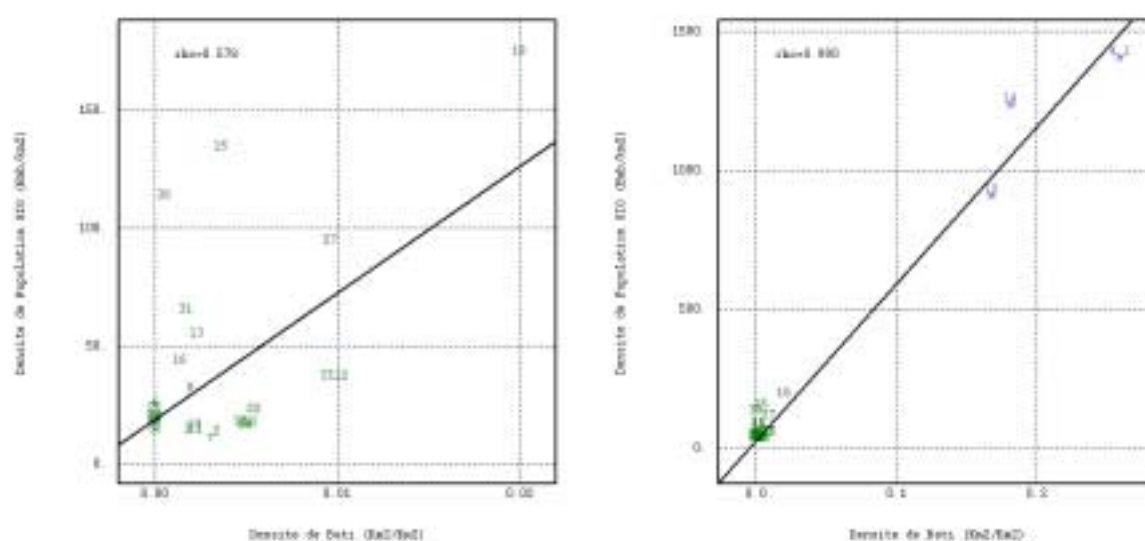


Figure 34 : Nuage de corrélation entre densité de population et densité de surface bâtie avec les trois points des villes (figure de droite), sans les points des villes (figure de gauche)

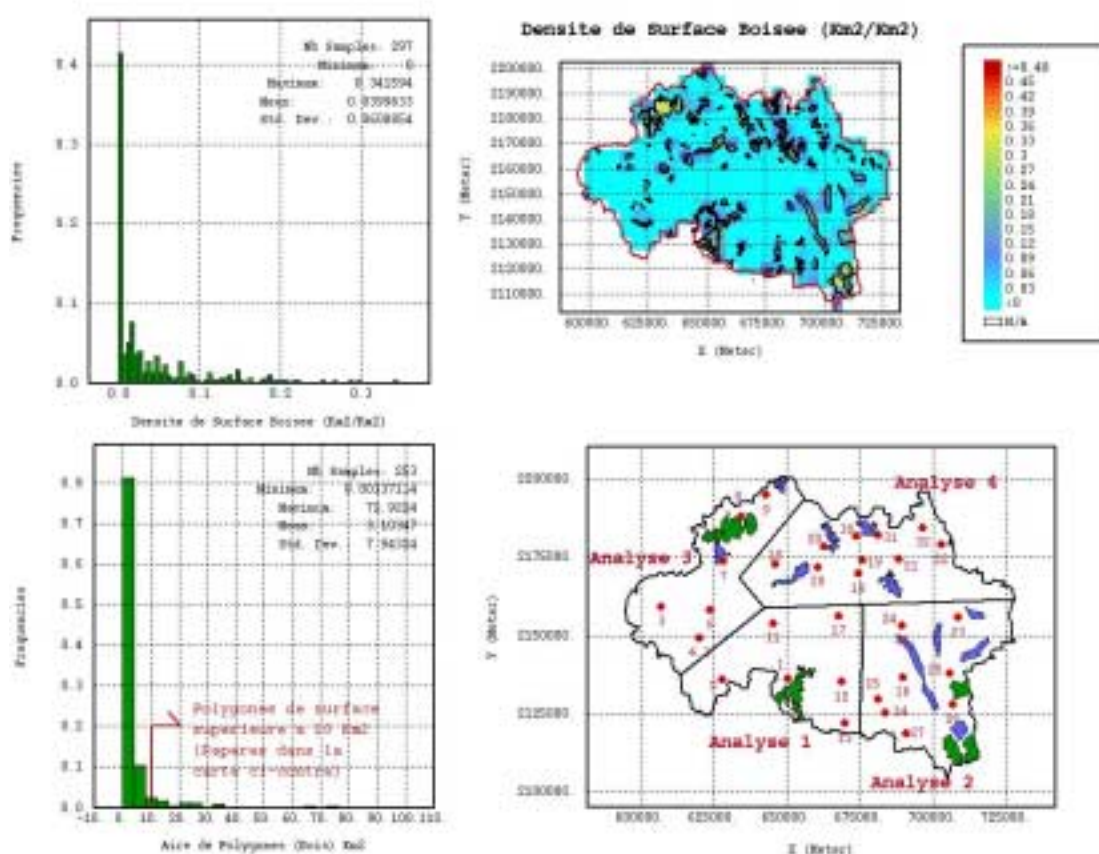
Il existe une corrélation positive entre la densité de surface bâtie et la densité de population (Figure 34). C'est pourquoi seule l'une de ces deux variables sera conservée dans la suite de l'étude.

Nous retenons la variable de densité de population qui pose moins de difficulté du point de vue mathématique. La variable *densité de bâti* s'annule en effet sur plusieurs mailles de la grille, ce qui risque de compliquer l'analyse.

6.4 ANALYSE DE LA VARIABLE DENSITE DE SURFACE BOISEE

Les 253 polygones qui représentent les zones boisées du département de l'Allier occupent une surface totale de 785 km², soit 11 % de la surface du département (qui est de 7 384 km2).

De petites dimensions pour la plupart, ils sont disséminés dans tout le domaine d'étude (Figure 35). Seulement 7% d'entre eux ont une superficie supérieure à 10 km², et 2% une superficie supérieure à 25 km².



En haut : à gauche: Histogramme de la variable densité de surface boisée

A droite : Carte de la variable *densité de surface boisée* avec implantation des polygones représentant les zones boisées

En bas : A gauche : Histogramme des aires des polygones

A droite : Sélection des polygones de surface supérieure à 10 Km2 (en bleu ceux dont la surface est comprise entre 10 et 25 km², en vert ceux dont la surface est supérieure à 25 km²). Les points indiquent la localisation des tubes passifs. Les traits délimitent les quatre zones analysées séparément.

Figure 35 : Histogramme et carte de densité de surface boisée, avec implantation des zones boisées

La variable *densité de surface boisée* représente la surface totale des polygones contenus dans chaque maille, rapportée à la surface de la maille. Elle varie entre 0 et 0.34 km²/(km² de maille).

Cette variable calculée par SIG est fonction de l'emplacement de la maille par rapport aux polygones et de la taille de la maille. Aussi peut-elle prendre une valeur nulle alors même que le tube est implanté non loin d'une zone boisée. Un simple étude des corrélations entre concentration et densité de surface boisée est donc insuffisante. Pour cette raison et afin de mener à bien l'analyse de l'influence des zones boisées sur la pollution par l'ozone, le département de l'Allier a été divisé en quatre zones géographiques définies d'après la localisation des principales surfaces boisées. Une analyse plus fine est conduite séparément dans chacune de ces zones.

6.4.1 Analyse de la zone 1

Dans cette zone sont localisés 6 tubes passifs : 1, 2, 11, 12, 13 et 17. Le but est de comparer la moyenne des concentrations mesurées par ces tubes à la distance qui les sépare de la grande surface boisée située dans cette région. Dans la Figure 35, on peut remarquer que le point le plus proche de la surface boisée est le numéro 1. La valeur de concentration en ce point est l'une des plus importantes de la zone, soit 90 µg/m³ (moyenne de la zone : 81 µg/m³). Les valeurs des autres tubes sont plus faibles, hormis celle du tube 11.

6.4.2 Analyse de la zone 2

Dans cette zone sont localisés 8 tubes passifs : 14, 15, 16, 23, 24, 25, 26 et 27. Les concentrations d'ozone les plus importantes sont relevées aux points 23, 25 et 26 (> 89 µg/m³), les plus proches des surfaces boisées. Les concentrations les plus faibles sont pour leur part relevées aux points 14, 15 et 16 (< 70 µg/m³, moyenne de la zone : 80 µg/m³), les plus éloignés des surfaces boisées.

Il faut noter que la densité de surface boisée aux points 15 et 27 est relativement importante alors que ceux-ci ne se trouvent près d'aucun grand massif forestier. Ceci est dû à la présence de zones boisées de petite surface dans les mailles les plus proches de ces points. Inversement, la densité de surface boisée en certains points, tels le 23, le 24, le 25 et le 26, est faible alors qu'ils sont situés près de vastes étendues de bois. Cela montre l'inconvénient de se limiter à une simple analyse des corrélations entre les concentrations et la variable *densité de surface boisée*.

6.4.3 Analyse de la zone 3

Dans cette zone sont localisés 6 tubes passifs : 3, 4.1, 6, 7, 8, et 9. Les tubes 7, 8 et 9 sont les plus proches des bois, mais parmi eux, seul le tube 7 présente une concentration d'ozone plus élevée (85 µg/m³, moyenne de la zone : 77.5 µg/m³). Au sud de cette zone, et à l'exception du tube de Montluçon (4.1), les concentrations mesurées sont assez importantes : tube 6 (87 µg/m³) et tube 3 (82 µg/m³). Ces tubes sont éloignés des zones boisées.

6.4.4 Analyse de la zone 4

Dans cette zone sont localisés 7 tubes passifs : 10, 18, 19, 20, 21, 22, 32. Le tube 20 est le plus proche d'un domaine boisé, mais la concentration d'ozone atteinte en ce point n'est que

de $67 \mu\text{g}/\text{m}^3$. La concentration la plus élevée a été relevée au point 10 ($87\mu\text{g}/\text{m}^3$, moyenne de la zone : $74 \mu\text{g}/\text{m}^3$), situé non loin d'un bois.

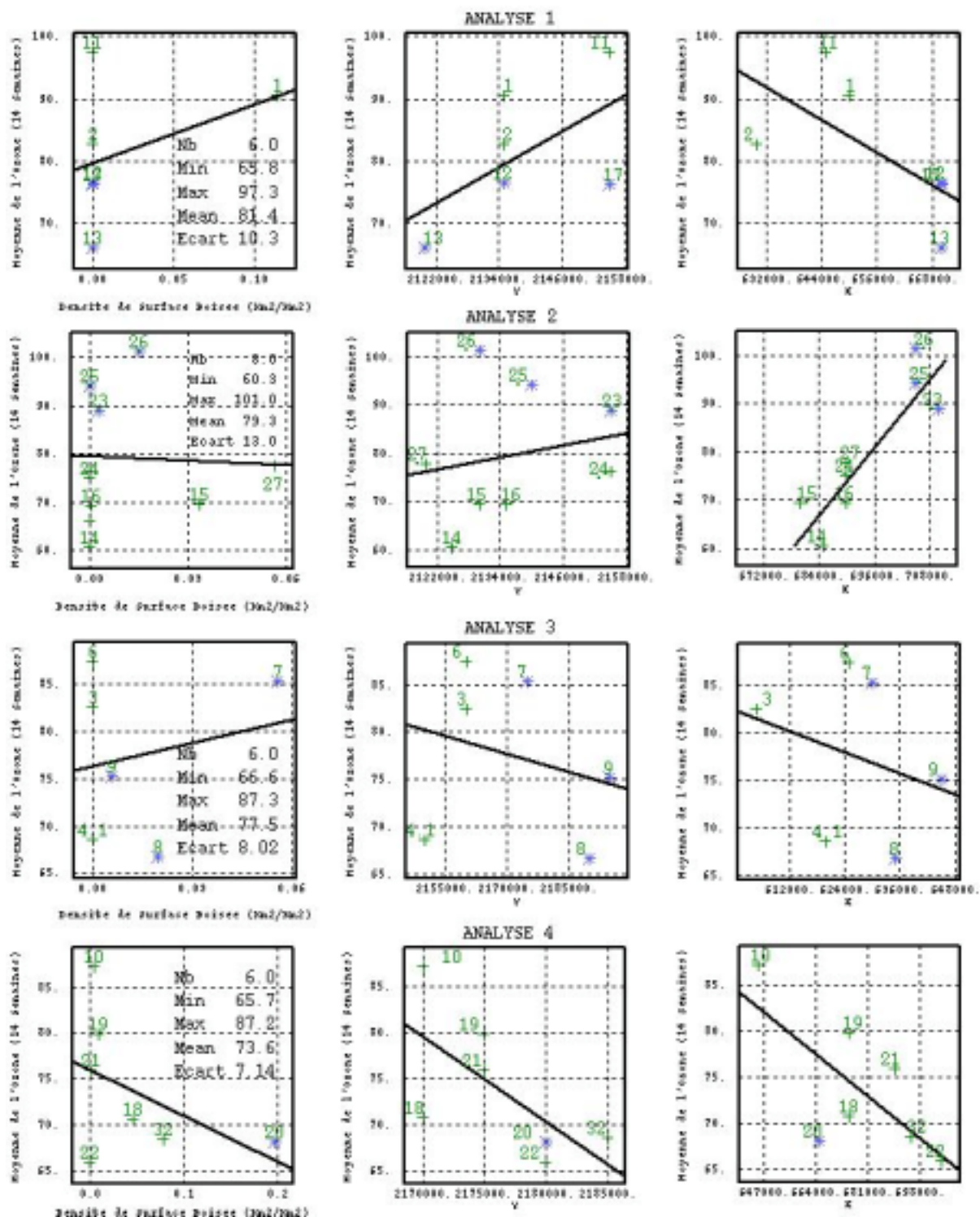


Figure 36 : Analyses de la variable *densité de surface boisée* dans les quatre zones du domaine d'étude

La question de savoir s'il existe une corrélation entre les concentrations d'ozone mesurées par les tubes passifs et la distance qui sépare ces tubes d'une surface boisée d'au moins 10 km² ne trouve pas de réponse immédiate. D'après l'étude qui vient d'être réalisée, il semblerait que la concentration moyenne d'ozone en un point proche d'une zone boisée ait plus de chance d'être élevée. Toutefois, en raison des nombreux facteurs qui influencent l'ozone, certains tubes situés à proximité d'étendues boisées mesurent de faibles concentrations (ex : tubes n° 8, 9, 20), tandis que des tubes plus éloignés (ex : tubes n° 11, 6, 3) mesurent des valeurs plus importantes.

Une dernière analyse consiste à savoir si l'aire des surfaces boisées est corrélée à la concentration d'ozone. Nous avons comparé à cette fin les concentrations moyennes d'ozone des quatre secteurs géographiques à l'aire boisée de ces secteurs (Figure 37). Aucune relation évidente n'apparaît. **La variable densité de surface boisée ne sera donc pas employée dans la suite de l'étude.**

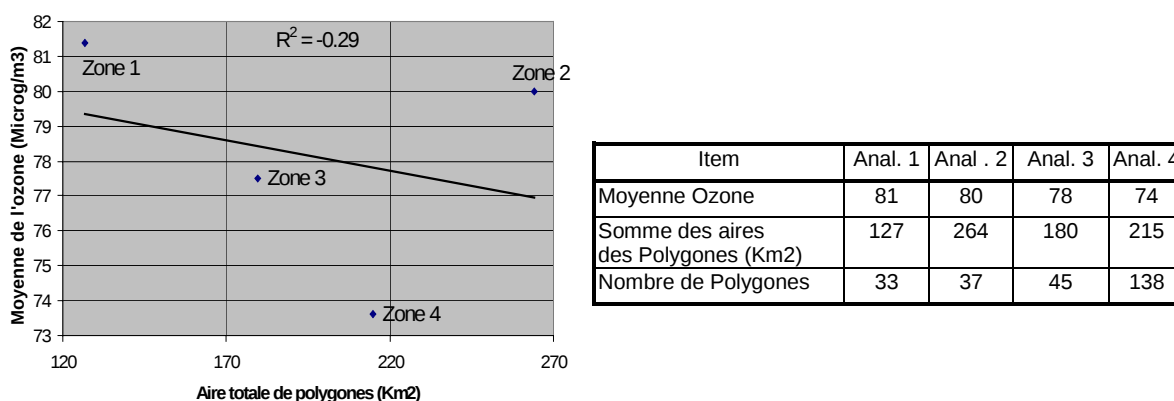


Figure 37 : Relation entre les concentrations moyennes d'ozone et l'aire totale des surfaces boisées dans les quatre zones géographiques

6.5 ANALYSE DU RELIEF

Les valeurs d'altitude mesurées par GPS sont connues en tout point d'échantillonnage. Nous disposons par ailleurs de l'altitude des chefs-lieux de communes. Afin d'obtenir une représentation du relief de l'Allier, et de construire une grille de la variable altitude, un **krigeage ordinaire** est réalisé à partir de ces données communales.

Le modèle de variogramme utilisé dans les équations de krigage, ainsi que sa validation croisée et les résultats de l'estimation sont présentés dans les figures et les tableaux qui suivent (Figure 38, Figure 39 et Figure 40, Tableau 5 et Tableau 6). La maille d'estimation est de 5 km.

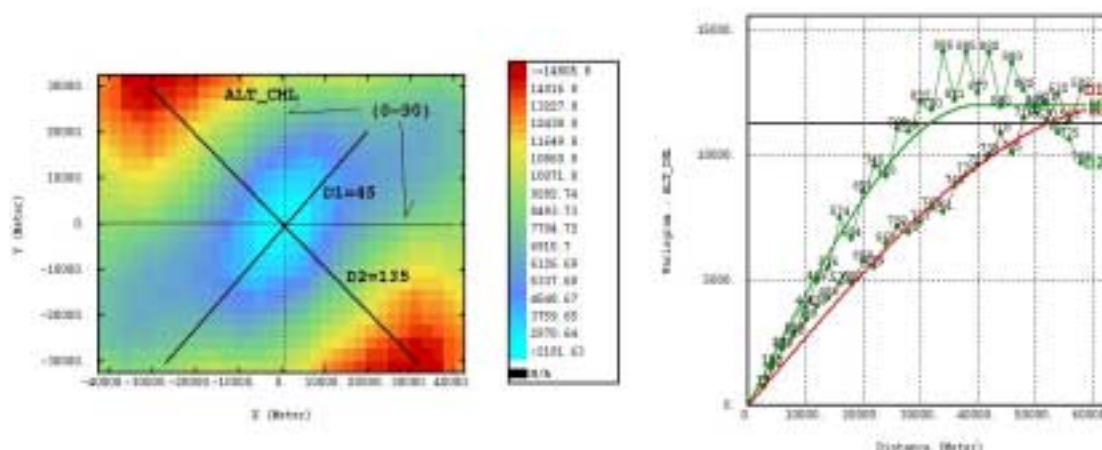


Figure 38 : Carte Variographique et Modèle bidirectionnel ajusté pour l'altitude de communes

<i>Modèle</i>	<i>Palier</i>	<i>Portée D1</i>	<i>Portée D2</i>
Sphérique	12 000	45°:66 000 Km	135 : 40 000 Km

Tableau 5 : Caractéristiques du modèle bidirectionnel ajusté

On est en présence d'une anisotropie dans la direction 45°. La variance des données d'altitude est d'environ 12000. Cette valeur a d'ailleurs été prise pour palier lors de l'ajustement du modèle de variogramme. L'altitude varie entre 200 m au nord du département et un peu moins de 1000 m au sud (premières élévations du Massif Central). Elle est en moyenne de 320 m.

Le modèle choisi s'ajuste très bien aux données dans le domaine d'étude. Le nuage de corrélation entre valeurs vraies et valeurs estimées par validation croisée est peu dispersé. Les écarts les plus importants (points bleus sur la Figure 39) sont associés aux fortes valeurs. Ils sont causés par l'effet de lissage du krigeage.

Si l'on voulait reproduire les pics situés au sud-est du département, il serait préférable de recourir à la technique géostatistique des simulations conditionnelles plutôt qu'à une estimation par krigeage ordinaire.

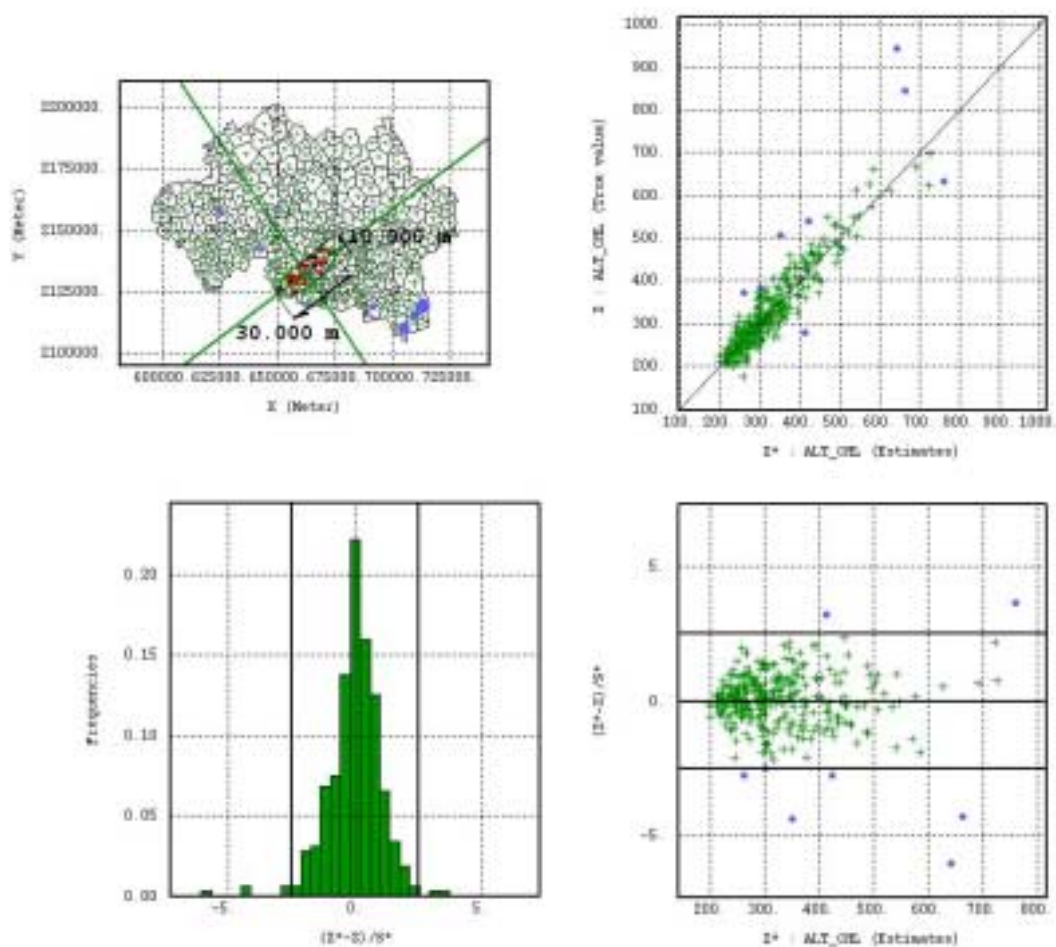


Figure 39 : Validation croisée du modèle de l'altitude des communes

Nb	Erreur		Erreur réduite	
	Moyenne (=0)	Variance	Moyenne (=0)	Variance (1)
320	-0.38	1676	-0.005	1.2

Tableau 6 : Statistiques de la validation croisée

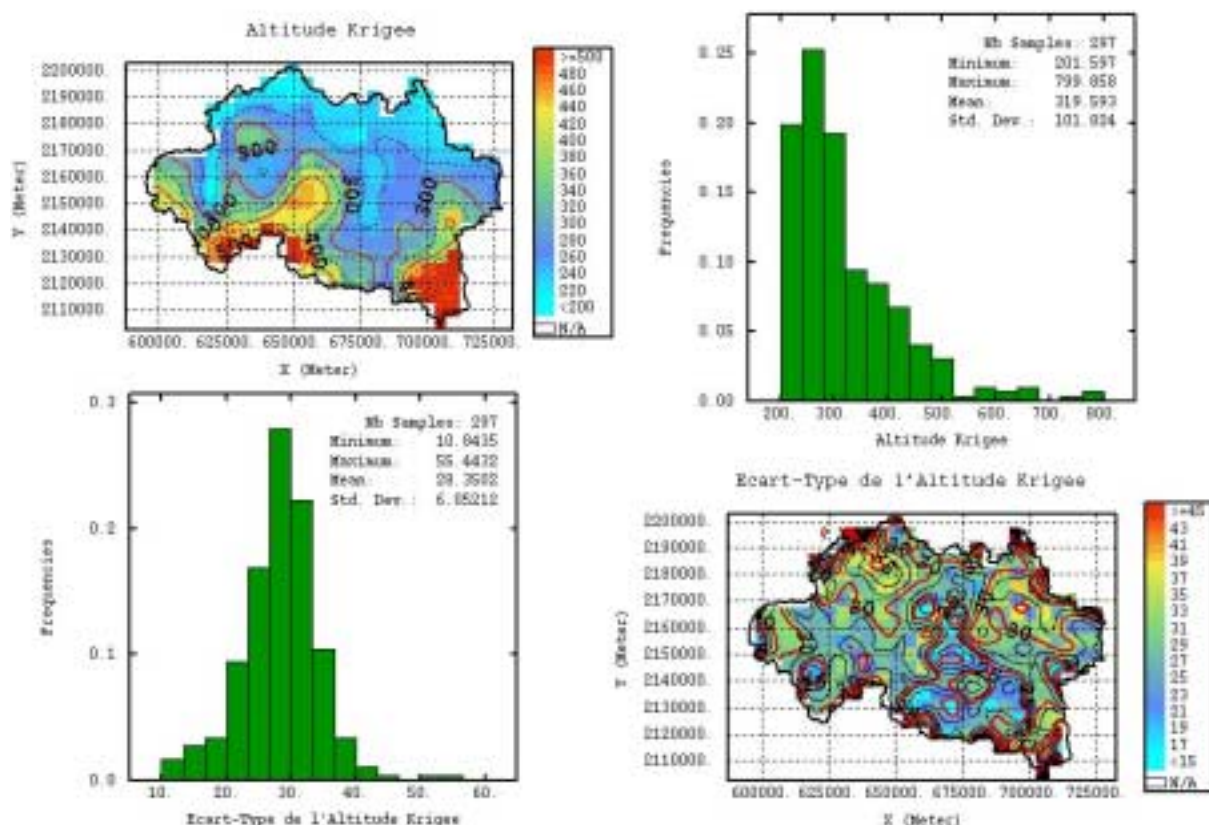


Figure 40 : Cartes d'estimation et écart-type de krigeage de l'altitude des communes

Les valeurs en rouge dans les cartes et les histogrammes de la Figure 40 indiquent les estimations faites hors du département. Ces zones qui sortent de notre domaine d'étude présentent le plus grand écart-type de krigeage, phénomène explicable par l'absence de données dans ces régions. A l'intérieur du département en revanche, les écarts-types sont faibles (moyenne de 28 m), en raison des nombreuses données.

Les données d'altitude estimées par krigeage et migrées vers les tubes sont en accord avec les données d'altitude mesurées par GPS (Figure 41). Trois points seulement s'écartent de la bissectrice. Il s'agit de tubes situés à l'extrême sud-est du département (tubes n°25, 26, et 27), là où l'effet de lissage du krigeage a été mis en évidence. Il existe par ailleurs une **corrélation linéaire positive entre l'altitude et l'ozone** (le coefficient de corrélation est de 0.61).

L'altitude krigeée est donc prise comme variable auxiliaire.

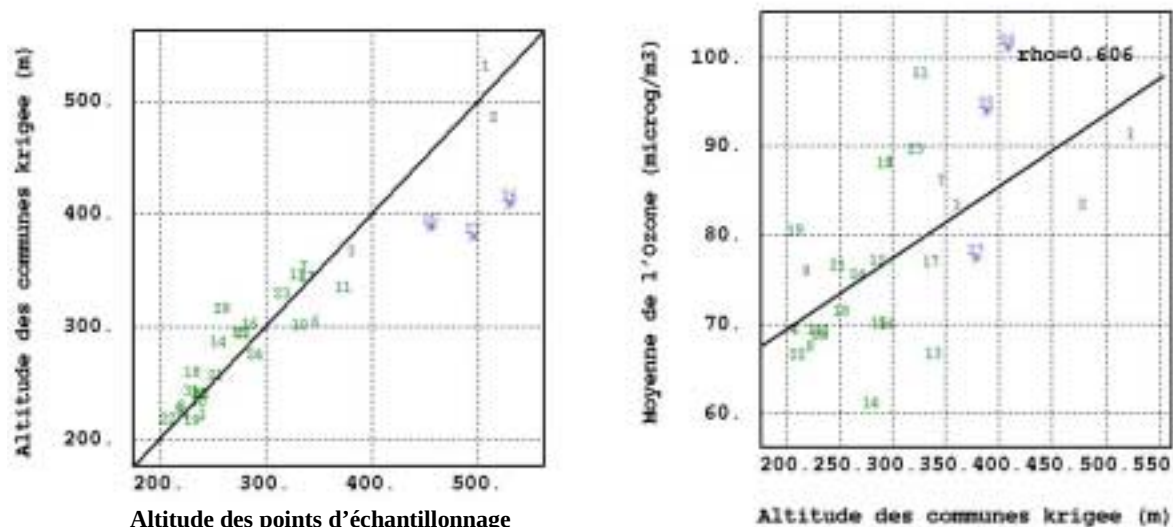


Figure 41 : Nuages de corrélation entre l'altitude krigée et l'altitude des points d'échantillonnage, entre l'altitude krigée et les concentrations d'ozone

6.6 MODELISATION MULTIVARIABLE DE L'OZONE

Les deux variables auxiliaires sélectionnées, *densité de population* et *altitude krigée*, sont prises en compte dans l'estimation par krigeage avec dérive externe.

La dérive externe définie à la suite de différents essais s'écrit :

$$\text{Dérive externe} = \{5.13 * (\text{Altitude réduite}) - \{1.94 * \text{Ln} (\text{Population réduite})\}$$

$\text{Altitude réduite} = \text{Altitude} / (\text{écart-type de l'altitude})$

$\text{Population réduite} = \text{Population} / (\text{écart-type de la population})$

La division par l'écart-type permet de rendre les variables adimensionnelles et d'en faire la somme.

Cette dérive présente les mêmes caractéristiques spatiales que la pollution d'ozone, avec des valeurs faibles près des agglomérations et en basse altitude, et des valeurs plus élevées dans les zones montagneuses (Figure 42).

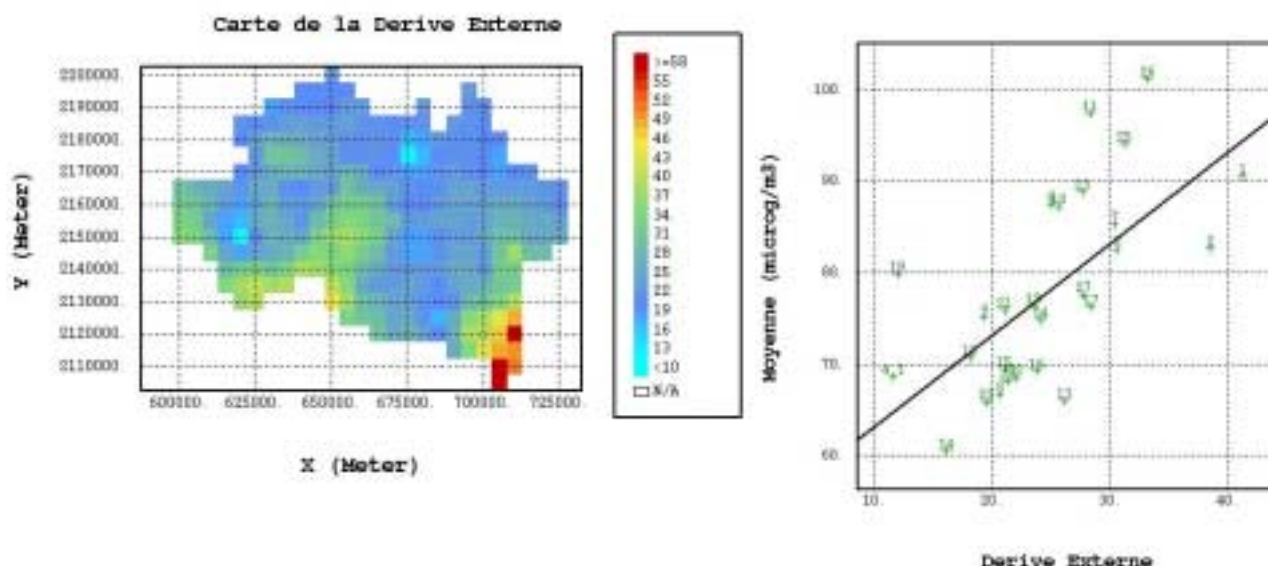


Figure 42 : Carte de la dérive externe et corrélation avec la variable de concentration

La corrélation linéaire avec la concentration moyenne d’ozone confirme la bonne qualité de cette dérive. Il reste à évaluer cette corrélation pour chacune des semaines de mesure sélectionnées, à savoir les semaines 1, 3, 6, 7, 8 et 13. Les nuages de corrélation entre les six séries de mesures et la dérive externe sont plus ou moins dispersés (Figure 43), mais l’existence d’une relation linéaire est toujours avérée.

Les modèles géostatistiques qui doivent être en accord à la fois avec les données expérimentales et avec la dérive externe sont alors ajustés. Du fait que les coefficients a et b de la dérive, et en conséquence les résidus, sont calculés implicitement par Isatis, l’utilisateur ne peut modéliser de façon directe le variogramme des résidus. Pour réaliser cette modélisation, l’utilisateur propose au logiciel différents modèles variographiques dont celui-ci ajuste le palier ou la pente et qu’il classe par ordre de préférence d’après les résultats numériques de la validation croisée. A la suite de plusieurs essais, et en testant différents voisinages, on retient le modèle qui conduit aux meilleurs résultats selon le critère de la validation croisée.

Les modèles ici obtenus sont des modèles linéaires de la forme : $\gamma(h) = m * |h|$. Le Tableau 7 indique les pentes (m) pour chaque semaine.

<i>Semaine</i>	<i>Pente du modèle linéaire</i>
1	0.00512
3	0.00653
6	0.00522
7	0.01266
8	0.00552
13	0.00601
Moyenne	0.00282

Tableau 7 : Caractéristiques des modèles ajustés en présence de la dérive externe

Les concentrations estimées par krigeage avec dérive externe sont proches de celles qui ont été estimées par krigeage ordinaire isotrope. Les cartes obtenues sont cependant moins lisses (Figure 44) et elles respectent mieux les caractéristiques du phénomène mises en évidence lors de l'analyse des données (concentrations plus élevées en altitude). Les écarts-types de krigeage (Figure 45) diminuent par rapport au krigeage ordinaire, ce qui traduit un gain de précision (voir la comparaison effectuée pour la moyenne des 14 séries dans la Figure 46).

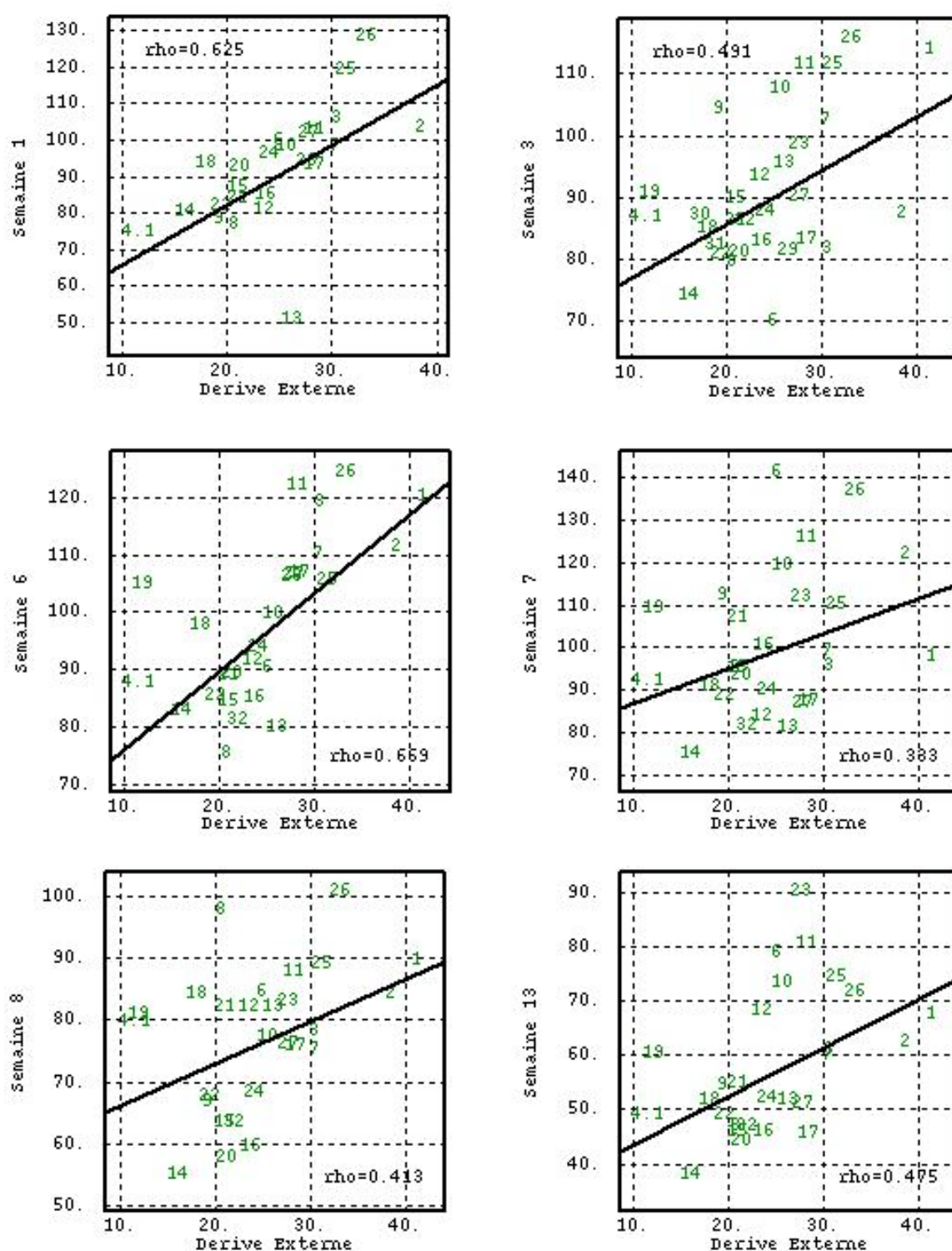


Figure 43 : Nuage de corrélation entre la dérive externe et les 6 séries de mesures sélectionnées

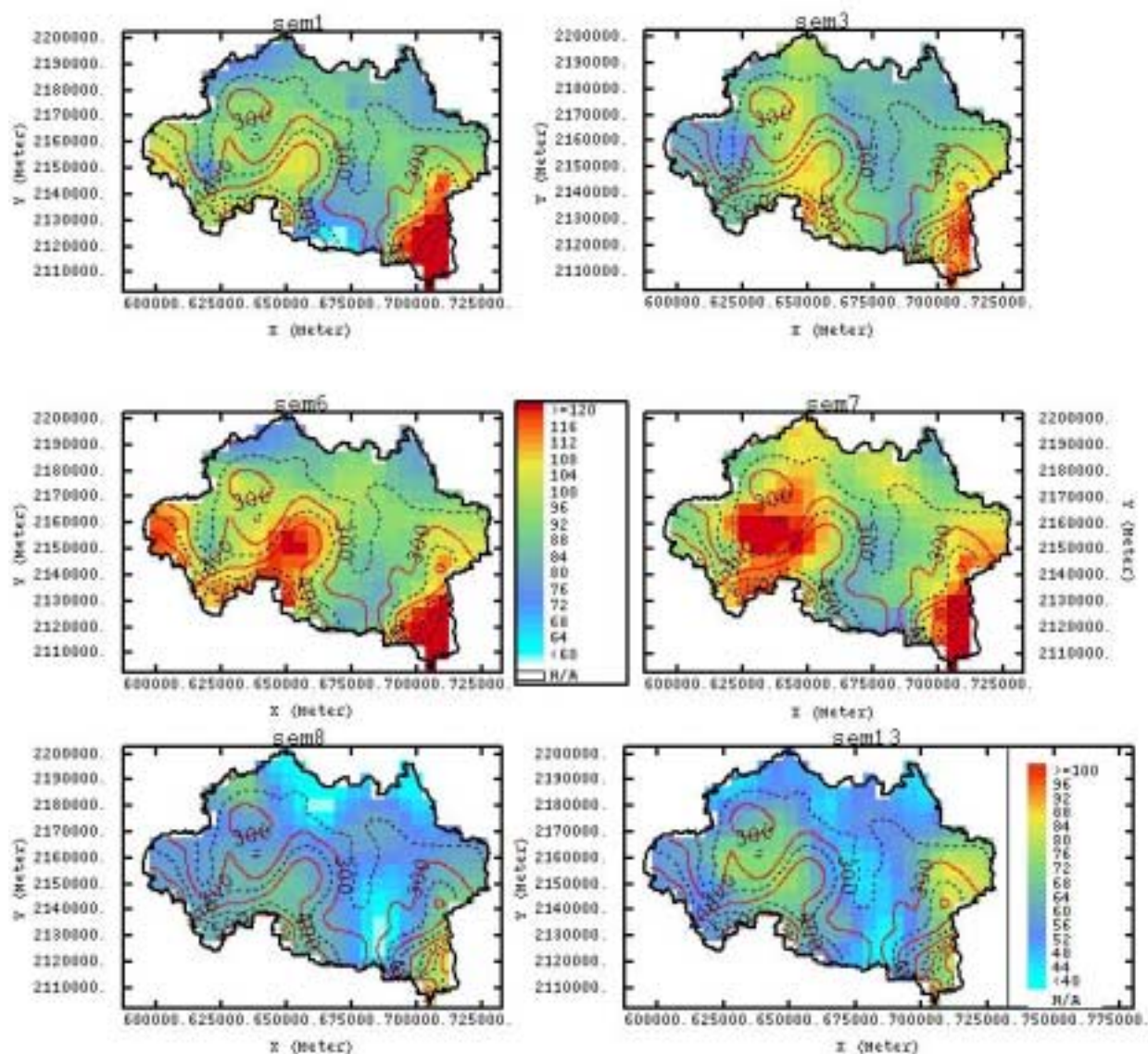


Figure 44 : Estimations par krigeage avec dérive externe.

Les courbes de niveau correspondent à l'altitude.

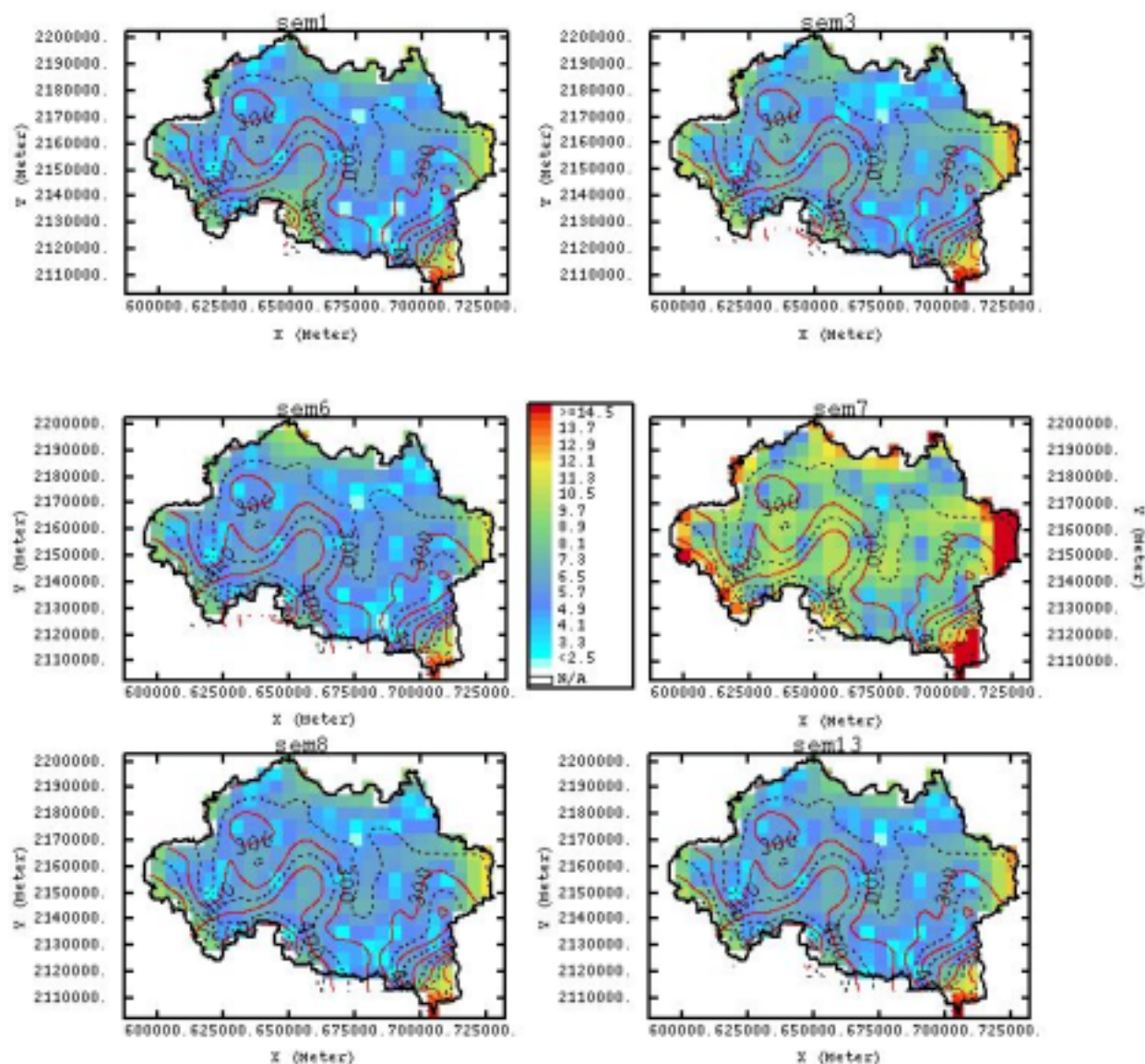


Figure 45 : Ecart-types du krigeage avec dérive externe

Les courbes de niveau correspondent à l'altitude.

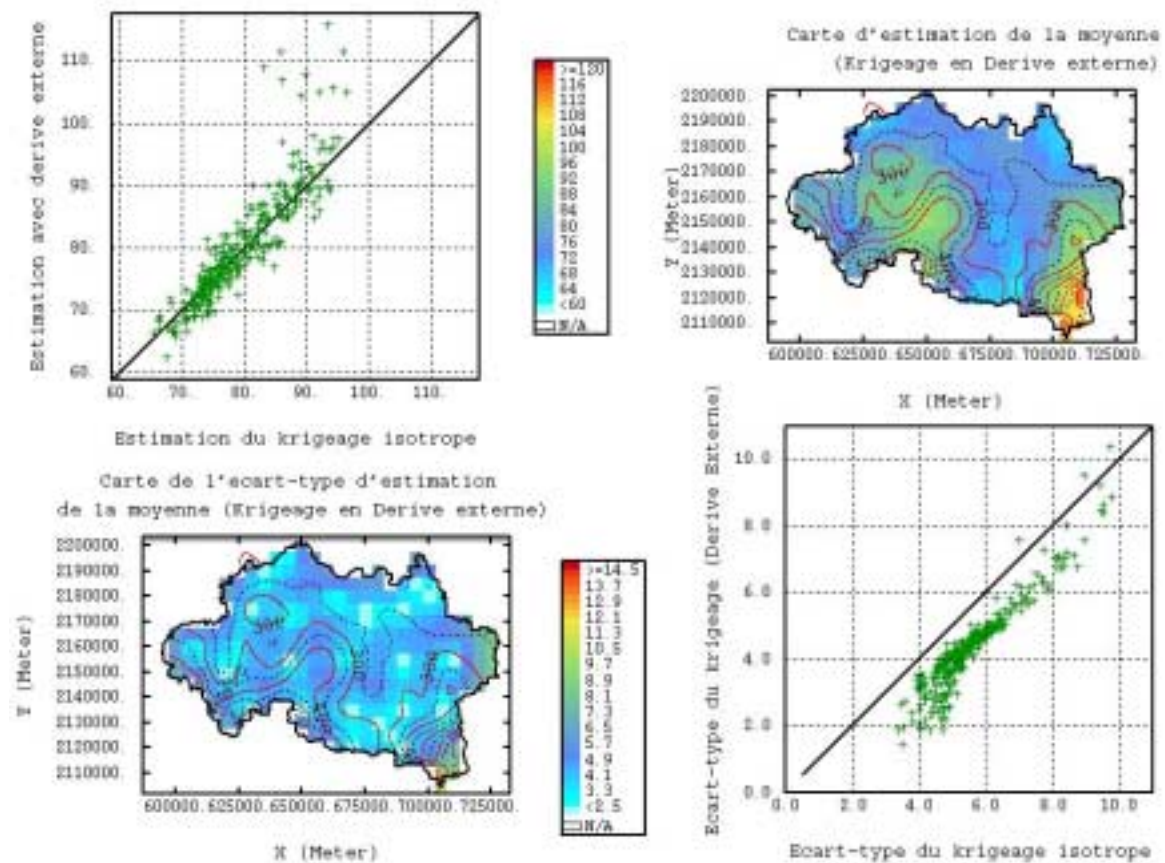


Figure 46 : Estimation de la moyenne des quatorze semaines de mesure par krigeage avec dérive externe

7. CONCLUSION

Ce travail avait pour fin d'illustrer sur un cas réel l'application des techniques géostatistiques à la cartographie de la qualité de l'air dans de vastes territoires. La distribution spatiale de l'ozone dans le département de l'Allier pendant la saison estivale a fait l'objet de cette étude. Les données issues d'une campagne de mesure qui s'est déroulée durant l'été 2001 ont été fournies par ATMO Auvergne.

Une analyse exploratoire des données d'échantillonnage passif a permis de dégager les principales caractéristiques du phénomène de pollution étudié et de modéliser sa structure spatiale. Cette analyse a été réalisée pour quelques semaines de mesure préalablement sélectionnées et pour la moyenne des quatorze semaines qui composaient la campagne. Les modèles ainsi établis ont été utilisés dans l'estimation par krigeage afin de dresser des cartes de concentration. L'intégration dans le krigeage de données liées aux concentrations (altitude et population) par l'intermédiaire d'une dérive externe a conduit à la réalisation de cartes de concentration plus détaillées.

Les résultats obtenus et leur cohérence avec les observations et les propriétés de l'ozone témoignent de l'efficacité des méthodes géostatistiques dans la cartographie régionale. Quelques difficultés ont été cependant rencontrées dans cette analyse. Elles concernent :

- l'ajustement du variogramme à l'origine. Ce problème est lié à celui de la variabilité à petite distance qui peut être évaluée par un échantillonnage approprié, et à celui de l'incertitude de mesure dont on peut estimer la variance à l'aide, là aussi, d'un échantillonnage adéquat.
- la recherche et la modélisation des anisotropies ;
- la prise en compte de données de type urbain dans une modélisation faite essentiellement à partir de sites ruraux ;
- l'interprétation des résultats en terme d'incertitude.

Les données disponibles n'ont pas ici permis d'approfondir ces questions. Celles-ci pourront être développées dans les travaux du LCSQA prévus pour l'année 2003 en tirant parti de l'expérience des AASQA sur ces sujets.

8. REFERENCES

- [1] ATMO AUVERGNE, 2001. Campagne de mesure de la pollution en ozone dans le département de l'Allier, été 2001.
- [2] ATMO PICARDIE, AIRPARIF, AIR NORMAND, OPAL'AIR, AREMA LM, AREMARTOIS, 2000. Campagne interrégionale d'étude de l'ozone et du dioxyde d'azote par tubes à diffusion passive, 26 juin-04 septembre 2000.
- [3] ASCOPARG, 2001. Mise en place d'une méthodologie pour la cartographie de l'ozone à l'échelle du département de l'Isère.
- [4] WACKERNAGEL H., 1992. Cours de géostatistique multivariable. Ecole des Mines de Paris, 62 p.