



Estimation des populations exposées aux dépassements de seuils réglementaires

1. Echelle urbaine

Laboratoire Central de Surveillance de la Qualité de l’Air

ESTIMATION DES POPULATIONS EXPOSEES AUX DEPASSEMENTS DE SEUILS REGLEMENTAIRES 1. ECHELLE URBAINE

Maxime Beauchamp, Laurent Létinois, Laure Malherbe (Ineris)

Décembre 2014

	Rédaction	Vérification	Approbation
NOM	Maxime Beauchamp Laurent Létinois Laure Malherbe	Bertrand Bessagnet	Laurence Rouïl
Qualité	Ingénieurs de l’Unité MOCA Direction DRC	Responsable de l’Unité MOCA Direction DRC	Responsable du Pôle DECI Direction DRC
Visa			



LE LABORATOIRE CENTRAL DE SURVEILLANCE DE LA QUALITE DE L'AIR

Le Laboratoire Central de Surveillance de la Qualité de l'Air est constitué des laboratoires de Mines Douai, de l'INERIS et du LNE. Il mène depuis 1991 des études et des recherches à la demande du Ministère chargé de l'environnement, et en concertation avec les Associations Agréées de Surveillance de la Qualité de l'Air (AASQA). Ces travaux en matière de pollution atmosphérique ont été financés par la Direction Générale de l'Énergie et du Climat (bureau de la qualité de l'air) du Ministère de l'Écologie, du Développement durable et de l'Énergie (MEDDE). Ils 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 au MEDDE et 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.

TABLE DES MATIERES

RESUME	7
REMERCIEMENTS ET COLLABORATIONS.....	9
INTRODUCTION	10
1. INTRODUCTION	11
2. PRESENTATION DES DONNEES	12
2.1 Air Pays de la Loire	12
2.2 ATMO Poitou-Charentes	13
2.3 ATMO Champagne-Ardenne	13
2.4 Lig’Air.....	13
3. DEFINITION DES DOMAINES D’ETUDE	14
4. PRESENTATION DES METHODES D’INTERPOLATION	15
4.1 Interpolation linéaire	15
4.2 Interpolation de Delaunay	16
4.3 Interpolation par krigeage	17
4.4 Choix techniques et méthodologiques	18
5. VALIDATION CROISEE.....	20
5.1 Scores	21
5.2 Principe du diagramme de Taylor	21
5.3 Diagramme de Taylor sur Nantes.....	22
5.4 Diagramme de Taylor sur Niort.....	24
5.5 Diagramme de Taylor sur Reims	26
5.6 Diagramme de Taylor sur Bourges.....	27
5.7 Diagramme de Taylor sur Orléans.....	28
5.8 Diagramme de Taylor sur Tours	28
5.9 Tests complémentaires sur l’interpolateur de Delaunay.....	29
6. VALIDATION PAR COMPARAISON A UNE SIMULATION DE REFERENCE.....	30
6.1 Données utilisées	30
6.2 Résultats.....	31
6.3 Lignes directrices pour un maillage efficace	34

7. APPLICATION DES DIFFERENTES METHODES D'INTERPOLATION.....	39
7.1 Expertise visuelle des cartographies.....	39
7.2 Temps de calcul et format de sortie	41
8. BILAN ET DISCUSSION.....	41
1. METHODES GEOSTATISTIQUES	43
2. INTERPOLATION OPTIMALE EN ASSIMILATION DE DONNEES	45
3. BILAN ET DISCUSSION.....	47
1. INTRODUCTION.....	48
2. SURFACE EN DEPASSEMENT	48
2.1 Cas de l'interpolateur linéaire	48
2.2 Cas de l'interpolateur de Delaunay	48
2.3 Résultats.....	49
3. POPULATION EXPOSEE	53
3.1 Les méthodes employées au sein des AASQA	54
3.2 Tests de sensibilité	55
3.2.1 Effet du maillage en entrée de l'interpolation.....	57
3.2.2 Effet de la résolution du maillage de sortie de l'interpolation.....	59
3.2.3 Effet du type de prise en compte du bâti	61
3.3 Bilan et discussion.....	61
CONCLUSION	63
REFERENCES	65
GLOSSAIRE.....	67
LISTE DES ANNEXES.....	69
ANNEXE 1 : TABLES DE SCORES EN VALIDATION CROISEE.....	71
ANNEXE 2 : NUAGES DE POINTS (VALIDATION CROISEE).....	77
ANNEXE 3 : CARTOGRAPHIE DES ECARTS AVEC LES INTERPOLATEURS LINEAIRES ET DE DELAUNAY (VALIDATION CROISEE).....	83
ANNEXE 4 : NUAGES DE POINTS (SUR GRILLE DE REFERENCE)	87
ANNEXE 5 : CARTOGRAPHIE DES ECARTS, AVEC LES INTERPOLATEURS LINEAIRES ET DE DELAUNAY (SUR GRILLE DE REFERENCE) – CAS DE BOURGES.....	92

Conformément aux Directives européennes sur la qualité de l'air et à leur transposition en droit français, les AASQA doivent caractériser les situations de dépassement de seuil dans les zones dont elles assurent la surveillance. Pour tout dépassement observé, il convient d'évaluer la surface en dépassement et d'estimer, selon le cas, la population exposée (seuil pour la protection de la santé) ou la superficie d'écosystème exposée (seuil pour la protection de la végétation).

La caractérisation des situations de dépassement en milieu urbain fait l'objet du présent rapport. Elle nécessite de disposer d'une information précise sur la distribution spatiale des concentrations dans l'agglomération. La modélisation de la dispersion est de plus en plus utilisée par les AASQA afin d'obtenir une telle information. Néanmoins, pour limiter les temps de calcul, les données ne sont généralement pas modélisées sur des grilles régulières mais sur des maillages adaptés aux variations des concentrations (plus lâches en situation de fond, et resserrés à proximité des axes routiers). C'est pourquoi il est important de définir une méthode d'interpolation qui spatialise intégralement les données du modèle.

L'interpolateur linéaire est largement utilisé. L'hypothèse de linéarité est arbitraire mais peut-être considérée comme valide, principalement à de faibles distances et lorsque les concentrations varient faiblement.

Des interpolateurs plus sophistiqués existent. Le krigeage, méthode d'estimation géostatistique, tient compte de la covariance spatiale des concentrations et permet d'interpoler après définition d'un modèle de covariance et d'un voisinage de points à prendre en compte. Apparue plus récemment, la méthode fondée sur la triangulation de Delaunay consiste à parcourir les triangles définis par les points de sortie du modèle et à interpoler linéairement les concentrations le long des arêtes. Elle permet de construire des polygones dans lesquels les concentrations sont comprises entre deux valeurs.

Afin d'être évalués et comparés, ces interpolateurs ont été appliqués à différents jeux de données de modélisation fournis par les AASQA. Ces données résultent de simulations effectuées sur les agglomérations de Bourges, Nantes, Niort, Orléans, Reims et Tours. Le NO₂, polluant pour lequel les données étaient les plus nombreuses, a été retenu pour cette analyse. Dans toutes les validations réalisées, l'interpolateur linéaire et l'interpolateur de Delaunay se révèlent les plus précis. Le krigeage ordinaire, dans son emploi usuel, est mal adapté à des données de typologie contrastée (fond/trafic) ; pour une utilisation appropriée, le modèle géostatistique doit être modifié, ce qui requiert des développements supplémentaires. Quelle que soit la méthode, la qualité de l'interpolation est liée à l'organisation spatiale du maillage de modélisation qui, outre une couverture régulière des zones de fond, doit permettre une bonne représentation des gradients de concentration près des routes et doit donc s'adapter au comportement du polluant considéré dans l'espace, qui n'est évidemment pas le même s'il s'agit de NO₂, d'O₃ ou de particules.

Dans les exemples étudiés, l'interpolation a été réalisée sans prise en compte des données de mesure. La deuxième partie du rapport offre un aperçu des méthodes qui existent en matière de correction des sorties de modèle par intégration des observations. Les méthodes dérivées de la géostatistique ainsi que les derniers travaux en matière d'assimilation de données en milieu urbain sont abordés sous forme de revue documentaire. Ces éléments pourront servir à la planification de travaux ultérieurs.

La troisième partie du rapport est consacrée à l'estimation des zones en dépassement et des populations exposées. La sensibilité des résultats aux paramètres de l'estimation est examinée. Si le maillage d'interpolation a peu d'influence dans la gamme de résolution testée (1m, 5m ou 10m), le mode de croisement entre zones de dépassement et population spatialisée a un effet notable. Un croisement maille à maille, si la population est préalablement répartie sur le maillage d'interpolation, ou, ce qui revient au même, un croisement au prorata de la surface de bâtiment intersectée est recommandé.

Les conclusions de ce travail seront reprises en 2016 dans une note de synthèse méthodologique et technique sur l'évaluation de la population et des écosystèmes exposés. Leur adaptation aux agglomérations et zones de grande taille sera examinée en 2016 avec les AASQA concernées.

REMERCIEMENTS ET COLLABORATIONS

Les auteurs remercient les associations Air Pays de la Loire, ATMO Champagne-Ardenne, ATMO Poitou-Charentes et Lig'Air pour les données fournies et leur collaboration. Les sujets traités dans ce rapport tiennent compte des échanges avec la Commission de Suivi *Emissions, Modélisation et Traitement de Données*. Que celle-ci soit remerciée pour la qualité des discussions.

INTRODUCTION

Les législations française et européenne sur la qualité de l'air ambiant définissent des valeurs seuils pour la protection de la santé humaine et de la végétation. En cas de dépassement de seuil, il convient de délimiter les zones géographiques touchées par le dépassement et d'estimer la population ou la surface de végétation exposée. Ces informations doivent être fournies chaque année à la Commission européenne lors de la déclaration annuelle des données de qualité de l'air.

Le LCSQA a été chargé par le MEDDE d'harmoniser les pratiques et d'élaborer une méthodologie nationale qui sera appliquée chaque année par l'ensemble des AASQA. La présente étude, réalisée en concertation avec la Commission de Suivi Emissions, Modélisation et Traitement de Données, s'inscrit dans cet objectif. En réponse aux besoins et priorités exprimés lors du séminaire LCSQA du 27/09/2013, elle a été orientée sur l'échelle urbaine et sur les trois problématiques suivantes :

- cartographie des concentrations à partir de données de simulation numérique ;
- délimitation des zones de dépassement ;
- calcul de la population exposée ;

A chaque étape, il s'agit de rechercher la plus grande précision d'estimation tout en tenant compte des contraintes de calcul. Les cas d'application retenus pour ce travail sont des agglomérations de taille moyenne (jeux de données fournis par les AASQA). Les grandes zones et agglomérations feront l'objet de compléments ultérieurs.

La cartographie des concentrations est traitée dans la partie 1. Afin de réduire les temps de calcul, les AASQA sont souvent conduites à modéliser les concentrations sur un ensemble irrégulier de points appelé couramment « maillage intelligent ». Ce maillage est plus lâche dans les zones de fond et plus resserré près des voies de circulation. Pour cartographier les concentrations, une interpolation à partir des données modélisées est alors nécessaire. Cette opération se révèle délicate, les données étant de typologie différente, et peut engendrer des artefacts non représentatifs de la pollution réelle. Trois méthodes d'interpolation répandues dans les AASQA sont évaluées : l'interpolation linéaire, l'interpolation de Delaunay et le krigeage.

La partie 2 dresse un bref état des lieux bibliographique sur la correction des modélisations urbaines à l'aide de données de mesure. Sont abordées notamment les techniques de krigeage multivariable (cokrigeage des observations et des sorties de modèles) et les développements récents en matière d'assimilation de données.

La partie 3 s'attache à la délimitation des zones de dépassement et au calcul de la population exposée. Les jeux de données, et en particulier celui de Bourges (projet PREVISION'AIR), sont utilisés pour des tests de sensibilité. Le problème de la répartition spatiale de la population, qui a été précédemment étudié (Létinois, 2015), n'est pas abordé.

Le détail des résultats est joint en annexe.

PARTIE 1 : INTERPOLATION DES CONCENTRATIONS A PARTIR DE RESULTATS DE MODELISATION

1. INTRODUCTION

Cette partie a pour objet d'évaluer la qualité relative de différents interpolateurs utilisés dans le post-traitement des sorties de modèle. Cette comparaison porte sur des interpolateurs d'usage fréquent ou bénéficiant d'un intérêt particulier pour les applications en Qualité de l'Air ; il s'agit de :

- L'interpolateur linéaire : il fait la moyenne pondérée des concentrations aux points voisins ; les pondérateurs dépendent de la surface des triangles déterminés par le point d'intérêt et par les points voisins ;
- L'interpolateur de Delaunay : il consiste à construire des polygones par triangulation de Delaunay et à interpoler les concentrations linéairement le long des arêtes des triangles ;
- Le krigeage : cet estimateur géostatistique linéaire effectue une moyenne pondérée des concentrations dans un certain voisinage ; les pondérateurs sont obtenus par la résolution d'un système d'équations qui fait intervenir un modèle de covariance spatiale.

Pour les raisons invoquées, les méthodes suivantes n'ont pas été incluses dans l'analyse :

- La méthode du plus proche voisin : elle a été jugée insuffisante pour reproduire de manière réaliste les gradients de concentration à une échelle spatiale fine.
- D'autres interpolateurs linéaires simples (ex. : pondération par l'inverse de la distance, de la distance au carré ou à une autre puissance) : ils ne présentaient pas de valeur ajoutée particulière par rapport aux interpolateurs linéaires considérés plus haut.
- L'interpolateur bilinéaire : s'il est souvent utilisé pour passer d'une grille de résolution spatiale donnée à une grille de plus haute résolution, il ne permet pas d'interpoler à partir d'un maillage irrégulier ; or à l'échelle urbaine, un maillage irrégulier est souvent employé.
- L'interpolateur d'Akima, méthode d'interpolation bivariée (à deux dimensions) pour des jeux de données irréguliers, fondée sur l'ajustement d'une surface dont l'équation est un polynôme local de degré 5. Peu utilisé, cet interpolateur n'a pas été conservé dans cette étude.

Le travail est réalisé sur différents cas d'études décrits en section 2.

Deux méthodes de validation de la qualité de l'interpolation ont été retenues pour la présente étude : la validation croisée et la validation sur une grille de référence de haute résolution (simulation HD). Les cartographies obtenues par les différentes interpolations sont ensuite comparées.

2. PRESENTATION DES DONNEES

Quatre AASQA ont participé à cette étude en réalisant des simulations par modélisation à l'échelle urbaine à l'aide de l'outil ADMS-Urban, capable de produire dans un temps de calcul raisonnable des simulations de type gaussien et dont la qualité est donc plus limitée que celle obtenue avec des modèles lagrangiens. Nous précisons que l'évaluation de la qualité de ce modèle n'est pas un des objectifs de ce rapport.

En fonction des AASQA, divers polluants (NO_2 , PM_{10} , PM_{25} , Benzène) ont été modélisés et différentes statistiques ont été calculées (moyenne annuelle, centiles, etc.). Par souci d'homogénéité dans l'étude, les cartographies et les procédures de validation ont été réalisées sur la moyenne annuelle du NO_2 . Cependant, les conclusions de l'étude peuvent se généraliser à d'autres polluants et d'autres variables : ce qui importe est la variabilité des données en zone urbaine et notamment l'augmentation des concentrations à proximité des axes routiers : le maillage doit donc s'adapter à la variabilité propre du polluant considéré.

Afin de répondre aux besoins de l'étude, deux types de simulations ont été élaborés par les AASQA.

Un premier jeu de données sur un maillage, dit intelligent, est constitué de données simulées sur l'ensemble d'une agglomération. Les statistiques sont déterminées à partir de toutes les heures de l'année simulée.

Le second jeu de données est élaboré à partir d'une simulation de haute définition (HD) qui se focalise sur un quartier précis de chaque agglomération. Selon l'AASQA, la grille de sortie de cette simulation possède une résolution spatiale de 1m ou 2m sur une zone variant de 1 km^2 à quelques km^2 . Pour ce type de grille, les AASQA ont été confrontées à des temps de calcul trop importants vu les ressources disponibles. Elles ont donc dû faire des compromis sur la simulation. Certaines simplifications ont été réalisées : réduction du nombre d'heures simulées (prise en compte d'une heure sur 3 ou 5 heures), réduction du nombre de sources.

De ce fait, il est à noter qu'en un même point, les simulations HD peuvent présenter des écarts de concentration de quelques microgrammes par rapport aux simulations sur la grille intelligente. Malgré ces écarts, on supposera pour la suite de l'étude que la répartition spatiale des concentrations est correctement représentée par les simulations HD.

2.1 Air Pays de la Loire

Air Pays de la Loire a fourni des données modélisées sur l'agglomération de Nantes sur l'année 2012.

Le maillage intelligent couvre une zone de l'agglomération de Nantes d'environ 332 km² et le nombre de points du maillage est de 38279.

Une simulation HD à 1m de résolution sur une zone de 1km² a été réalisée sur le périphérique de Beaujoire. Certaines sources d'émissions ont été retirées.

2.2 ATMO Poitou-Charentes

ATMO Poitou-Charentes a fourni des données modélisées sur l'agglomération de Niort sur l'année 2012.

Le maillage intelligent couvre une zone de l'agglomération de Niort d'environ 1373 km² et le nombre de points du maillage est de 92604.

Une simulation HD à 2m de résolution sur une zone de 0.5 km² a été réalisée sur une zone concernée par un dépassement en NO₂. Seulement 1 heure sur 3 a été simulée.

2.3 ATMO Champagne-Ardenne

ATMO Champagne-Ardenne a fourni des données modélisées sur l'agglomération de Reims sur l'année 2012.

Le maillage intelligent couvre une zone de l'agglomération de Reims d'environ 336 km² et le nombre de points du maillage est de 36516.

Une simulation HD à 2m de résolution sur une zone de 0.66 km² a été réalisée sur une zone concernée par un dépassement en NO₂. Seulement 1 heure sur 5 a été simulée.

2.4 Lig'Air

Lig'Air a fourni des données modélisées sur les agglomérations d'Orléans (2010), Tours (2010) et Bourges (2012).

Les maillages intelligents sur les zones d'Orléans et Tours couvrent respectivement 979 km² et 2285 km² pour un nombre de points de 31271 et de 32020.

En ce qui concernant Bourges, la simulation couvre environ 1680 km² et le nombre de points du maillage intelligent est de 23043. Cette simulation est issue des travaux réalisés dans le cadre du projet *PREVISION'AIR* (projet inter-AASQA relatif à la prévision de la qualité de l'air à une échelle urbaine), où 6 maillages différents ont été réalisés sur l'agglomération de Bourges (voir Figure 14). Sur les différents jeux de données disponibles sur Bourges, nous avons volontairement choisi un jeu de données de faible résolution spatiale (maillage n°3 de *PREVISION'AIR*) afin d'observer les résultats obtenus en validation croisée avec les différents interpolateurs pour ce type de maillage. Les jeux de données additionnels seront utilisés en section 6 pour évaluer la qualité des interpolateurs en fonction du maillage.

Une simulation HD à 1m de résolution a été générée sur une zone de 4 km². Seulement 1 heure sur 5 a été simulée.

3. DEFINITION DES DOMAINES D'ETUDE

Les méthodes d'interpolation sélectionnées ont été appliquées sur les jeux de données complets (chapitre 2). Cependant, afin de mieux distinguer l'influence de la méthode d'interpolation sur la cartographie finale, des zones plus restreintes ont été définies là où les concentrations sont les plus élevées. Les nombres de points associés aux jeux de données et à leurs fenêtres de zoom sont présentés dans le **Tableau 1** et les coordonnées limites des fenêtres de zoom sont fournies dans le Tableau 2. Pour des raisons pratiques liées au temps de calcul, les évaluations n'ont été réalisées que sur ces fenêtres de zoom (chapitres 5, 5.9, 6).

Tableau 1 : Nombre de points et surface couverte par les maillages de modélisation « intelligents »

ville	Système de coordonnées	Type de données	Nombre de points	Surface couverte (km ²)	Nombre de points moyen / km ²
Bourges (maillage n°1)	LAMBERT 93	<i>toutes données*</i>	–	–	–
		<i>zoom</i>	806	4.19	192.4
Bourges (maillage n°2)	LAMBERT 93	<i>toutes données*</i>	–	–	–
		<i>zoom</i>	1044	4.29	243.4
Bourges (maillage n°3)	LAMBERT 93	<i>toutes données</i>	23043	1680	13.7
		<i>zoom</i>	1606	4.17	385.1
Bourges (maillage n°4)	LAMBERT 93	<i>toutes données*</i>	–	–	–
		<i>zoom</i>	1255	4.32	290.5
Bourges (maillage n°5)	LAMBERT 93	<i>toutes données*</i>	–	–	–
		<i>zoom</i>	831	4.26	195.1
Bourges (maillage n°6)	LAMBERT 93	<i>toutes données*</i>	–	–	–
		<i>zoom</i>	1035	4.25	243.5
Nantes	LAMBERT 93	<i>toutes données*</i>	38279	332	115.3
		<i>zoom</i>	702	4	175.5
Niort	LAMBERT 93	<i>toutes données</i>	92604	1373	67.4
		<i>zoom</i>	12447	25	497.9
Orléans	LAMBERT 2	<i>toutes données</i>	31271	979	31.9
		<i>zoom</i>	1438	2.95	487
Reims	LAMBERT 2	<i>toutes données</i>	35976	336	107.1
		<i>zoom</i>	13977	25	559
Tours	LAMBERT 2	<i>toutes données</i>	32020	2285	14
		<i>zoom</i>	951	2.39	398

* Les maillages 1, 2, 4, 5 et 6 de *PREVISION'AIR* sur Bourges ne sont utilisés que pour la validation sur grille de référence haute résolution : seules sont fournies les informations relatives à la fenêtre de zoom.

Tableau 2 : Fenêtres de zoom choisies pour les différents jeux de données

ville	Système de coordonnées	xmin	xmax	ymin	ymax
Bourges	LAMBERT 93	652000	657000	6663000	6667000
Orléans	LAMBERT 2	566578	568858	2322174	2323474
Nantes	LAMBERT 93	350000	352000	6685000	6687000
Niort	LAMBERT 93	430000	435000	6582000	6587000
Reims	LAMBERT 93	720000	725000	2470000	2475000
Tours	LAMBERT 2	475519	477679	2266603	2267723

4. PRESENTATION DES METHODES D'INTERPOLATION

4.1 Interpolation linéaire

Soient s_i les points de la grille à partir de laquelle est réalisée l'interpolation.

On considère le triangle formé par les sommets s_1 , s_2 et s_3 , dans lequel se situe le point d'intérêt s .

La valeur de la concentration est supposée dépendre de ses coordonnées spatiales:

$$Z(s)=Z(x,y)=\alpha x+\beta y+\gamma$$

La solution est obtenue à partir d'une combinaison linéaire des observations aux sommets du triangle en résolvant le système :

$$\begin{cases} \alpha x_1 + \beta y_1 + \gamma = z_1 \\ \alpha x_2 + \beta y_2 + \gamma = z_2 \\ \alpha x_3 + \beta y_3 + \gamma = z_3 \end{cases}$$

La résolution de ce système fournit les coefficients nécessaires au calcul de tout point inclus dans le triangle.

On montre que la solution peut s'écrire :

$$Z(s) = \frac{\sum_{i=1}^3 |S_{-i}| z_i}{|S|}$$

où $|S|$ est l'aire formée par le triangle composé des sommets s_1 , s_2 et s_3 et $|S_{-1}|$ (respectivement $|S_{-2}|$ et $|S_{-3}|$) est l'aire formée par le triangle composé des sommets s , s_2 et s_3 (respectivement s , s_1 , s_3 et s , s_1 , s_2).

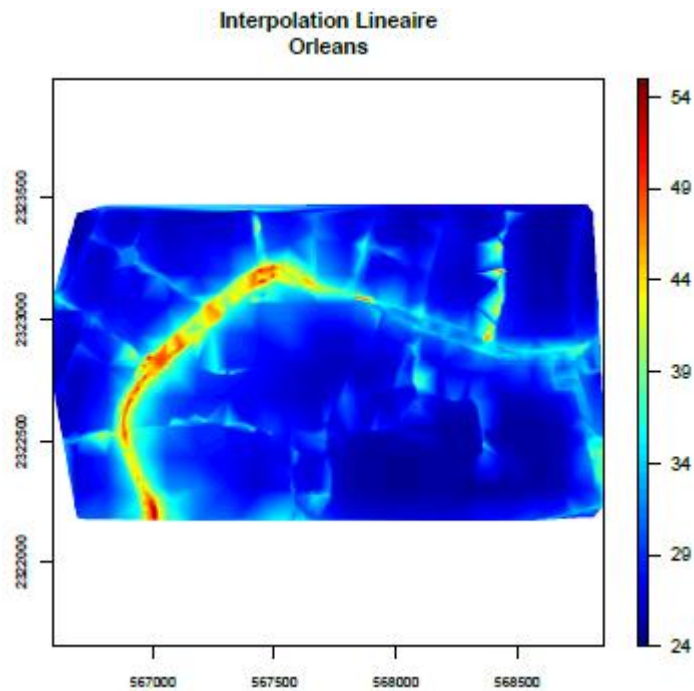


Figure 1 : Interpolation linéaire de la moyenne annuelle en NO₂ (µg.m-3) sur Orléans

Chaque observation reçoit un poids égal à la proportion de surface du triangle qui lui est opposé. Ainsi, plus le point à interpoler est proche d'un site d'observation, plus la valeur de la concentration est proche de la valeur observée en ce site.

Il n'est pas possible d'interpoler en dehors de l'enveloppe convexe formée par les points de données.

4.2 Interpolation de Delaunay

L'interpolateur de Delaunay repose sur la triangulation de Delaunay qui définit un maillage en construisant des triangles aussi compacts que possible : le cercle dans lequel chaque triangle est circonscrit ne contient que les points définissant ce seul triangle.

Une première variante de l'interpolateur consiste à générer une grille régulière où la valeur interpolée en chaque point est calculée à partir des 3 valeurs du triangle dans lequel il est situé, pondérées par l'inverse de la distance. Plus récemment, des chercheurs de l'IRSTV¹, en partenariat avec Air Pays de la Loire, ont développé une autre variante plus adaptée à la pollution en zone urbaine qui a été intégrée dans le logiciel OrbisGIS².

Une triangulation de Delaunay (Figure 2b) est construite à partir des données produites sur le maillage intelligent en sortie du modèle urbain (Figure 2a). Ces dernières constituent les sommets des triangles.

¹ Institut de Recherche en Sciences et Techniques de la Ville

² <http://www.orbisgis.org/>

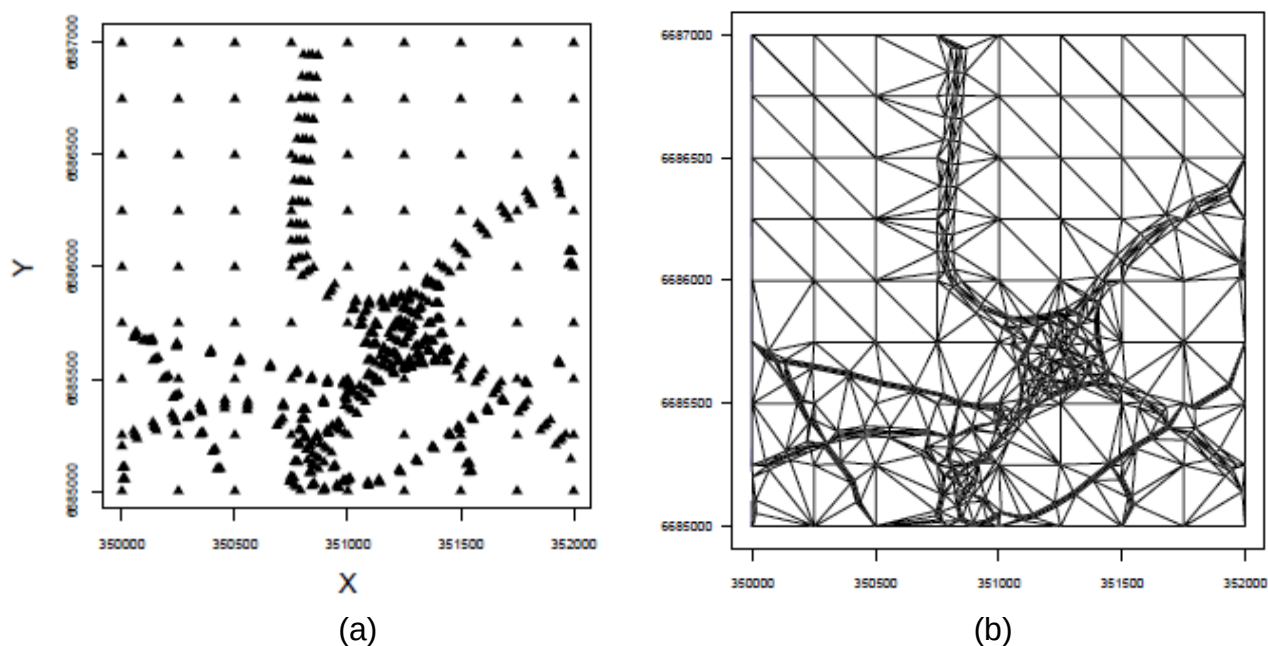


Figure 2 : Données en sortie du modèle urbain (zoom sur Nantes) et triangulation de Delaunay associée

Une isovaleur est une valeur numérique qui, pour une variable donnée (la moyenne annuelle du NO₂ ici), définit un ensemble de polygones joignant l'ensemble des positions repérées où la variable est égale à cette valeur.

Pour chaque triangle de la triangulation de Delaunay, il est possible d'évaluer le long de chacune des arêtes la position où la concentration atteint une isovaleur donnée en supposant la variation des concentrations comme linéaire entre les deux sommets qui composent l'arête. L'enveloppe convexe des positions associées à deux isovaleurs **iso1** et **iso2** forme un polygone (triangle ou quadrilatère) à l'intérieur duquel la valeur de concentration est comprise entre **iso1** et **iso2**.

La cartographie finale des concentrations est construite en décomposant la triangulation de Delaunay sous forme de polygones avec une échelle d'isovaleurs définie par une valeur minimale, maximale et un pas régulier (voir les figures en section 8).

4.3 Interpolation par krigeage

Le krigeage est un interpolateur issu des méthodes d'estimation géostatistique. Il existe différentes sortes de krigeage qui permettent de s'adapter aux données disponibles et dans le cas de plusieurs variables, aux relations entre celles-ci. Dans cette étude, nous n'utiliserons que le krigeage ordinaire (dit à moyenne inconnue).

La valeur estimée en un point est la moyenne pondérée des observations dans un voisinage restreint (voisinage glissant en distance ou en nombre de points) ou de toutes les observations (voisinage unique).

Les pondérateurs (λ_α) associés aux observations Z_α sont calculés en imposant des contraintes de non-biais et de minimisation de la variance de l'erreur d'estimation. Pour plus de détails, on pourra se reporter à la formation LCSQA en Cartographie³.

L'interpolation est donc entièrement contrôlée par le choix du modèle de variogramme. Une hypothèse de stationnarité est introduite afin de se libérer des contraintes de localisation et ainsi regrouper l'information sur les variables en fonction de la distance qui les séparent plutôt que de leur localisation. Classiquement, on calcule un variogramme empirique sur les données en moyennant, par classe de distance, les écarts quadratiques entre les observations. Puis on ajuste sur cette courbe un modèle de variogramme admissible (sphérique, exponentiel, gaussien, etc.) ou une combinaison linéaire de ces modèles.

En milieu urbain, la valeur des concentrations est fortement influencée par les sources d'émission routières, ce qui invalide l'hypothèse de stationnarité du processus aléatoire dans le krigeage. Le calcul classique du variogramme n'est donc pas totalement adapté à cette problématique. Des travaux sont actuellement en cours, en collaboration avec ARMINES (Centre de géosciences/géostatistique), pour adapter les méthodes géostatistiques à l'estimation des niveaux de pollution en milieu urbain. Ces développements n'étant pas terminés, le krigeage est ici appliqué dans ses hypothèses classiques, sans distinction entre les points de modélisation proches du trafic et les points de modélisation représentatifs de la pollution de fond. Parmi les différentes structures possibles, une structure exponentielle (avec effet de pépité) est choisie pour l'ajustement de la courbe empirique, et cela pour l'ensemble des jeux de données.

Les cartographies finales étant élaborées sur une grille régulière, un krigeage de bloc est utilisé, qui estime la valeur moyenne de la variable dans la maille, à la différence du krigeage ponctuel, qui estime la valeur de la variable au centre de la maille. Le krigeage de bloc fournit des estimations plus représentatives des concentrations attendues sur les mailles d'une grille régulière : plus la maille est lâche, plus son estimation par krigeage ponctuel s'écarte de celle obtenue par krigeage de bloc.

4.4 Choix techniques et méthodologiques

Les interpolations sont réalisées à l'aide du logiciel R (<https://cran.r-project.org/>). L'interpolation linéaire est disponible dans la librairie **akima**, via la fonction **interp**. L'option **linear=TRUE** doit être choisie. Aucune librairie de R ne permet d'effectuer intégralement une interpolation de Delaunay mais la triangulation de Delaunay existe dans le package **rgeos**. Le LCSQA a développé un script qui utilise la structure de données retournée par cette librairie pour parcourir les arêtes des triangles et construire une carte des concentrations à partir d'un ensemble d'isovaleurs fournies par l'utilisateur. Ce script pourra être fourni aux AASQA sur demande.

Enfin, il existe plusieurs librairies dans R pour réaliser des interpolations par krigeage. La librairie **gstat** a été utilisée ici. Toutefois, le krigeage nécessite une étape de modélisation du variogramme et un choix du voisinage qui ne peut être fait que par l'utilisateur. Comme il est expliqué dans la présentation des méthodes, seul un variogramme de

³ Formation LCSQA en Cartographie et Méthodes Géostatistiques. <http://www.lcsqa.org/formations>.

structure exponentielle avec effet de pépite est utilisé dans cette étude. En ce qui concerne le voisinage de krigeage, il peut être utile, lorsque l'on dispose d'un jeu de données important, de limiter les observations utilisées dans l'estimation à un ensemble réduit, défini sur un critère de distance par rapport au point ou sur un critère de nombre maximal d'observations à utiliser. Les observations étant de typologie différente, le critère de distance ne nous est pas apparu optimal car même pour une distance faible, des points dont le niveau de concentration s'explique par des sources de pollution différentes de celles qui influencent le point à estimer peuvent se retrouver dans le voisinage. Nous avons préféré restreindre ce dernier aux N voisins les plus proches du point à estimer. La décision sur le nombre de voisins à prendre en compte peut être prise d'après plusieurs critères (résultats de validation croisée, pente de la régression mesure-estimation, etc.).

Notre décision s'appuie sur l'observation de la décroissance des pondérateurs de krigeage. Sur la totalité des jeux de données utilisés, les poids de krigeage au-delà du 10^e voisin le plus proche sont généralement très faibles (Figure 3). Nous avons donc réalisé des krigeages avec des voisinages respectifs de 2, 3, 4, 5 et 10 voisins sur des grilles à 2 mètres de résolution.

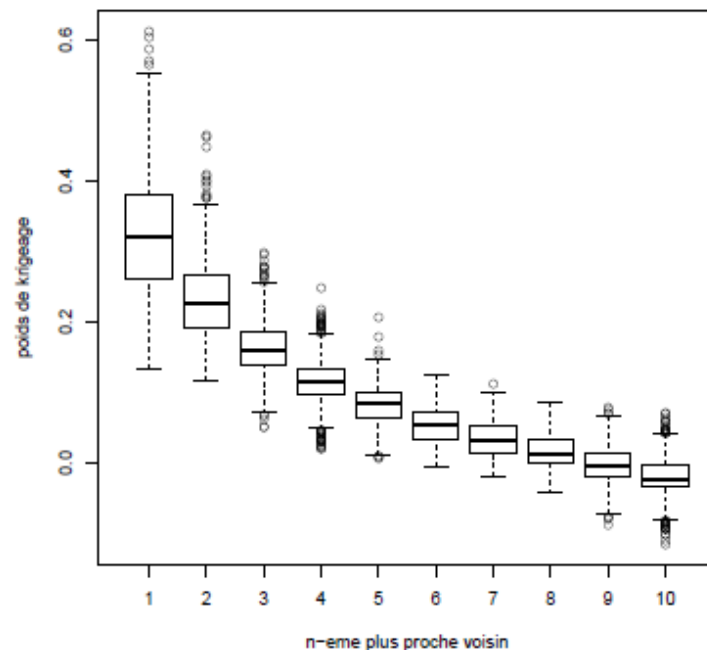


Figure 3 : Décroissance des pondérateurs de krigeage obtenue sur les données de Nantes avec un variogramme exponentiel

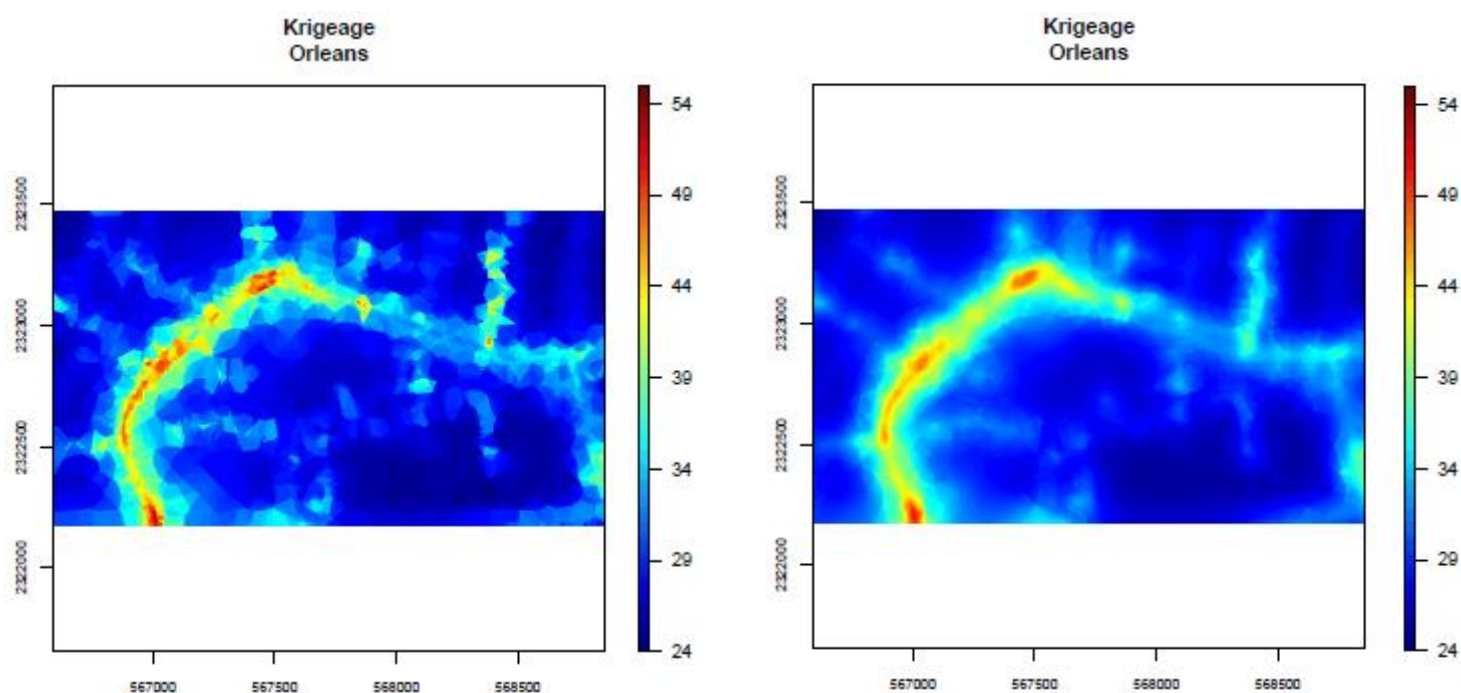


Figure 4 : Influence du nombre de voisins sur la cartographie des concentrations de la moyenne annuelle NO₂ (µg.m-3). A gauche, krigeage de blocs à 2 voisins. A droite, krigeage de blocs à 10 voisins.

La **Figure 4** montre l'influence du nombre de voisins sur la cartographie des concentrations pour la ville d'Orléans. Plus le nombre de voisins augmente, plus la carte est lisse mais certaines valeurs très fortes localement sont alors sous-estimées. La carte qui utilise un faible nombre de voisins apparaît en revanche plus discontinue.

Des approches pour rendre continue une cartographie utilisant un voisinage glissant ont déjà été développées (Rivoirard and Romary, 2011 ; Gribov and Krivoruchko, 2004) mais elles se sont surtout intéressées à des voisinages glissants en distance. Pour des voisinages en nombre, des approches (non testées ici) de type Babak and Deusch (2008) qui font la moyenne de plusieurs krigeages pourraient être testées dans des travaux ultérieurs.

5. VALIDATION CROISEE

La première méthode d'évaluation de la qualité de l'interpolation dans cette étude est la validation croisée. Dans chaque fenêtre de zoom, chacune des données de simulation est retirée et réestimée par interpolation à partir des données voisines. Ce calcul est réalisé de la façon différente selon l'interpolateur :

- **Interpolation de Delaunay** : la valeur estimée du point en validation croisée est la valeur moyenne du polygone dans lequel le point se retrouve à l'issue de l'interpolation. Par exemple, s'il se retrouve dans un polygone dont les concentrations sont comprises entre 40 et 41, sa valeur sera prise à 40.5. Il est important de préciser que contrairement aux autres méthodes, les scores obtenus sur l'interpolation de Delaunay dépendent fortement du pas choisi entre les isovaleurs utilisées pour construire les polygones : plus il est fin et plus les résultats de validation seront représentatifs des qualités et défauts de la méthode. Nous avons fixé ce pas à 1, ce qui semble un bon compromis entre temps de

calcul et qualité de la cartographie. Il est évident qu'un pas plus faible améliorerait les scores du modèle mais demanderait des temps de calcul plus longs.

- **Interpolation par krigeage** : des grilles respectives de 1m, 5m, 10m, 20m et 50m de résolution ont été produites et des krigeages de bloc sont réalisés. La valeur du point estimée en validation croisée est la valeur moyenne du bloc dans lequel le point se situe. A titre indicatif, les résultats du krigeage ponctuel, qui est a priori plus favorable à la validation croisée que le krigeage de bloc, sont également donnés. La valeur du point estimée en validation croisée est la valeur fournie par le krigeage en ce point précis. Tous les krigeages ont été réalisés avec un nombre maximal de voisins égal à 2, 3, 4, 5 ou 10, comme expliqué en 4.4.
- **Interpolation linéaire** : le point est retiré et réestimé linéairement.

Les nuages de points entre les valeurs modélisées originales et les valeurs interpolées sont présentés en annexe 2. Des cartes d'écarts, calculés comme la différence entre données de modélisation et valeurs interpolées (en validation croisée) sont également présentées en annexe 3.

5.1 Scores

L'évaluation de la qualité d'un interpolateur se fait usuellement par le calcul de scores. Dans cette étude de modélisation, les scores sont calculés en considérant les données de modélisation originales comme valeurs de référence. Six scores sont calculés :

- l'écart-type des valeurs interpolées σ_M , qui peut ainsi être comparé à l'écart-type des données de modélisation originales σ_D ;
- le coefficient de corrélation entre données de modélisation originales et valeurs interpolées (R) ;
- la racine carrée de l'erreur moyenne quadratique ;
- la racine carrée de l'erreur moyenne quadratique centrée (E') ;
- le biais moyen normalisé ;
- l'erreur moyenne normalisée.

Les formules pour le calcul de ces scores sont fournies dans le glossaire. Les tableaux de scores complets sont présentés en annexe 1 tandis qu'une représentation synthétique de trois de ces scores (σ_M , R et E') est fournie à l'aide du diagramme de Taylor dans les sections 5.3 à 5.8.

5.2 Principe du diagramme de Taylor

Il existe une relation géométrique qui relie trois statistiques couramment utilisées pour évaluer la qualité d'un modèle, à savoir le coefficient de corrélation (R), l'erreur quadratique centrée (E') et les écarts-types des données de référence (ici données de simulation) (σ_D) et des données de modèle (ici valeurs interpolées) (σ_M) :

$$E'^2 = \sigma_M^2 + \sigma_D^2 - 2R\sigma_M\sigma_D$$

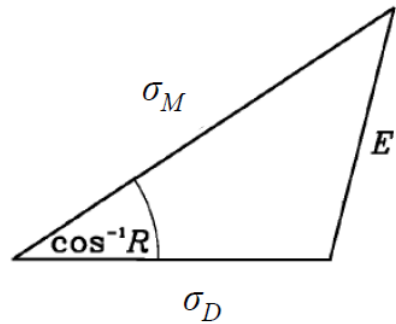


Figure 5 : Relation géométrique reliant R , E' , σ_D et σ_M .

Taylor a intégré cette relation dans un graphique, appelé diagramme de Taylor (2001, *Summarizing multiple aspects of model performance in a single diagram*), en utilisant les coordonnées polaires. Ce diagramme permet d'observer le degré de correspondance modèle-mesure.

Le point associé au modèle considéré est placé sur le diagramme de telle sorte que sa distance par rapport à l'origine soit égale à σ_M . Sa position azimutale renseigne sur la valeur du coefficient de corrélation R entre le modèle et les observations.

Une estimation parfaite se situerait donc sur l'axe des abscisses du diagramme de Taylor, à la valeur $\sigma_M = \sigma_D$ et $R=1$.

L'intégration de cette relation est montrée explicitement sur le diagramme de Taylor appliqué au jeu de données de Nantes, dans le cas du modèle linéaire. Pour faciliter la lecture des scores associés à chaque modèle, les arcs de cercles en traits pointillés représentent les lignes d'écart-type σ_M et les arcs de cercle en traits pleins représentent les lignes de RMSE centrées (E').

5.3 Diagramme de Taylor sur Nantes

L'analyse du diagramme de Taylor sur Nantes révèle que les meilleurs interpolateurs sont les plus simples, à savoir l'interpolateur linéaire et l'interpolateur de Delaunay. Les coefficients de corrélation correspondants sont d'environ 0.95. L'interpolateur de Delaunay présente une erreur quadratique centrée (13.20 contre 13.27) et un écart-type (42.34 contre 42.31) quasi-identiques à ceux observés sur l'interpolateur linéaire. Dans les deux cas, l'écart-type des valeurs interpolées surestime l'écart-type des données de modélisation originales.

En ce qui concerne les krigeages réalisés, on observe que le voisinage choisi n'a finalement que peu d'impact en matière de corrélation et qu'il peut surtout faire fluctuer légèrement la valeur de l'écart-type. La résolution de la grille de krigeage, i.e. la dimension des blocs, a quant à elle plus d'impact sur l'erreur quadratique centrée qui augmente avec celle-ci. De plus, les résultats sur la grille à 50m de résolution sont assez dégradés par rapport aux autres résolutions (1m, 5m, 10m, 20m) qui ont des scores assez proches.

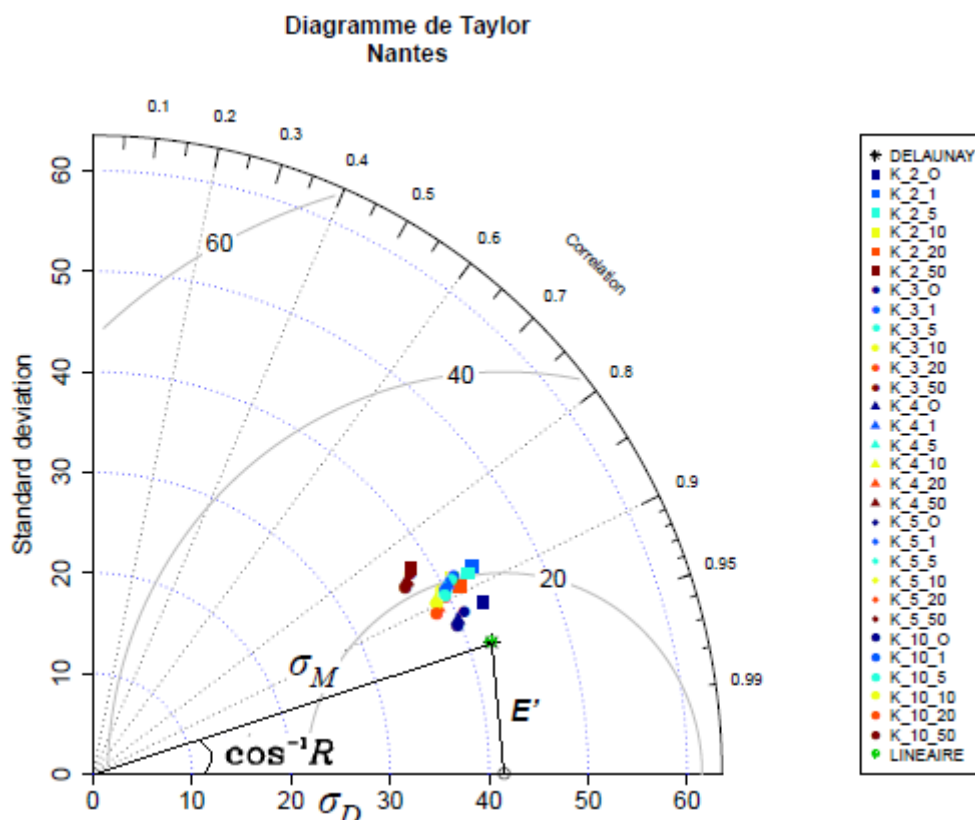


Figure 6 : Diagramme de Taylor obtenu sur la fenêtre de zoom de Nantes (voir le glossaire pour les abréviations des différents krigeages)

La Figure 7 présente les nuages de points entre les données de modélisation originales et les valeurs interpolées en validation croisée pour le krigage ponctuel, l'interpolation linéaire et l'interpolation de Delaunay. Les nuages de points ne sont pas fournis pour les sections 5.4 à 5.8 mais sont disponibles en annexe 2 (Figure 35). Il existe un ensemble de données de modélisation dont la concentration, comprise entre 30 et 50 $\mu\text{g.m}^{-3}$, est fortement surestimée par l'interpolation, quelle que soit la méthode utilisée. La Figure 41 de l'annexe 3 montre que ces surestimations sont situées à faible distance d'axes routiers près desquels, à des distances encore plus réduites, se trouvent d'autres données de modélisation avec des niveaux de concentrations très élevées. Ces dernières influencent fortement la concentration interpolée en validation croisée par des interpolateurs linéaires et conduisent à de fortes surestimations. Elles sont quant à elles logiquement sous-estimées lors de la validation croisée.

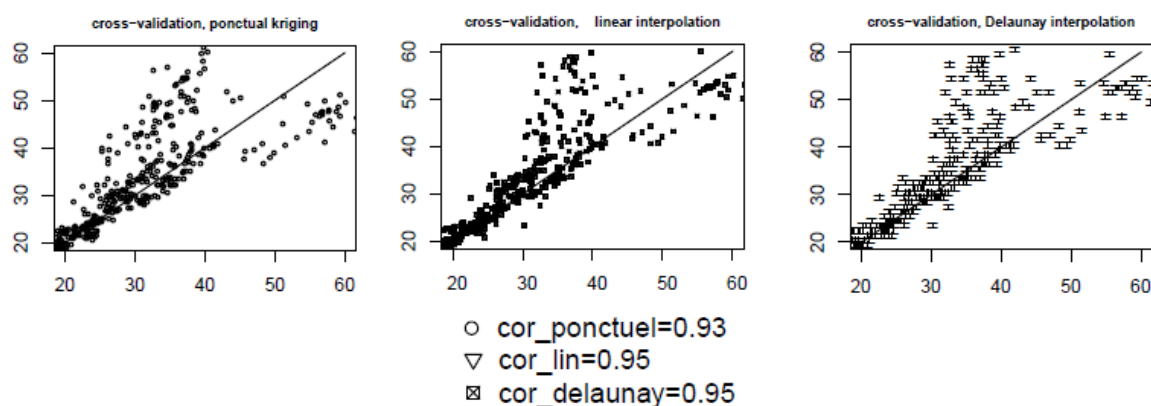


Figure 7 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom de Nantes (krigeage ponctuel, Delaunay, linéaire)

Abscisses : données de modélisation du NO₂ en µg.m-3

Ordonnées : valeurs interpolées en validation croisée

5.4 Diagramme de Taylor sur Niort

Sur Niort, le meilleur interpolateur est l'interpolateur linéaire ($R=0.92$; $E'=1.87$; $\sigma_M=4.40$). L'interpolateur de Delaunay présente une erreur quadratique centrée (1.97) de moins bonne qualité que celle observée sur l'interpolateur linéaire mais une corrélation quasi-identique (0.91 contre 0.92). Cependant, l'écart-type des valeurs interpolées (4.46) sous-estime moins fortement l'écart-type des données de modélisation originales.

Les krigeages ponctuels à 3 ou 4 voisins sont assez proches par leurs scores de l'interpolateur de Delaunay, à l'exception de l'écart-type des données originales qu'ils sous-estiment plus fortement. Les krigeages de blocs voient leur qualité se dégrader avec l'augmentation de la surface des blocs. Un lissage (diminution de σ_M) lié à l'augmentation du nombre de voisins est également visible. Le krigeage à 5m de résolution semble générer des anomalies qui le distinguent du schéma régulier de dégradation observé avec les autres résolutions : sa corrélation est beaucoup plus faible (0.4) et sa RMSE beaucoup plus forte (~5).

Les nuages de points obtenus avec les différents krigeages (Figure 35, annexe 2) sont sensiblement plus dispersés que ceux obtenus avec l'interpolateur linéaire et l'interpolateur de Delaunay.

La carte des écarts (Figure 42, annexe 3) montre qu'en moyenne, l'interpolateur linéaire engendre des sous-estimations moins fortes que l'interpolateur de Delaunay.

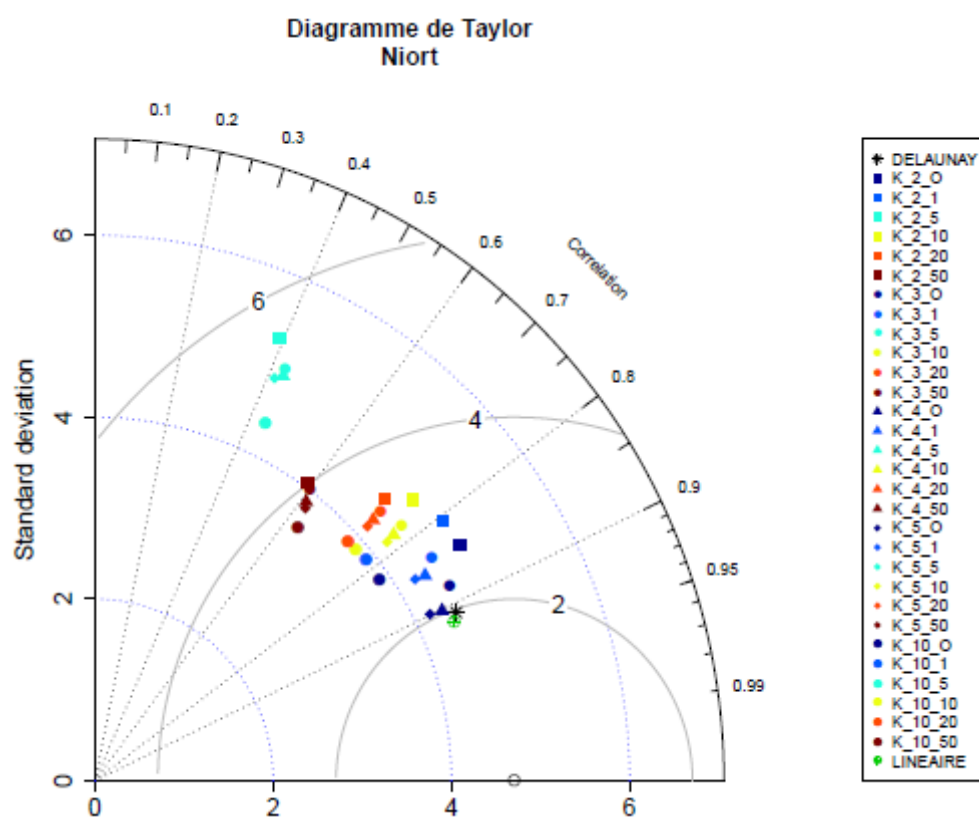


Figure 8 : Diagramme de Taylor obtenu sur la fenêtre de zoom de Niort

5.5 Diagramme de Taylor sur Reims

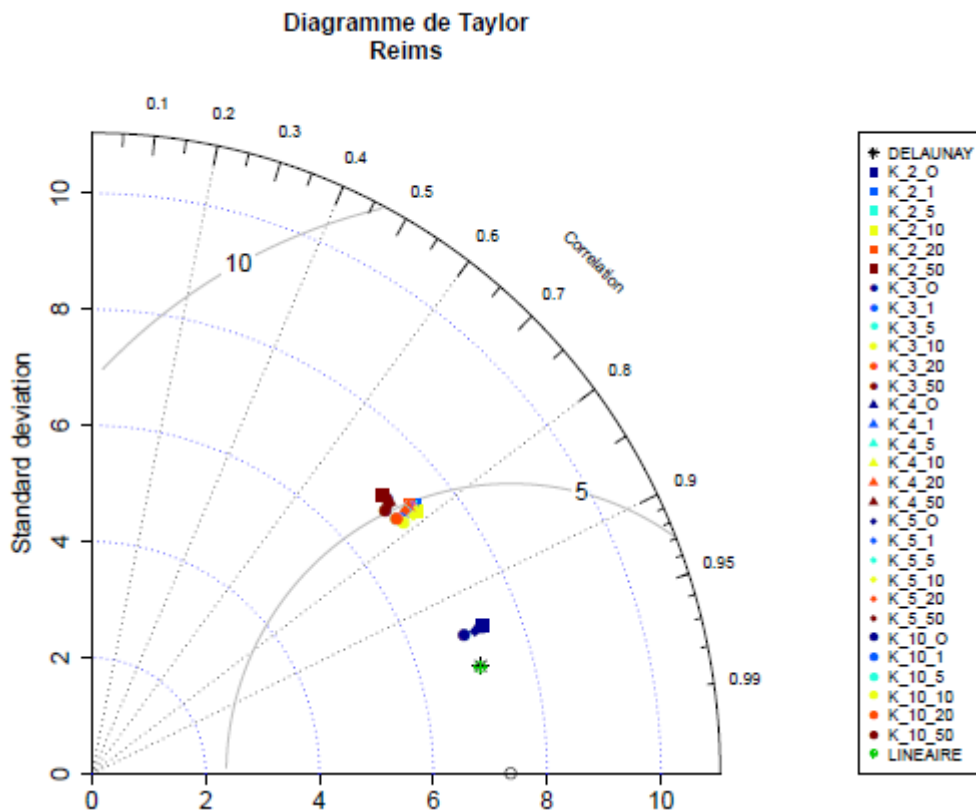


Figure 9 : Diagramme de Taylor obtenu sur la fenêtre de zoom de Reims

Les meilleurs interpolateurs sont une nouvelle fois l'interpolateur linéaire et l'interpolateur de Delaunay qui présentent une meilleure RMSE centrée et un meilleur coefficient de corrélation que le krigeage ponctuel.

Sur Reims, les krigeages de bloc ont tous des scores assez proches quelle que soit la résolution et le nombre de voisins ($R \sim 0.75$; $E' \sim 5$; $\sigma_M \sim 7$). Ces résultats sont significativement moins bons que ceux du krigeage ponctuel qui montre un écart-type similaire mais une RMSE centrée deux fois plus faible (2.6) et une corrélation bien supérieure (0.94).

L'analyse des nuages de points (Figure 37, annexe 2) montre que l'interpolateur linéaire est légèrement plus resserré autour de la première bissectrice que ne l'est l'interpolateur de Delaunay.

5.6 Diagramme de Taylor sur Bourges

Pour rappel, le maillage retenu pour Bourges est un maillage régulier de 50 m de résolution différent des maillages intelligents utilisés sur les autres agglomérations.

Le diagramme de Taylor révèle des résultats sensiblement différents de ceux obtenus sur les autres jeux de données. Les meilleurs interpolateurs sont regroupés au sein d'un ensemble compact qui rassemble le krigeage ponctuel et le krigeage de bloc avec un nombre de voisins supérieur à 4. Ces interpolateurs sous-estiment l'écart-type des observations (~ 2) et présentent une corrélation supérieure ou égale à 0.8. L'interpolateur linéaire et l'interpolateur de Delaunay sont toujours quasi-équivalents par leurs scores mais leur corrélation est légèrement inférieure à 0.8. Ces résultats sont à mettre en relation avec les faibles concentrations sur la zone (peu de points ont une concentration supérieure à $40 \mu\text{g.m}^{-3}$) et la faible densité du maillage. De plus, les nuages de points montrent que les krigeages, malgré leurs meilleurs scores, sous-estiment plus fortement les données de modélisation originales que l'interpolateur linéaire et l'interpolateur de Delaunay, ce qui est confirmé par les calculs d'erreur et de biais moyen normalisé (Tableau 13 de l'annexe 1).

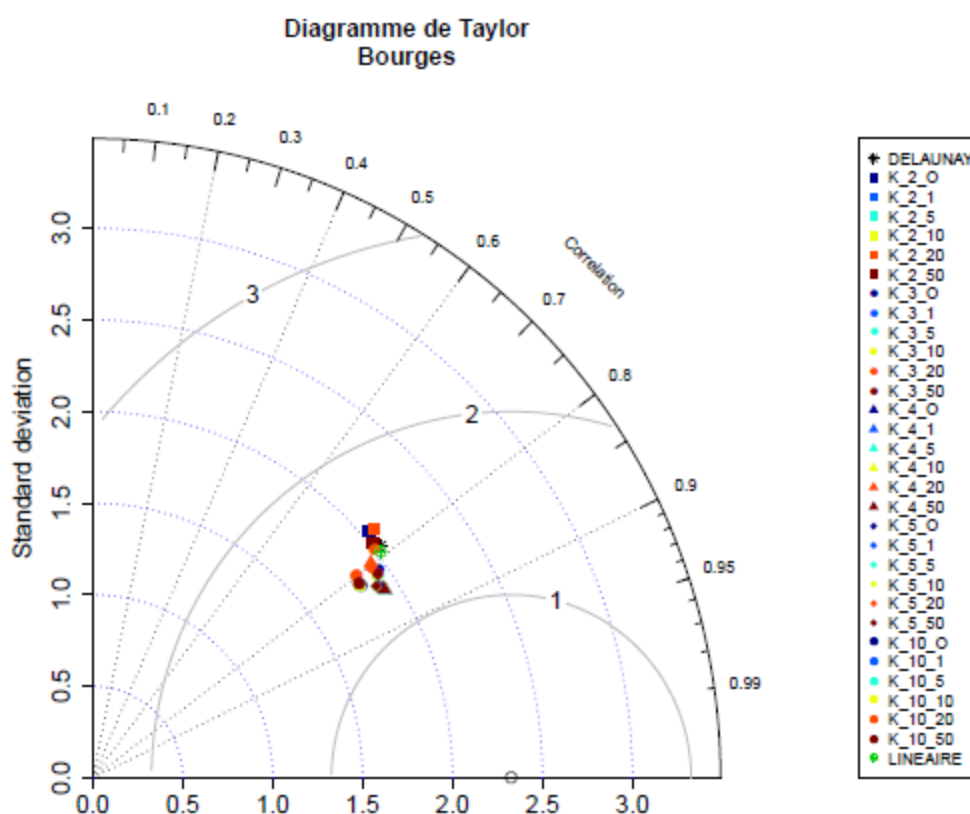


Figure 10 : Diagramme de Taylor obtenu sur la fenêtre de zoom de Bourges

5.7 Diagramme de Taylor sur Orléans

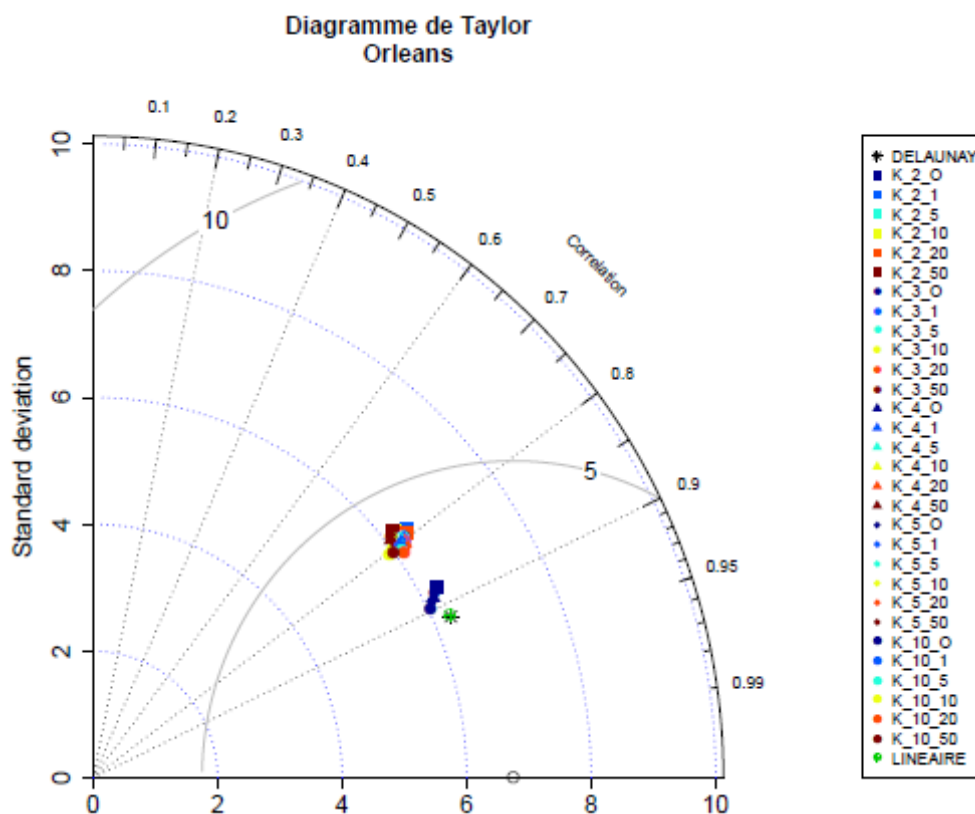


Figure 11 : Diagramme de Taylor obtenu sur la fenêtre de zoom d'Orléans

L'interpolation linéaire et l'interpolateur de Delaunay ont des scores quasi-identiques ($R \sim 0.91$; $E' \sim 2.74$; $\sigma_M \sim 6.29$). Le krigeage ponctuel à 10 voisins présente des scores assez proches mais de qualité légèrement moindre ($R \sim 0.90$; $E' \sim 2.95$; $\sigma_M \sim 6.03$).

Les scores obtenus par krigeage de bloc sur Orléans sont représentés de manière très condensée sur le diagramme. Ni le voisinage ni la résolution ne semblent avoir un impact significatif sur la qualité des résultats même si la résolution à 50m présente des scores légèrement plus faibles que la résolution à 1m. Les corrélations observées pour ces krigeages ($R \sim 0.8$) sont sensiblement plus faibles que les corrélations observées pour les krigeages ponctuels (0.8 contre 0.9), l'interpolateur linéaire et l'interpolateur de Delaunay.

Tant sur les nuages de points (Figure 39, annexe 2) que sur les cartes d'écarts (Figure 45, annexe 2), l'interpolateur linéaire et l'interpolateur de Delaunay présentent des résultats similaires.

5.8 Diagramme de Taylor sur Tours

Les résultats concordent avec les observations précédentes: l'interpolateur linéaire présente les meilleurs résultats de validation croisée ($R = 0.94$; $E' = 3.41$; $\sigma_M = 9.07$), tout comme l'interpolateur de Delaunay : ils sont confondus sur le diagramme. Le krigeage a un coefficient de corrélation et une RMSE centrée dégradés par rapport aux méthodes

linéaires (R inférieur ou égal à 0.9 et RMSE centrée supérieure à 4.2) et un écart-type stable autour de 9. Cette dégradation est moins prononcée pour le krigeage ponctuel. Comme dans le cas de Niort, elle est la plus sensible pour le krigeage de bloc sur des mailles de 5 m de résolution. L'effet de lissage (diminution de σ_M) lié à l'augmentation du nombre de voisins se distingue également.

Tant sur les nuages de points (Figure 40, annexe 2) que sur les cartes d'écarts (Figure 46, annexe 3), l'interpolateur linéaire et l'interpolateur de Delaunay présentent des résultats similaires.

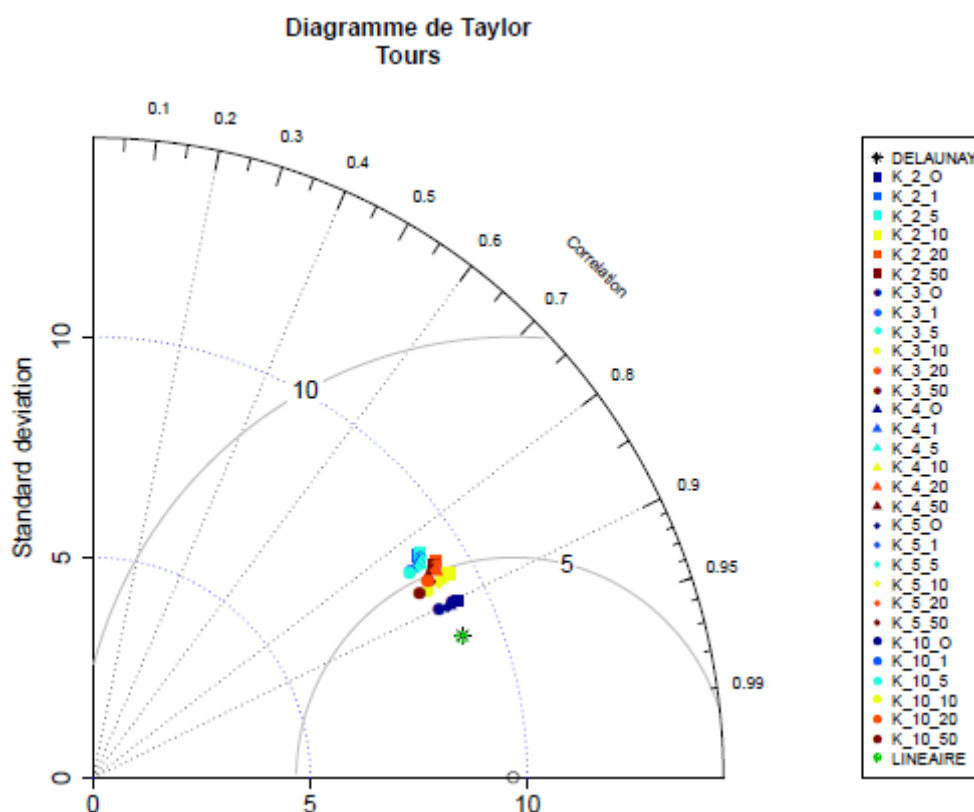


Figure 12 : Diagramme de Taylor obtenu sur la fenêtre de zoom de Tours

5.9 Tests complémentaires sur l'interpolateur de Delaunay

Comme précisé en introduction de la section 5, la qualité de l'interpolateur de Delaunay est liée au pas choisi entre les isovaleurs. Le pas avait été fixé à 1 pour un bon compromis entre qualité des scores et temps de calcul.

Nous testons ici sur les données de Niort l'influence sur les scores d'un pas plus faible. Les pas testés sont de 0.2, 0.4, 0.6, 0.8 et 1.

Le diagramme de Taylor en Figure 13 montre que les scores obtenus avec les différents pas ne sont pas significativement différents. Le pas le plus faible (0.2) est celui qui aboutit aux meilleurs résultats d'interpolation mais ces résultats sont quasi-identiques à ceux obtenus avec le pas majorant (1).

En conséquence, il semble raisonnable de conserver un pas de 1 dans l'échelle des concentrations, ce qui permet de construire une cartographie fine des niveaux de pollution, sans alourdir les temps de calculs.

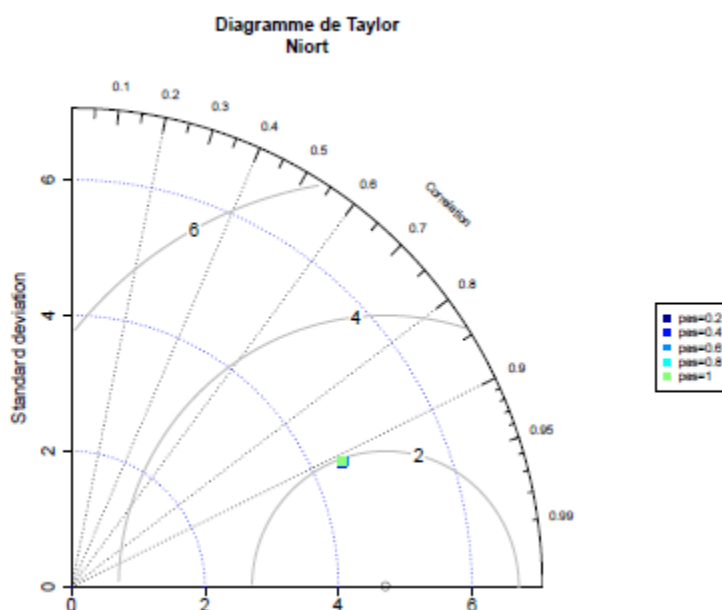


Figure 13 : Interpolation de Delaunay réalisée sur les données de Niort pas de 0.2, 0.4, 0.6, 0.8 et 1 entre les isovaleurs

6. VALIDATION PAR COMPARAISON A UNE SIMULATION DE REFERENCE

La seconde méthode de validation utilisée dans cette étude est la comparaison des interpolateurs à l'aide d'une grille de référence de haute résolution (simulation HD). Le principe de cette méthode est de ne retenir qu'un certain nombre de points de la grille HD, d'appliquer ensuite une interpolation à partir des points retenus sur l'ensemble du maillage pour enfin comparer les résultats d'interpolation en tout point de la grille HD avec la simulation de référence.

6.1 Données utilisées

Pour cette validation, les AASQA ont réalisé des simulations complémentaires sur des grilles de haute résolution (cf. chapitre 2). Sont concernées les villes de Nantes, Niort, Reims et Bourges. La configuration des calculs a été adaptée par les AASQA afin de limiter les temps de calcul. Les résultats de simulation en un même point sur un maillage intelligent (utilisé dans le chapitre 5) et sur une grille de référence ne sont donc pas identiques, ce qui risque de biaiser l'évaluation. En conséquence, pour cette partie, les valeurs de la simulation HD ont été extraites sur chaque point du maillage intelligent. Ce sont ces valeurs qui sont ensuite utilisées pour appliquer les différents interpolateurs (interpolation linéaire, krigeage et interpolation de Delaunay) et comparer les concentrations estimées à la simulation de référence. Au vu des résultats du chapitre précédent, la validation de l'interpolation par krigeage se limite au krigeage ponctuel à 5 voisins.

La présentation des résultats est plus détaillée pour le cas de Bourges qui compte six maillages intelligents différents. Cette multiplicité permet d'évaluer l'effet de la définition du maillage en complément de la méthode d'interpolation.

6.2 Résultats

La **Figure 15** montre le diagramme de Taylor obtenu à partir des 6 maillages d'entrée de *PREVISION'AIR* (ADMS) (voir 2.4) dont le nombre de points varie de 18000 à 64000 (Figure 14). Chacun de ces maillages est utilisé pour interpoler selon les différentes méthodes la valeur de concentration du NO₂ aux points de la grille de référence.

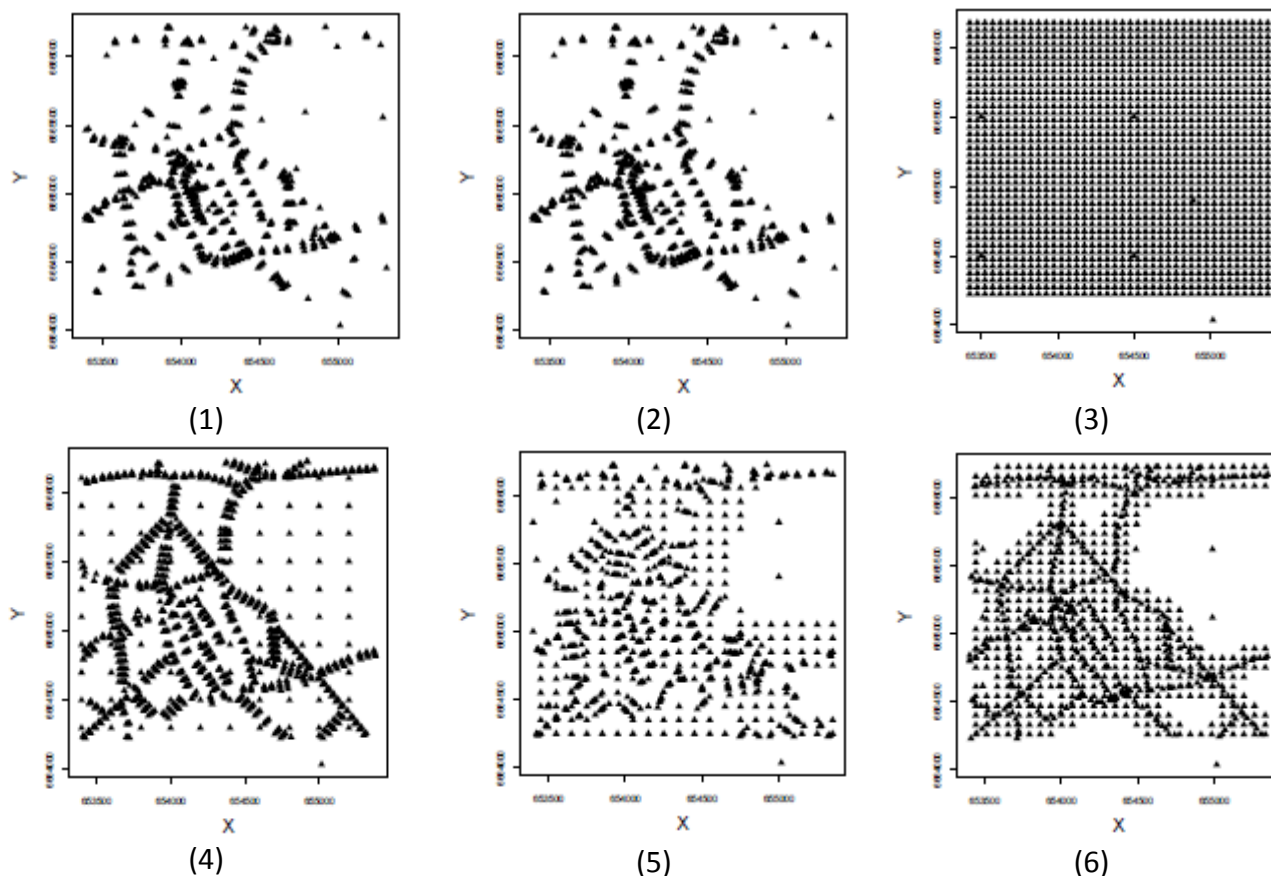


Figure 14 : Maillages issus de *PREVISION'AIR*, utilisés pour l'interpolation sur la grille de référence de Bourges

Pour Bourges, à maillage fixé, l'interpolation linéaire et l'interpolation de Delaunay ont des résultats quasi-identiques. Il faut cependant noter que l'interpolation linéaire a une RMSE centrée plus faible et une meilleure corrélation quel que soit le maillage utilisé en entrée. En revanche, l'écart-type modélisé par l'interpolateur de Delaunay est toujours plus proche de celui observé sur les données, à l'exception du maillage n°6 où il le surestime plus fortement que l'interpolateur linéaire. Encore une fois, les écarts entre méthodes d'interpolation restent très faibles tant en valeur absolue que relative. Le krigeage fournit des résultats assez instables dont la qualité dépend fortement du maillage initial choisi en entrée : ainsi, comme cela avait déjà été observé en validation croisée, les résultats obtenus avec le krigeage sont meilleurs (d'après la RMSE et la corrélation) que ceux obtenus avec l'interpolateur linéaire et l'interpolateur de Delaunay en utilisant le maillage n°3 en entrée. Ce n'est pas le cas avec les autres mailles d'entrée.

Pour Bourges, à méthode d'interpolation fixée, c'est le maillage n°2 qui présente les meilleurs résultats, tant pour l'interpolateur linéaire que pour l'interpolateur de Delaunay. Les nuages de points des données de modélisation originales-valeurs interpolées et les cartes d'écarts montrent que la sous-estimation des concentrations à proximité des axes routiers est la moins forte avec le maillage n°2 (Figure 48 et Figure 57), même si les valeurs élevées (supérieures à $50 \mu\text{g.m}^{-3}$ sur l'agglomération de Bourges) restent assez fortement sous-estimées.

Ainsi l'interpolateur linéaire avec le maillage n°2 fournit-il dans l'ensemble les meilleurs résultats. Toutefois, c'est l'interpolateur de Delaunay avec le même maillage d'entrée qui fournit les deuxièmes meilleurs scores et non pas l'interpolateur linéaire avec un autre maillage d'entrée. Cela tend à démontrer que la qualité du maillage défini dans le modèle urbain, plus encore que le type d'interpolateur utilisé, influe sur les résultats finaux obtenus après interpolation sur une grille régulière. Il est ainsi important de construire un maillage qui permette de restituer au mieux la décroissance des concentrations perpendiculairement aux axes routiers. Le maillage n°2 sur Bourges, qui compte de nombreux points distants de quelques dizaines de mètres des axes, en est un exemple.

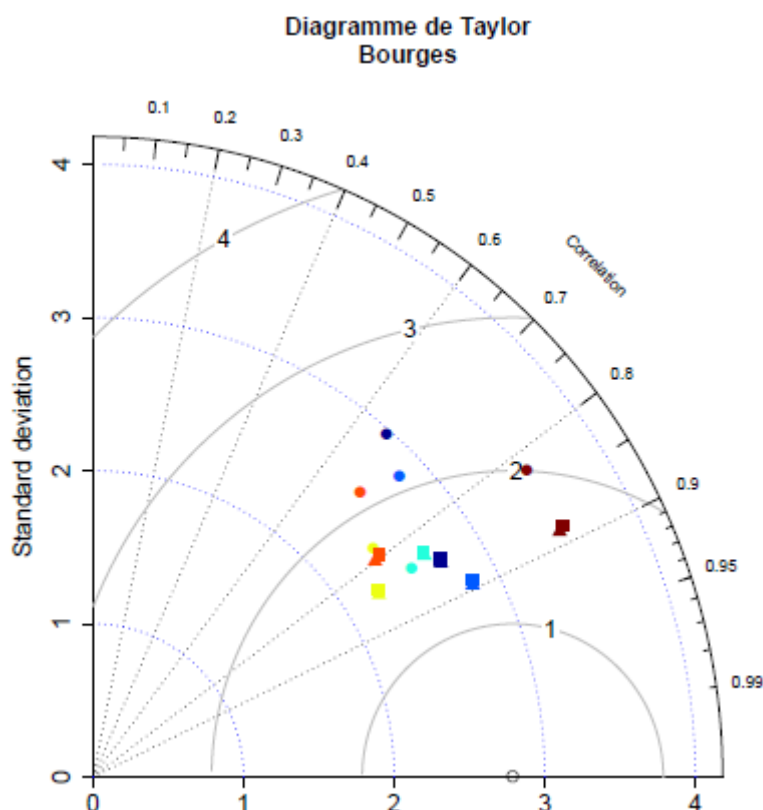


Figure 15 : Diagramme de Taylor obtenu sur les données de références du maillage HD de Bourges.

Données d'entrée : les 6 maillages de *PREVISION'AIR*

Méthodes utilisées : interpolation linéaire, de Delaunay et krigeage à 5 voisins

Des exercices d'inter-comparaisons similaires réalisés sur d'autres jeux de données seraient utiles afin de confirmer ces observations, mais aussi afin de définir une stratégie de définition du maillage en sortie du modèle urbain, dans l'optique d'optimiser les temps de calculs.

Sur Nantes et Reims, les conclusions sont les mêmes que sur Bourges : comme l'indiquent les valeurs de corrélation et de RMSE centrée, l'interpolateur linéaire est plus efficace que l'interpolateur de Delaunay, d'une manière assez peu significative cependant. Sur Niort, la RMSE centrée est meilleure avec l'interpolation de Delaunay et les corrélations sont identiques. Les écarts-types obtenus avec l'interpolation linéaire sur Niort et Nantes, s'ils sous-estiment eux aussi les écarts-types réellement observés sur les données (7.52 sur Nantes, 4.05 sur Niort), le font de manière bien moins prononcée que l'interpolation de Delaunay (6.49 contre 5.74 sur Nantes, 3.96 contre 3.37 sur Niort,). Sur Reims, où les données présentent un écart-type de 5.88, l'interpolation de Delaunay et l'interpolation linéaire le surestiment de manière quasi-identique (5.97 et 5.96).

Le Tableau 3 récapitule les scores obtenus dans les différents cas d'étude.

Tableau 3 : Scores obtenus sur les grilles de référence avec l'interpolateur linéaire et l'interpolateur de Delaunay

ville	méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Bourges maillage 1	<i>Delaunay</i>	2.71	0.854	1.84	1.51	7.41	9.4
	<i>linéaire</i>	2.70	0.86	1.82	1.48	7.41	9.32
Bourges maillage 2	<i>Delaunay</i>	2.83	0.89	1.63	1.31	6.8	7.84
	<i>linéaire</i>	2.81	0.90	1.61	1.28	6.81	7.68
Bourges maillage 3	<i>Delaunay</i>	2.64	0.83	1.58	1.58	0.41	4.66
	<i>linéaire</i>	2.63	0.84	1.56	1.56	0.41	3.64
Bourges maillage 4	<i>Delaunay</i>	2.25	0.84	1.59	1.50	3.61	6.37
	<i>linéaire</i>	2.24	0.85	1.56	1.47	3.58	6.01
Bourges maillage 5	<i>Delaunay</i>	2.39	0.80	1.72	1.69	2.04	6.27
	<i>linéaire</i>	2.34	0.80	1.70	1.67	2.21	5.88
Bourges maillage 6	<i>Delaunay</i>	3.53	0.88	1.96	1.68	7.12	7.8
	<i>linéaire</i>	3.49	0.89	1.92	1.64	7.07	7.34
Niort	<i>Delaunay</i>	3.37	0.93	1.37	1.33	2.33	4.66
	<i>linéaire</i>	3.95	0.93	1.57	1.62	2.06	4.27
Nantes	<i>Delaunay</i>	6.50	0.92	3.33	2.99	5.5	8.65
	<i>linéaire</i>	6.49	0.92	3.31	2.97	5.5	8.61
Reims	<i>Delaunay</i>	5.97	0.95	2.23	2.21	0.95	4.03
	<i>linéaire</i>	5.96	0.96	2.21	2.19	0.97	3.96

Les nuages de points entre les valeurs modélisées originales et les valeurs interpolées sur la grille de référence sont présentés en annexe 4. Des cartes d'écarts, calculés comme la différence entre les valeurs interpolées et les données de modélisation (sur la grille de référence) sont également présentées en annexe 5.

6.3 Lignes directrices pour un maillage efficace

Dans la section précédente, nous avons montré que le maillage n°2 sur Bourges était le plus performant et qu'il surestimait le moins fortement les concentrations moyennes annuelles de NO₂ à proximité des axes routiers.

Il n'est toutefois pas inutile de chercher à améliorer ce maillage, qui présente des défauts facilement explicables : des surestimations ou sous-estimations importantes.

Ces écarts peuvent s'expliquer différemment selon l'endroit où l'on se place dans la carte (Figure 19a), en proximité routière ou en fond :

Des transects sont définis en proximité routière le long des axes pour tenir compte de la décroissance des concentrations. La Figure 16 présente un exemple de transect. Ces transects sont constitués de 4 points. Une paire de points intérieurs délimite la largeur de la rue (bordure de bâtiment). Deux autres points sont situés à l'extérieur de la rue à égale distance de chaque côté de celle-ci. Deux effets sont constatés sur l'utilisation de ce type de transect : une sous-estimation des concentrations au milieu de la rue et une surestimation à l'extérieur de la rue. Les concentrations les plus fortes situées dans la rue ne sont donc pas prises en compte par la paire de points située en bordure de bâtiment. Perpendiculairement aux axes routiers, la décroissance naturelle du NO₂ n'est pas linéaire mais tendrait plutôt généralement vers un phénomène exponentiel (voir par exemple Baldwin et al., 2015) : de fortes concentrations à proximité de l'axe et une décroissance rapide en s'éloignant de l'axe.

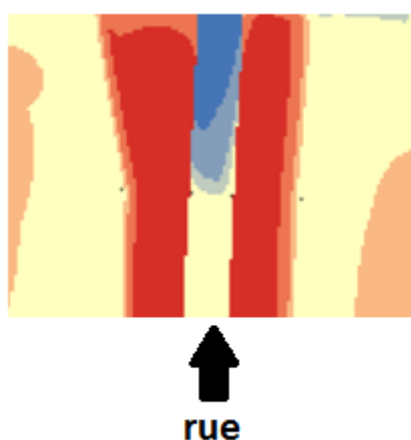


Figure 16 : Modèle de transect utilisé sur la grille 2 de Bourges

En situation de fond, des surfaces importantes de sur-estimation apparaissent. Celles-ci s'expliquent un réseau de points trop lâche sur ces zones.

Une démarche d'amélioration a été entreprise sur la grille 2. Elle vise à améliorer la cartographie tout en limitant le nombre de points récepteurs du maillage pour ne pas augmenter excessivement les temps de calcul.

Une première amélioration (grille améliorée V1) vise à ajouter un point au centre de l'axe traversant le transect et à ajouter un point intermédiaire entre les côtés de l'axe et les points extrêmes du transect.

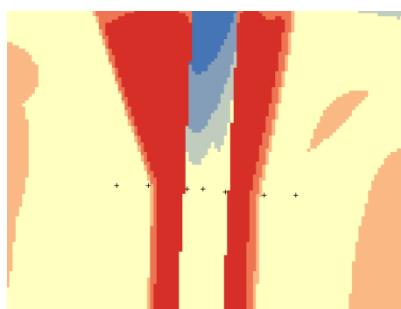


Figure 17 : a) Exemple de prise en compte de la décroissance exponentielle du NO₂ perpendiculairement aux axes routiers dans le maillage grille 2 améliorée V1

Cette première amélioration supprime la sous-estimation située sur l'axe routier et réduit la zone de surestimation autour de l'axe routier, qui reste toutefois assez importante du fait de la non-prise en compte de la décroissance exponentielle des concentrations lorsqu'on s'éloigne de l'axe. Afin de mieux représenter cette décroissance, les points intermédiaires qui ont été ajoutés sont rapprochés de l'axe, là où elle est la plus forte (plus on s'éloigne de l'axe et plus la décroissance tend vers un phénomène linéaire) (grille améliorée V2).

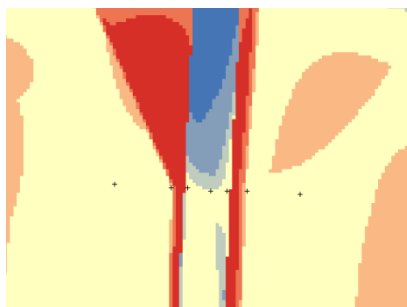
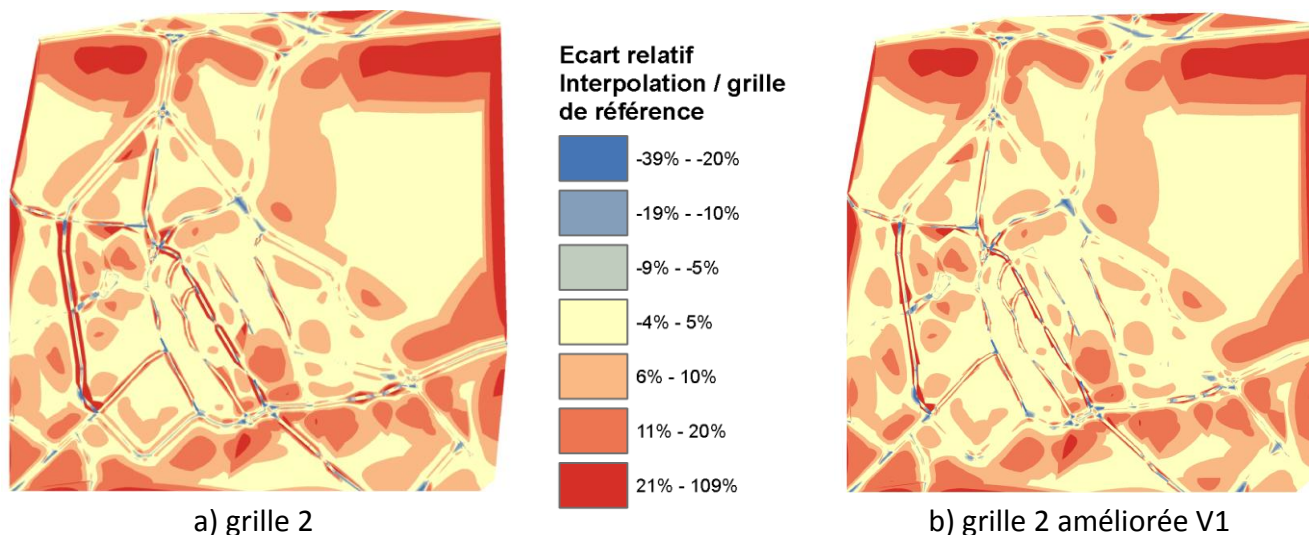
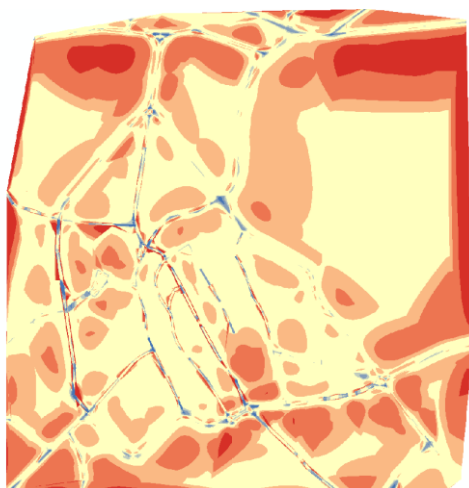


Figure 18 : Exemple de prise en compte de la décroissance exponentielle du NO₂ perpendiculairement aux axes routiers dans le maillage grille 2 améliorée V2

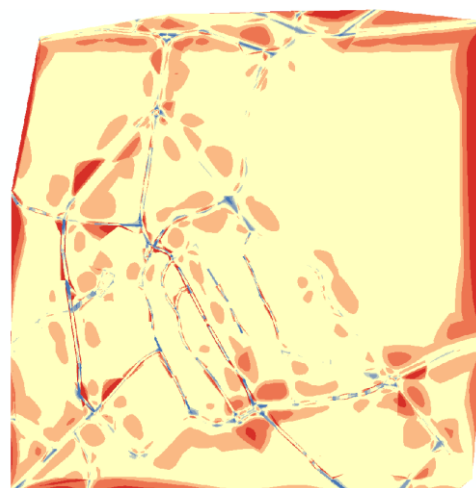
La Figure 18 montre l'impact du positionnement des nouveaux points à très faible distance des axes sur un exemple de route. Les écarts relatifs observés sur la moyenne annuelle de NO₂ entre les sorties de l'interpolation linéaire et la grille HD de référence sont ainsi fortement réduits en bordure d'axe entre les versions V2 et V1.

L'impact de ces améliorations sur les écarts relatifs est présenté en Figure 19. On s'aperçoit que la grille initiale conduit à une surestimation assez importante des concentrations de fond urbain. Ce problème est résolu par l'ajout de points supplémentaires dans ces zones de fond (grille V3).





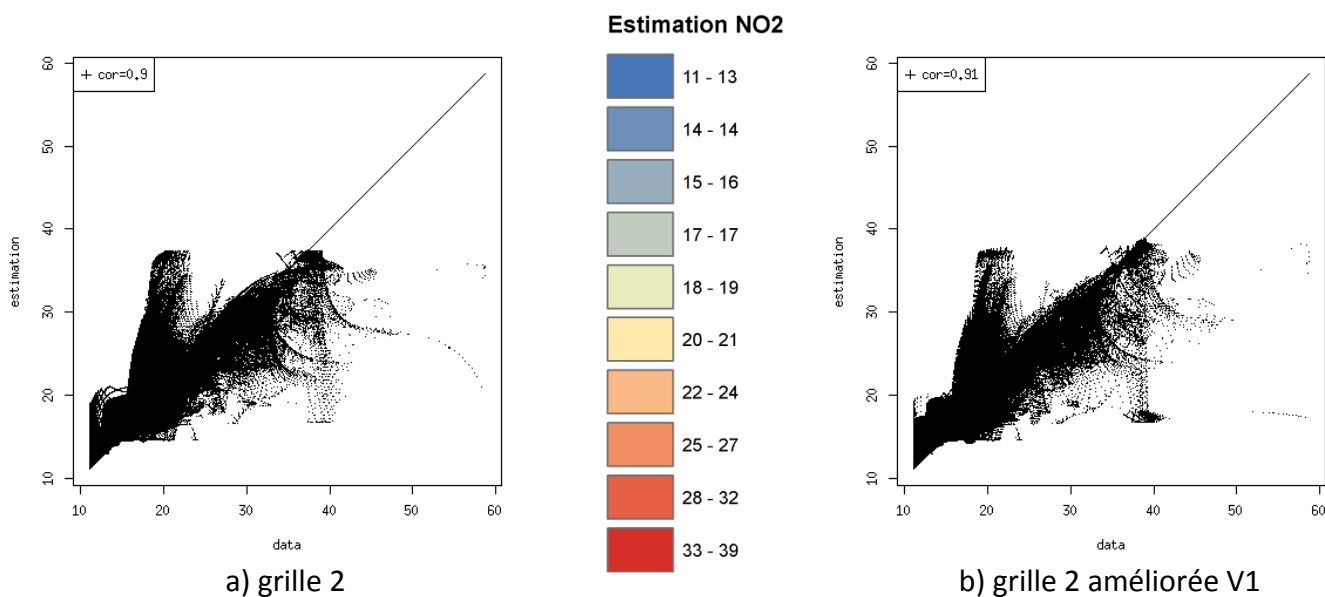
c) grille 2 améliorée V2

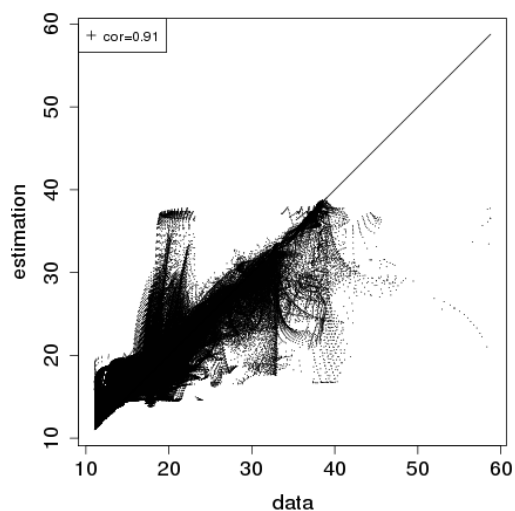


d) grille 2 améliorée V3

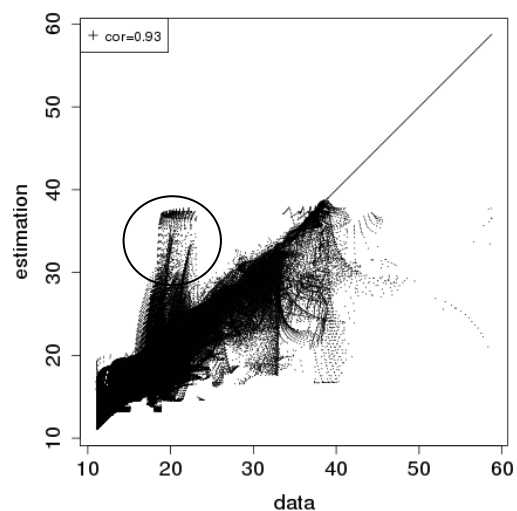
Figure 19 : Evolution des cartes de l'écart relatif sur la moyenne annuelle NO_2 ($\mu\text{g.m}^{-3}$) entre les sorties de l'interpolation linéaire et la grille HD de référence selon les améliorations de la grille 2 sur Bourges

L'analyse des nuages de points en Figure 20 montre une réduction importante de la surestimation autour des axes routiers et des concentrations de fond (voir la zone identifiée sur le nuage associé à la grille V3). La corrélation est passée de 0.90 à 0.93 entre la version initiale du maillage et sa version 3.





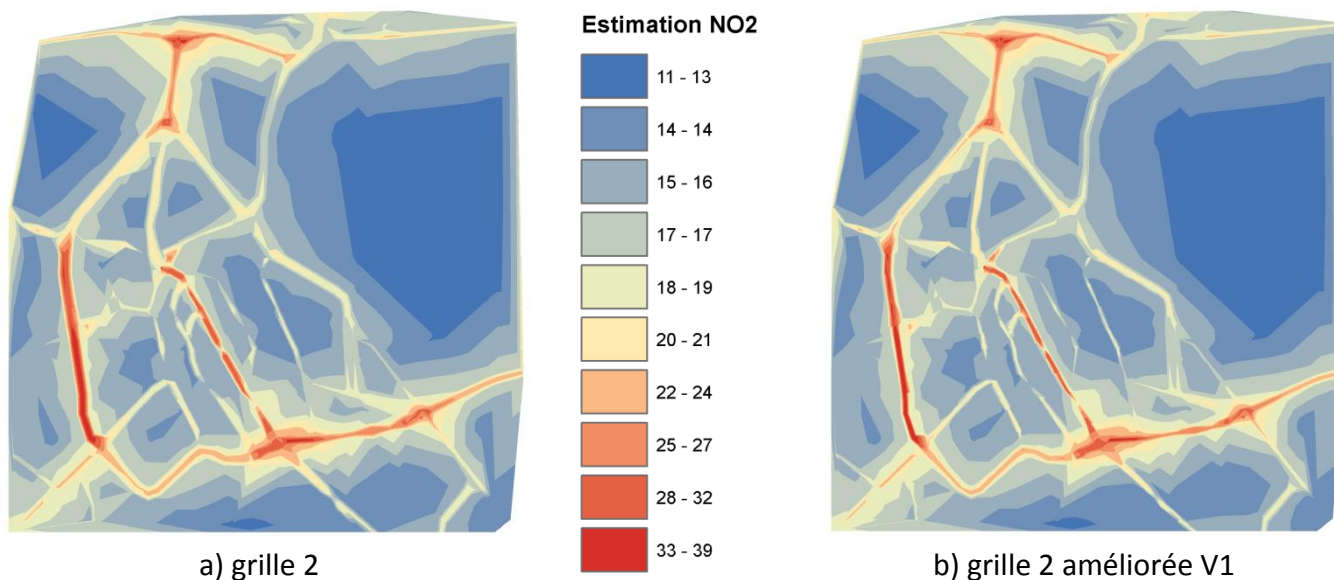
c) grille 2 améliorée V2

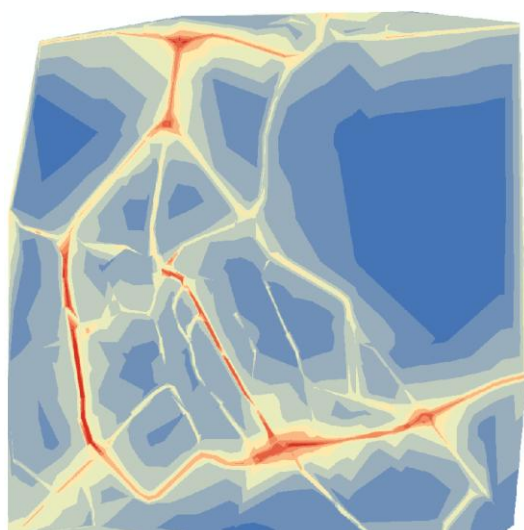


d) grille 2 améliorée V3

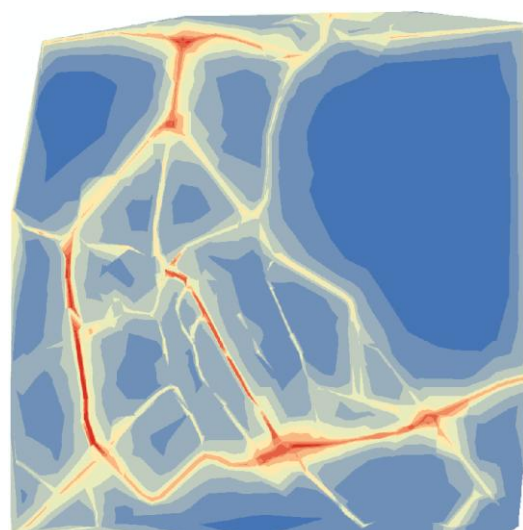
Figure 20 : Evolution des nuages de points et de la corrélation selon l'amélioration de la grille 2 sur Bourges

La Figure 21 présente l'évolution des cartes de la concentration moyenne annuelle de NO_2 selon les différentes versions de la grille 2 sur Bourges. On y retrouve les améliorations identifiées précédemment : une diffusion moins forte et plus réaliste des concentrations autour des principaux axes routiers et des niveaux de fond urbain plus faibles.





c) grille 2 améliorée V2



d) grille 2 améliorée V3

Figure 21 : Evolution des cartes de la concentration moyenne annuelle de NO_2 ($\mu\text{g.m}^{-3}$) selon l'amélioration de la grille 2 sur Bourges

7. APPLICATION DES DIFFERENTES METHODES D'INTERPOLATION

Les méthodes d'interpolation décrites et évaluées dans les chapitres précédents sont appliquées aux différents jeux de données fournis par les AASQA, sur le domaine complet et sur les fenêtres de zoom définies dans le Tableau 2. Sont ici présentés les résultats obtenus pour l'agglomération de Nantes. L'intégralité des cartes produites seront fournies au format PDF sur le site du LCSQA, en annexe de ce rapport.

7.1 Expertise visuelle des cartographies

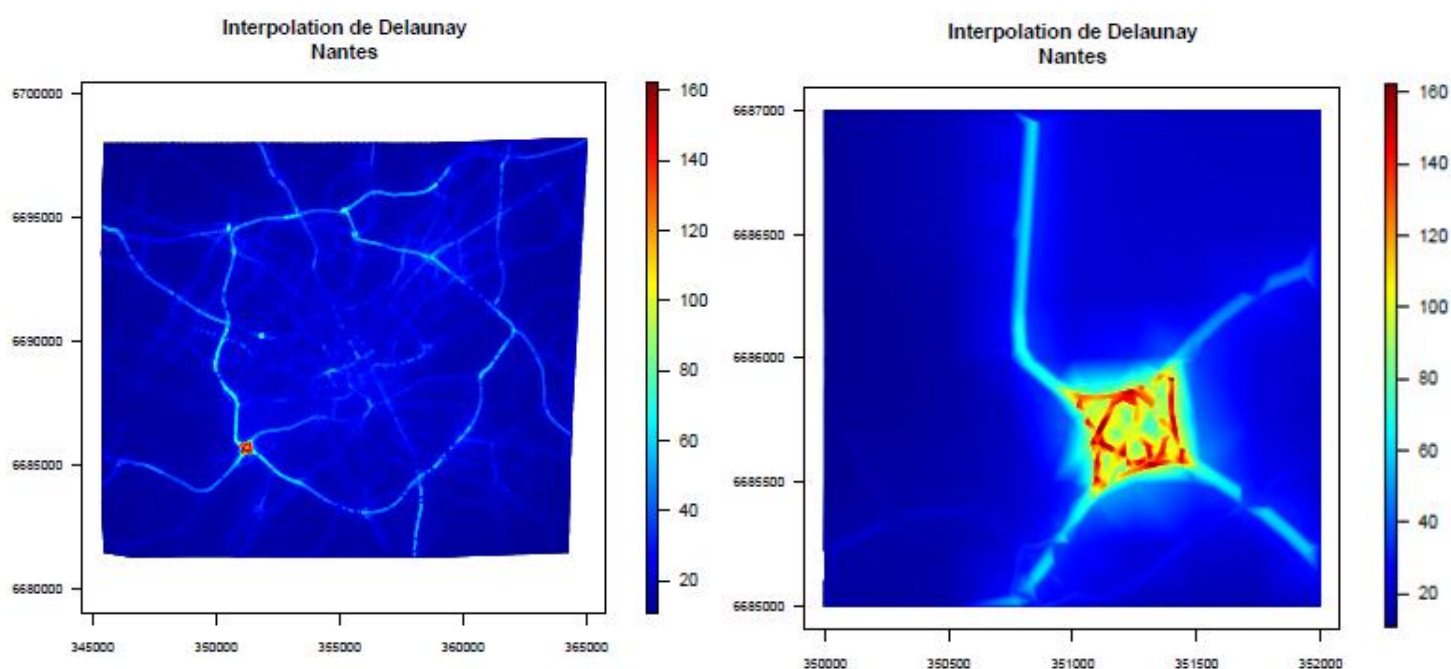


Figure 22 : Interpolation de Delaunay de la moyenne annuelle de NO_2 ($\mu\text{g.m}^{-3}$) réalisée sur les données de Nantes. Données complètes (gauche) et fenêtre de zoom (droite)

Les cartographies produites par krigeage des données de modélisation originales présentent dans la plupart des cas des discontinuités assez marquées liées au voisinage glissant utilisé. Dans le cas de Nantes (Figure 23), la carte produite paraît peu réaliste, même avec un nombre de voisins fixé à 10 dans l'interpolation. Ce n'est pas toujours le cas : sur Orléans (Figure 4) et Tours, les cartographies peuvent être plus lisses et une adaptation du krigeage spécifique à la zone urbaine offre donc des perspectives intéressantes.

Les cartographies produites avec l'interpolateur linéaire et l'interpolateur de Delaunay sont assez semblables : il est même difficile à première vue de les différencier sur les données d'Orléans et Tours (Figure 27 et Figure 28). En revanche, sur Bourges, Nantes, Niort, et Reims, le tracé des cartographies associées aux deux interpolateurs retenus montre que la décroissance des concentrations estimées perpendiculairement aux axes routiers dans les zones de très forte concentration semble moins rapide avec l'interpolateur de Delaunay. Il est délicat de dire laquelle des deux cartes est la meilleure tant les nuages de points en validation croisée et sur grille de référence sont similaires avec les deux interpolateurs. Les cartes d'écarts n'apportent pas non plus d'information complémentaire pour départager les méthodes. En fonction des axes routiers et des agglomérations, ce peut être l'un ou l'autre des deux interpolateurs qui sous-estime le moins les niveaux de concentrations. Ainsi, les deux méthodes d'interpolation aboutissant à des cartographies réalistes et des scores satisfaisants, le choix définitif reposera sur des critères additionnels tels que les temps de calcul ou les outils à disposition.

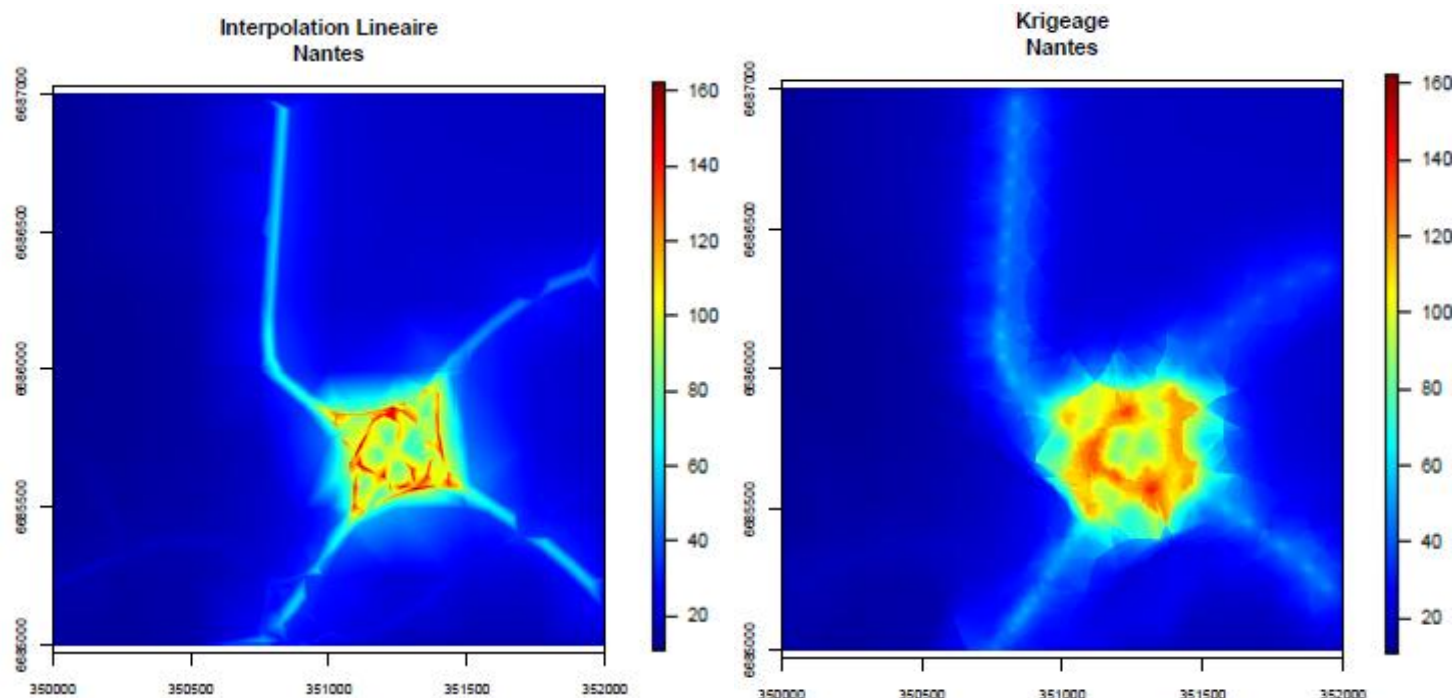


Figure 23 : Interpolation linéaire et krigeage de bloc (2m de résolution) à 10 voisins de la moyenne annuelle de NO_2 ($\mu\text{g.m}^{-3}$) sur la fenêtre de zoom de Nantes.

7.2 Temps de calcul et format de sortie

L'ensemble des codes informatiques utilisés dans cette étude sont écrits en langage R. L'objectif étant de comparer les méthodes d'interpolation, ils n'ont pas été optimisés pour être efficaces en temps de calcul. Si l'interpolation linéaire est très rapide (fonction R implémentée en langage C), l'interpolation de Delaunay peut nécessiter un temps de calcul contraignant en fonction du domaine étudié. On préconisera plutôt l'utilisation de logiciels SIG type OrbisGIS où l'implémentation de cet interpolateur s'avère plus efficace.

Les codes R développés par le LCSQA pourront toutefois être fournis sur demande. Les interpolations obtenues en sortie de ces codes sont fournies au format CSV en cas de maillage régulier (linéaire et krigeage) ou au format vectoriel et raster pour l'interpolation de Delaunay.

8. BILAN ET DISCUSSION

Dans les modélisations qu'elles réalisent à l'échelle urbaine, les contraintes rencontrées (temps de calcul, ressources informatiques, limites du logiciel...) imposent souvent aux AASQA de restreindre le nombre de points de calcul. Pour cartographier les concentrations et délimiter les zones de dépassement, une étape d'interpolation est alors nécessaire. Dans cette première partie, nous avons exploité un grand nombre de jeux de données de modélisation fournis par les AASQA afin de déterminer quelle méthode d'interpolation parmi l'interpolation linéaire, l'interpolation de Delaunay et le krigeage était la plus appropriée. La qualité de l'interpolation a été évaluée par validation croisée et lorsque c'était possible, par comparaison à une simulation de référence. Les cartographies obtenues par ces différentes techniques ont été également comparées.

Les résultats de validation croisée montrent que le krigeage n'est pas le plus adapté à l'interpolation de données en zone urbaine, où les sources de concentration et donc les typologies des données d'entrée sont très variables (fond urbain, fond périurbain, trafic). L'hypothèse de stationnarité du processus aléatoire qui intervient dans le krigeage n'est pas valide avec ce type de données. Toutefois, l'estimation géostatistique ne se limite pas au krigeage et des travaux sont actuellement en cours afin de pouvoir intégrer de manière adéquate la structure spatiale propre aux zones urbaines dans l'interpolation. De plus, une estimation géostatistique adaptée au contexte urbain, en prenant en compte directement la structure spatiale des données, pourrait permettre de réduire la résolution des maillages en sortie des modèles urbains, et donc les temps de calcul.

L'interpolateur linéaire et l'interpolateur de Delaunay sont, d'après la littérature scientifique et les résultats de ces travaux, les méthodes les plus efficaces et les moins coûteuses en temps de calcul. C'est aussi ce que nous constatons avec la présente étude. La validation croisée et la validation sur grille de référence ont montré que les scores produits par les deux méthodes étaient assez proches, tout comme l'aspect final des cartographies produites sur des grilles régulières. Les scores de l'interpolation linéaire sont généralement de meilleure qualité que ceux de l'interpolateur de Delaunay mais les écarts sont très faibles. Les nuages de points entre les données de modélisation originales et les valeurs interpolées ainsi que les cartes d'écarts confirment ce constat et ne permettent pas de privilégier de manière exclusive l'une ou l'autre de ces deux méthodes.

Enfin, plus encore que l'interpolateur employé, c'est le maillage utilisé en entrée de l'interpolation qui régit principalement la qualité de la cartographie finale : il est ainsi très important de construire un maillage qui permette de retranscrire la décroissance des concentrations perpendiculairement aux axes routiers. Le maillage est dépendant du contexte local et doit être construit en conséquence. Une AASQA qui souhaite optimiser son maillage en fonction de critères liés à la qualité de l'interpolation et au temps de calcul peut appliquer la méthode de validation sur une grille de référence. Des simulations haute résolution sur quelques périodes représentatives de l'année lui permettront d'établir un cas de référence à partir duquel elle pourra concevoir un maillage intelligent, qui minimise les écarts et limite les temps de calcul sur l'ensemble de l'année.

PARTIE 2 : ASSIMILATION DE DONNEES, UN ETAT DE L'ART

Les modèles déterministes comportent, tout comme les modèles statistiques, une erreur, du fait de la simplification plus ou moins grande des phénomènes physico-chimiques qui définissent la valeur de concentration d'un polluant en un point de l'espace, sous forme d'équations.

C'est pourquoi il se développe de plus en plus de méthodes de correction des modèles afin de prendre en compte les mesures existantes dans la cartographie finale et ainsi disposer d'une information optimale qui combine l'ensemble des données disponibles.

1. METHODES GEOSTATISTIQUES

Comme cela a été dit dans la première partie, la géostatistique usuelle n'est pas adaptée à l'interpolation de données de qualité de l'air de typologies différentes (fond et trafic). En revanche, il existe des modèles de krigeage non-stationnaire incluant une dérive externe et des méthodes de krigeage multivarié (cokrigeage)⁴.

Dans le cas du krigeage en dérive externe, on utilise les sorties du modèle urbain aux points de mesure (interpolés aux points de mesure si besoin) afin de construire un modèle statistique qui explique la mesure par le modèle urbain. On étudie ensuite la structure spatiale du résidu de ce modèle statistique qui est utilisé dans la résolution du système de krigeage. Si le modèle statistique a du sens, la modélisation de la structure spatiale (variographique) des résidus par une fonction isotrope (qui ne dépend que de la distance) est évidemment tronquée car la distance entre deux points n'est pas le seul critère qui explique la variation des résidus.

Dans le cas du cokrigeage, on utilise les sorties du modèle urbain comme variable secondaire dans le système de cokrigeage, tandis que les mesures constituent la variable principale. Le système de cokrigeage nécessite de modéliser à la fois la structure variographique des mesures, du modèle urbain, ainsi que la structure croisée. La même remarque que pour le modèle de dérive externe peut être faite, à savoir que la modélisation spatiale de la variation des concentrations par des fonctions isotropes n'est pas adaptée aux concentrations en zone urbaine.

Ces méthodes sont donc imparfaites pour corriger les sorties de modèles urbains mais elles ont toutefois été utilisées en considérant que le grand nombre de données en sortie du modèle permettrait de compenser les faiblesses liées à l'ajustement de la structure spatiale. Notons que le LCSQA, en collaboration avec le Centre de Géosciences (géostatistique) de Mines ParisTech, cherche à mettre en place un nouveau modèle de calcul de covariance, cohérent avec la physique du phénomène, et qui dissocie la pollution de fond et la pollution de proximité automobile. Ces travaux sont en cours et

⁴ Voir formation LCSQA en cartographie : <http://www.lcsqa.org/formations>

n'ont pas encore abouti à une méthode opérationnelle.

Nous présentons ici un cas d'application sur la communauté urbaine de Strasbourg (CUS) étudié par l'ASPA en 2011. Environ 60 sites de mesure ont été échantillonnés durant cette campagne, principalement en zone habitée avec quelques points localisés à proximité d'axes de circulation. Pour chacun de ces sites, une moyenne annuelle du NO₂ a été reconstituée à partir des moyennes estivales et hivernales estimées au cours des périodes suivantes : du 15 juin au 13 juillet 2011 pour l'été et du 9 décembre 2010 au 6 janvier 2011 pour l'hiver.

En parallèle une simulation ADMS-Urban a été réalisée sur l'année 2011. Les coefficients de corrélation entre mesures et modèles urbains se sont révélés assez bons (0.779 après suppression des valeurs extrêmes de quelques sites trafic) pour utiliser le modèle comme variable explicative dans un modèle statistique ou comme variable secondaire dans le cokrigage. On notera que le cokrigage réalisé par l'ASPA est ici colocalisé, dans le sens où l'interpolateur n'utilise la variable secondaire que sur le point à estimer et là où la variable principale est connue.

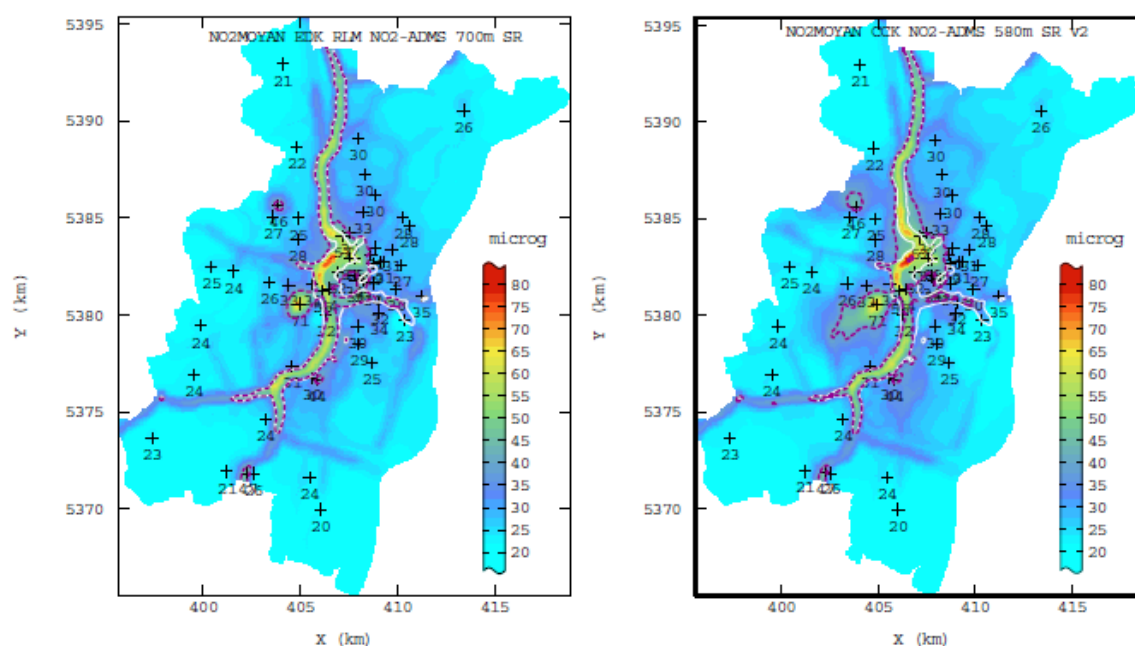


Figure 24 : Estimation de la concentration moyenne en NO₂ par krigeage des mesures avec la simulation ADMS-Urban en dérive externe (à gauche) et par cokrigage colocalisé des mesures avec la simulation ADMS-Urban (à droite). Localisation et valeurs des points de mesure (+). Isolignes au-dessus de 40 µg.m⁻³ déduite de l'estimation par krigeage (pointillés violets) et de la simulation ADMS (lignes blanches)

Le niveau de détail des cartographies finales permet de bien faire ressortir les axes routiers. Malgré leurs imperfections, ces approches permettent en partie de corriger la tendance à la sous-estimation d'ADMS-Urban sur certains secteurs du centre-ville de la CUS. Si le rendu cartographique apparaît quelque peu meilleur avec la méthode de dérive externe, celle-ci présente toutefois des écarts-type de krigeage supérieurs à la méthode du cokrigage colocalisé.

2. INTERPOLATION OPTIMALE EN ASSIMILATION DE DONNEES

Des travaux récents ont été réalisés afin d'adapter ce qui est fait en assimilation de données à grande échelle à la correction des concentrations de polluants modélisées en milieu urbain par des modèles déterministes.

Une étude de 2013 conduite sur l'agglomération clermontoise par des chercheurs de l'INRIA et du CEREAS (Tilloy et al., 2013) a permis de corriger les concentrations de NO₂ modélisées toutes les 3h par le modèle ADMS-Urban (30.971 points récepteurs) sur toute l'année 2008. Les neuf stations de mesure du réseau de la qualité de l'air ont été utilisées pour corriger le modèle avec un estimateur de type BLUE. Nous présentons ici un résumé du travail effectué dans cette étude.

On rappelle préalablement le vocabulaire utilisé en assimilation de données. Les notations sont reprises de l'article cité précédemment.

$\mathbf{c}^b$ est l'ébauche, vecteur d'état correspondant aux sorties du modèle urbain.

$\mathbf{c}^t$ est le vecteur d'état réel (aux points récepteurs du modèle urbain), qui est inaccessible.

$\mathbf{c}^a$ est le vecteur analyse, issu de la combinaison de l'ébauche et des observations.

$\mathbf{o}$ est le vecteur des observations fourni par les mesures réalisées aux stations fixes du réseau de mesure.

On rappelle qu'en assimilation de données, le vecteur $\mathbf{o}$ se trouve dans un espace dit d'observations, qui est distinct de l'espace des états dans le sens où il est impossible d'observer intégralement l'état du système physique. Afin de pouvoir comparer les observations avec l'état du système, on introduit un opérateur $\mathbf{H}$, que l'on appelle opérateur d'observation, allant de l'espace des états dans l'espace d'observations. Pour un état $\mathbf{x}$ donné, $\mathbf{H}(\mathbf{x})$ appartient à l'espace des observations. En dimension finie, dans un cadre linéaire, on peut imaginer que chaque ligne de la matrice $\mathbf{H}$ représente un opérateur d'interpolation entre les points de grille du modèle discret et les points d'observation. Ce n'est pas le cas ici car un point récepteur a été ajouté sur chaque station dans le modèle. En conséquence $\mathbf{H}_{ij}=1$ lorsque la colonne j correspond au point récepteur localisé à la station i et 0 sinon.

Ici, $\mathbf{H}\mathbf{c}^b$ peut être vue comme la valeur simulée par le modèle au point d'observation. La différence $\mathbf{o}-\mathbf{H}\mathbf{c}^b$ est appelée innovation.

On considère maintenant que l'erreur de simulation $\mathbf{c}^b - \mathbf{c}^t$ est sans biais et de variance $\mathbf{B}$.

On considère également que l'erreur d'observation $\mathbf{o} - \mathbf{H}\mathbf{c}^t$ est sans biais et de variance $\mathbf{R}$.

Le meilleur estimateur linéaire (**BLUE**), dont l'erreur $\mathbf{c}^a - \mathbf{c}^t$ est sans biais et dont la variance est de trace minimale (Bouttier and Courtier, 1999) est défini par :

$$\mathbf{c}^a = \mathbf{c}^b + \mathbf{K}(\mathbf{o} - \mathbf{H}\mathbf{c}^b) \text{ avec } \mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$$

Comme précisé dans l'article de Tilloy et al. (2013), s'il est possible dans des contextes de pollution de fond à grande distance de modéliser les covariances de l'erreur d'état par une fonction dépendant de la distance géographique, ce n'est plus possible en milieu urbain où les covariances dépendent également du réseau routier. Ces commentaires rejoignent ceux déjà énoncés lors de l'interpolation géostatistique par krigeage des sorties du modèle urbain et de ses éventuelles corrections (voir Partie1-4.3 et Partie2-1).

La matrice de covariance **R** est considérée comme diagonale et tous ses éléments diagonaux sont égaux à v_0 , la variance des erreurs d'observation. Ainsi, on considère qu'il n'existe aucune corrélation entre les erreurs observées à 2 stations différentes.

La construction de la matrice de covariance **B** est plus complexe et constitue l'élément novateur de l'approche. La covariance est vue comme une fonction qui décroît exponentiellement non seulement le long de la route mais aussi lorsqu'on s'éloigne de la route selon la formule :

$$B_{ij} = v_c \exp\left(-\frac{d_{ij}}{L_d}\right) \exp\left(-\frac{|P_i - P_j|}{L_p(i,j)}\right)$$

où v_c est une variance, d_{ij} est la plus courte distance à parcourir le long du réseau routier pour rejoindre le point j à partir du point i (ou leurs projetés orthogonaux sur la route si les deux points récepteurs ne sont pas sur une route) et P_i est la distance du point i à la route la plus proche.

Enfin L_d et L_p sont des distances caractéristiques (strictement positives) le long et perpendiculairement à la route. L_p est corrigée de la sorte : $L_p(i,j) = L_p + \alpha \min(P_i, P_j)$ afin que la corrélation entre deux points récepteurs mesurant la pollution de fond reste significative à grande distance. Pour plus de détails concernant les améliorations futures susceptibles d'être apportées à la modélisation de la covariance des erreurs d'états, on se reportera à la section 3.3 de Tilloy et al.

La détermination des paramètres de l'assimilation se fait comme suit :

- v_0 est calculée comme la somme des incertitudes liées à la mesure (calibration, conditions de température et de pression, etc.) et de l'erreur de représentativité de la mesure (qui n'est pas réalisée à la même hauteur que la simulation).
- (v_c, L_d, L_p, α) sont déterminés d'après un test de diagnostic du χ^2 et les sorties cartographiques.

L'application de la méthode montre des cartographies assimilées qui corrigent la sous-estimation des niveaux de pollution d'ADMS-Urban à proximité des routes (Figure 25).

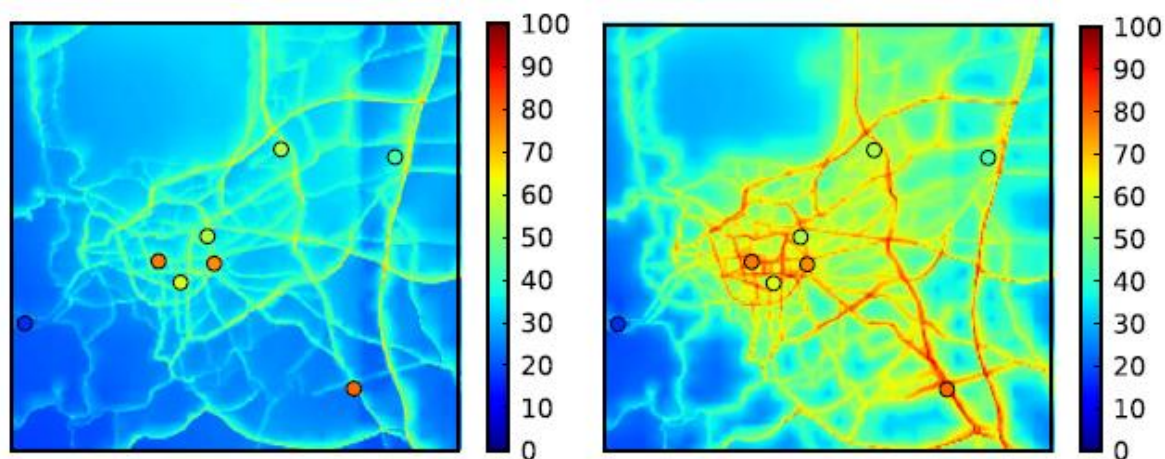


Figure 25 : Cartes obtenues avant (gauche) et après assimilation (droite) sur Clermont le 10 juillet 2008 à 06h UTC.

Les scores obtenus en validation croisée sur les stations de mesure montrent également un gain important dans la qualité de l'estimation.

Pour plus de détails sur les avantages et les inconvénients de la méthode, on se reportera aux sections Résultats et Discussions de l'article de Tilloy et al.

A l'avenir, il est important que la mise en œuvre de ce type de méthodes s'accompagne d'une validation sur un plus grand nombre de points (campagnes de mesures).

3. BILAN ET DISCUSSION

S'agissant de l'utilisation de données de tubes, les discussions en *Commission de Suivi Emissions, Modélisation et Traitement de Données* ont montré que les AASQA avaient des approches diverses. Certaines utilisent ces données pour évaluer ou ajuster leurs résultats de modélisation, d'autres les considèrent comme trop imprécises pour ces usages. Afin de répondre objectivement à ces différences d'appréciation, une étude métrologique sur la performance des tubes sera conduite en 2015 par le LCSQA. A l'issue de cette étude, des compléments méthodologiques sur l'exploitation conjointe de données de modélisation urbaine et de données d'échantillonnage passif pourront être apportés si le besoin est avéré.

En ce qui concerne la possibilité de combiner la modélisation urbaine et la mesures fixe, elle reste très dépendante du nombre de stations incluses dans le domaine d'étude.

PARTIE 3 : SURFACES EN DEPASSEMENT ET EXPOSITION DE LA POPULATION

1. INTRODUCTION

La première partie a montré que l'interpolation par krigeage n'était pas encore adaptée à l'estimation en zone urbaine. Dans cette dernière partie, nous nous intéresserons donc uniquement à l'interpolateur linéaire et à l'interpolateur de Delaunay. Ces deux méthodes seront utilisées pour cartographier les concentrations afin d'estimer les surfaces en dépassement et les populations exposées à ces dépassements.

2. SURFACE EN DEPASSEMENT

Remarques :

- Dans toute la suite, le terme de *supérieur* doit s'entendre au sens de *strictement supérieur*.
- Comme pour la mesure, la comparaison avec la valeur limite se fait après arrondi de la valeur estimée selon la même précision numérique que cette dernière, i.e. à l'entier (Guide IPR, partie I, 2013).

2.1 Cas de l'interpolateur linéaire

L'interpolation linéaire est effectuée sur une grille régulière (avec une résolution fixée à 2m suite à un compromis entre temps de calcul et précision de l'interpolateur). La zone en dépassement est donc l'ensemble des mailles dans lesquelles la valeur interpolée est strictement supérieure à la valeur limite. La superficie associée est simplement calculée comme le produit du nombre de mailles par leur surface (4m²).

2.2 Cas de l'interpolateur de Delaunay

Nous avons vu en section 4.2 comment était définie l'interpolation de Delaunay. Plus généralement, soient une isovaleur **iso1** et un triangle de la triangulation de Delaunay dont la concentration est supérieure à iso1 en au moins un sommet. L'enveloppe convexe des positions associées à cette isovaleur **iso1** et aux sommets concernés forme un polygone (triangle ou quadrilatère, voir **Figure 26**) à l'intérieur duquel la valeur de concentration est supérieure à **iso1**.

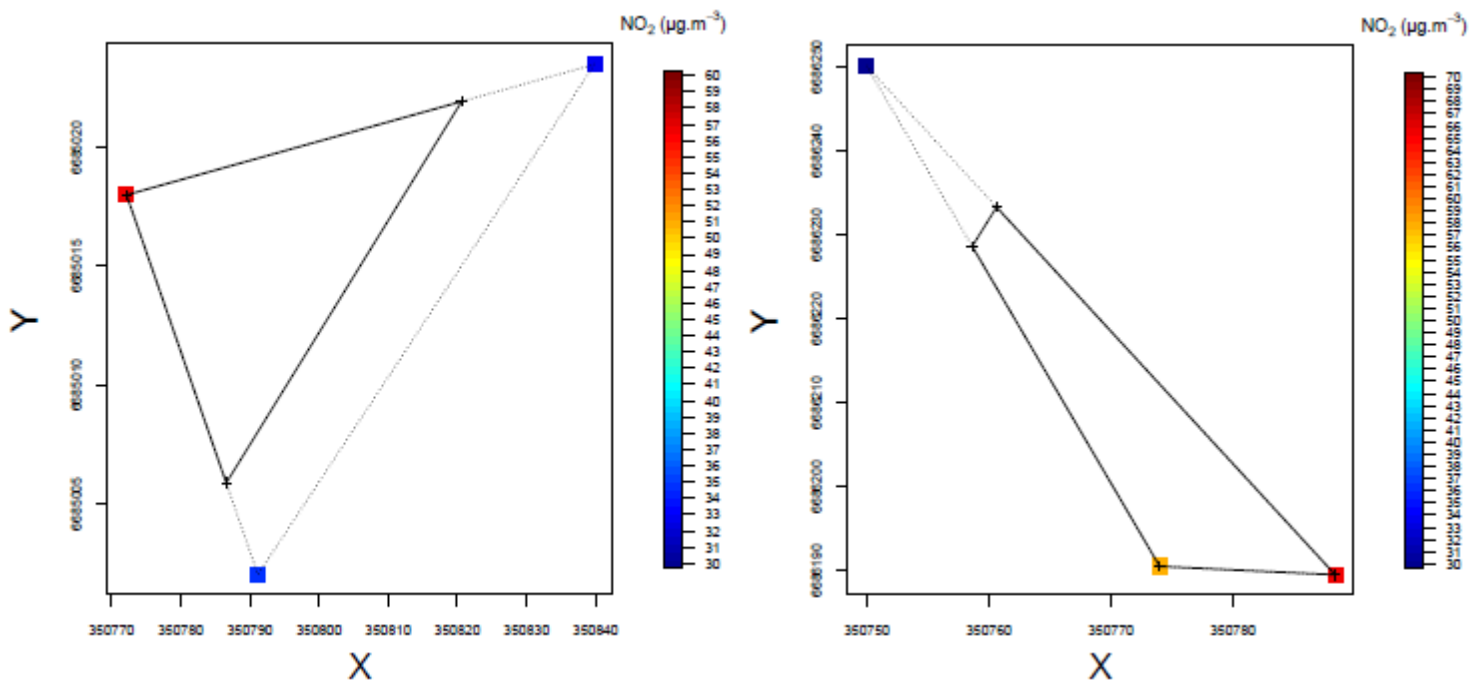


Figure 26 : Exemples de construction de polygones dont la moyenne annuelle NO_2 ($\mu\text{g.m}^{-3}$) est supérieure à la valeur limite (i.e. à l'isovaleur $40 \mu\text{g.m}^{-3}$) avec l'interpolateur de Delaunay. Les points de couleurs sont les sommets du triangle ; les croix sont les positions (obtenues par interpolation linéaire) où est atteinte la valeur limite le long des arêtes

La surface en dépassement sur l'ensemble du domaine est calculée comme la somme de la surface des polygones à l'intérieur desquels la valeur de concentration est supérieure à la valeur limite.

2.3 Résultats

1. Les fenêtres de zoom ont été définies afin d'identifier dans chaque jeu de données les zones où les concentrations étaient les plus fortes. En conséquence, celles-ci sont généralement centrées autour d'axes routiers de grande importance. Nous avons vu en section 7.1 de la partie 1 que les cartographies produites avec l'interpolateur linéaire et l'interpolateur de Delaunay étaient assez semblables, si ce n'est une tendance de l'interpolateur de Delaunay à faire décroître moins rapidement les fortes concentrations perpendiculairement aux axes routiers. Si la différence ne semble pas toujours significative du point de vue des cartographies, elle ne l'est pas non plus en ce qui concerne la surface estimée en dépassement. Ainsi, sur les fenêtres de zoom des six villes étudiées, l'écart relatif entre la surface estimée en dépassement de la valeur limite NO_2 n'excède jamais 1%, sauf pour les données de Niort où l'écart atteint 5% (voir

Tableau 4). Aucune tendance générale ne se dégage : sur Niort et Orléans, la surface estimée par Delaunay est supérieure à celle qui est estimée par l'interpolateur linéaire et la situation inverse est observée sur les quatre autres villes. Mais les écarts n'étant pas significatifs (<5%), on peut conclure à une équivalence des deux interpolateurs sur la question du calcul des surfaces en dépassement de valeurs limites.

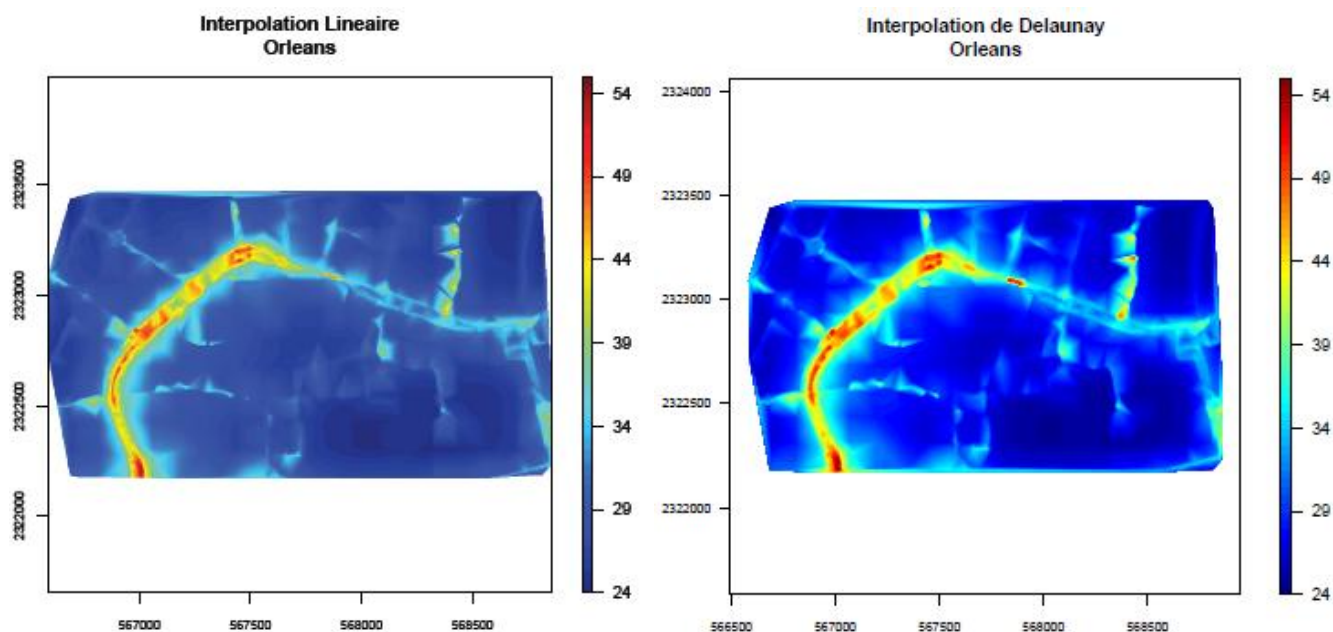


Figure 27 : Interpolation linéaire (à gauche) et interpolation de Delaunay (à droite) de la moyenne annuelle NO_2 ($\mu\text{g.m}^{-3}$) sur la fenêtre de zoom d'Orléans

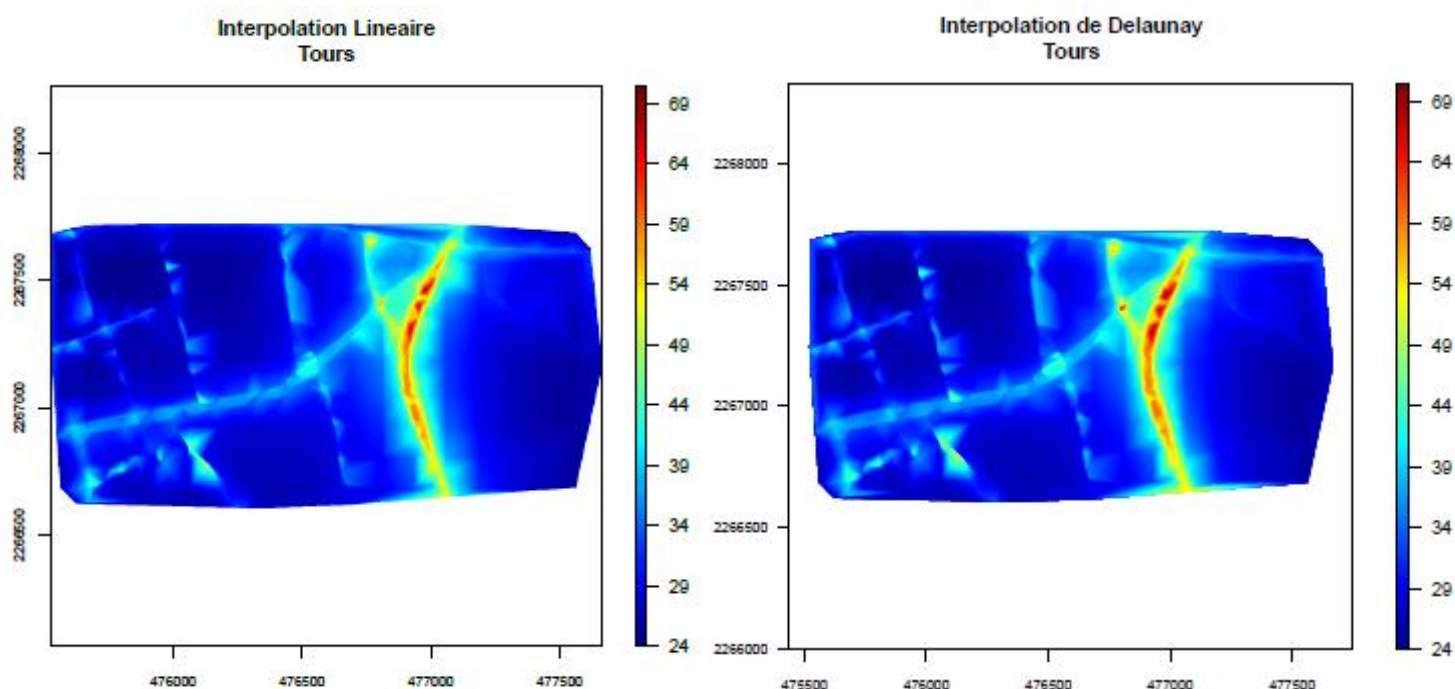


Figure 28 : Interpolation linéaire (à gauche) et interpolation de Delaunay (à droite) de la moyenne annuelle NO_2 ($\mu\text{g.m}^{-3}$) sur la fenêtre de zoom de Tours

Tableau 4 : Surfaces en dépassement estimées avec l'interpolateur linéaire (résolution du maillage d'interpolation : 2m) et l'interpolateur de Delaunay sur les fenêtres de zoom

ville	méthode	surface en dépassement (m ²)
Bourges	<i>linéaire</i>	48
	<i>Delaunay</i>	47.17
Nantes	<i>linéaire</i>	518836
	<i>Delaunay</i>	518819
Niort	<i>linéaire</i>	2580
	<i>Delaunay</i>	2717
Orléans	<i>linéaire</i>	107024
	<i>Delaunay</i>	107036
Reims	<i>linéaire</i>	430960
	<i>Delaunay</i>	430833
Tours	<i>linéaire</i>	185920
	<i>Delaunay</i>	185916

Dans la partie 6.3, le maillage numéro 2 utilisé sur la ville de Bourges a été optimisé pour réduire les surestimations en fond et à proximité des axes de circulation, et corriger les sous-estimations sur les axes. Outre l'effet sur les concentrations, il est utile d'évaluer l'effet de cette amélioration sur l'estimation des zones en dépassement. Les concentrations n'étant pas suffisamment élevées sur cette ville, un seuil de 20 µg/m³ est retenu pour simuler un dépassement.

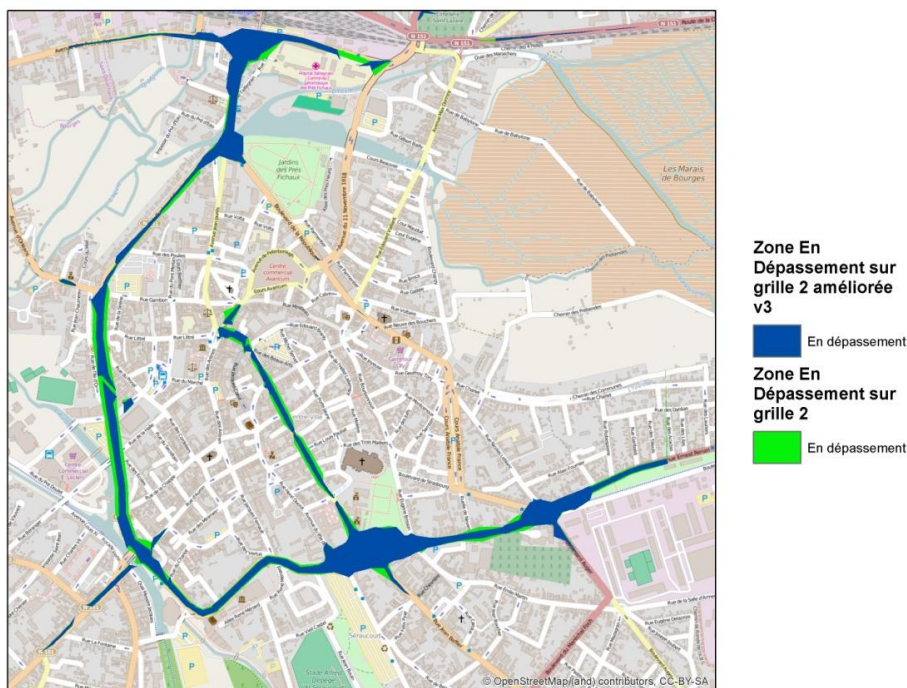


Figure 29 : Comparaison des zones en dépassement sur la ville de Bourges estimées à partir de l'interpolation linéaire de la grille 2 (en vert) et de la grille 2 améliorée V3 (en bleu)

L'amélioration réduit la zone en dépassement autour des axes et grands carrefours (Figure 29). La comparaison avec la simulation de référence HD (Figure 30 et Figure 31) montre que la vitesse de décroissance des concentrations autour des axes est mieux appréhendée mais reste sous-estimée autour de certains carrefours. D'autre part, certains axes mineurs présentent des dépassements sur la simulation HD qui ne sont pas reproduits par l'interpolation. Cet exemple sur Bourges montre qu'il serait probablement utile de continuer les travaux d'optimisation de la grille.

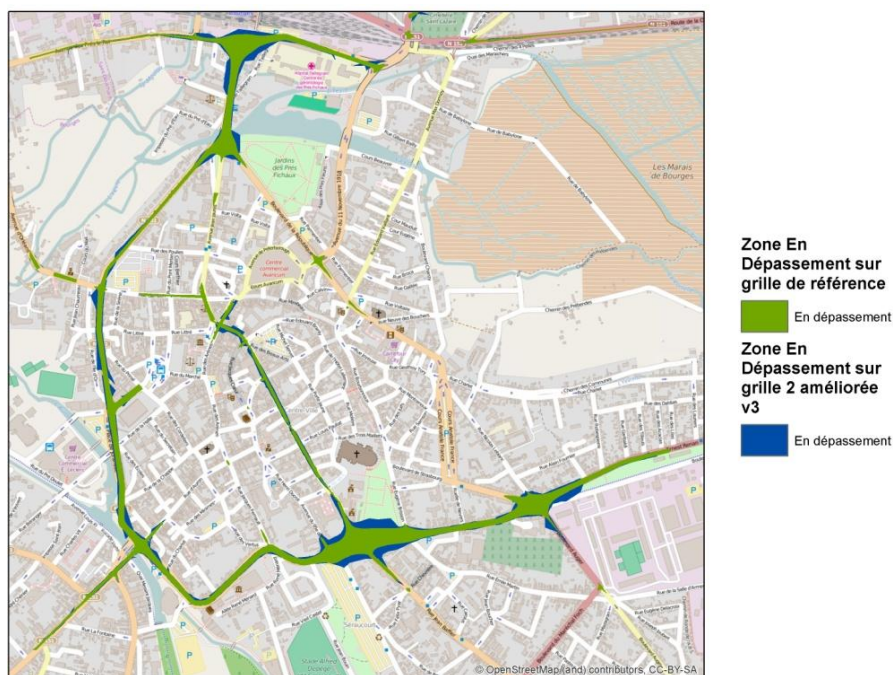


Figure 30 : Comparaison des zones en dépassement sur la ville de Bourges estimées à partir de la simulation de référence (en bleu) et de la grille 2 améliorée V3 (en bleu) (interpolation linéaire)

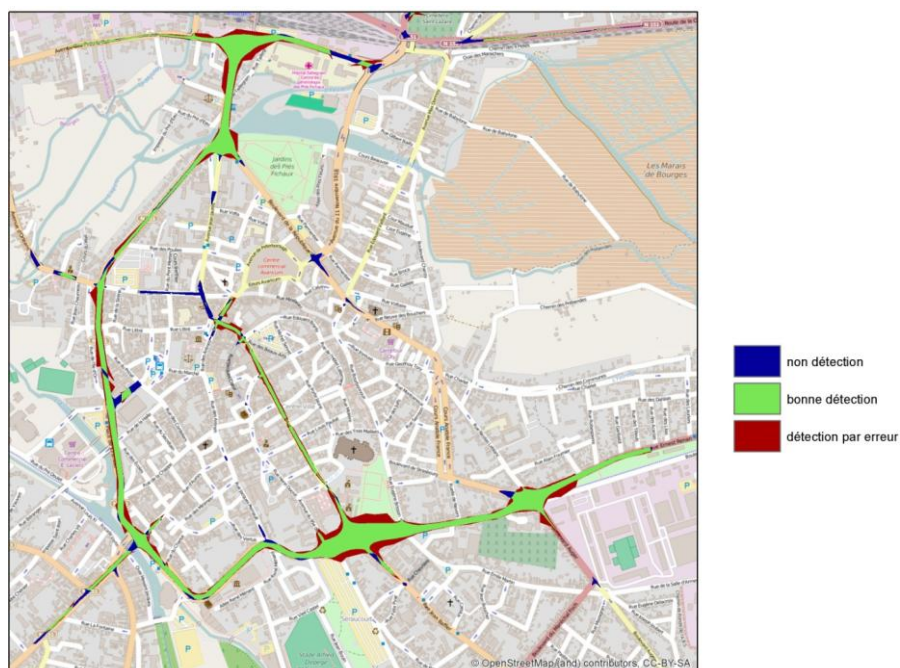


Figure 31 : Comparaison entre la zone de dépassement estimée à partir de la grille de référence et celle estimée à partir de l'interpolation sur la grille 2 améliorée V3 (interpolation linéaire) : bonne et mauvaise détection

3. POPULATION EXPOSEE

La précédente partie a permis d'établir les surfaces en dépassement sur les différentes agglomérations. Il reste maintenant à estimer les populations soumises à ces dépassements. Cette dernière étape est illustrée avec l'exemple de Bourges. Une étude d'optimisation du maillage a été menée pour cette ville (maille 2 « amélioré », cf. §6.3). Les résultats d'interpolation et d'estimation de la zone en dépassement sur ce maillage sont utilisés dans la présente partie pour estimer les populations exposées.

Un exemple simplifié (Figure 32) présente une surface fictive en dépassement couvrant un certain nombre de bâtiments habités. Trois bâtiments sont complètement inclus dans la zone. Un quatrième (numéro 81539) est partiellement couvert par celle-ci et deux autres sont localisés à l'extérieur.

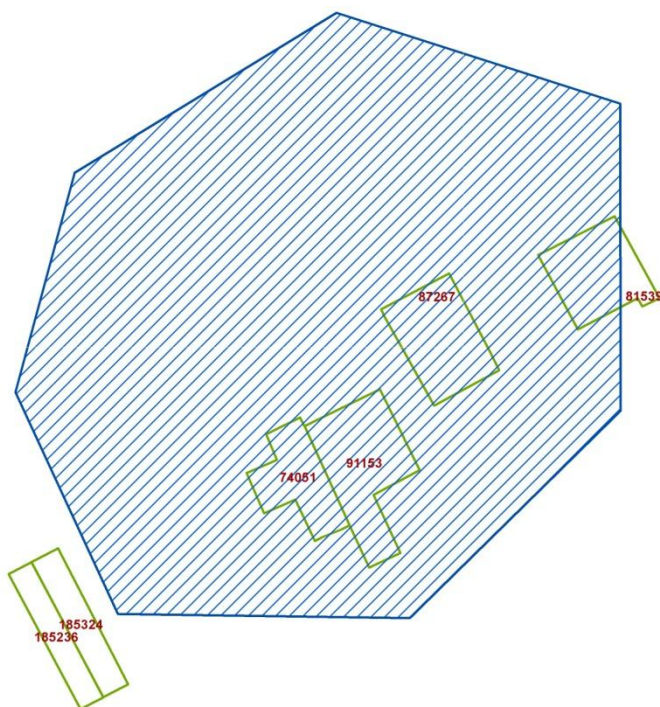


Figure 32 : Exemple du croisement d'une surface en dépassement (en bleu) avec des bâtiments habités (en vert)

3.1 Les méthodes employées au sein des AASQA

Les données de population exploitées dans les AASQA se présentent sous deux formes différentes :

- La population est placée directement dans les bâtiments d'habitation.
- La population est placée dans les bâtiments d'habitation pour ensuite être spatialisée sur un maillage régulier. Les populations des bâtiments croisant le maillage sont cumulées de la manière suivante : si un bâtiment est totalement inclus dans une maille, la totalité de ses habitants est comptabilisée dans celle-ci ; s'il se trouve à la frontière de 2, 3 ou 4 mailles, sa population est répartie dans chacune des mailles proportionnellement à la surface de chevauchement sur chacune d'elles.

Le choix de l'une ou l'autre forme est principalement guidé par un critère opérationnel : pour une étendue géographique importante de la zone de modélisation et une résolution spatiale fine, l'utilisation du maillage facilite le croisement entre les concentrations modélisées et les données de population.

Différentes approches sont employées dans les AASQA pour croiser les cartes de concentration avec les données de population et estimer les populations soumises à des dépassements :

Méthode 1. A partir des résultats de modélisation bruts, si ces derniers sont de haute résolution, ou des données interpolées, dans le cas le plus fréquent, la zone en dépassement est estimée sous la forme d'un polygone (objet vectoriel). La population est soit spatialisée dans les bâtiments de la zone concernée, soit placée dans un maillage régulier, selon la règle de proportionnalité précédemment décrite. La zone en dépassement est alors croisée avec la population à l'aide d'un outil SIG. Il est supposé que les populations de tous les bâtiments ou mailles entièrement inclus dans la zone sont exposées. Pour les bâtiments ou mailles dont le contour recoupe la frontière de la zone, l'une des deux hypothèses suivantes est employée :

- 1A. Une première hypothèse suppose que seule une partie de la population du bâtiment (ou maille) est exposée. La population retenue est proportionnelle à la surface en dépassement couvrant le bâtiment (ou la maille).
- 1B. Une seconde hypothèse suppose que toute la population des bâtiments (ou mailles) croisant la zone se trouve exposée au dépassement. Les AASQA justifient cette hypothèse en estimant que celle-ci intègre l'incertitude des simulations.

Méthode 2. Les concentrations sont calculées sur un maillage régulier (résultats de modélisation bruts ou après interpolation). Ce même maillage est utilisé pour spatialiser la population. Le croisement entre la zone en dépassement et la population est alors réalisé simplement maille par maille : la population d'une maille est retenue si celle-ci est en dépassement. Cette méthode peut être vue comme un cas particulier de la méthode 1A puisque seule la population des portions de bâtiments incluses dans les mailles en dépassement est comptabilisée. Un cas particulier (méthode 3) a été relevé pour une AASQA. Celle-ci est contrainte par des temps de calcul très importants du fait d'une large étendue spatiale et d'une résolution très fine. Elle a à sa disposition des populations spatialisées sur des bâtiments qu'elle souhaite appliquer sur une grille régulière. L'application des populations se déroule en plusieurs étapes :

- chaque maille est représentée par un point situé en son centre ; on considère qu'une maille croise un bâtiment si ce dernier inclut son centre.
- pour chaque bâtiment est compté le nombre de points se trouvant dans son enceinte ;
- la population du bâtiment est ensuite répartie dans chacune des mailles qui le croisent au prorata de ce nombre de points.

3.2 Tests de sensibilité

Les résultats précédents sur le maillage 2 de la ville de Bourges sont repris ici pour tester différents paramètres et proposer des recommandations sur le croisement de la population avec des surfaces en dépassement. Ces résultats ont montré qu'il existe peu de différence entre l'interpolateur linéaire et celui de Delaunay. La suite des travaux porte uniquement sur l'interpolateur linéaire.

Une méthodologie nationale, dite « méthodologie MAJIC », destinée à spatialiser finement la population a été mise au point par le LCSQA⁵ et elle constitue désormais la référence à utiliser pour le rapportage. Elle a été appliquée sur la ville de Bourges pour élaborer un jeu de données de population spatialisé sur chacun des bâtiments de la zone d'étude. Pour faciliter la suite des calculs, les populations ont été ensuite réparties sur le maillage HD selon une règle de proportionnalité. Cette règle suppose que la population dans un bâtiment est uniformément répartie dans celui-ci et peut donc être répartie dans chacune des mailles qui croisent ce bâtiment au prorata des surfaces d'intersection.

Le croisement direct de la population maillée avec la simulation HD (méthode 2) définit une **population exposée de référence** qui sera utilisée pour la suite de l'étude (Tableau 5).

Tableau 5 : Définition de la population exposée de référence

grille HD de référence	
Surface en dépassement (km ²)	0.133
Surface en dépassement (%*)	3.3%
Population exposée (Nbre Habitants)	180
Population exposée (%*)	1.1%
Résolution du maillage de modélisation	1 m (maillage HD)
Population	Spatialisée sur le maillage HD selon la règle de proportionnalité
Méthode de croisement	Méthode 2

* : Pourcentage calculé sur la surface et la population de la fenêtre de zoom étudiée (cf Tableau 8).

Rappel : les niveaux de concentration de NO₂ étant relativement faibles, le seuil ici considéré pour le dépassement est de 20 µg.m⁻³.

⁵ Létinois L., 2015. Méthodologie de répartition spatiale de la population. Rapport LCSQA. (Mise à jour du document de 2013.)

Dans les tests qui suivent, l'influence de différents paramètres sur l'estimation de la population exposée est testée : la qualité du maillage en entrée de l'interpolation, la résolution du maillage de sortie de l'interpolation et la méthode de croisement (règle de prise en compte du bâti).

La situation initiale pour réaliser ces tests est définie dans le tableau suivant.

Tableau 6 : Situation initiale avant étude de sensibilité

Surface en dépassement (km ²)	0.155
Surface en dépassement (%)	3.9%
Population exposée (Nbre Habitants)	467
Population exposée (%)	2.9%
Maillage en entrée de l'interpolation	Grille 2 améliorée v3
Données après interpolation	Grille de 1 m de résolution (maillage HD)
Population	Spatialisée sur le maillage HD selon la règle de proportionnalité
Méthode de croisement	Méthode 2 (la grille d'interpolation coïncidant avec la grille de population)

La comparaison des résultats avec les données du Tableau 5 (indiquées en bleu dans les tableaux suivants) permet d'évaluer l'écart par rapport à la réalité supposée. La comparaison avec les données du Tableau 6 (indiquées en vert dans les tableaux suivants) permet d'évaluer l'écart par rapport à une configuration de calcul optimisée.

3.2.1 Effet du maillage en entrée de l'interpolation

De nombreux tests ont déjà été présentés dans les parties précédentes sur l'amélioration du maillage en entrée de l'interpolation. Les résultats obtenus pour les différentes optimisations du maillage (voir 6.3) sont réutilisés pour estimer la population exposée.

Tableau 7 : Impact de la qualité du maillage d'entrée de l'interpolation sur le calcul des populations exposées

	Grille HD de référence	Grille 2	Grille 2 améliorée	Grille 2 améliorée v2	Grille 2 améliorée v3
Surface en dépassement (km ²)	0.133	0.190	0.165	0.152	0.155
Surface en dépassement (%)	3.3%	4.8%	4.1%	3.8%	3.9%

Population exposée (Nbre Habitants)	180	704	521	416	467
Population exposée (%)	1.1%	4.4%	3.3%	2.6%	2.9%

En 6.3 (Partie 1) et 2.3 (Partie 2), il a été montré que l'optimisation du maillage 2 permettait de restreindre la surface en dépassement et de se rapprocher ainsi de la surface « réelle ». La tendance sur la population exposée est identique, même si l'on peut constater une légère dégradation entre la version 2 et la version 3 de la grille. Cette dégradation s'explique par le fait que la ville étudiée présente une faible surface en dépassement et donc peu d'habitations concernées. Une faible modification des frontières peut inclure ou exclure deux ou trois bâtiments comportant une population importante. La Figure 33 présente un extrait de la ville de Bourges sur lequel sont présentées les zones en dépassement respectivement propre à la version 2 (en rouge), propre à la version 3 (en bleu) et commune à la version 2 et à la version 3. La version 3 touche deux bâtiments sur la partie gauche de la carte contenant environ 50 personnes. Ces bâtiments sont la source principale de l'écart constaté entre ces deux versions.

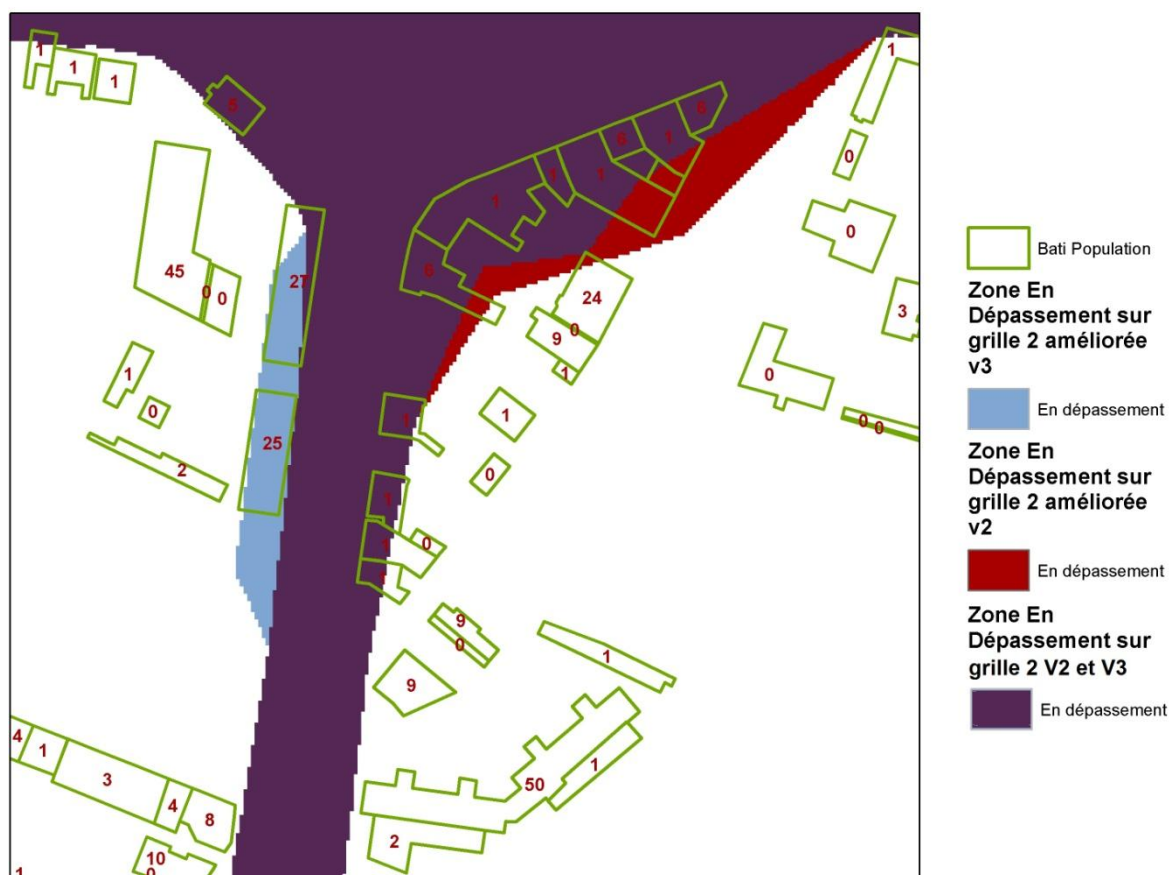


Figure 33 : Comparaison des zones en dépassement pour la grille 2 de la ville de Bourges en version 2 et version 3

3.2.2 Effet de la résolution du maillage de sortie de l'interpolation

Remarque : les conclusions de ce test ne s'appliquent qu'à l'interpolateur linéaire puisque l'interpolation de Delaunay a pour résultat un réseau de polygones.

Un premier test est effectué en dégradant la résolution du maillage de sortie de l'interpolateur linéaire.

La population précédemment répartie sur le maillage de 1m est agrégée sur les grandes mailles avant croisement maille à maille (méthode 2).

Pour rappel, l'interpolateur linéaire estime la concentration au centre de la maille (cf. Partie 3 - 2.1). La dégradation de la résolution augmente légèrement les surfaces en dépassement et les populations exposées (partie gauche du **Tableau 8**) par rapport à la grille 2 améliorée de résolution métrique. La dégradation d'une résolution de 5 mètres à 10 mètres présente peu d'effet sur les zones exposées dans notre cas d'étude.

Tableau 8 : Effet de la résolution du maillage de sortie sur l'exposition de la population.

	Grille de référence	Grille 2 améliorée v3 – résolution après interpolation : 1 m	Grille 2 améliorée v3 - résolution après interpolation : 5 m	Grille 2 améliorée v3 – résolution après interpolation : 10 m
Surface en dépassement (km2)	0.133	0.155	0.155	0.154
Surface en dépassement (%)	3.3%	3.9%	4.1%	3.8%
Population exposée (Nbre Habitants)	180	467	472	480
Population exposée (%)	1.1%	2.9%	3.0%	3.0%

La Figure 34 présente une vue sur un quartier de la ville de Bourges. La zone en dépassement de la simulation de référence est superposée avec celle estimée à partir d'un maillage d'interpolation de 10m de résolution. On peut remarquer sur cette carte que les zones étroites de dépassement ne sont pas nécessairement bien évaluées et les zones plus larges peuvent être surévaluées du fait d'une définition plus grossière. Les zones étroites, principalement sur les axes, couvrent peu de population et ont donc peu d'effet sur les populations exposées si ces zones sont sous-évaluées. A l'inverse, les zones plus larges couvrent des populations plus importantes car elles débordent des voies de circulation. Leur élargissement du fait de la dégradation de la résolution provoque une surestimation des populations exposées.

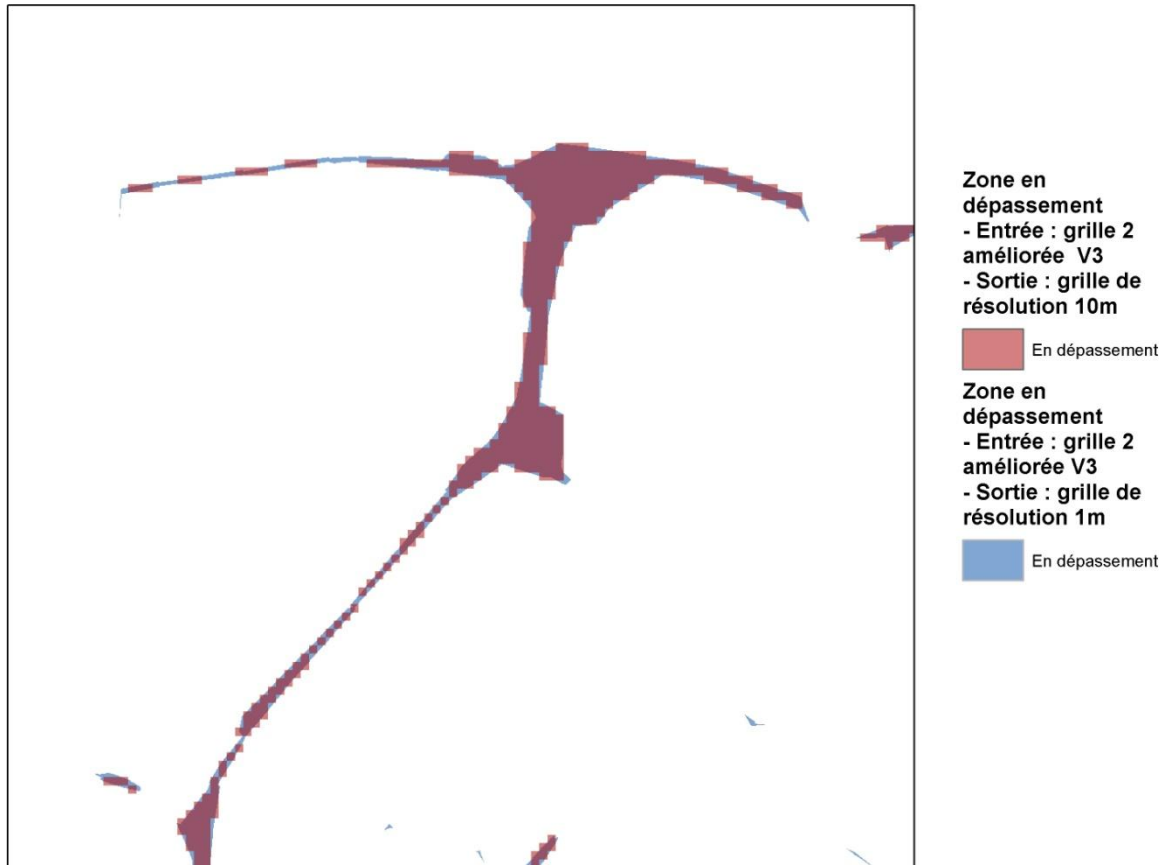


Figure 34 : comparaison des zones en dépassement entre la configuration optimisée et une interpolation sur une grille de 10 m de résolution

3.2.3 Effet du type de prise en compte du bâti

Comme il est expliqué au paragraphe 3.1, deux méthodes sont principalement utilisées par les AASQA pour croiser une zone en dépassement avec les populations résidentes. Une première (méthode 1A) estime que la population exposée d'un bâtiment croisant une zone est déterminée au prorata de la surface de croisement entre celui-ci et la zone en dépassement, elle suppose donc que la population dans un bâtiment est uniformément répartie dans celui-ci. Une seconde (méthode 1B) considère que la population totale de tout bâtiment croisant la zone est exposée. On a vu aussi que le croisement maille à maille (méthode 2, cf. §3.1) pratiqué dans le cas de référence et la situation initiale était en fait un cas particulier de la méthode 1A.

Ces deux méthodes 1A et 1B ont été appliquées sur la simulation HD et la grille n°2 améliorée. Les résultats de la comparaison entre les deux méthodes sont présentés dans le Tableau 9. L'application de la méthode 1B augmente la population exposée de l'ordre de 510 % sur la simulation HD et de l'ordre de 240 % pour la grille 2 améliorée. Le nombre de bâtiments concernés par la simulation HD étant moins important, le calcul d'exposition de la population est donc plus sensible à la méthode employée.

Tableau 9 : Comparaison du calcul d'exposition des populations en fonction du type de prise en compte du bâti (bâtiment entier ou morceau de bâtiment)

	méthode 1A* (grille de référence)	méthode 1B** (grille de référence)	méthode 1A* (grille d'interpolation V3)	méthode 1B** (grille d'interpolation v3)
Population exposée (Nbre Habitants)	180	922	467	1100
Population exposée (%)	1.1%	5.8%	2.9%	6.9%

* méthode 1A (équivalente à la méthode 2) : lorsqu'un bâtiment recoupe une maille en dépassement, sa population est comptabilisée au prorata de la surface d'intersection** méthode 1B : lorsqu'un bâtiment recoupe une maille en dépassement, sa population est prise en compte intégralement.

3.3 Bilan et discussion

Les résultats de la Partie 1 ont montré qu'en l'état actuel des méthodes, il était possible de ne s'intéresser qu'à l'interpolateur linéaire et à l'interpolateur de Delaunay. Les surfaces en dépassement estimées par ces deux interpolateurs présentent peu de différence. Ceci conforte les résultats obtenus dans la Partie 1 et permet d'orienter des tests de sensibilité sur une seule méthode : l'interpolateur linéaire.

L'effet des différentes améliorations du maillage 2 de Bourges sur les surfaces en dépassement et les populations associées a été testé. L'amélioration du maillage permet de réduire les écarts avec la simulation de référence, en particulier :

- une couverture suffisamment homogène des zones de fond limite les surestimations constatées dans ces zones après interpolation;
- perpendiculairement aux axes routiers, les points de modélisation sont disposés de manière à éviter les sous-estimations et surestimations observées respectivement au centre et en bordure des voies.

L'utilisation d'une résolution moins fine (5m ou 10 m) d'un maillage de sortie d'interpolation dégrade légèrement les résultats (moins de 5 % d'écart). Ces résultats concernent la zone de Bourges. Il serait utile de mener des tests complémentaires afin de pouvoir les généraliser et d'évaluer une taille maximale de maille acceptable.

Le dernier test réalisé porte sur la méthode de croisement entre une zone en dépassement et des données de population. Il met en évidence des écarts importants entre les deux méthodes le plus fréquemment employées par les AASQA.

L'application de la règle qui retient la population totale de tout bâtiment croisant la surface en dépassement paraît largement surestimer les populations exposées dans le cas de Bourges. De plus, cette règle n'est pas applicable si l'AASQA choisit, pour faciliter les traitements, de spatialiser les populations sur un réseau maillé. Sans justification complémentaire, cette règle ne paraît pas appropriée sauf éventuellement sur des zones de taille réduite, caractérisées par un habitat de type maison : il paraît alors difficile de découper les quelques maisons concernées en fonction de la surface en dépassement.

CONCLUSION

D'après un bilan récent des pratiques (LCSQA, décembre 2013), la modélisation de la dispersion est l'approche la plus répandue pour délimiter les zones de dépassement à l'échelle urbaine et en déduire le nombre d'habitants exposés. Les dépassements de valeurs limites observés en milieu urbain (PM_{10} , NO_2) le sont généralement à proximité du trafic, dans des zones où les gradients de concentration sont plus marqués. Pour une description précise de la situation de dépassement, une estimation à fine résolution des concentrations et de leur variabilité spatiale est donc requise. Du fait des temps de calcul et des limitations éventuelles des outils de modélisation, il n'est cependant pas toujours possible de réaliser les simulations sur un maillage dense et une étape supplémentaire d'interpolation est alors nécessaire.

Cette étude repose sur l'exploitation de différents jeux de données correspondant à des agglomérations de taille moyenne. Les principales méthodes d'interpolation utilisées dans les AASQA ont été évaluées et comparées. Dans toutes les situations traitées, l'interpolateur linéaire et l'interpolateur de Delaunay produisent des résultats comparables et offrent la meilleure précision. Dans sa configuration usuelle, le krigeage est mal adapté à l'interpolation à partir de données de typologie contrastée. Pour une application satisfaisante du krigeage, le modèle géostatistique doit être modifié (des développements sont en cours).

D'autre part, la qualité de l'interpolation se révèle étroitement liée à l'organisation spatiale du maillage de simulation. Il est délicat de fournir des recommandations précises à ce sujet, la disposition des points de modélisation dépendant de la ville et de sa topographie. Des orientations générales issues des tests effectués sont néanmoins fournies. En complément, les modélisateurs sont invités à ajuster leurs maillages à l'aide de simulations de référence conduites sur des périodes de test. Ils pourront s'inspirer de la démarche adoptée dans cette étude.

Dans les cas traités, l'interpolation a été réalisée sans prise en compte de données de mesure. Actuellement, les données de mesure fixe ou indicative situées dans le domaine d'étude servent principalement au calage des modèles et à l'évaluation des modélisations. Conditionnée par le nombre de points de mesure disponibles et exigeant de ce fait la plus grande vigilance, l'introduction de ces données dans l'élaboration des cartographies reste encore l'objet de travaux exploratoires. Deux exemples issus d'une analyse bibliographique, l'un fondé sur le krigeage multivarié et l'autre sur l'interpolation optimale, sont présentés,

S'agissant du maillage de sortie, c'est-à-dire le maillage sur lequel les concentrations sont interpolées, l'utilisation d'une maille de 1m, 5m ou 10m n'induit pas de différence notable dans les estimations finales (surface de dépassement et nombre d'habitants exposés). L'expérience des AASQA serait utile pour compléter ces résultats et en extraire des préconisations.

La dernière question concerne le mode de croisement entre zones de dépassement et population spatialisée. Différentes méthodes ont été testées. Pour des raisons pratiques, et en accord avec les résultats des tests, un croisement maille à maille, lorsque la population est préalablement répartie sur le maillage d'interpolation, ou, ce qui est équivalent, un croisement au prorata de la surface de bâtiment intersectée est recommandé.

Les conclusions de ce travail seront reprises en 2016 dans une note de synthèse méthodologique et technique sur l'évaluation de la population et des écosystèmes exposés. Les problèmes spécifiquement posés par les agglomérations et zones de grande taille seront examinés en 2016 avec les AASQA concernées.

RÉFÉRENCES

- Akima, H., 1978. A method of bivariate interpolation and smooth surface fitting for irregularly distributed data points. *ACM Trans. on Math. Software* 4: 148-159.
- Babak O., Deutsch C.V., 2008. Two new approaches to avoid large kriging weights to end samples in stringsof data. In: Ortiz JM, Emery X (eds) *Geostats2008, Proceedings of the 8th international geostatistics congress, FCFM—Gecamin*, vol. 1, pp 369–378.
- Baldwin N., Owais Gilani, Suresh Raja, Stuart Batterman, Rajiv Ganguly, Philip Hopke, Veronica Berrocal, Tom Robins, Sarah Hoogterp, 2015. Factors Affecting Pollutant Concentrations in the Near-Road Environment. *Atmospheric Environment*, vol. 115: 223-235.
- Gribov A., Krivoruchko K., 2004. Geostatistical mapping with continuous moving neighbourhood. *MathGeol* 36(2): 267–281.
- Guide IPR, Commissien Européenne, 2013. Guidance on the Commission Implementing Decision laying down rules for Directives 2004/107/EC and 2008/50/EC of the European Parliament and of the Council as regards the reciprocal exchange of information and reporting on ambient air (Decision 2011/850/EU).
Part I: http://ec.europa.eu/environment/air/quality/legislation/pdf/IPR_guidance1.pdf
- Lawson C.L., 1972. Generation of a triangular grid with application to contour plotting. *Tech. Memo. 299, Sect. 914, Jet Propulsion Lab., Caltech, Pasadena, California*.
- Rivoirard J., Romary T., 2011. Continuity for kriging with moving neighborhood. *International association for Mathematical Geosciences*.
- Tilloy A., Mallet V., Poulet D., Pesin C. et Brocheton F., 2013. *BLUE-based NO₂ data assimilation a turban scale*. *J. Geophys. Res. Atmos.*, vol.118, n°4 : 2013-2040 (doi:10.1002/jgrd.50233).

GLOSSAIRE

Abréviations	Libellés	Formule*
K_N1_0	Krigeage ponctuel à N1 voisins	-
K_N1_N2	Krigeage de bloc régulier de N2 mètres de résolution à N1 voisins	-
stdev	écart-type (standard deviation)	$\sigma_x = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$ $\sigma_y = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}$
cor	corrélation	$\frac{\sigma_{xy}}{\sigma_x \sigma_y}$ où σ_{xy} est la covariance entre x et y
RMSE	Racine carrée de l'erreur moyenne quadratique (root-mean-square error)	$\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}$
cRMSE	Racine carrée de l'erreur moyenne quadratique centrée (centered RMSE)	$\sqrt{\frac{1}{n} \sum_{i=1}^n ((x_i - \bar{x}) - (y_i - \bar{y}))^2}$
nmb	Biais moyen normalisé (normalized mean bias)	$\frac{(\bar{x} - \bar{y})}{\bar{x}}$
nme	Erreur moyenne normalisée (normalized mean error)	$\frac{ \bar{x} - \bar{y} }{\bar{x}}$

* dans le développement des formules de calcul des scores, **x**=($x_1, \dots, x_n$) désigne les données de modélisation et **y**=($y_1, \dots, y_n$) les valeurs interpolées.

LISTE DES ANNEXES

Annexes	titres
Annexe 1	Tables de scores en validation croisée
Annexe 2	Nuages de points (validation croisée)
Annexe 3	Cartographie des écarts, avec l'interpolateur linéaire et l'interpolateur de Delaunay (validation croisée)
Annexe 4	Nuages de points (sur grille de référence)
Annexe 5	Cartographie des écarts, avec l'interpolateur linéaire et l'interpolateur de Delaunay (sur grille de référence)

ANNEXES

ANNEXE 1 : TABLES DE SCORES EN VALIDATION CROISEE

Tableau 10 : Scores en validation croisée sur la fenêtre de zoom de Nantes

méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Delaunay	42.3370001363558	0.950966165655905	13.4063790786327	13.1969314615179	4.48	15.03
krigeage_nmax_2_res_0	42.968820434935	0.917334140002224	17.7759669216614	17.2006976602695	7.61	20.13
krigeage_nmax_2_res_1	43.5430167014457	0.879710170916868	21.3339869691532	20.909543275739	7.18	23.01
krigeage_nmax_2_res_5	42.8327123691582	0.884323749453727	20.7128703949984	20.2782395685622	7.16	22.96
krigeage_nmax_2_res_10	41.1619824698759	0.880950065860365	20.4475181868118	20.1116881439625	6.26	23.06
krigeage_nmax_2_res_20	41.5040426526216	0.892752943104097	19.4277636386693	19.168046327156	5.37	21.86
krigeage_nmax_2_res_50	38.0997218061171	0.843304146520713	22.4313842640172	22.430871583601	-0.26	24.06
krigeage_nmax_3_res_0	40.7802839387885	0.918619048090151	16.7125701195238	16.5593780076427	3.83	20.04
krigeage_nmax_3_res_1	41.4087238475776	0.879119021337476	20.4483984702959	20.3258948142511	3.79	22.82
krigeage_nmax_3_res_5	41.0181984393969	0.882771659936715	20.0725932554815	19.9240558120078	4.13	22.95
krigeage_nmax_3_res_10	39.6829565405351	0.887618280826656	19.3590805566793	19.2555965890439	3.39	22.38
krigeage_nmax_3_res_20	39.6548023366194	0.897809048127666	18.4216720196621	18.3645582576876	2.46	21.52
krigeage_nmax_3_res_50	37.7357612513919	0.849411589785518	21.9641526128371	21.9520614066197	-1.24	24.16
krigeage_nmax_4_res_0	39.9139154176386	0.923503169359813	15.999437490112	15.9373460776881	2.39	19
krigeage_nmax_4_res_1	40.4257297983342	0.884396920603303	19.7011420988734	19.6603407905357	2.15	21.69
krigeage_nmax_4_res_5	40.0328247872271	0.888537474304679	19.2915525554581	19.2343442830786	2.52	21.79
krigeage_nmax_4_res_10	38.7664707128845	0.894662285050764	18.566529024026	18.5360999134122	1.8	21.27
krigeage_nmax_4_res_20	38.5899339200495	0.903652758494463	17.7374744538223	17.7299464184188	0.88	20.55
krigeage_nmax_4_res_50	36.9933636829792	0.855896720431357	21.466252583998	21.4181932011083	-2.43	23.07
krigeage_nmax_5_res_0	39.874121184141	0.925532159035202	15.794081169856	15.7220863755419	2.55	18.62
krigeage_nmax_5_res_1	40.3016733653982	0.886970851435879	19.4570038803512	19.4170543595057	2.11	21.24
krigeage_nmax_5_res_5	39.9397273291795	0.891469290708694	19.0196477359598	18.9654422569051	2.43	21.35
krigeage_nmax_5_res_10	38.8261519082958	0.896585502127531	18.4153468607029	18.3750102040788	2.07	21.03
krigeage_nmax_5_res_20	38.490771958394	0.905045075291043	17.6038348094953	17.5980396635534	0.77	20.3
krigeage_nmax_5_res_50	37.0590685787315	0.859709822663887	21.1808167925602	21.1493025889744	-1.96	22.62
krigeage_nmax_10_res_0	39.6151069815177	0.928834504323321	15.4161279185144	15.3496905506267	2.42	18.28
krigeage_nmax_10_res_1	40.0007935510825	0.890271528161583	19.1135030802833	19.0794129888736	1.93	20.89
krigeage_nmax_10_res_5	39.6781190966942	0.89495197069042	18.6744210084275	18.6207679755696	2.4	21.02
krigeage_nmax_10_res_10	38.5630534282799	0.899770314295607	18.113310023295	18.073239886639	2.04	20.8
krigeage_nmax_10_res_20	38.2226039702164	0.908957228818592	17.2325854125259	17.2289872018061	0.6	19.95
krigeage_nmax_10_res_50	36.5450108071496	0.863583226074703	20.8973443288854	20.8399376198683	-2.62	22.3
Linéaire	42.3137654135273	0.949863742160856	13.5662448746881	13.2721136221779	4.78	15.09

Tableau 11 : Scores en validation croisée sur la fenêtre de zoom de Niort

Méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Delaunay	4.45831163290719	0.909487280922257	1.9715476154534	1.96596032513375	0.85	5.41
krigeage_nmax_2_res_0	4.85069760555389	0.845230745998331	2.72635287517959	2.65894716006167	3.03	7.62
krigeage_nmax_2_res_1	4.84325679759721	0.807305490014447	3.01877458582121	2.96327521502813	2.9	8.23
krigeage_nmax_2_res_5	5.28960531930656	0.391645124440559	5.98230724720822	5.52765773180648	11.5	21.13
krigeage_nmax_2_res_10	4.71956957221223	0.756298954149599	3.31089111871713	3.28542234727896	2.06	8.3
krigeage_nmax_2_res_20	4.49010259246423	0.723379690951143	3.42607647316869	3.42008771817332	1.02	8.92
krigeage_nmax_2_res_50	4.04844686785756	0.588569031937395	4.00574587311776	4.00574587311776	-0.08	10.94
krigeage_nmax_3_res_0	4.51971256465473	0.881009519509684	2.29216055284092	2.25321991825032	2.12	6.55
krigeage_nmax_3_res_1	4.5051453663829	0.838813599325064	2.64575131106459	2.61725046566048	1.95	7.22
krigeage_nmax_3_res_5	5.0065859979594	0.424129254589725	5.61569229926284	5.21104596026556	10.52	20.02
krigeage_nmax_3_res_10	4.45077110017352	0.774866104977535	3.09160152671718	3.07603641070778	1.56	7.82
krigeage_nmax_3_res_20	4.3716824654018	0.733963912801465	3.32761175619993	3.31918664735805	1.19	8.88
krigeage_nmax_3_res_50	4.01754365759702	0.601166180125191	3.93700393700591	3.93598780485916	0.43	10.89
krigeage_nmax_4_res_0	4.3206076551216	0.901478917147441	2.05888319241282	2.03297811104793	1.64	6.02
krigeage_nmax_4_res_1	4.33740628913117	0.853959069795551	2.48394846967484	2.46373699895098	1.58	6.86
krigeage_nmax_4_res_5	4.92730375788473	0.428462679204009	5.55688761808263	5.14616361962968	10.54	20.04
krigeage_nmax_4_res_10	4.30763977749953	0.777920359348438	3.03315017762062	3.02076149339864	1.37	7.96
krigeage_nmax_4_res_20	4.24007503283032	0.73622400939674	3.27963412593539	3.27108544675922	1.19	8.99
krigeage_nmax_4_res_50	3.87716146979779	0.611316014357094	3.8501948002666	3.84798648646276	0.67	10.86
krigeage_nmax_5_res_0	4.18364654100945	0.89889199136872	2.0820662813657	2.05621010599598	1.64	6.47
krigeage_nmax_5_res_1	4.22166590742673	0.85099652305639	2.49519538313135	2.47487373415292	1.6	7.27
krigeage_nmax_5_res_5	4.86783153093815	0.414102858187339	5.56336229271472	5.17638870256089	10.25	20.15
krigeage_nmax_5_res_10	4.20180660836333	0.780412921642032	2.99382698230876	2.98311917294633	1.27	8.09
krigeage_nmax_5_res_20	4.14654777598706	0.737857310004057	3.24946149384786	3.2402160421799	1.24	9.09
krigeage_nmax_5_res_50	3.81659794366619	0.618904305351877	3.80092094103521	3.79697247817258	0.88	10.9
krigeage_nmax_10_res_0	3.87979707151706	0.823115782842647	2.69685001436862	2.66477015894429	2.09	9.31
krigeage_nmax_10_res_1	3.8929396857554	0.780559493163366	2.97119504576862	2.94210808774933	2.09	9.9
krigeage_nmax_10_res_5	4.36895740491771	0.437902705110862	4.99479729318418	4.81144468948776	6.74	18.7
krigeage_nmax_10_res_10	3.88232558904106	0.754109642746586	3.12489999839995	3.10080634674273	1.95	10.08
krigeage_nmax_10_res_20	3.86460687508805	0.734960656934897	3.2335738742141	3.20889389042393	2	10.36
krigeage_nmax_10_res_50	3.59595400119909	0.632778912612235	3.70189140845595	3.68700420395746	1.66	11.47
Linéaire	4.39370837417366	0.91689878133713	1.88015956769632	1.87456661658102	0.73	4.8

Tableau 12 : Scores en validation croisée sur la fenêtre de zoom de Reims

Méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Delaunay	7.08146693441742	0.96494300372095	1.94010309004444	1.93313217344288	0.62	4.73
krigeage_nmax_2_res_0	7.31616148442823	0.937081896916259	2.67805899860328	2.60326717799	2.18	6.25
krigeage_nmax_2_res_1	7.3299629277675	0.774808882488333	4.96397018524487	4.9288943993557	2.04	9.99
krigeage_nmax_2_res_5	7.25252679338443	0.771166047618175	4.96970824093326	4.94337941088887	1.77	10.02
krigeage_nmax_2_res_10	7.27340066759207	0.78314432140136	4.84778299844372	4.81881728227996	1.84	9.9
krigeage_nmax_2_res_20	7.24655739782297	0.768935461802957	4.98056221726022	4.96548084277847	1.34	10.16
krigeage_nmax_2_res_50	7.01789463649574	0.728402393780181	5.30933140800233	5.30744759748036	0.5	10.96
krigeage_nmax_3_res_0	7.28385681690547	0.9373707720034	2.61725046566048	2.59229627936314	1.25	6.16
krigeage_nmax_3_res_1	7.26655861196904	0.775748506248039	4.91365444450462	4.89806084078179	1.36	9.89
krigeage_nmax_3_res_5	7.22549591760051	0.772880008355992	4.92767693746252	4.91629942131274	1.16	9.92
krigeage_nmax_3_res_10	7.27781589492302	0.785699503239739	4.80749414976243	4.7917637671321	1.35	9.77
krigeage_nmax_3_res_20	7.24756262510302	0.77264328653018	4.93315314986267	4.92585018042571	0.93	10.04
krigeage_nmax_3_res_50	7.03982694010465	0.738275245964469	5.22216430227928	5.21718314802155	0.79	10.78
krigeage_nmax_4_res_0	7.21953930486559	0.938367854450383	2.57740179250345	2.56281095674262	0.95	5.85
krigeage_nmax_4_res_1	7.21684668774048	0.777250885271212	4.87473076179598	4.86610727378672	1.01	9.62
krigeage_nmax_4_res_5	7.18733429953472	0.771911384861409	4.92097551304617	4.91497711083175	0.85	9.69
krigeage_nmax_4_res_10	7.25046538435921	0.784683574894003	4.8039567025526	4.7945802736006	1.04	9.56
krigeage_nmax_4_res_20	7.20643739109371	0.771125632232184	4.93224087003058	4.92930015316576	0.59	9.88
krigeage_nmax_4_res_50	7.02480767568422	0.743283199431813	5.16613975807856	5.16265435604593	0.65	10.59
krigeage_nmax_5_res_0	7.16661923175302	0.939344113871984	2.55244980362004	2.53673017879316	0.99	5.79
krigeage_nmax_5_res_1	7.16526255803113	0.777379533274952	4.85746847648032	4.84912363216283	0.99	9.6
krigeage_nmax_5_res_5	7.14427652156837	0.77186103618946	4.90897137901618	4.90244836790761	0.88	9.7
krigeage_nmax_5_res_10	7.1888080093872	0.785350445641297	4.77776098188262	4.7685427543433	1.03	9.52
krigeage_nmax_5_res_20	7.1294384429553	0.771780336453071	4.90173438692877	4.89887742243057	0.58	9.83
krigeage_nmax_5_res_50	6.98892130115959	0.747566068568616	5.11370707021824	5.10901164610143	0.76	10.5
krigeage_nmax_10_res_0	6.95870551834207	0.939183870279932	2.53692727526825	2.52764712727074	0.76	6.01
krigeage_nmax_10_res_1	6.97515829962563	0.778407058840658	4.79061582680139	4.78497648897045	0.81	9.68
krigeage_nmax_10_res_5	6.93278314135127	0.77330723537835	4.83166637921121	4.82825020064205	0.63	9.74
krigeage_nmax_10_res_10	6.97552199353668	0.786144963757524	4.7060599231204	4.70127642241977	0.74	9.51
krigeage_nmax_10_res_20	6.91443256631832	0.773981260931075	4.81715683780381	4.81643021334266	0.3	9.78
krigeage_nmax_10_res_50	6.85927709295602	0.751970030054737	5.03199761526176	5.02881695829148	0.62	10.37
Linéaire	7.06785407186623	0.965328676333744	1.93287350853593	1.92223827867411	0.7	4.69

Tableau 13 : Scores en validation croisée sur la fenêtre de zoom de Bourges

Méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Delaunay	2.03553797741438	0.784441502527676	1.45808093053849	1.45808093053849	0.01	4.62
krigeage_nmax_2_res_0	2.04149722849953	0.750937502059571	1.56108936323325	1.56076904121013	0.15	4.54
krigeage_nmax_2_res_1	2.01654351065677	0.772428028537263	1.49030198282093	1.49030198282093	0.04	4.3
krigeage_nmax_2_res_5	2.01647361755369	0.772412542526692	1.49030198282093	1.49030198282093	0.04	4.3
krigeage_nmax_2_res_10	2.01657287598046	0.772420203232192	1.49030198282093	1.49030198282093	0.04	4.3
krigeage_nmax_2_res_20	2.07048929866287	0.754385081934187	1.55627761019684	1.55595629758679	0.13	4.82
krigeage_nmax_2_res_50	2.01715492959885	0.772378105016564	1.49063744753713	1.49063744753713	0.04	4.3
krigeage_nmax_3_res_0	1.95426671656796	0.812639129481491	1.35351394525509	1.35351394525509	0.08	3.86
krigeage_nmax_3_res_1	1.93385585230349	0.819334868898228	1.33041346956501	1.33041346956501	0	3.79
krigeage_nmax_3_res_5	1.93385520780317	0.81933475693998	1.33041346956501	1.33041346956501	0	3.79
krigeage_nmax_3_res_10	1.93391651174542	0.819334410262795	1.33041346956501	1.33041346956501	0	3.79
krigeage_nmax_3_res_20	2.00444538130382	0.781335760153344	1.46047937335657	1.46047937335657	0.13	4.48
krigeage_nmax_3_res_50	1.93437537090115	0.819102032794546	1.3311649033835	1.3311649033835	0.01	3.79
krigeage_nmax_4_res_0	1.92344671769624	0.845564138503166	1.23895116933639	1.23895116933639	0.06	3.55
krigeage_nmax_4_res_1	1.91808259317322	0.842764648986934	1.24939985593084	1.24939985593084	0.03	3.56
krigeage_nmax_4_res_5	1.91808046474269	0.842764424520171	1.24939985593084	1.24939985593084	0.03	3.56
krigeage_nmax_4_res_10	1.91811205131688	0.842762975730563	1.24939985593084	1.24939985593084	0.03	3.56
krigeage_nmax_4_res_20	1.94046090440965	0.796440750361732	1.40605832026982	1.40570267126445	0.12	4.21
krigeage_nmax_4_res_50	1.91854772682731	0.84261659866769	1.24979998399744	1.24979998399744	0.03	3.57
krigeage_nmax_5_res_0	1.90543407323651	0.832439368871246	1.28569047597001	1.28569047597001	0.08	3.74
krigeage_nmax_5_res_1	1.89122845901824	0.832994403334124	1.28413394939936	1.28413394939936	0.04	3.73
krigeage_nmax_5_res_5	1.89116778414197	0.832987496085058	1.28413394939936	1.28413394939936	0.04	3.73
krigeage_nmax_5_res_10	1.89127691712705	0.832969021201757	1.28413394939936	1.28413394939936	0.04	3.73
krigeage_nmax_5_res_20	1.92158901231948	0.803253148172645	1.38311243216161	1.38275088139549	0.1	4.21
krigeage_nmax_5_res_50	1.8928081883583	0.832414292476366	1.28607931326182	1.28607931326182	0.04	3.74
krigeage_nmax_10_res_0	1.82436452163367	0.819140143733181	1.33266649991661	1.33266649991661	0.02	3.99
krigeage_nmax_10_res_1	1.81967584053076	0.815566479449571	1.34424700111252	1.34424700111252	-0.01	4
krigeage_nmax_10_res_5	1.81965630114737	0.815559325543702	1.34424700111252	1.34424700111252	-0.01	4
krigeage_nmax_10_res_10	1.82023562586106	0.815446673581006	1.34461890511773	1.34461890511773	-0.01	4.01
krigeage_nmax_10_res_20	1.83247489677154	0.799907941706147	1.39247980236699	1.39247980236699	0.07	4.3
krigeage_nmax_10_res_50	1.81980384583183	0.814221654241991	1.34833230325465	1.34833230325465	-0.01	4.02
Linéaire	2.01825168944516	0.792366312139766	1.42618371888057	1.42618371888057	-0.06	3.65

Tableau 14 : Scores en validation croisée sur la fenêtre de zoom d'Orléans

Méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Delaunay	6.28543235259779	0.91430509109349	2.73550726557251	2.73532447800988	0.13	5.34
krigeage_nmax_2_res_0	6.28167672987117	0.878303710413184	3.25760648329414	3.23094413445977	1.22	6.9
krigeage_nmax_2_res_1	6.38531985185266	0.788175181964297	4.29965114863985	4.27200187265877	1.43	8.82
krigeage_nmax_2_res_5	6.34230627520603	0.793756764922118	4.21698944746131	4.20511593181448	0.93	8.52
krigeage_nmax_2_res_10	6.20938684678113	0.778750987820186	4.33900910347051	4.32203655699486	1.13	8.88
krigeage_nmax_2_res_20	6.35363035912989	0.793477751882014	4.2203080456289	4.21058190752775	0.84	8.9
krigeage_nmax_2_res_50	6.20797235805165	0.775534871457596	4.35660418215839	4.35270031130102	0.55	9.31
krigeage_nmax_3_res_0	6.19895403049056	0.883622169195555	3.19374388453426	3.15277655408689	1.5	6.98
krigeage_nmax_3_res_1	6.28858145443686	0.794543625224533	4.21746369278978	4.18449519058154	1.55	8.8
krigeage_nmax_3_res_5	6.2524733234101	0.795161999975986	4.18748134324202	4.17025179095939	1.12	8.68
krigeage_nmax_3_res_10	6.11984569318135	0.788752246707307	4.23060279392902	4.20618592076004	1.34	8.81
krigeage_nmax_3_res_20	6.2587137946551	0.802287033940459	4.11630902630014	4.09926822737913	1.1	8.82
krigeage_nmax_3_res_50	6.11397769346652	0.782577687050407	4.27188482990822	4.26473914794328	0.73	9.19
krigeage_nmax_4_res_0	6.16082554000272	0.88744434465489	3.14197390186488	3.09919344346235	1.52	6.89
krigeage_nmax_4_res_1	6.25934655529892	0.797199469278607	4.18915265895145	4.15114442051827	1.66	8.76
krigeage_nmax_4_res_5	6.21358885029731	0.800613839655296	4.12565146370849	4.10694533686535	1.16	8.58
krigeage_nmax_4_res_10	6.07056276523298	0.795703828142761	4.15367307331716	4.12880127882174	1.34	8.67
krigeage_nmax_4_res_20	6.23437436422035	0.803134791495992	4.1047533421632	4.0854620301748	1.17	8.74
krigeage_nmax_4_res_50	6.07788567675603	0.786757351257032	4.22587269093616	4.21746369278978	0.78	9.14
krigeage_nmax_5_res_0	6.10371795808202	0.890593360776054	3.09014562763634	3.05401375242483	1.39	6.87
krigeage_nmax_5_res_1	6.21138749089779	0.797570825720935	4.16833300013327	4.13714877663349	1.5	8.76
krigeage_nmax_5_res_5	6.17273051791819	0.804062057878851	4.07970587175105	4.06374211780226	1.06	8.5
krigeage_nmax_5_res_10	6.02711523215123	0.796255891621573	4.13847798109402	4.11582312545134	1.27	8.66
krigeage_nmax_5_res_20	6.20881797820415	0.809984465172091	4.03088079704672	4.0097381460639	1.22	8.62
krigeage_nmax_5_res_50	6.08296114577042	0.79180659481969	4.17684091150238	4.16929250592951	0.74	9.03
krigeage_nmax_10_res_0	6.02984883302006	0.897935653687833	2.97254100055828	2.95296461204668	1	6.69
krigeage_nmax_10_res_1	6.12976791224873	0.805064414459561	4.06337298312621	4.04561490999823	1.12	8.57
krigeage_nmax_10_res_5	6.10652608971615	0.811518267443519	3.98421886948998	3.97555027637684	0.78	8.33
krigeage_nmax_10_res_10	5.92241744177912	0.804078830394403	4.03583944180142	4.02330709740134	0.94	8.44
krigeage_nmax_10_res_20	6.14119104064377	0.814365725711428	3.96282727355104	3.95183501679916	0.87	8.42
krigeage_nmax_10_res_50	5.99808741330472	0.804566029615518	4.03422855078886	4.02901973189509	0.6	8.74
Linéaire	6.2762130715991	0.912988639898238	2.74335560946808	2.74317334486904	0.13	5.31

Tableau 15 : Scores en validation croisée sur la fenêtre de zoom de Tours

Méthode	stdev	cor	RMSE	cRMSE	nmb	nme
Delaunay	9.09641609265436	0.935001677423302	3.43336569564036	3.4348216838724	0.41	6.42
krigeage_nmax_2_res_0	9.30126766375614	0.902124912213015	4.29231872069165	4.20107129194447	2.38	8.58
krigeage_nmax_2_res_1	9.03695533486089	0.829727600034354	5.66453881617913	5.47731686138387	3.9	11.07
krigeage_nmax_2_res_5	9.08445510530955	0.827300716811767	5.68612345979227	5.52512443298791	3.63	11.13
krigeage_nmax_2_res_10	9.3974998851268	0.8705390464887	4.91334916324904	4.84685465018294	2.17	9.83
krigeage_nmax_2_res_20	9.29660276272443	0.847447817758786	5.30103763427501	5.23774760751222	2.2	10.48
krigeage_nmax_2_res_50	9.21361882864172	0.851800797629754	5.16459098090062	5.14606645895678	1.18	10.5
krigeage_nmax_3_res_0	9.18240065087964	0.899225167830263	4.34039168739412	4.24652799354955	2.42	8.8
krigeage_nmax_3_res_1	9.03990959406111	0.832835853598748	5.62316636780382	5.4282593895281	3.96	11.11
krigeage_nmax_3_res_5	9.01199374018727	0.836561454133255	5.51479827373586	5.36320799522077	3.47	10.98
krigeage_nmax_3_res_10	9.20422542236929	0.871361013342403	4.86723740945518	4.79531020894373	2.25	9.88
krigeage_nmax_3_res_20	9.2647385632078	0.853122489116782	5.20326820373503	5.1333225108111	2.3	10.19
krigeage_nmax_3_res_50	9.13145848148177	0.854644929116885	5.10822865580624	5.08173198821032	1.4	10.37
krigeage_nmax_4_res_0	9.09390276061288	0.902582068882836	4.26919196101557	4.16677333196803	2.51	8.75
krigeage_nmax_4_res_1	8.96009179840928	0.837439820375976	5.54905397342646	5.34013108453341	4.07	10.96
krigeage_nmax_4_res_5	8.9491133011599	0.841671611131821	5.42300654618819	5.26953508385702	3.46	10.91
krigeage_nmax_4_res_10	9.20682236300878	0.872332613886344	4.85613014652614	4.77776098188262	2.35	9.88
krigeage_nmax_4_res_20	9.18546235857119	0.860132416703147	5.0649777886976	4.9951976937855	2.26	10.05
krigeage_nmax_4_res_50	8.95736229853373	0.861121999180544	4.97121715478212	4.94226668645066	1.45	10.25
krigeage_nmax_5_res_0	9.03861801176988	0.903916730507184	4.22516271876007	4.13388437187109	2.36	8.72
krigeage_nmax_5_res_1	8.87325125664697	0.843035507407971	5.43626710160566	5.23583804180381	3.95	10.85
krigeage_nmax_5_res_5	8.86301926193207	0.843428513728804	5.37196425900247	5.22800153022166	3.33	10.86
krigeage_nmax_5_res_10	9.08412042255721	0.875448853377851	4.77577218887166	4.70127642241977	2.27	9.8
krigeage_nmax_5_res_20	9.18139714355992	0.858473479382817	5.09519381378177	5.02374362403178	2.29	10.06
krigeage_nmax_5_res_50	8.95399599063765	0.868191283576902	4.85592421687159	4.81694924199955	1.66	10.05
krigeage_nmax_10_res_0	8.82355231217929	0.902220647905057	4.19964284195692	4.15740303554996	1.6	8.71
krigeage_nmax_10_res_1	8.67290374033524	0.841843478219622	5.36591091987185	5.23020076096511	3.24	10.79
krigeage_nmax_10_res_5	8.63775324733118	0.843187690925022	5.29915087537617	5.20509365909971	2.68	10.8
krigeage_nmax_10_res_10	8.80406354345791	0.877176041809264	4.67835441154259	4.63971981912701	1.62	9.72
krigeage_nmax_10_res_20	8.90171573466629	0.864777011734036	4.90836021498015	4.87134478352744	1.63	9.89
krigeage_nmax_10_res_50	8.61038314416976	0.873984607206647	4.69659451091959	4.68604310692934	0.85	9.71
Linéaire	9.07685267848885	0.935648139732728	3.40690475358499	3.40602407507639	0.38	6.34

ANNEXE 2 : NUAGES DE POINTS (VALIDATION CROISEE)

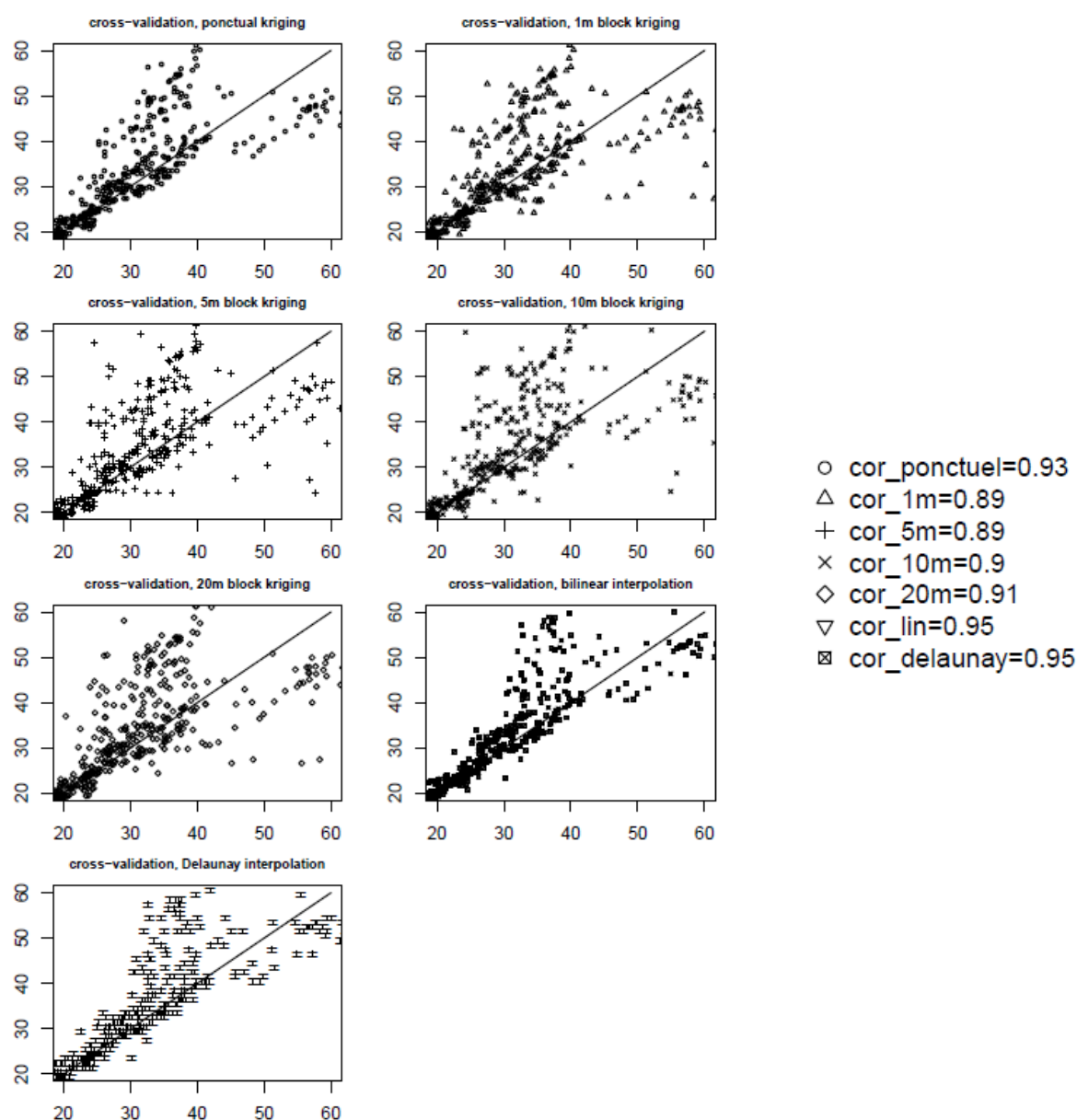


Figure 35 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom de Nantes (Delaunay, linéaire, krigeage ponctuel à 5 voisins et krigeage de bloc à 5 voisins sur les grilles de 1m, 5m, 10m et 20m de résolution).

Abcisses : données de modélisation du NO_2 en $\mu\text{g.m}^{-3}$

Ordonnées : valeurs interpolées en validation croisée

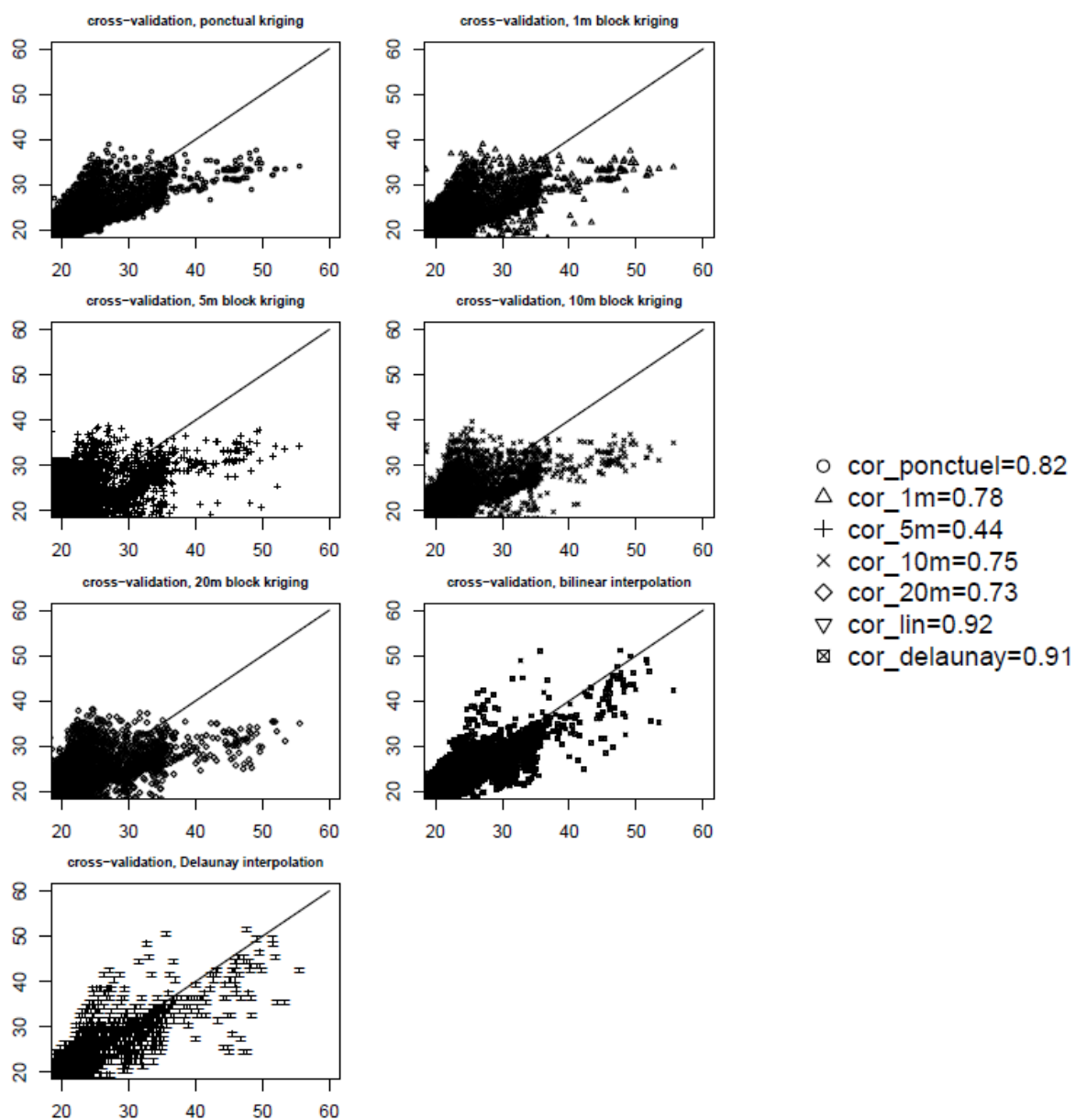


Figure 36 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom de Niort (Delaunay, linéaire, krigeage ponctuel à 5 voisins et krigeage de bloc à 5 voisins sur les grilles de 1m, 5m, 10m et 20m de résolution).

Abscisses : données de modélisation du NO_2 en $\mu\text{g.m}^{-3}$

Ordonnées : valeurs interpolées en validation croisée

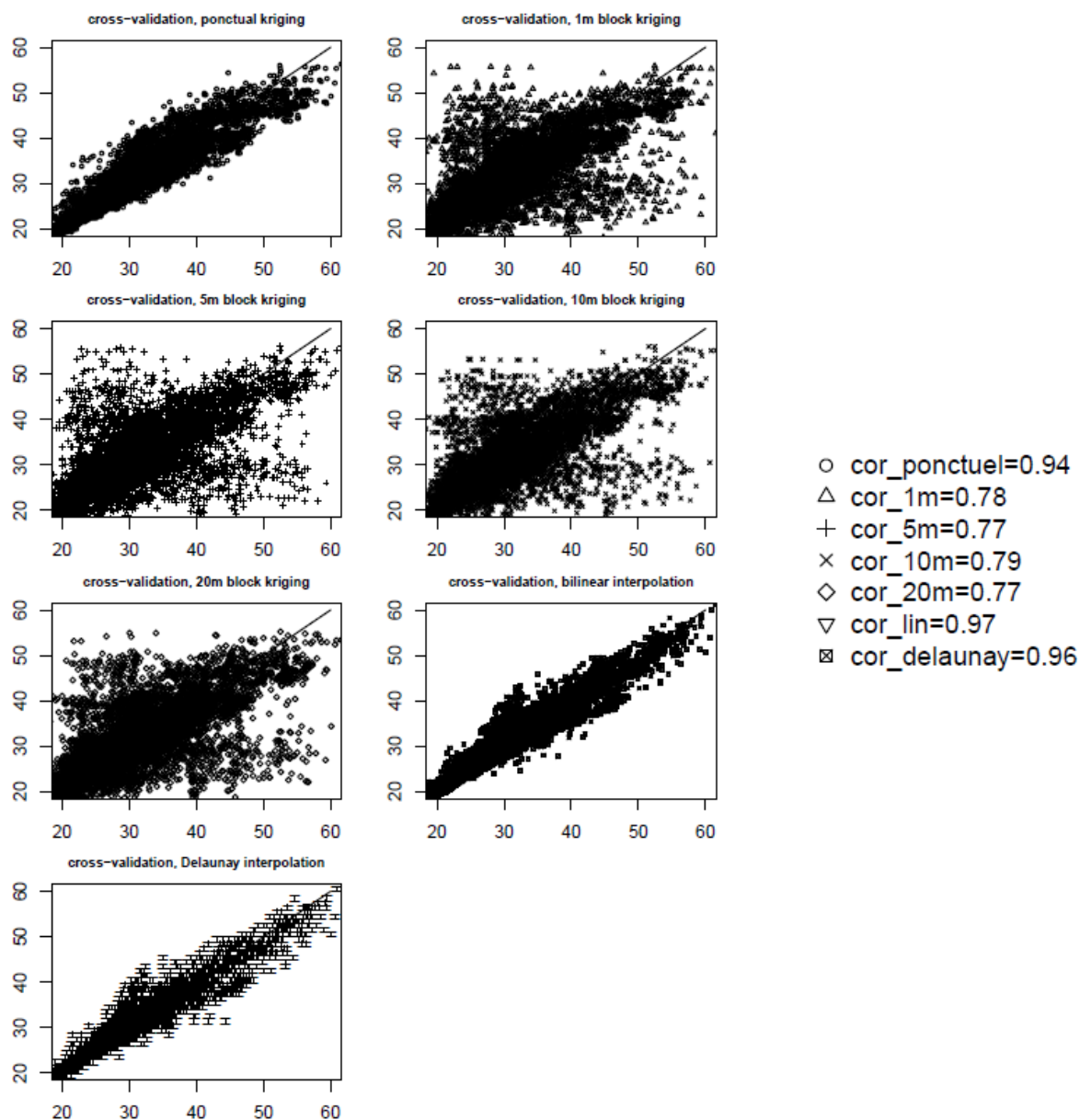


Figure 37 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom de Reims (Delaunay, linéaire, krigeage ponctuel à 5 voisins et krigeage de bloc à 5 voisins sur les grilles de 1m, 5m, 10m et 20m de résolution).

Abcisses : données de modélisation du NO_2 en $\mu\text{g.m}^{-3}$

Ordonnées : valeurs interpolées en validation croisée

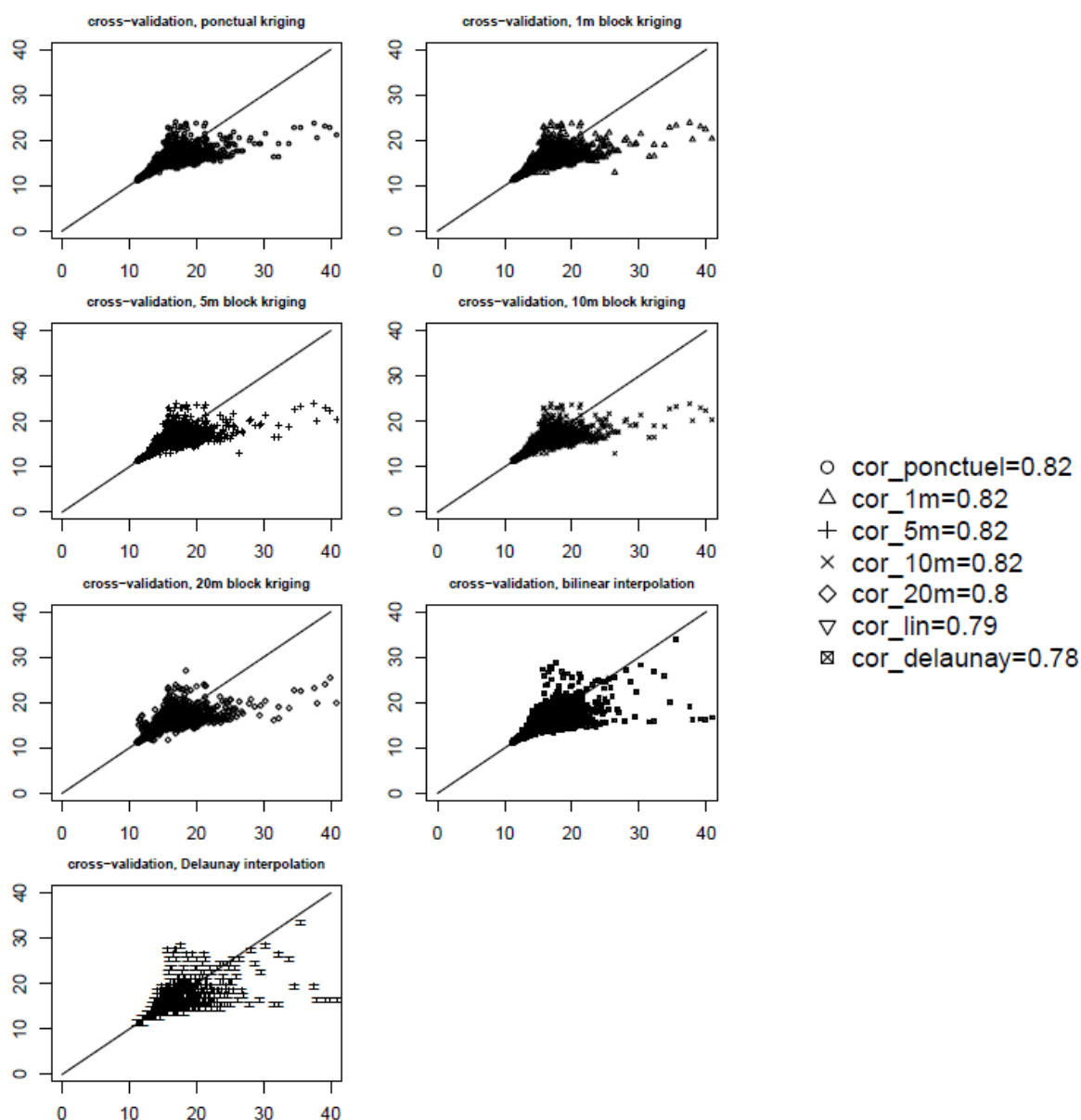


Figure 38 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom de Bourges - maillage n°3 de PREVISION'AIR - (Delaunay, linéaire, krigeage ponctuel à 5 voisins et krigeage de bloc à 5 voisins sur les grilles de 1m, 5m, 10m et 20m de résolution).

Abscisses : données de modélisation du NO_2 en $\mu\text{g.m}^{-3}$

Ordonnées : valeurs interpolées en validation croisée

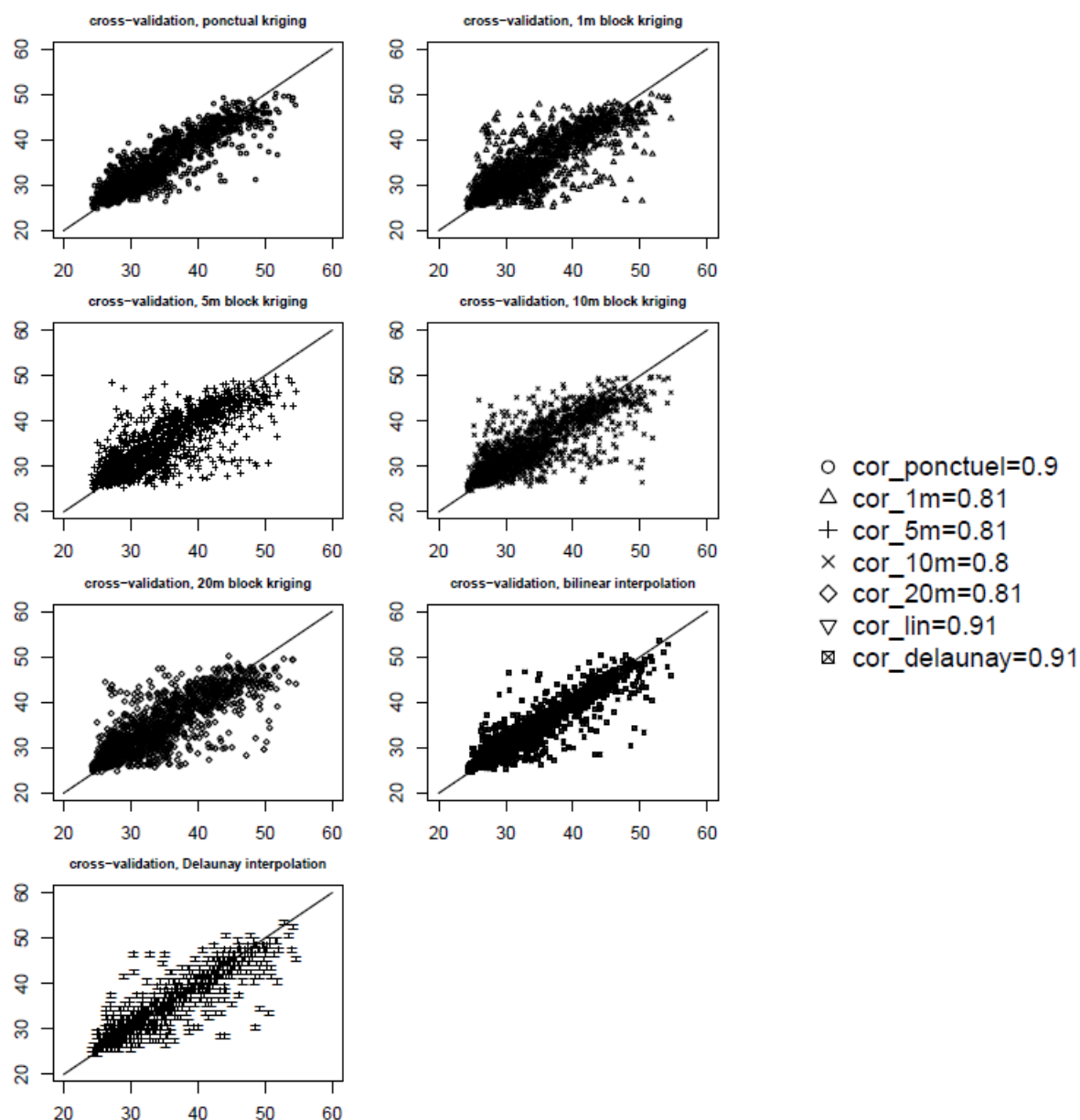


Figure 39 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom d'Orléans (Delaunay, linéaire, krigeage ponctuel à 5 voisins et krigeage de bloc à 5 voisins sur les grilles de 1m, 5m, 10m et 20m de résolution).

Abcisses : données de modélisation du NO_2 en $\mu\text{g.m}^{-3}$

Ordonnées : valeurs interpolées en validation croisée

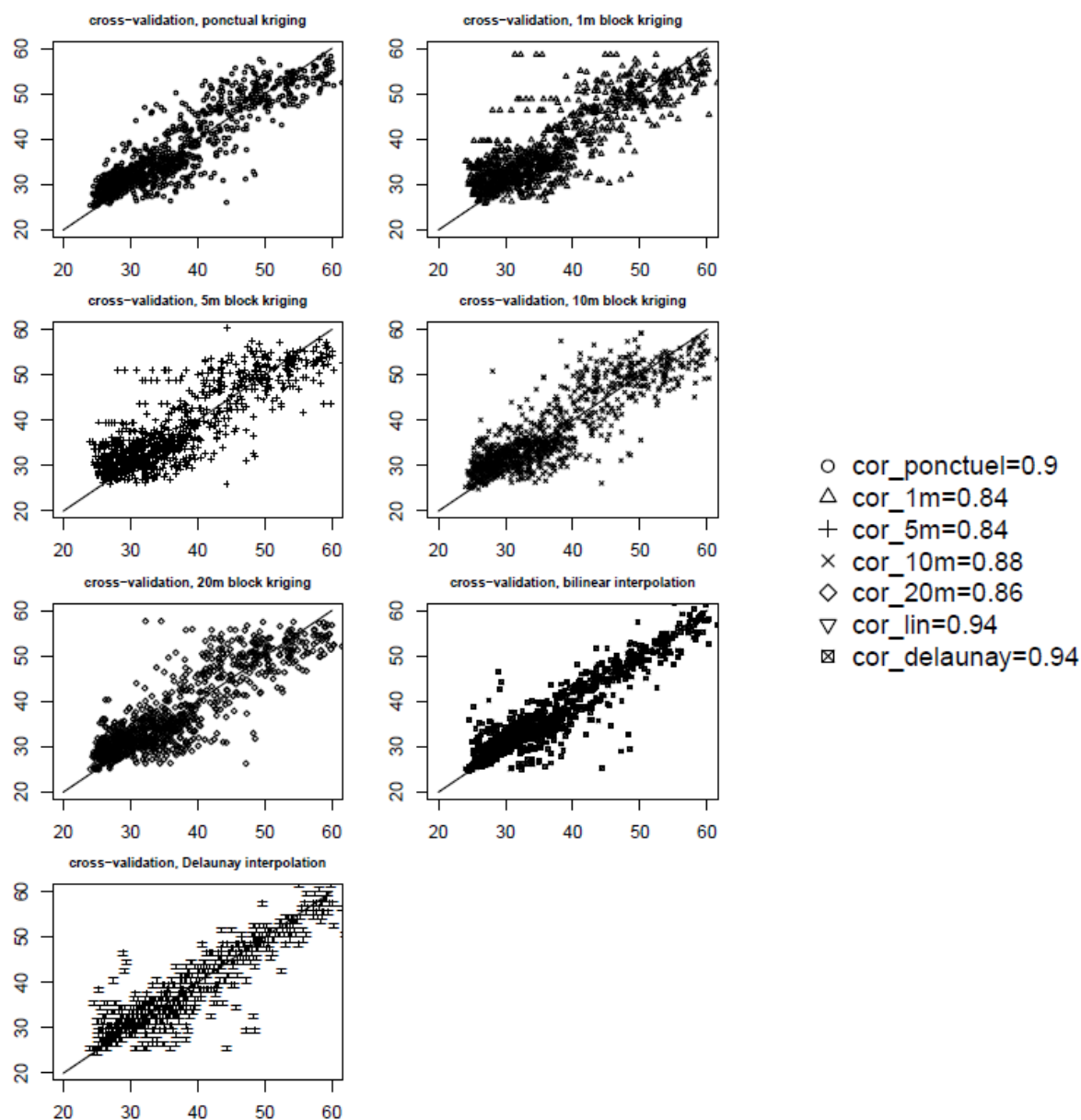


Figure 40 : Nuages de points (données de modélisation et valeurs interpolées) sur la fenêtre de zoom de Tours (Delaunay, linéaire, krigeage ponctuel à 5 voisins et krigeage de bloc à 5 voisins sur les grilles de 1m, 5m, 10m et 20m de résolution).

Abscisses : données de modélisation du NO_2 en $\mu\text{g.m}^{-3}$

Ordonnées : valeurs interpolées en validation croisée

ANNEXE 3 : CARTOGRAPHIE DES ECARTS AVEC LES INTERPOLATEURS LINEAIRES ET DE DELAUNAY (VALIDATION CROISEE)

Dans cette annexe, l'écart est calculé comme la différence entre la valeur interpolée et la donnée de modélisation (valeur interpolée – données de modélisation): une valeur négative (resp. positive) de l'écart caractérise une sous-estimation (resp. surestimation) de la valeur du modèle par l'interpolateur. Les résultats obtenus avec l'interpolateur linéaire sont présentés à gauche et les résultats obtenus avec l'interpolateur de Delaunay sont présentés à droite.

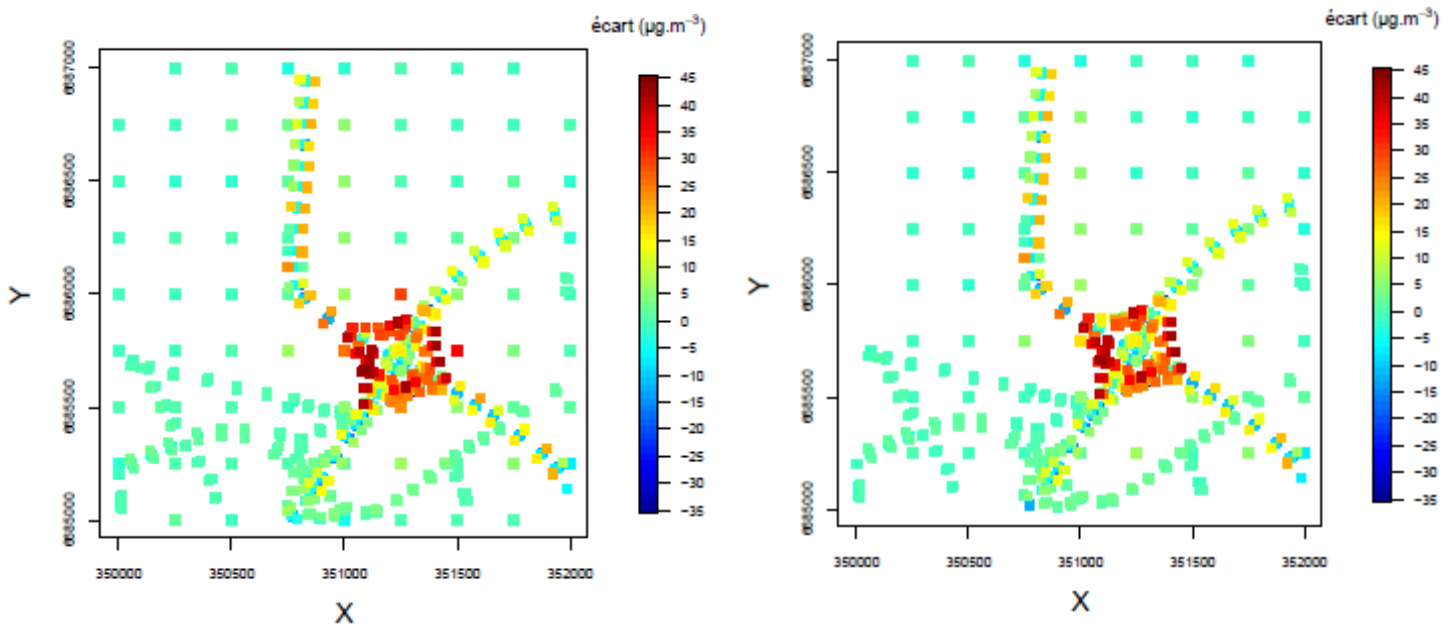


Figure 41 : Ecart par rapport aux données de modélisation originales pour les interpolateurs linéaires et de Delaunay sur les données de Nantes

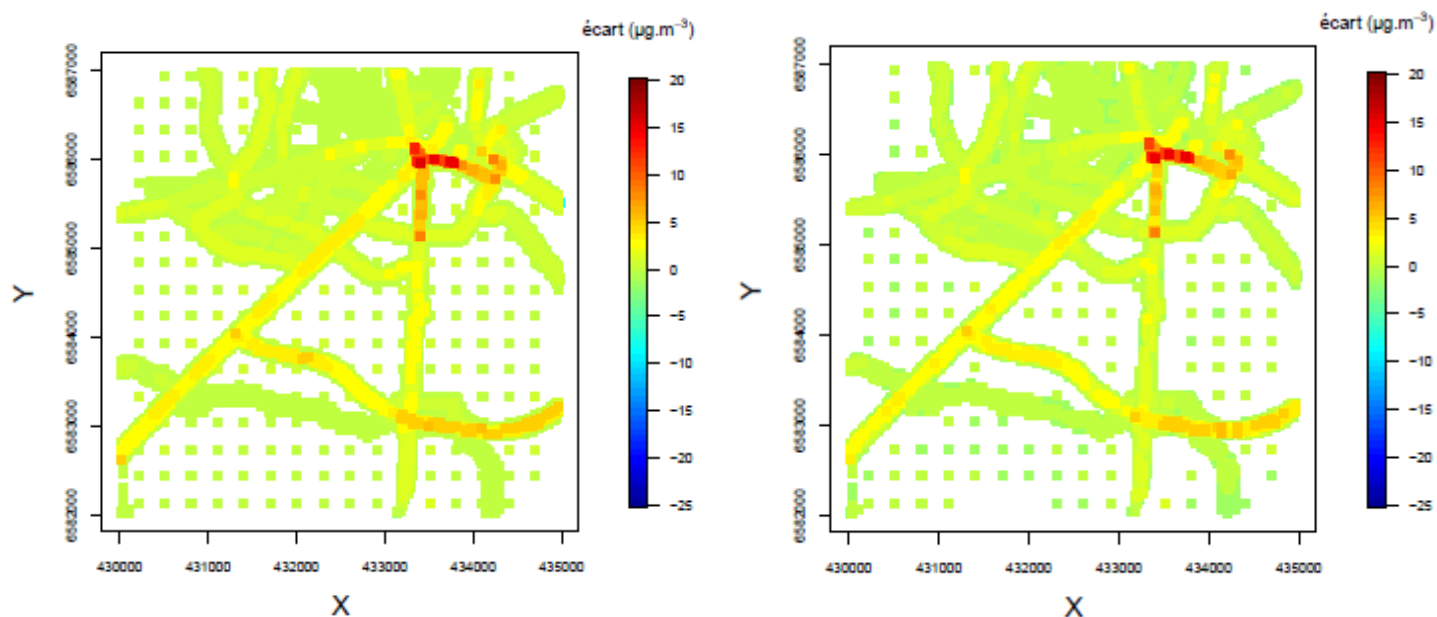


Figure 42 : Ecart par rapport aux données de modélisation originales pour les interpolateurs linéaires et de Delaunay sur les données de Niort

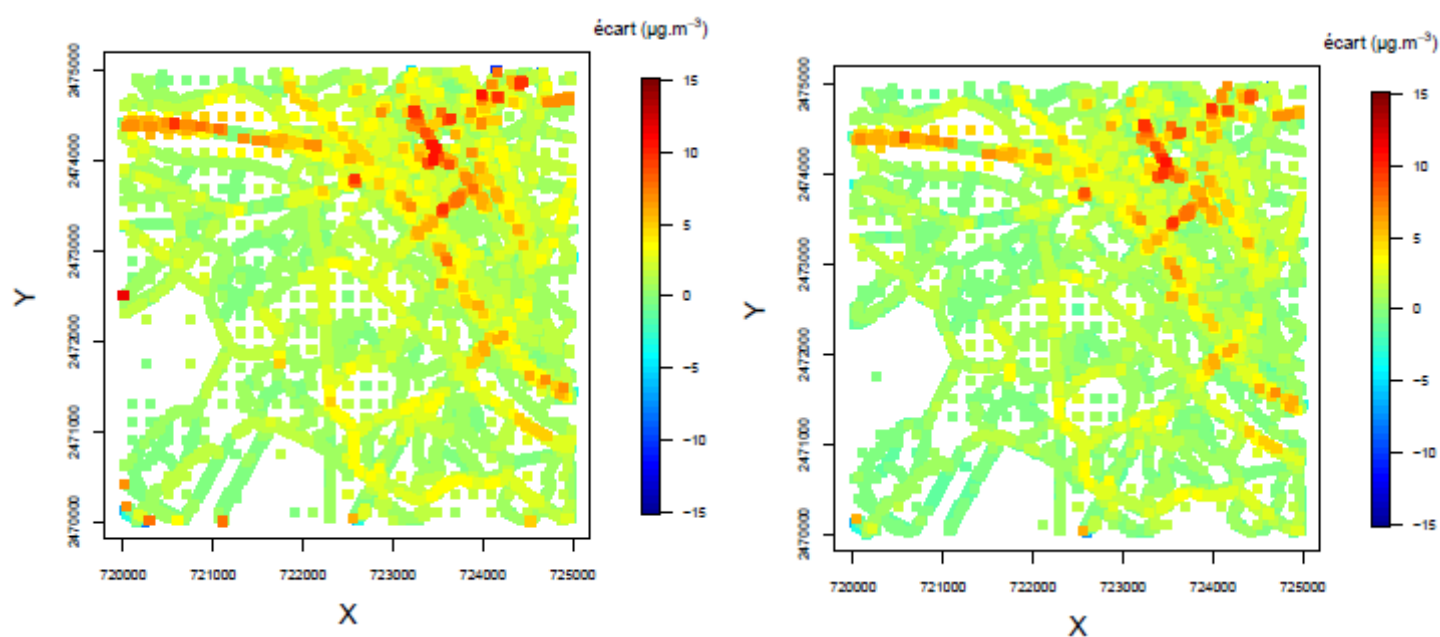


Figure 43 : Ecart par rapport aux données de modélisation originales pour les interpolateurs linéaires et de Delaunay sur les données de Reims

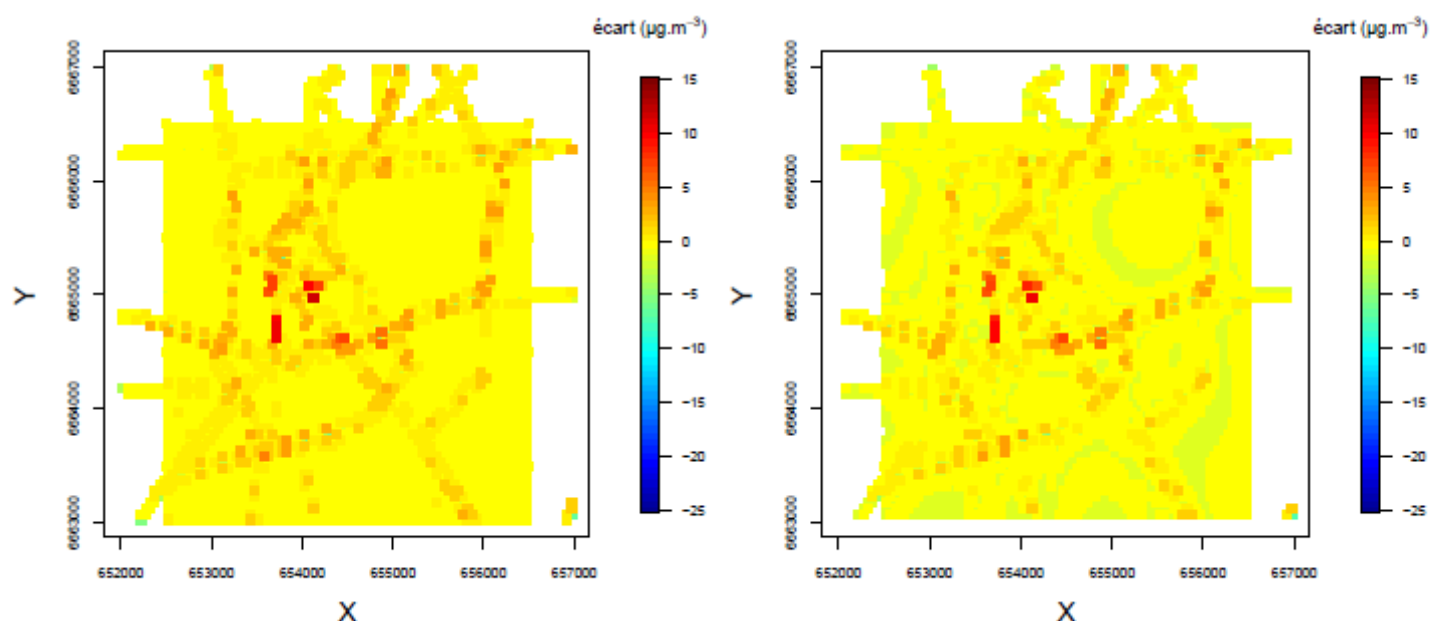


Figure 44 : Ecart par rapport aux données de modélisation originales pour les interpolateurs linéaires et de Delaunay sur les données de Bourges (maillage n°3 de PREVISION'AIR)

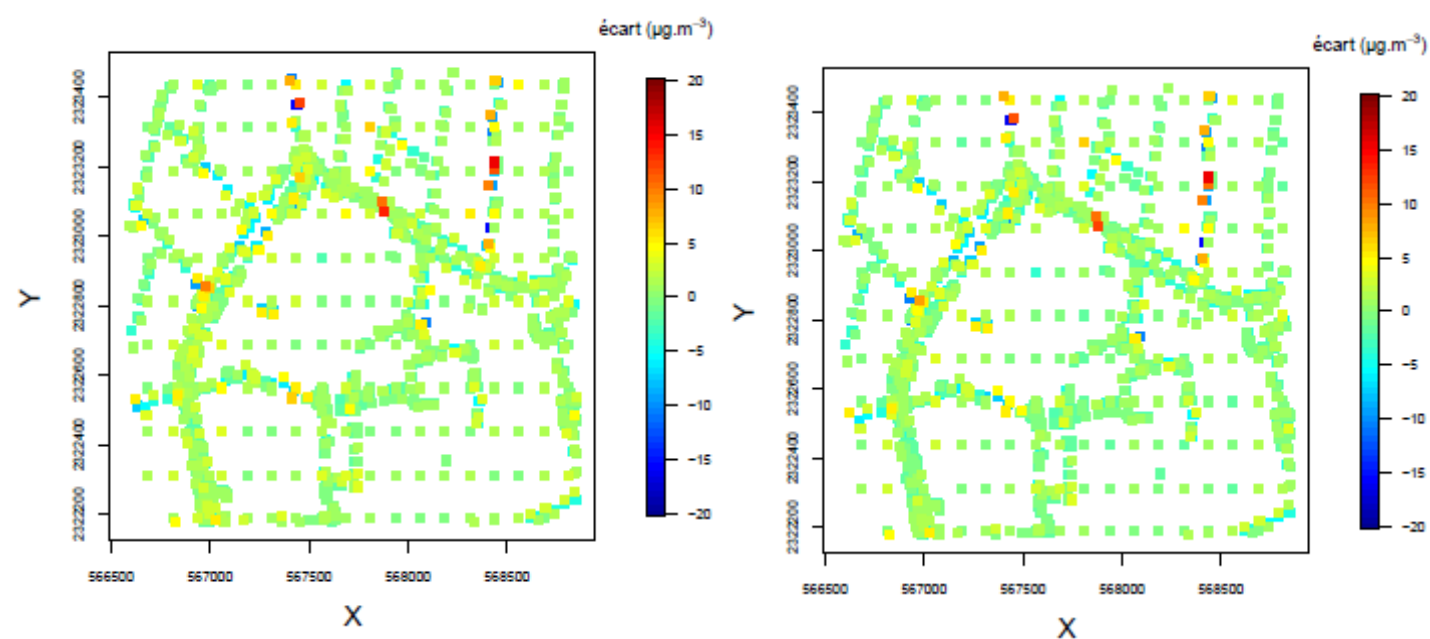


Figure 45 : Ecart par rapport aux données de modélisation originales pour les interpolateurs linéaires et de Delaunay sur les données d'Orléans

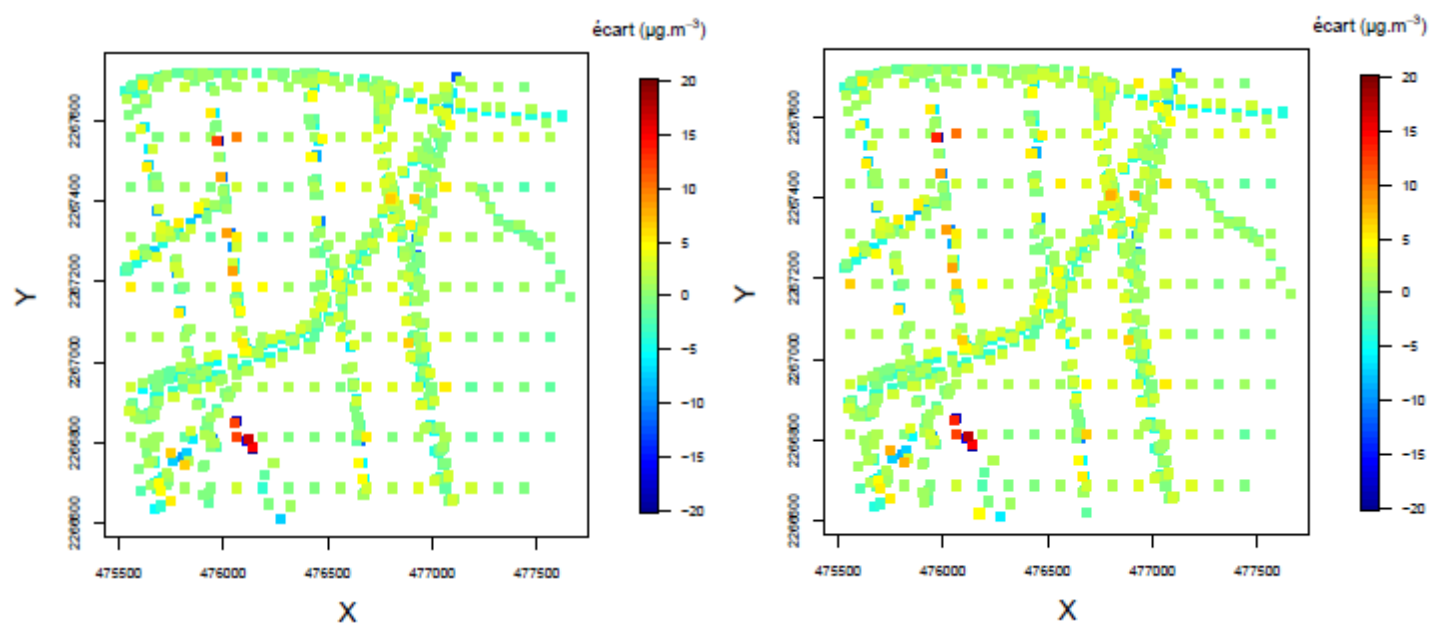


Figure 46 : Ecart par rapport aux données de modélisation originales pour les interpolateurs linéaires et de Delaunay sur les données de Tours

ANNEXE 4 : NUAGES DE POINTS (SUR GRILLE DE REFERENCE)

Dans cette annexe, l'écart est calculé comme la différence entre données de modélisation et valeurs interpolées : une valeur négative (resp. positive) de l'écart caractérise une sous-estimation (resp. surestimation) de la valeur du modèle par l'interpolateur.

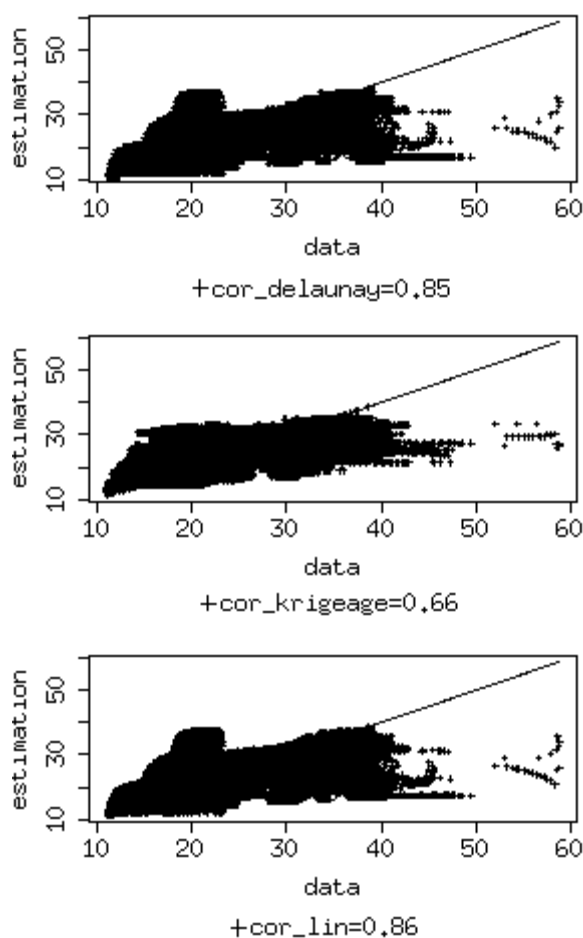


Figure 47 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Bourges (Delaunay, krigeage à 5 voisins et linéaire). Les valeurs sont interpolées à partir des données de modélisation fournies sur le maillage n°1 de PREVISION'AIR

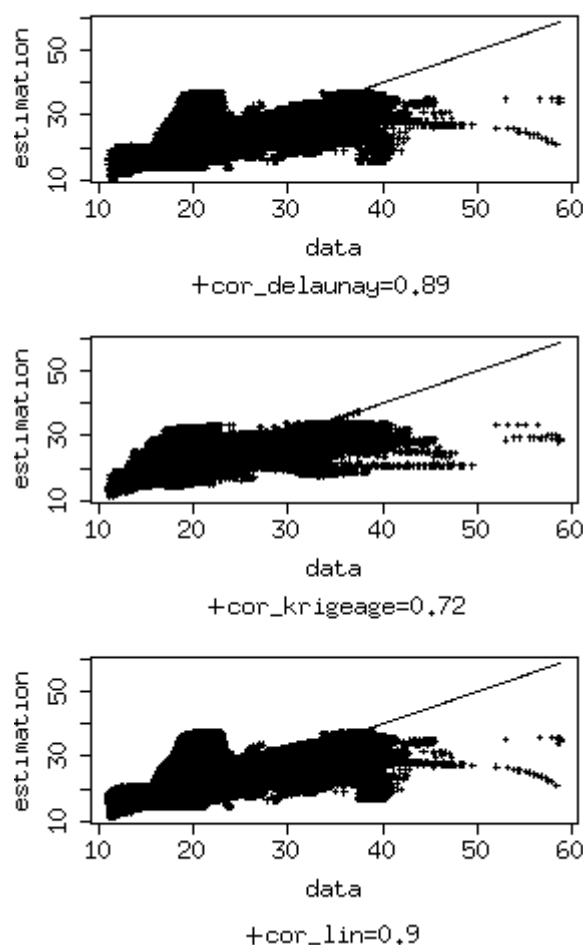


Figure 48 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Bourges (Delaunay, krigeage à 5 voisins et linéaire). Les valeurs sont interpolées à partir des données de modélisation fournies sur le maillage n°2 de PREVISION'AIR

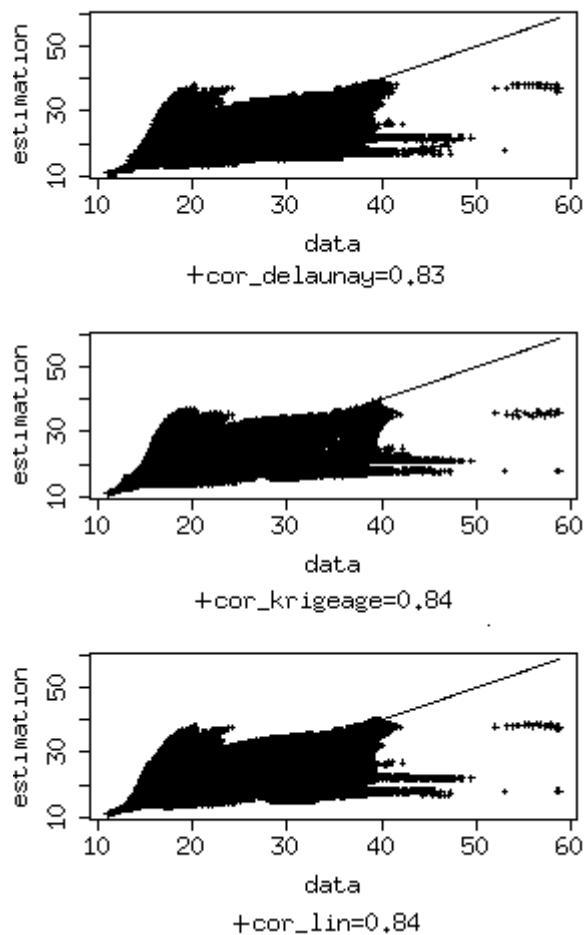


Figure 49 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Bourges (Delaunay, krigeage à 5 voisins et linéaire). Les valeurs sont interpolées à partir des données de modélisation fournies sur le maillage n°3 de PREVISION'AIR

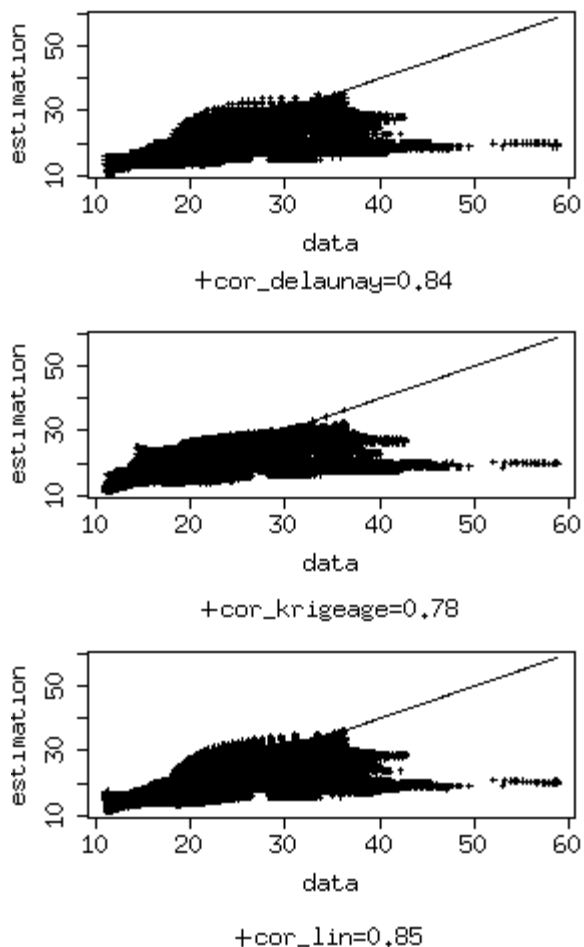


Figure 50 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Bourges (Delaunay, krigeage à 5 voisins et linéaire). Les valeurs sont interpolées à partir des données de modélisation fournies sur le maillage n°4 de PREVISION'AIR

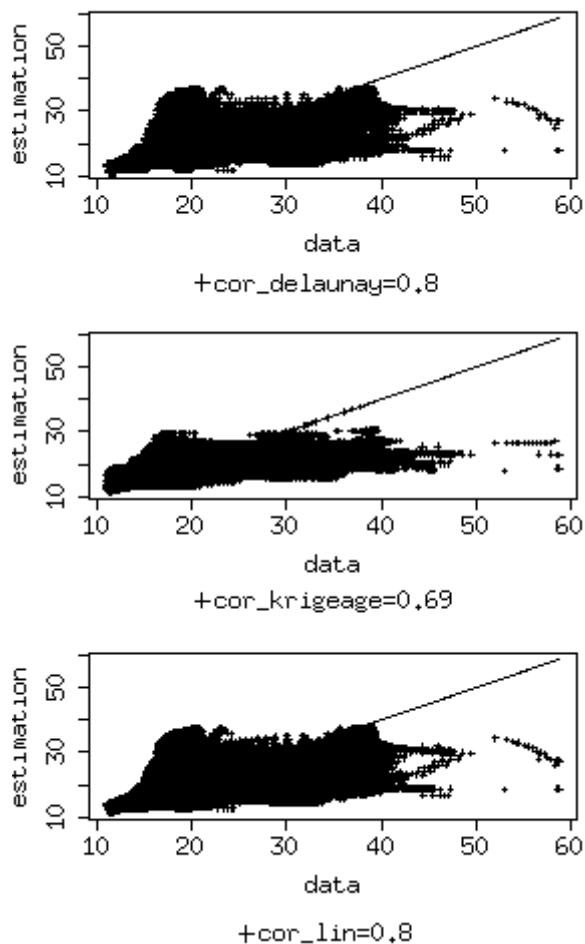


Figure 51 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Bourges (Delaunay, krigeage à 5 voisins et linéaire). Les valeurs sont interpolées à partir des données de modélisation fournies sur le maillage n°5 de PREVISION’AIR

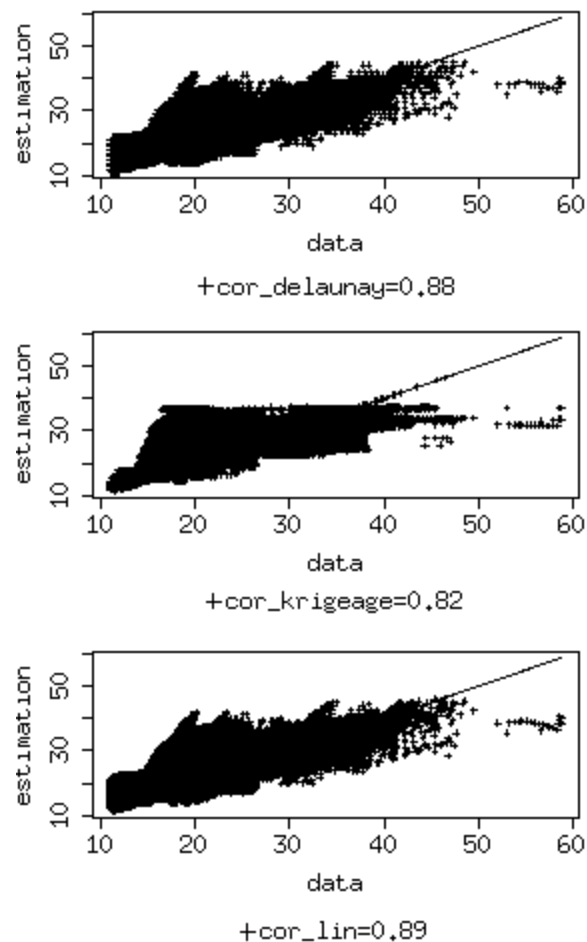


Figure 52 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Bourges (Delaunay, krigeage à 5 voisins et linéaire). Les valeurs sont interpolées à partir des données de modélisation fournies sur le maillage n°6 de PREVISION’AIR

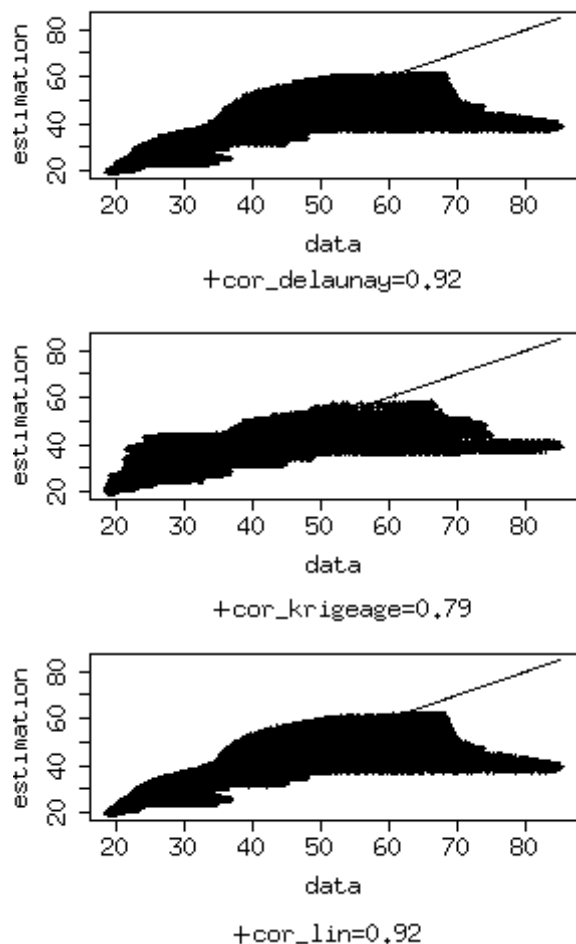


Figure 53 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Nantes (Delaunay, krigage à 5 voisins et linéaire).

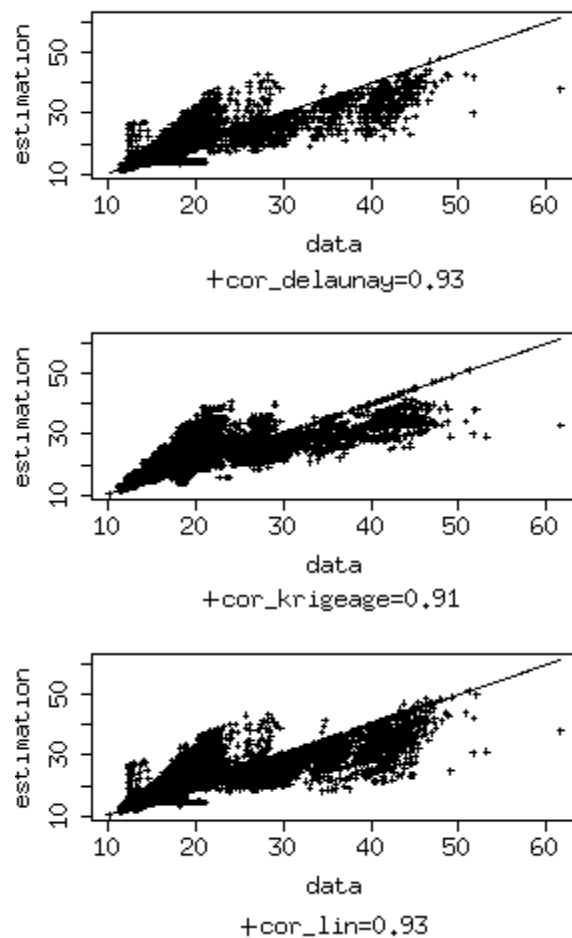


Figure 54 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Niort (Delaunay, krigage à 5 voisins et linéaire).

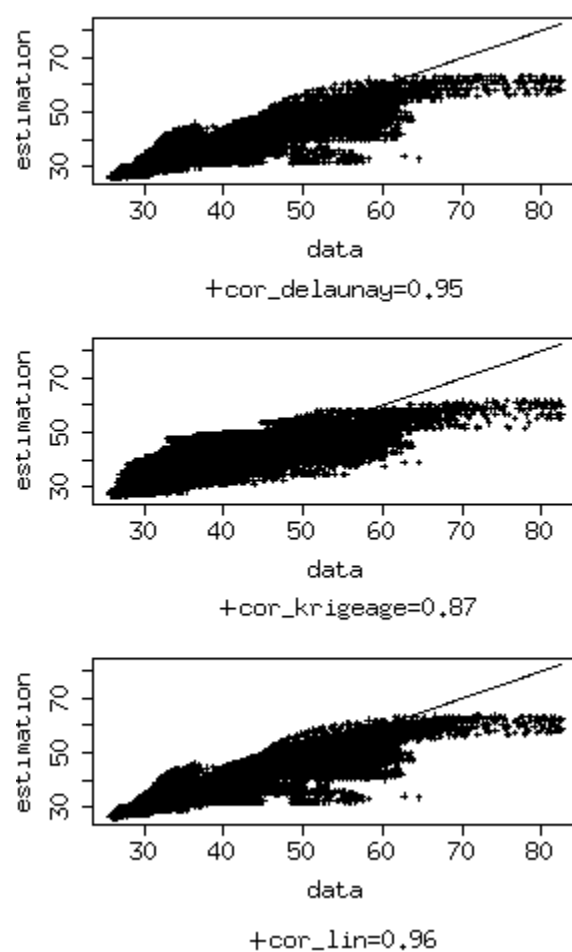


Figure 55 : Nuages de points (données de modélisation et valeurs interpolées) sur la grille de référence de Reims (Delaunay, krigeage à 5 voisins et linéaire).

ANNEXE 5 : CARTOGRAPHIE DES ECARTS, AVEC LES INTERPOLATEURS LINEAIRES ET DE DELAUNAY (SUR GRILLE DE REFERENCE) - CAS DE BOURGES

Dans cette annexe, l'écart est calculé comme la différence entre la valeur interpolée et la donnée de modélisation (valeur interpolée – donnée de modélisation) : une valeur négative (resp. positive) de l'écart caractérise une sous-estimation (resp. surestimation) de la valeur du modèle par l'interpolateur. Les résultats obtenus avec l'interpolateur linéaire sont présentés à gauche et les résultats obtenus avec l'interpolateur de Delaunay sont présentés à droite.

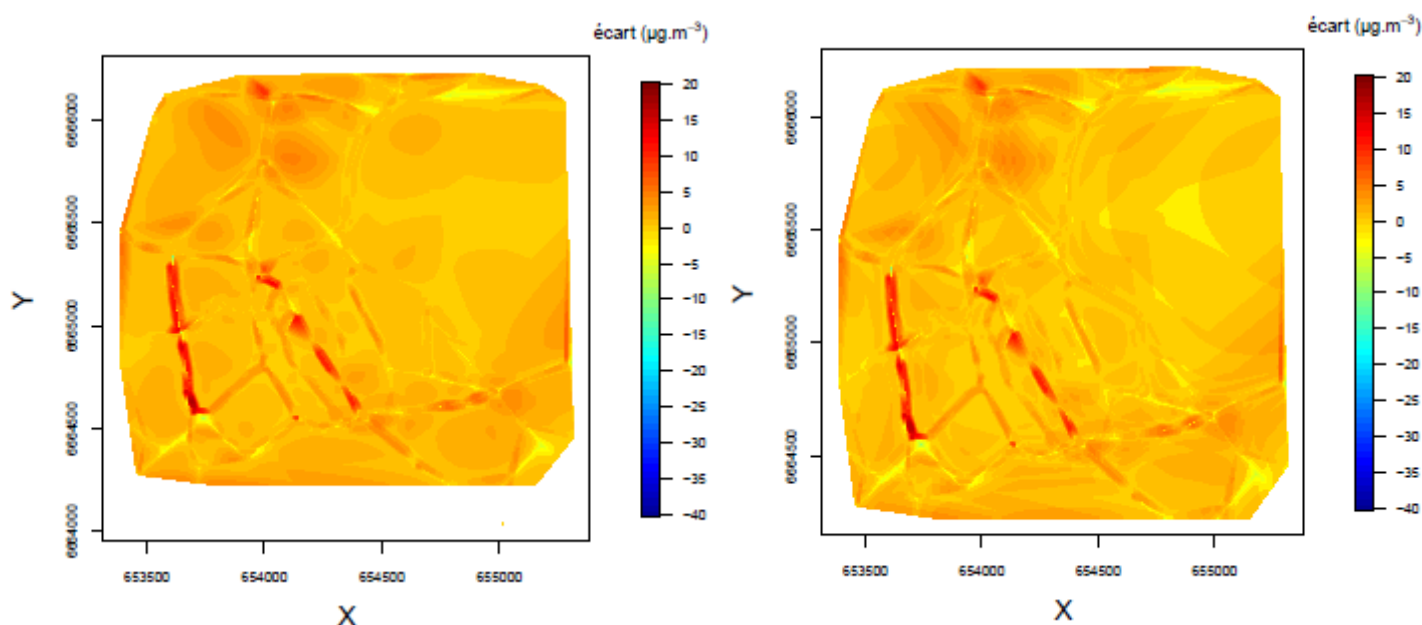


Figure 56 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur de Delaunay sur la grille de référence. Les valeurs sont interpolées à partir des données de modélisation sur le maillage n°1 de PREVISION'AIR.

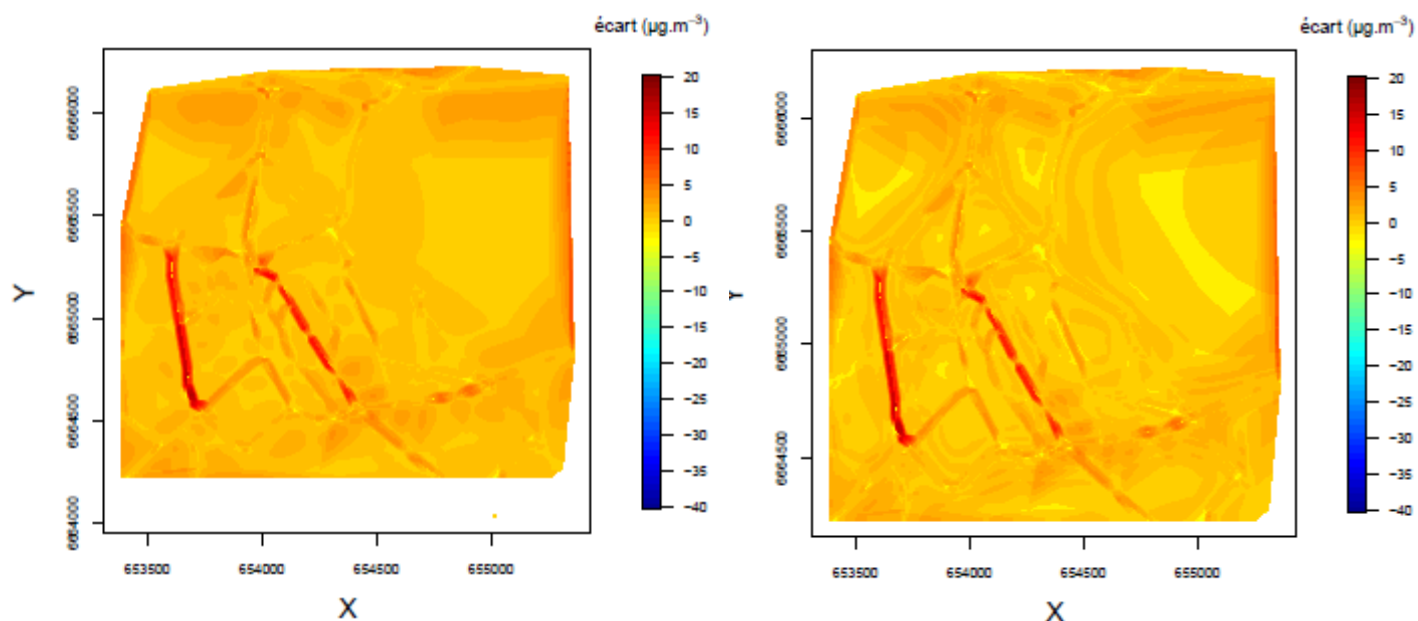


Figure 57 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur de Delaunay sur la grille de référence de Bourges. Les valeurs sont interpolées à partir des données de modélisation sur le maillage n°2 de PREVISION'AIR.

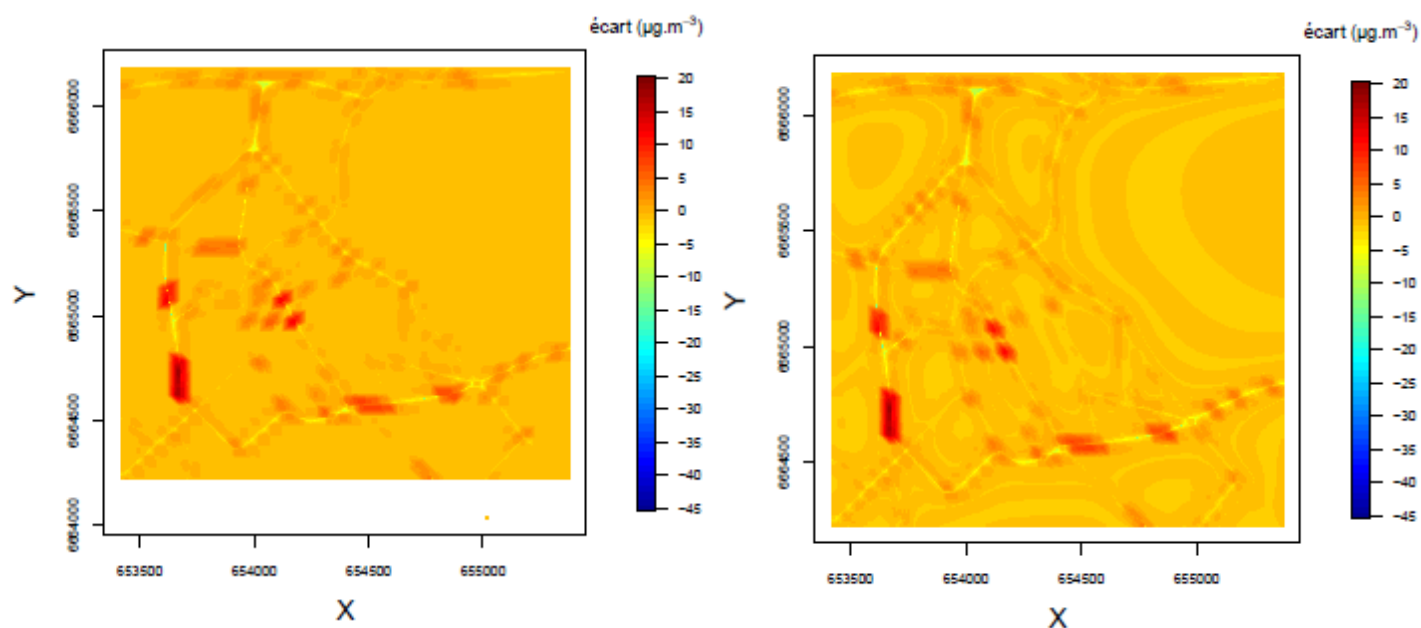


Figure 58 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur de Delaunay sur la grille de référence de Bourges. Les valeurs sont interpolées à partir des données de modélisation sur le maillage n°3 de PREVISION'AIR.

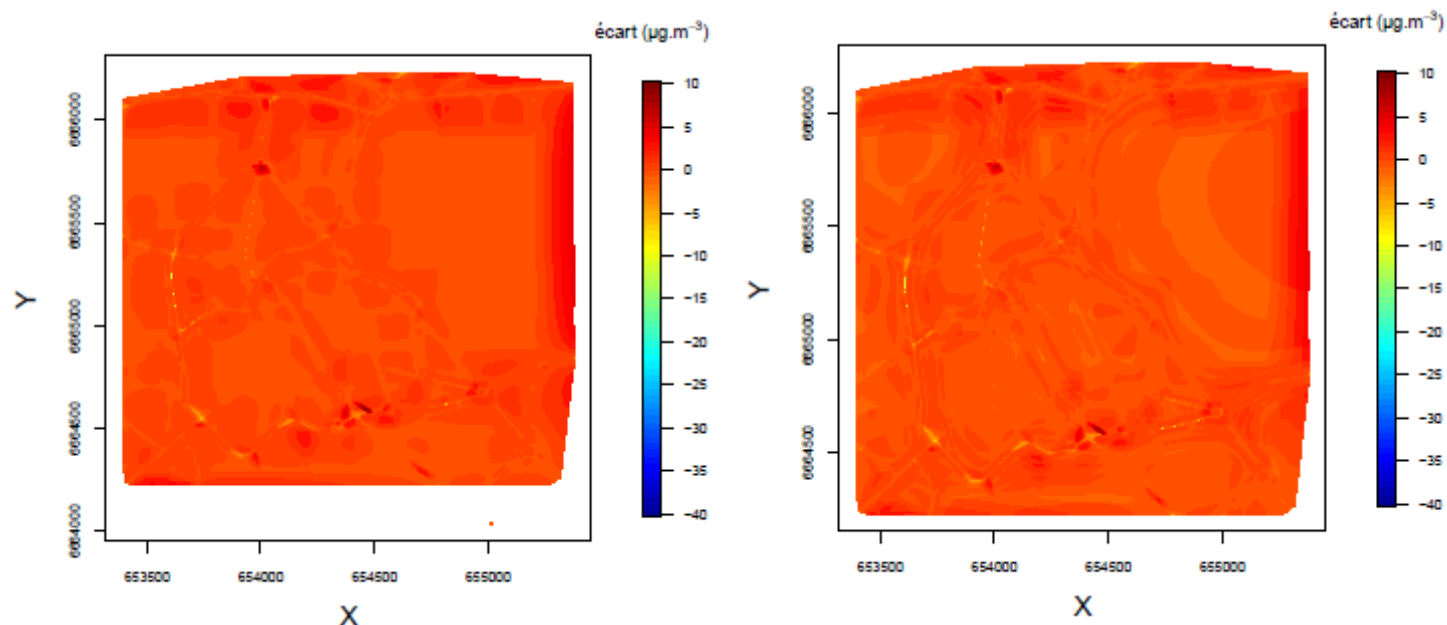


Figure 59 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur de Delaunay sur la grille de référence de Bourges. Les valeurs sont interpolées à partir des données de modélisation sur le maillage n°4 de PREVISION'AIR.

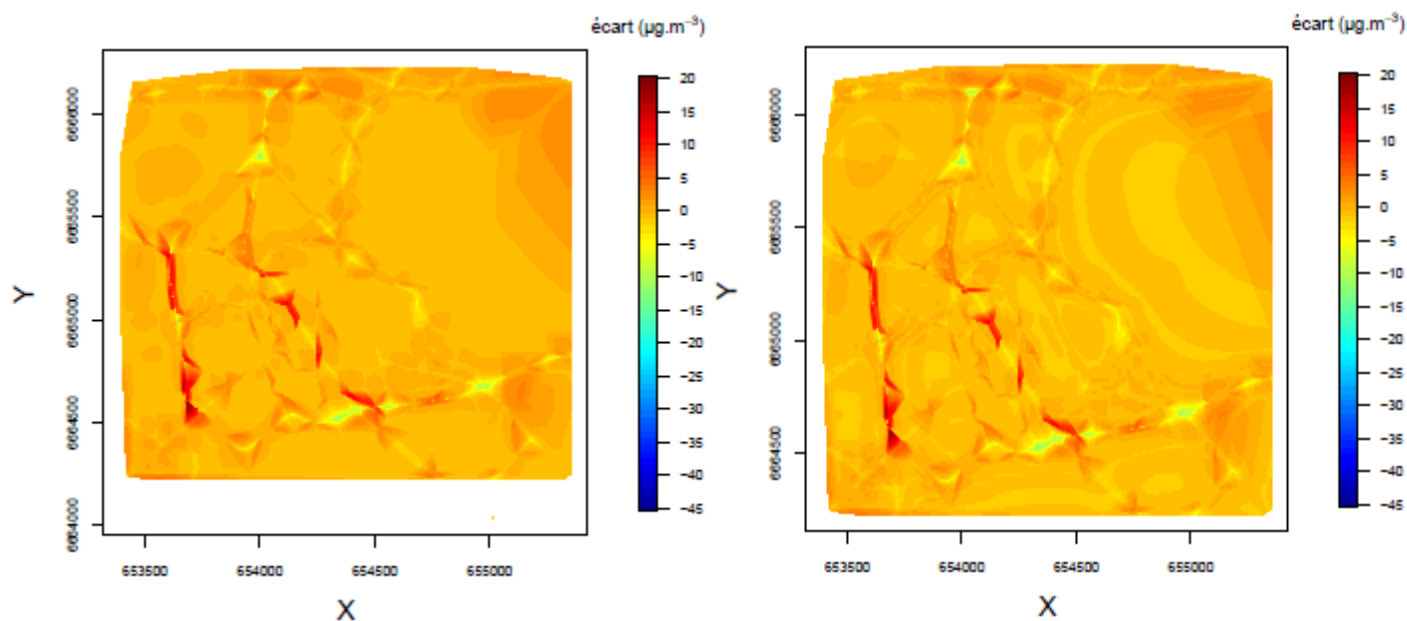


Figure 60 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur de Delaunay sur la grille de référence de Bourges. Les valeurs sont interpolées à partir des données de modélisation sur le maillage n°5 de PREVISION'AIR.

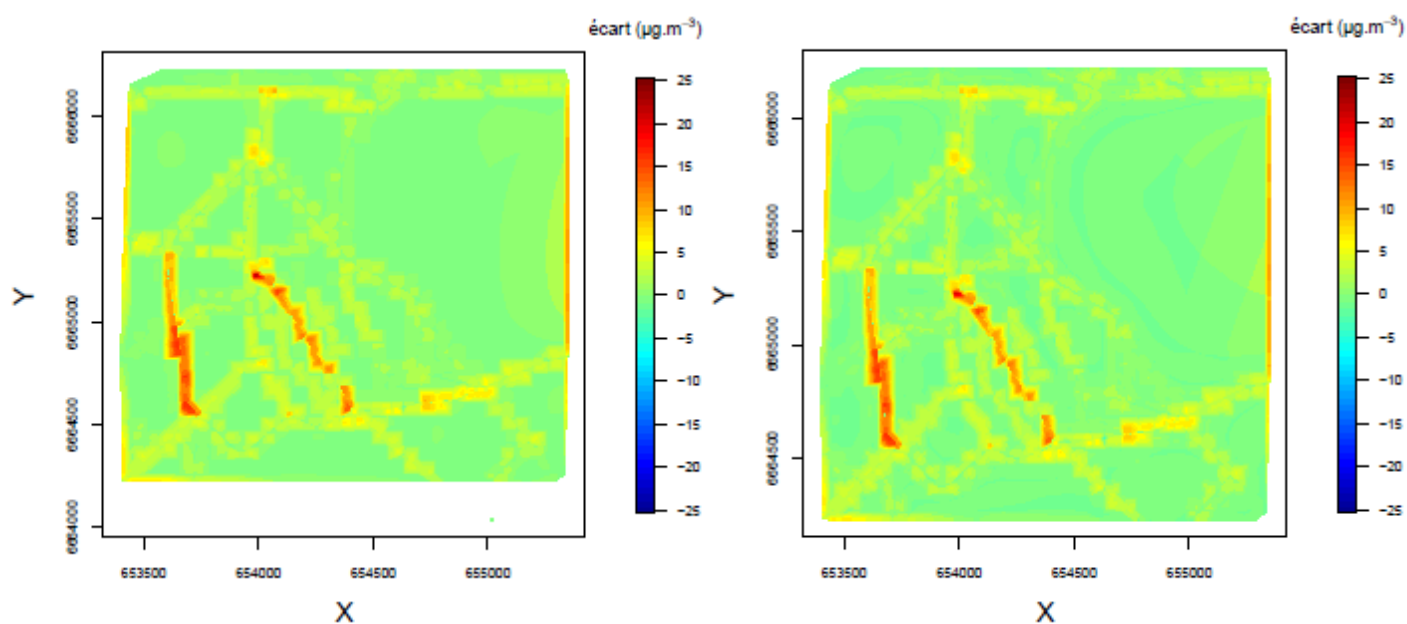


Figure 61 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur de Delaunay sur la grille de référence de Bourges. Les valeurs sont interpolées à partir des données de modélisation sur le maillage n°6 de PREVISION'AIR.

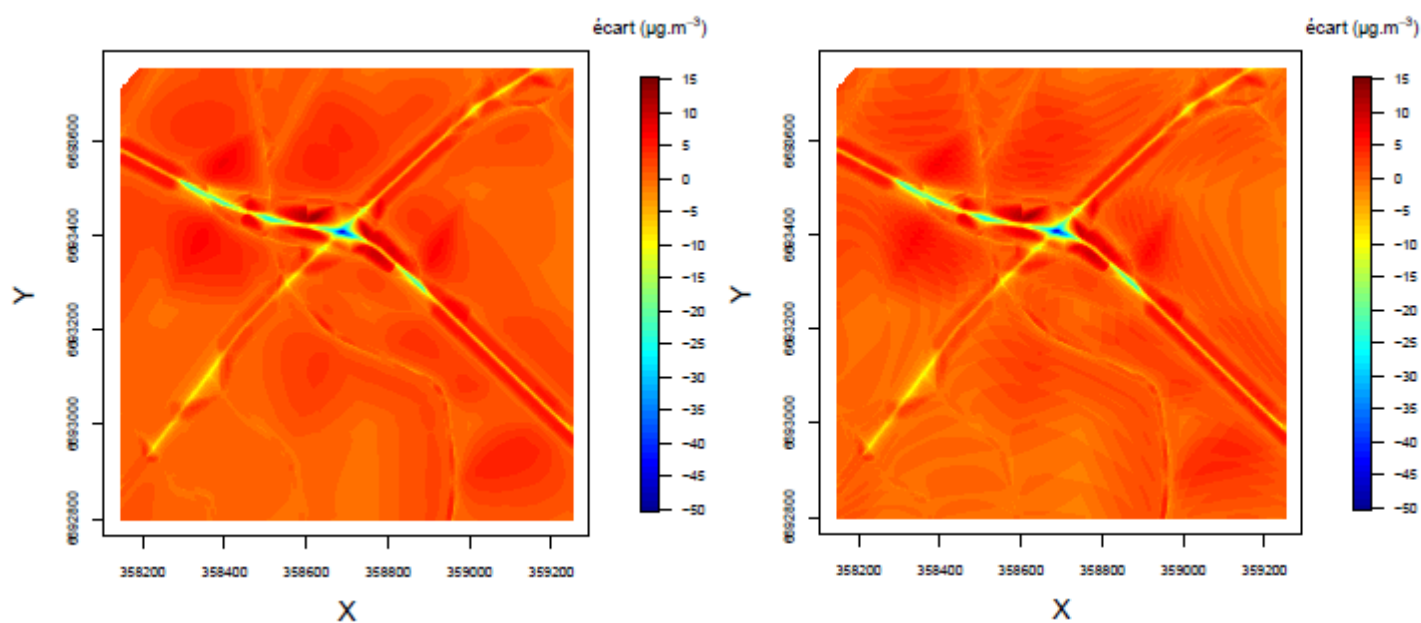


Figure 62 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur les interpolateurs linéaires et de Delaunay sur la grille de référence de Nantes

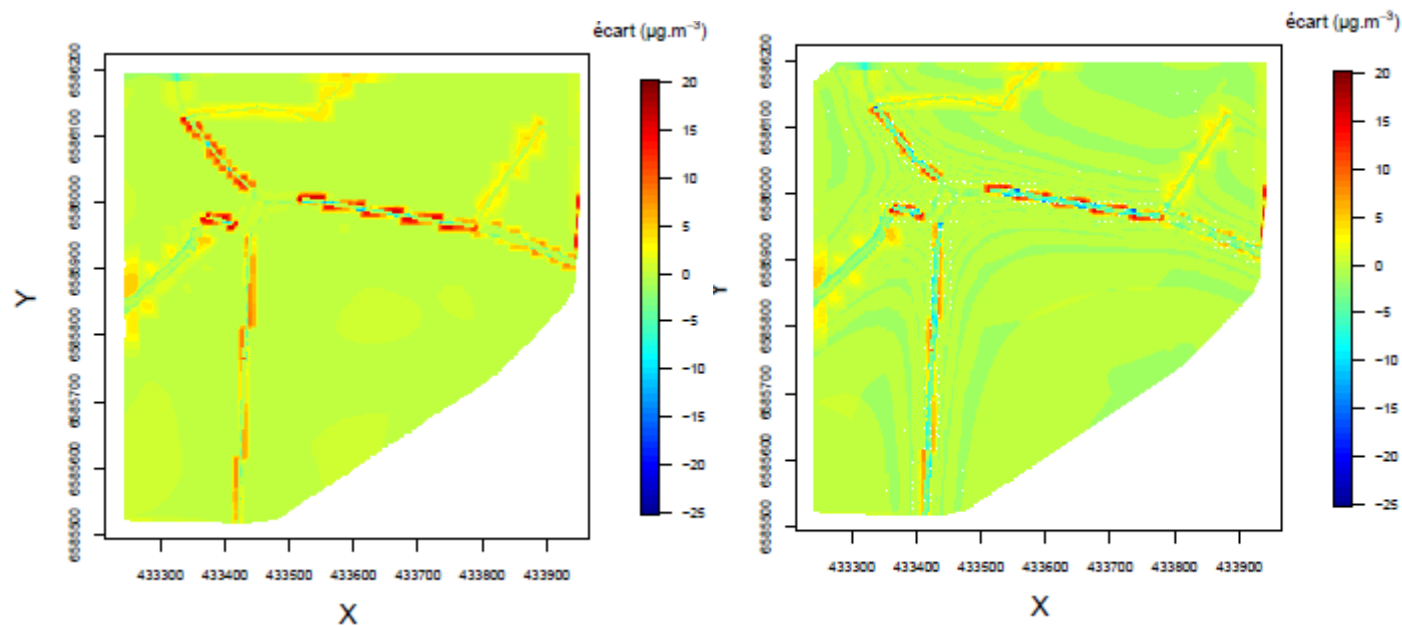


Figure 63 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur avec les interpolateurs linéaires et de Delaunay sur la grille de référence de Niort

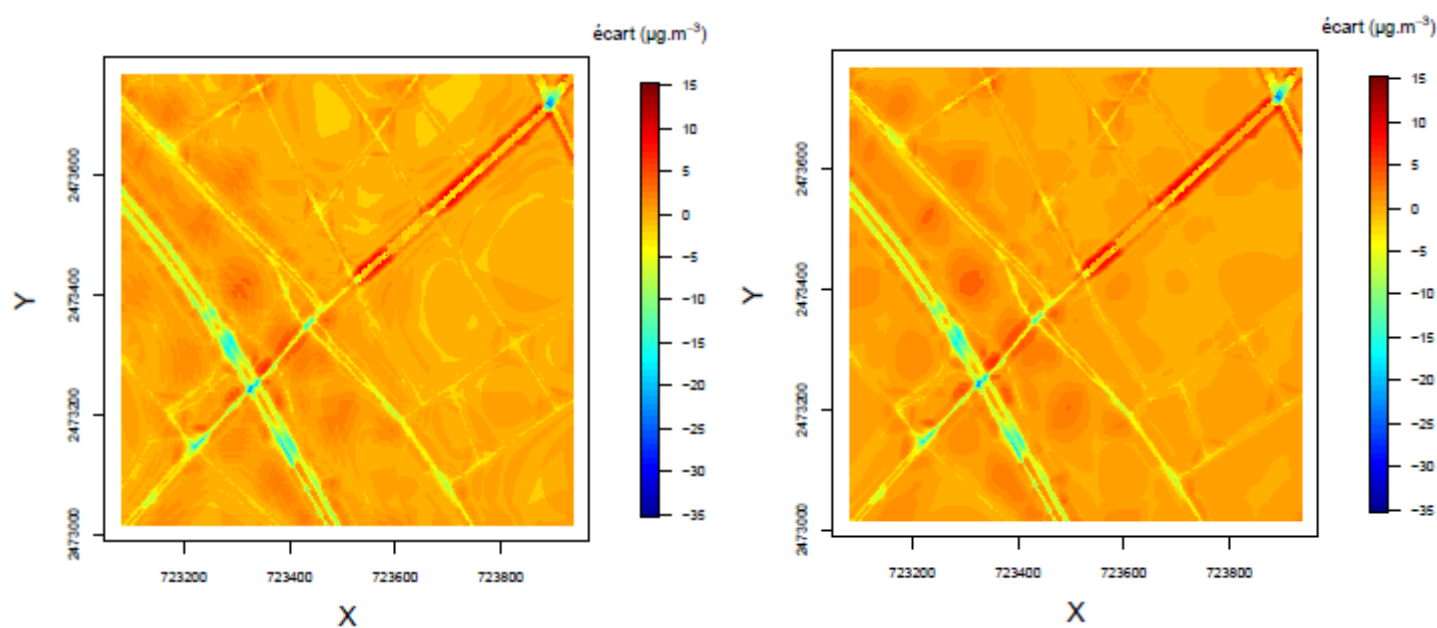


Figure 64 : Ecart par rapport aux données de modélisation originales pour l'interpolateur linéaire et l'interpolateur avec les interpolateurs linéaires et de Delaunay sur la grille de référence de Reims



direction et secrétariat du LCSQA

INERIS - parc technologique Alata - BP 2 - F60550 Verneuil-en-Halatte
tél. 03 44 55 64 04 - www.lcsqa.org