

Note de synthèse

ANNEXE TECHNIQUE AU RAPPORT LCSQA 2014 : ESTIMATION DE L'EXPOSITION DES POPULATIONS AUX DÉPASSEMENTS DE SEUILS RÉGLEMENTAIRES

Interpolation des sorties de modèles urbains par krigeage avec dérive polynomiale

Maxime Beauchamp, Laure Malherbe

SYNTHÈSE

Cette annexe complète le rapport de 2014 intitulé *Estimation des populations exposées aux dépassements de seuils réglementaires - Echelle urbaine*. Dans ce rapport, sont comparés différents interpolateurs linéaires permettant de cartographier les concentrations à partir de résultats de modélisation urbaine : l'interpolateur linéaire classique, l'interpolateur de Delaunay et le krigeage ordinaire. Les scores obtenus ont montré le comportement satisfaisant des deux premières méthodes, la troisième, dans sa configuration usuelle, n'étant pas adaptée aux caractéristiques des données en milieu urbain.

Des développements ont été réalisés en 2015 par le LCSQA pour construire un interpolateur des sorties de modèles urbains plus performant que les interpolateurs linéaires précédemment étudiés. La méthode développée permet de prendre en compte l'intégralité du cadastre des émissions routières via la caractérisation de la décroissance exponentielle des concentrations perpendiculairement aux axes routiers. La décomposition en série entière de la fonction exponentielle permet de se ramener à une dérive externe polynomiale qui est alors introduite dans un krigeage. Les cartographies et les scores sont fortement améliorés.

Dans ce qui suit, la méthode d'interpolation développée utilise souvent des éléments bibliographiques spécifiques au NO₂. Des points importants tels que le type de modèle statistique proposé pour les données d'observations ou la distance d'influence d'un axe routier sur les niveaux de pollution sont donc en particulier hérités de cette littérature. On considère toutefois qu'ils peuvent être réutilisés pour l'interpolation des données de particules : en effet, l'influence du réseau routier étant similaire mais d'importance moindre sur les niveaux de particules que sur ceux de dioxyde d'azote, l'utilisation du modèle développé pour le NO₂ reste valable pour les PM₁₀. Les distances d'influence des axes sont alors probablement majorées sans que cela ait un quelconque impact sur la cartographie, les données inutiles étant simplement ignorées lors de l'ajustement de la fonction de dérive par le krigeage.

1.1 Modèle statistique

Dans les zones urbaines, la concentration $Z(x)$ d'un polluant (NO₂ et PM₁₀) en un point x peut être représentée comme la somme d'une composante de fond et d'une contribution liée au trafic automobile dans un proche voisinage de x . Pour simplifier, les deux composantes sont considérées comme spatialement indépendantes (voir par exemple des modélisations similaires dans Font et al., 2014). On note $Z(x)$ la concentration totale en x , $Y(x)$ la concentration de fond, $S(x)$ l'incrément de pollution lié au trafic routier et $w(x)$ un résidu de pollution non expliqué par la somme de $Y(x)$ et $S(x)$.

$$Z(x) = Y(x) + S(x) + w(x) \quad (1)$$

L'indépendance entre Y et S peut être vérifiée en moyennant les valeurs de S par classe de concentrations de fond urbain. $Y (= Z - S - w)$ peut être obtenu soit en retirant directement les sources liées au trafic lors de la modélisation urbaine, soit en krigant les valeurs de concentration modélisées sur les points récepteurs en situation de fond urbain.

En fonction du type de donnée interpolée, il est possible de négliger les transformations chimiques, en particulier dans le cas de moyennes temporelles, et la pollution liée au trafic $S(x)$ peut alors être vue comme la somme des contributions $S^k(x)$ en provenance des routes R^k localisées dans un proche voisinage de x , défini par une distance maximale par rapport à x .

En accord avec la littérature (voir par exemple Zou et al. (2006), Gilbert et al. (2003), Roordeknappe et al. (1999)) ainsi que les campagnes de mesures régulièrement effectuées par les AASQA, l'influence directe du trafic routier peut être considérée comme **non significative** sur les niveaux de NO₂ au-delà de 400m.

$$Z(x) = Y(x) + \sum_{\substack{k \\ d(R^k, x) \leq 400}} S^k(x) + w(x) \quad (2)$$

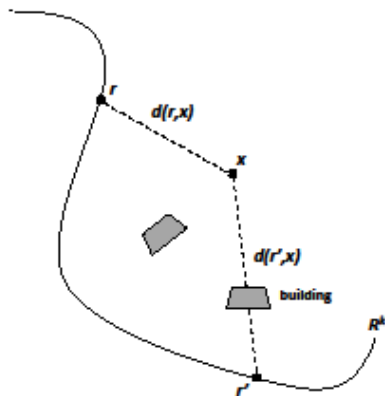


Figure 1 : Schéma explicatif pour la formulation du terme $S^k(x)$

Le terme $S^k(x)$ apporté par la route R^k en x peut également être vu comme l'addition de termes ponctuels apportés par chaque position r de R^k (voir Figure 1), qui décroissent selon une fonction f de la distance $d(r, x)$ entre le point r de R^k et x , et de l'effet d'écran $b(r, x)$ lié à la présence de bâtiments entre le point x et le point r de R^k . Dans ce qui suit, le modèle statistique a été simplifié et l'effet d'écran relatif aux bâtiments n'a pas été considéré, il impliquerait une modélisation complexe dépendant des hauteurs, surfaces, volumes et formes des bâtiments.

Finalement, on a :

$$S^k(x) = \int_{R^k} f(S(r), d(r, x)) \cdot \mathbf{1}_{d(r,x) < 400} dr \quad (3)$$

où $S(r) = cE(r)$ désigne la concentration liée au trafic routier au point r de la route R^k . Cette concentration est communément prise comme une fonction linéaire ($c \in \mathbb{R}^+$) des émissions de NO_x — notées $E(r)$ — issues du trafic routier au même point. Pour simplifier les notations, l'indicatrice $\mathbf{1}_{d(r,x) < 400}$ n'est pas rappelée dans les prochaines équations.

Cependant, il n'est évidemment pas possible de quantifier les valeurs de toutes les sources ponctuelles puisque dans les inventaires d'émissions, les quantités de NO_x apportées par les routes R^k sont spécifiées sous forme d'émissions linéiques dans un fichier au format SIG.

En conséquence, le terme $E(r)$ s'écrit : $E(r) = E_k / |R^k|$. E_k est la quantité totale annuelle de NO_x émise par la route R^k (en $\text{kT} \cdot \text{an}^{-1}$) et $|R^k|$ est la longueur de cette route.

$$S^k(x) = \int_{R^k} f\left(\frac{cE_k}{|R^k|}, d(r, x)\right) dr \quad (4)$$

1.2 Choix de la fonction de décroissance

D'après la littérature (Taylor et al. (2015), Baldwin et al. (2015), Faus-Kessler et al. (2008), Pleijel et al. (2004)), il est pertinent de considérer que le NO_2 lié au trafic routier décroît de façon exponentielle en fonction de la distance à la route.

On propose donc la forme suivante pour la fonction f :

$$f(r, x) = \frac{cE_k}{|R^k|} e^{-\alpha d(r,x)} \quad (5)$$

Seguret et al. (2003) ont proposé un autre type de modélisation dans une étude portant sur la modélisation du NO₂ à proximité d'un grand axe automobile dans la vallée de la Thur (France), sous forme de fonctions polynomiales et exponentielles cohérentes avec les données expérimentales (Monn et al., 1997). Mais cette étude se focalisait sur une unique route. Dans notre cas, plusieurs milliers de brins routiers sont considérés et l'introduction d'un tel modèle complexifierait de manière considérable l'estimation du variogramme et le calcul de la dérive.

Plutôt que de fixer la valeur du paramètre α dans la fonction exponentielle, l'idée est d'approcher la fonction de décroissance en utilisant un développement en série, qui est finalement assimilé à une dérive polynomiale, i.e une somme de fonctions puissance avec signes alternés. Dans cette fonction de dérive, les coefficients sont inconnus et ainsi localement estimés lors de la résolution du système de krigeage (14) en voisinage glissant.

Puisque les fonctions puissance x^u augmentent quand $x > 0$, la seule façon d'obtenir une forme décroissante est d'alterner les signes, ce qui peut induire un manque de robustesse : la monotonie de la fonction de décroissance n'est pas garantie, de même que sa positivité.

Ainsi, les artefacts produits par cette approximation ainsi que le nombre de fonctions puissance requis pour atteindre un niveau de performance satisfaisant sont-ils sujets à discussion. Néanmoins, les performances de la méthode se sont avérées satisfaisantes (voir section 3.2).

Puisque cette dérive « pseudo-exponentielle » n'a pas une signification physique claire, d'autres fonctions de décroissance peuvent être choisies, par exemple l'utilisation de régresseurs qui décroissent strictement en fonction de la distance à la route, intégrés sous la forme d'un krigeage universel.

1.2.1 Un développement en série sous forme de “pseudo-exponentielle”

$$S(x) = \sum_{k=1}^N S^k(x) = \sum_{k=1}^N \int_{R^k} \frac{cE_k}{|R^k|} e^{-\alpha d(r,x)} dr \quad (6)$$

Le paramètre $\theta = (c, \alpha)$ est inconnu. L'idée est d'utiliser le développement en série de la fonction exponentielle pour obtenir $S(x)$, sous forme de combinaison de variables explicatives que sont $E_k/|R^k|$ et $d(r, x)$, élevées à la puissance u ($u \in [0, +\infty[$):

$$e^{-\alpha d(r,x)} = \sum_{u=0}^{+\infty} \frac{(-1)^u \alpha^u (d(r,x))^u}{u!} \quad (7)$$

Ainsi, la fonction exponentielle peut être utilisée sous forme de dérive externe sans fixer la valeur de θ :

$$\begin{aligned}
S(x) &= \sum_{k=1}^N S^k(x) = \sum_{k=1}^N \int_{R^k} \frac{cE_k}{|R^k|} e^{-\alpha d(r,x)} dr = \sum_{k=1}^N \int_{R^k} \sum_{u=0}^{+\infty} \frac{cE_k}{|R^k|} \frac{(-\alpha d(r,x))^u}{u!} dr \\
&= \sum_{u=0}^{+\infty} c\alpha^u \sum_{k=1}^N \int_{R^k} \frac{E_k}{|R^k|} \frac{(-1)^u (d(r,x))^u}{u!} dr
\end{aligned} \tag{8}$$

Le système de krigeage est construit sur une hypothèse de non-biais et un critère de variance optimale de l'erreur d'estimation. k_x remplace désormais la notation k puisque l'indice qui identifie la route R^k diffère d'un point x à un autre.

On note :

$$S_u^{k_x}(x) = \int_{R^{k_x}} \frac{E_{k_x}}{|R^{k_x}|} \frac{(-1)^u (d(r,x))^u}{u!} dr \tag{9}$$

$$S_u(x) = \sum_{k_x=1}^N \int_{R^{k_x}} \frac{E_{k_x}}{|R^{k_x}|} \frac{(-1)^u (d(r,x))^u}{u!} dr \tag{10}$$

Les conditions de non-biais conduisent à :

$$\sum_x \lambda_x \sum_{u=0}^{+\infty} c\alpha^u S_u(x) = \sum_{u=0}^{+\infty} c\alpha^u S_u(x_0) = \sum_{u=0}^{+\infty} \alpha^u \left(\sum_x \lambda_x S_u(x) - S_u(x_0) \right) = 0 \tag{11}$$

Elles consistent à trouver les coefficients du polynôme $a_0 + a_1\alpha + \dots + a_N\alpha^N$. Puisque α est inconnu, les coefficients sont identiquement nuls et on a :

$$\sum_x \lambda_x S_u(x) - S_u(x_0) = 0 \quad \forall u \tag{12}$$

Ces conditions de non-biais sont admissibles mais si elles ont l'avantage de ne pas fixer α , elles ne lient pas non plus les dérivées entre elles, d'où la notion de pseudo-exponentielle puisque la fonction de dérive peut finalement fortement s'écarter de l'idée originale qui a introduit la formulation exponentielle.

Le critère d'optimalité sur la variance d'estimation est classiquement obtenu en annulant les dérivées partielles du Lagrangien ψ en fonction des poids λ_x .

$$\frac{\partial \psi}{\partial \lambda_x} = \mathbf{0} \Rightarrow \sum_y \lambda_y G_{xy} - \sum_{u=0}^{+\infty} \mu_u S_u(x) = C_{x0} \tag{13}$$

où :

$$\psi = \mathbb{V}(\varepsilon) - 2 \sum_{u=0}^{+\infty} \mu_u (\sum_x \lambda_x S_u(x) - S_u(x_0))$$

L'annulation des dérivées partielles de ψ en fonction des paramètres de Lagrange redonne les conditions de non-biais.

Le système matriciel est :

$$\begin{cases} \mathbf{GL} = \mathbf{G}_0 \\ Z_0^* = \lambda \mathbf{Z}_\alpha \\ \sigma_{\text{EDK}}^2 = G_{00} - \mathbf{L}^T \mathbf{G}_0 \end{cases} \quad (14)$$

où :

$\mathbf{G} = \begin{bmatrix} \mathbf{\Gamma} & \mathbf{F} \\ \mathbf{F}^T & \mathbf{\Gamma} \end{bmatrix}$ est la matrice $\mathbf{\Gamma}$ des valeurs du variogramme entre les points d'observation avec des lignes et colonnes additionnelles liées à la dérive $\mathbf{F}$.

$\mathbf{\Gamma}$ est la matrice des valeurs du variogramme entre les points d'observation : $\mathbf{\Gamma} = \begin{bmatrix} \gamma_{11}^R & & \gamma_{1n}^R \\ \vdots & \ddots & \vdots \\ \gamma_{n1}^R & & \gamma_{nn}^R \end{bmatrix}_{n \times n}$, où γ^R est le variogramme ajusté sur les résidus $\widehat{w}$ de la régression en validation croisée des observations Z sur $(\widehat{Y} + S)$ (voir équation 1).

$\mathbf{F} = (F_{\alpha u})$ est la matrice de dérive. Sa taille est proportionnelle à l'ordre de troncature u_{MAX} de la série. La valeur de u_{MAX} est pragmatiquement fixée en accord avec la convergence des scores de performance (voir section 3.2) :

$$\mathbf{F} = \begin{bmatrix} \sum_{k_1} \int_{R^{k_1}} \frac{E_{k_1}}{|R^{k_1}|} \frac{(-1)^0 (d(r, s_1))^0}{0!} dr & \dots & \sum_{k_1} \int_{R^{k_1}} \frac{E_{k_1}}{|R^{k_1}|} \frac{(-1)^{u_{\text{MAX}}} (d(r, s_1))^{u_{\text{MAX}}}}{u_{\text{MAX}}!} dr \\ \vdots & \ddots & \vdots \\ \sum_{k_N} \int_{R^{k_N}} \frac{E_{k_N}}{|R^{k_N}|} \frac{(-1)^0 (d(r, s_N))^0}{0!} dr & \dots & \sum_{k_N} \int_{R^{k_N}} \frac{E_{k_N}}{|R^{k_N}|} \frac{(-1)^{u_{\text{MAX}}} (d(r, s_N))^{u_{\text{MAX}}}}{u_{\text{MAX}}!} dr \end{bmatrix}_{n \times (u_{\text{MAX}}+1)}$$

$\mathbf{L}$ est le vecteur des poids optimaux du krigeage et des paramètres de Lagrange :

$$\mathbf{L} = \begin{bmatrix} \lambda \\ \mu \end{bmatrix}_{n+u_{\text{MAX}}+1}$$

$\mathbf{G}_0$ est le vecteur des valeurs du variogramme entre les points d'observation et le site d'estimation, avec des termes additifs correspondant à la fonction de dérive (F_{0u}) en ce point.

$$\mathbf{G}_0 = \begin{bmatrix} \gamma_{x0}^R \\ \sum_{k_0} \int_{R^{k_0}} \frac{E_{k_0}}{|R^{k_0}|} \frac{(-1)^0 (d(r, s_0))^0}{0!} dr \\ \vdots \\ \sum_{k_0} \int_{R^{k_0}} \frac{E_{k_0}}{|R^{k_0}|} \frac{(-1)^{u_{MAX}} (d(r, s_0))^{u_{MAX}}}{u_{MAX}!} dr \end{bmatrix}_{n+u_{MAX}+1}$$

1.2.2 Autre possibilité

Afin de rester dans le même cadre théorique que celui utilisé pour le développement de l'exponentielle en série, un krigeage universel avec des régresseurs $f_u(r, x)$ tels que $f(r, x) = \sum_u f_u(r, x) = cE_k/|R^k|$ quand $d(r, x) = 0$ et $f(r, x) \rightarrow 0$ quand $d(r, x) \rightarrow +\infty$ peut être utilisé. Une telle fonction peut être :

$$f(r, x) = \frac{cE_k}{|R^k|} \sum_{u=1}^{u_{MAX}} \frac{1}{u + d(r, x)^u} \quad (15)$$

Elle conduit au même système de krigeage que celui introduit pour la dérive pseudo-exponentielle où $F_{\alpha u}$, la u -ème dérive au point α devient:

$$F_{\alpha u} = \sum_{k_\alpha} \int_{R^{k_\alpha}} \frac{E_{k_\alpha}}{|R^{k_\alpha}|} \frac{1}{u + d(r, s_\alpha)^u} dr \quad (16)$$

Une telle fonction de dérive a l'avantage d'être cohérente physiquement, en imposant une monotonie dans le processus de diffusion perpendiculairement aux axes routiers. Du point de vue des scores de performance, l'utilisation de ce type de dérive montre des résultats similaires, quoique légèrement moins bons que ceux obtenus avec la dérive polynomiale inspirée de la fonction exponentielle.

1.3 Considérations pratiques

En pratique, le terme intégral le long de la route est inconnu. Il est toutefois possible de l'approcher en décomposant la route sous la forme de petits segments :

$$\int_{R^k} \frac{E_k}{|R^k|} \frac{(-1)^u (d(r, x))^u}{u!} dr = \sum_{l=1}^p E_k \frac{|R_l^k|}{|R^k|} \frac{(-1)^u (d(r, x))^u}{u!} \quad (17)$$

où R_l^k est la l -ème section de la route R^k et r est désormais le point milieu de ce segment.

Finalement, comme dans tout modèle de dérive externe, le paramètre θ n'apparaît pas dans le système de krigeage. Néanmoins, le variogramme des résidus de la régression de S sur f doit être modélisé. Comme la dérive est, après manipulation des équations de la section 1, décrite sous la forme de régresseurs linéaires dans le système matriciel, le variogramme est ici directement estimé en utilisant une procédure itérative d'ajustement de la dérive dans laquelle les résidus proviennent d'une validation croisée de type k-fold. Il pourrait également l'être via un ajustement indirect par FAI-k (voir Chiles and Delfiner, 2012). Une autre option serait de recourir à des méthodes de krigeage bayésien (Omre, 1987) pour estimer la valeur de θ .

2. DONNEES

Deux domaines sont considérés dans cette étude.

LIG'Air a fourni un jeu de données test sur l'agglomération d'Orléans pour l'année 2010. Des données modélisées à partir du logiciel ADMS-Urban fournissent des moyennes annuelles du NO_2 sur un ensemble de plus de 3000 points récepteurs, disposés selon un maillage dit intelligent (Figure 2), qui est resserré à proximité des axes routiers et plus lâche en situation de pollution de fond urbain. Le cadastre des émissions routières de NO_x a également été fourni par l'AASQA (Figure 3). Il se base sur les données COPERT IV (**CO**mputer **P**rogram to calculate **E**mission from **R**oad **T**ransport).

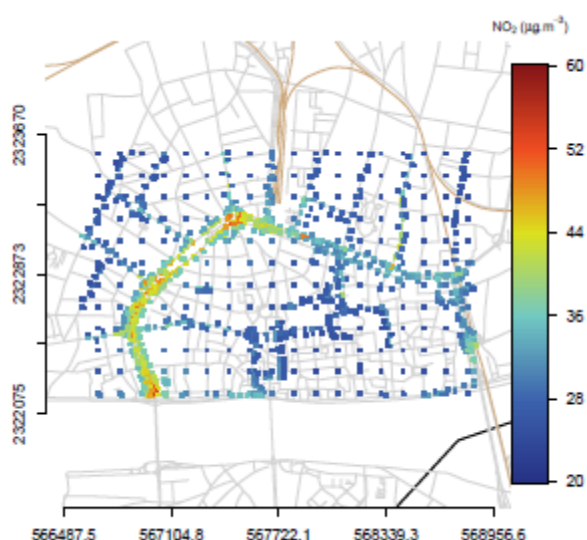


Figure 2 : Moyenne annuelle de NO_2 modélisée par LIG'Air avec ADMS-Urban sur un ensemble de points récepteurs pour l'année 2010

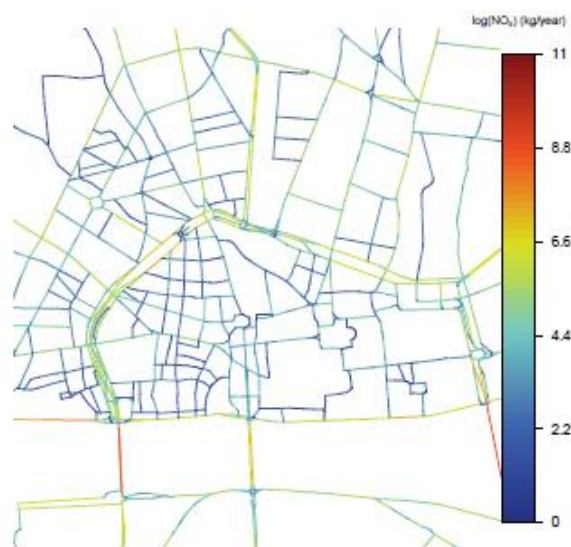


Figure 3 : Cadastre des émissions routières de NO_x (en échelle log) sur l'agglomération d'Orléans

Dans le cadre du forum européen FAIRMODE (fairmode.jrc.ec.europa.eu), un exercice d'inter-comparaison sur l'évaluation de la représentativité spatiale des stations mesure a été organisé en 2016 par le JRC (Centre Commun de Recherche de la Commission Européenne). Le cas test défini pour cet exercice concerne l'agglomération d'Anvers (Belgique). Le jeu de données fourni aux participants comprend en particulier des données de modélisation urbaine obtenues par VITO avec les logiciels IFDM et OSPM (rues canyons) (Figure 4, Figure 6). Il inclut également des données d'émission (Figure 5, Figure 7).

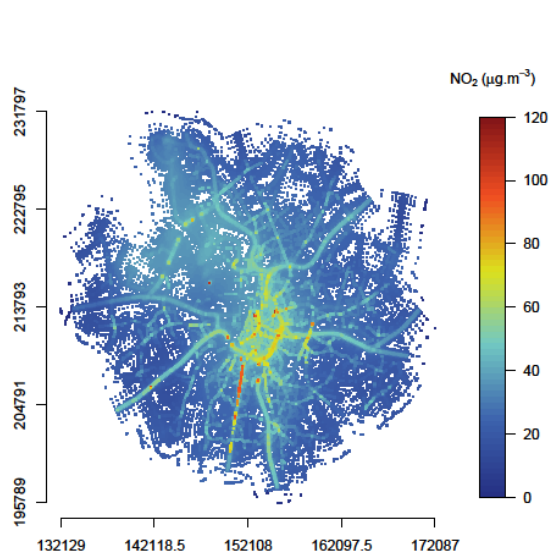


Figure 4 : Moyenne annuelle de NO₂ modélisée par VITO avec IFDM et OSPM sur un ensemble points récepteurs pour l'année 2012

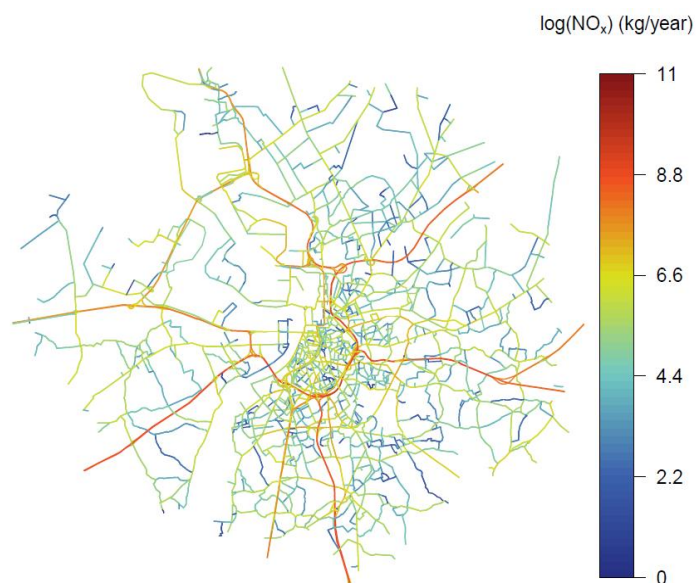


Figure 5 : Cadastre des émissions routières de NO_x (en échelle log) sur l'agglomération d'Anvers

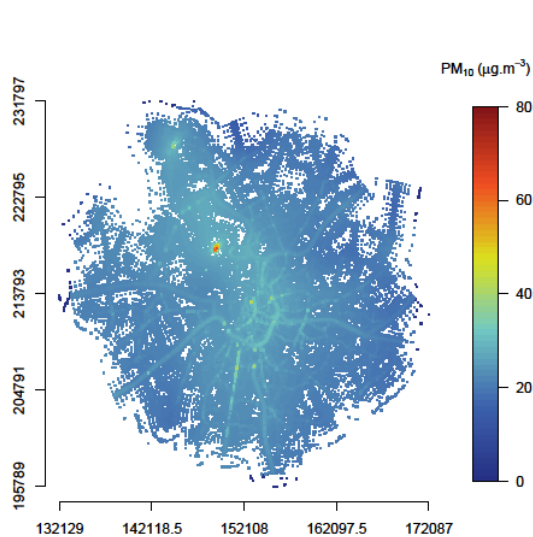


Figure 6 : Moyenne annuelle de PM₁₀ modélisée par VITO avec IFDM et OSPM sur un ensemble de points récepteurs pour l'année 2012

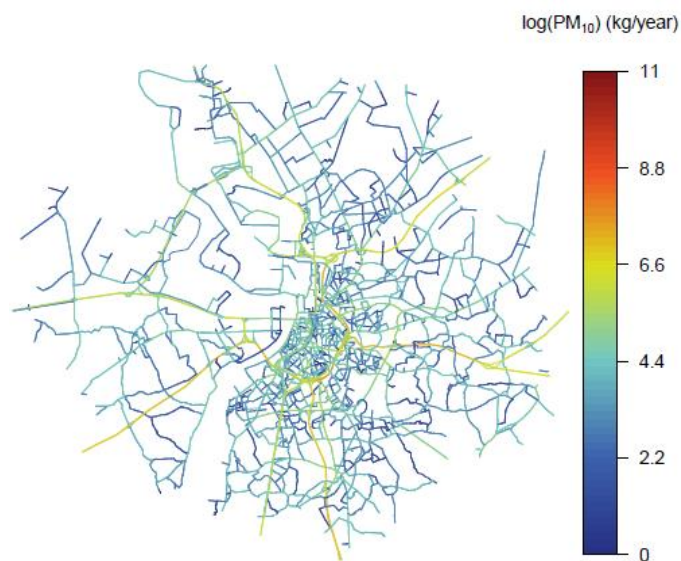


Figure 7 : Cadastre des émissions routières de PM₁₀ (en échelle log) sur l'agglomération d'Anvers

3. EVALUATION DES CARTOGRAPHIES ET DES PERFORMANCES

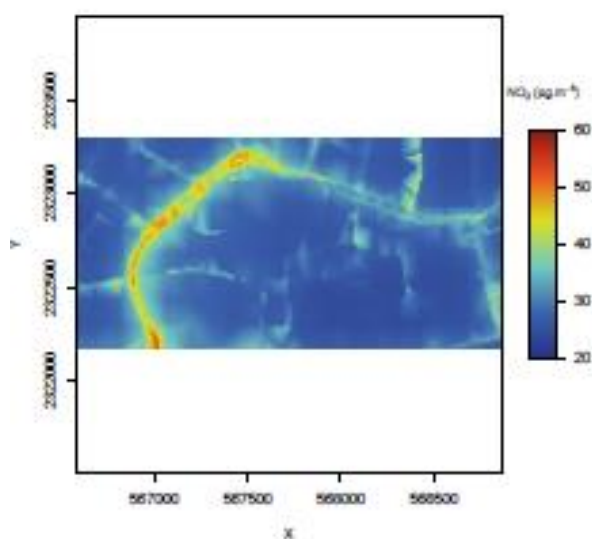
Dans le domaine de la qualité de l'air, il est courant d'utiliser l'interpolation dite « linéaire » pour interpoler les sorties de modèles urbains. Plus récemment, a été développée l'interpolation de Delaunay, qui se base sur un maillage de Delaunay construit sur les points récepteurs du modèle, associé à une hypothèse de décroissance linéaire des concentrations le long des arêtes des triangles du maillage. On peut se référer notamment aux travaux conduits par Air Pays de la Loire, en partenariat avec l'IRSTV (voir par exemple Fortin et al., 2014). Le rapport LCSQA intitulé *Estimation des populations exposées aux dépassements de seuils réglementaires - Echelle urbaine* (Beauchamp et al., 2014) a montré des similarités fortes du point de vue des performances et du rendu cartographique entre cet interpolateur et l'interpolateur linéaire.

Enfin, le krigeage (ordinaire) est parfois utilisé par les AASQA pour interpoler les sorties de modèles urbains, bien que les hypothèses classiques du krigeage ne soient pas adaptées pour gérer les fortes non-stationnarités liées aux concentrations à proximité du trafic. Des krigeages à partir de campagnes par tubes à diffusion passive, dans lesquels les sorties de modèles urbains sont utilisées en dérive externe, ont aussi déjà été utilisés (ASPA, 2014).

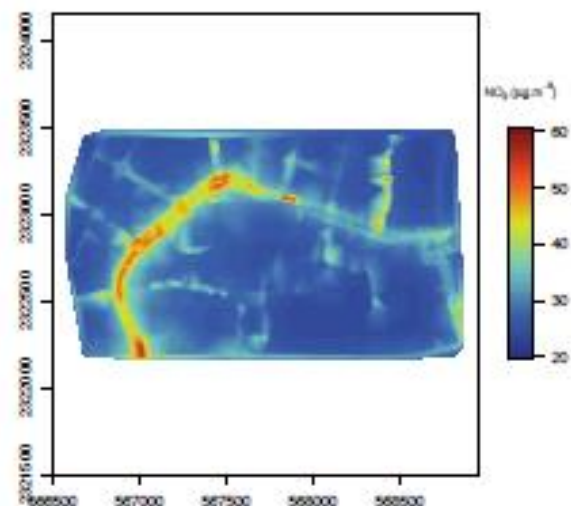
3.1 Cartographies

La Figure 8 présente le rendu cartographique de ces trois interpolateurs sur les données d'Orléans. Elle fait clairement apparaître le lissage trop prononcé du krigeage ordinaire et la grande similarité entre l'interpolation linéaire et l'interpolation de Delaunay. Des différences peuvent toutefois être constatées entre ces deux dernières approches, principalement liées à des effets de bords dans le cas de Delaunay et à l'influence de la résolution utilisée pour l'interpolation linéaire (5m sur l'agglomération d'Orléans).

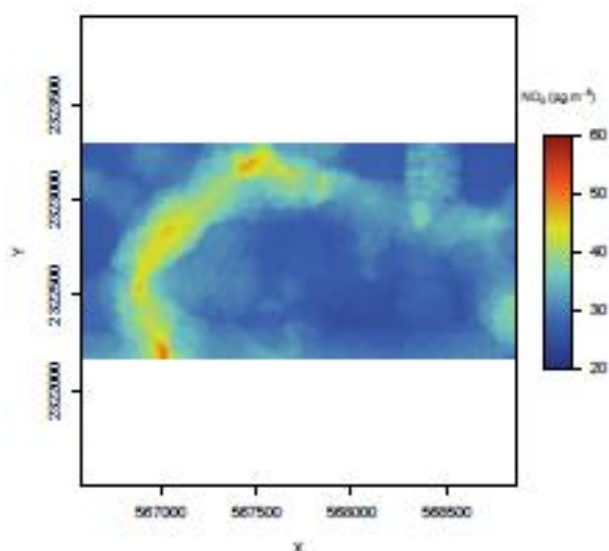
La construction de la cartographie (Figure 8d) qui utilise le modèle de krigeage avec dérive externe polynomiale introduit en section 1 a nécessité le développement d'un programme informatique couplant C++ pour la programmation du système linéaire associée à ce nouveau krigeage, R pour la manipulation des données d'entrée et de sortie, et OGR/GDAL pour la gestion des fichiers SIG associés au cadastre des émissions. Le langage R, classiquement utilisé par le LCSQA et un grand nombre d'AASQA pour des applications géostatistiques, n'a pu être utilisé seul pour la mise en place de cette méthode car il n'est pas adapté à la manipulation de ce type de méthodes numériques, spécialement lorsque le krigeage a recours à un voisinage glissant qui implique l'inversion de grandes matrices à chaque estimation. Ces programmes se basent néanmoins sur des outils standards et ils pourront être mis à la disposition des AASQA sur demande. Ils permettent, dans un temps CPU très raisonnable de construire une cartographie qui intègre la totalité de l'information disponible sur le réseau routier : en chaque point du maillage, la distance à chacune des routes du cadastre est calculée, et le système linéaire associé est construit et résolu afin de fournir une estimation de la concentration moyenne annuelle en NO₂.



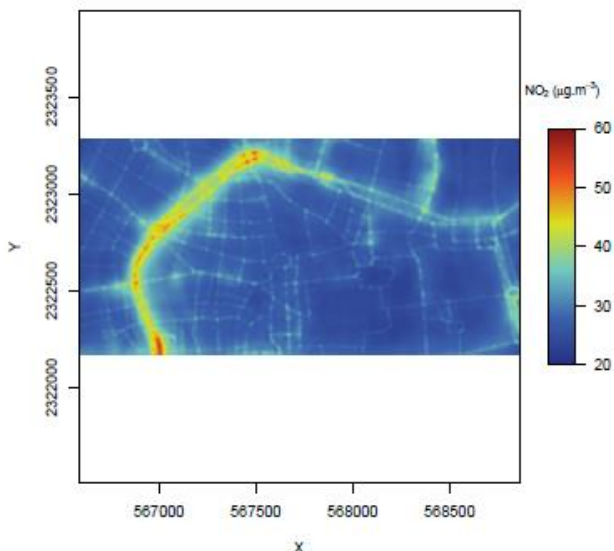
(a) Interpolation linéaire



(b) Interpolation de Delaunay



(c) Krigeage ordinaire



(d) Krigeage avec dérive polynomiale

Figure 8 : Cartographie de la moyenne annuelle de NO₂ sur l'agglomération d'Orléans (année 2010). Spatialisation sur une grille de 5m de résolution des données modélisées par ADMS-Urban avec 4 méthodes d'interpolation différentes (linéaire, Delaunay, krigeage ordinaire, krigeage avec dérive polynomiale).

Sur cette cartographie, même les axes routiers qui ne sont pas représentés par les points récepteurs ADMS apparaissent, grâce à la prise en compte du cadastre des émissions dans l'interpolation. On attirera l'attention sur le fait que la répartition des points récepteurs doit prendre en compte le plus de situations différentes dans la zone d'interpolation. Il est préférable d'utiliser moins de points récepteurs mais représentatifs de plus de situations de pollution que l'inverse (voir section 3.2).

En termes de surface en dépassement de la valeur limite ($40\mu\text{g.m}^{-3}$ en moyenne annuelle NO_2), la méthode d'interpolation proposée a l'avantage d'être moins dépendante de l'échantillonnage spatial des données que ne le sont les interpolateurs linéaires et de Delaunay. Ainsi, alors que la décroissance des concentrations est modélisée de manière physiquement cohérente par la dérive polynomiale et son comportement pseudo-exponentiel à de faibles distances, un trop faible échantillonnage spatial en transect des axes aboutit à une surestimation des concentrations à proximité de l'axe avec des hypothèses de variation linéaire entre les points ; C'est particulièrement visible en comparant les Figures 8b) et 8d).

Les cartes produites à partir des données disponibles sur l'agglomération d'Anvers montrent le même type de comportement.

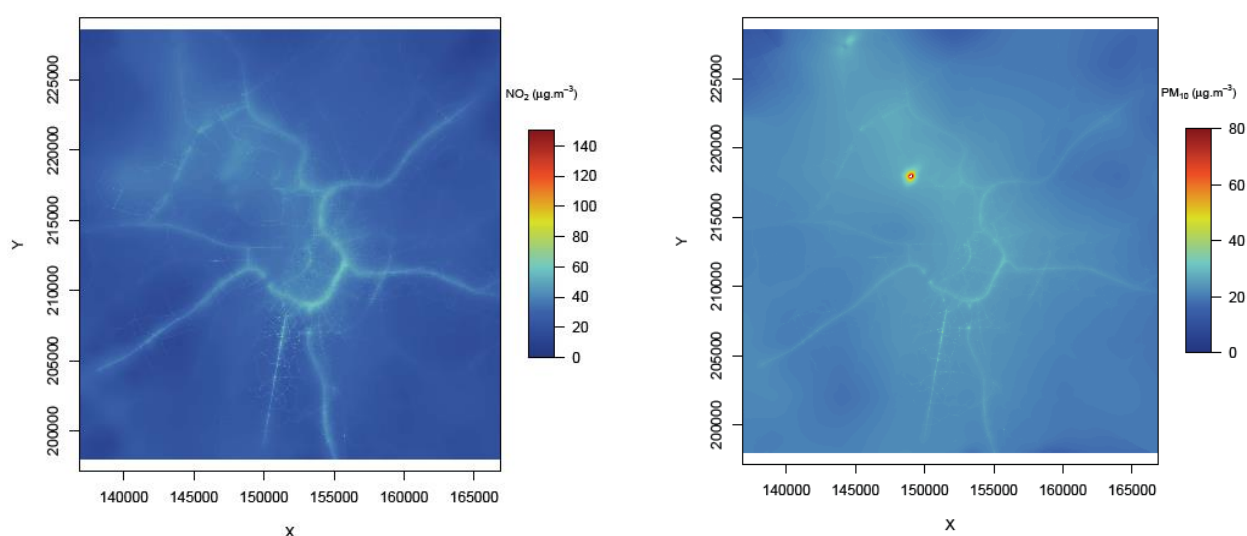


Figure 9 : Cartographie des moyennes annuelles de NO_2 et PM_{10} sur l'agglomération d'Anvers (année 2012). Spatialisation sur une grille de 20m de résolution des données modélisées par IFDM-OSPM avec la nouvelle approche par krigeage, qui introduit une dérive polynomiale.

3.2 Scores de performance

Des scores de performance calculés à partir d'une validation croisée sur les données d'Orléans ont permis d'identifier la nouvelle approche comme étant la meilleure pour interpoler les sorties d'un modèle urbain (Figure 10 et Figure 11).

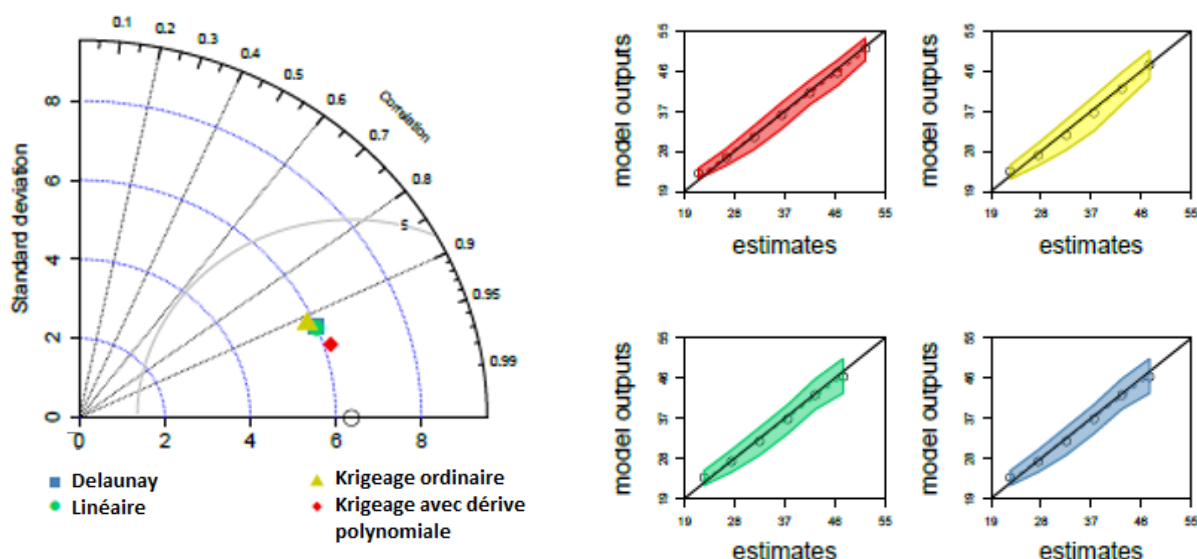


Figure 10 : Diagramme de Taylor et nuages de points valeurs interpolées – valeurs modélisées (par classes de concentrations estimées)

La Figure 11 permet aussi de constater que même en réduisant fortement le nombre de points récepteurs dans le modèle (jusqu'à environ 500 sur Orléans), le krigeage avec dérive polynomiale présente une corrélation et une RMSE quasi-équivalentes à celles qui sont obtenues avec le jeu de points récepteurs initiaux. De plus, ces scores sont toujours nettement supérieurs à ceux obtenus avec les trois autres interpolateurs, quel que soit le nombre de points récepteurs utilisés.

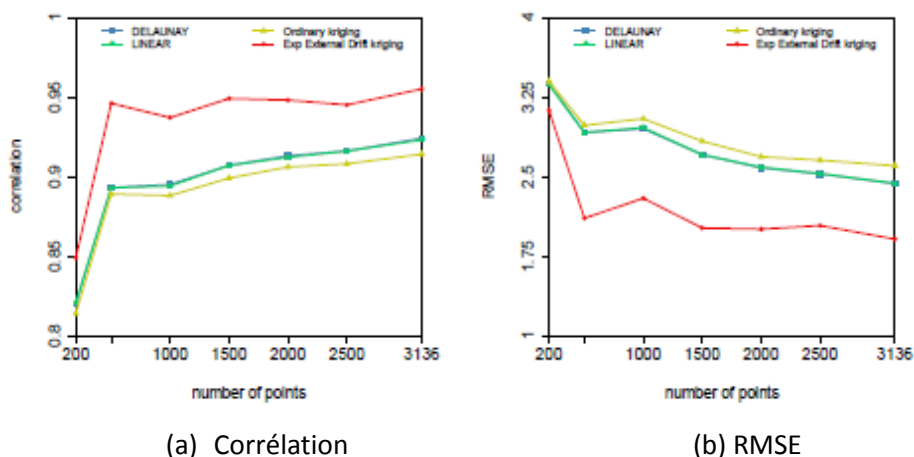


Figure 11 : Corrélation et RMSE calculées pour les 4 interpolateurs testés et différents nombres de points récepteurs

3.3 Points de vigilance

En plus des sorties de modélisation, FAIRMODE avait également mis à disposition un jeu de données simulant une campagne d'échantillonnage par tubes NO_2 à partir de 341 points récepteurs sur lesquels un bruit aléatoire (mais cohérent avec ce type de données de mesure) a été ajouté aux sorties de modélisation. Une cartographie a pu être construite en utilisant le krigeage avec dérive polynomiale. L'intérêt de cet exercice est de montrer qu'il est possible de spatialiser des données de campagnes en une seule étape qui ne dissocie pas les typologies des sites de mesures (fond et proximité), contrairement à la méthodologie développée antérieurement pour l'exploitation des données de tubes (Beauchamp et al., 2011 ; Malherbe et Cárdenas, 2005).

La méthode a toutefois montré ses limites car peu de transects figuraient parmi ces 341 points. Ainsi, la décroissance n'est pas correctement modélisée par la dérive, ce qui aboutit à une surestimation des niveaux de concentration à proximité des axes routiers à faible valeur d'émission (voir Figure 13). Un meilleur échantillonnage spatial des tubes passifs aurait probablement permis d'améliorer la qualité de la cartographie et de se rapprocher des résultats obtenus en Figure 9.

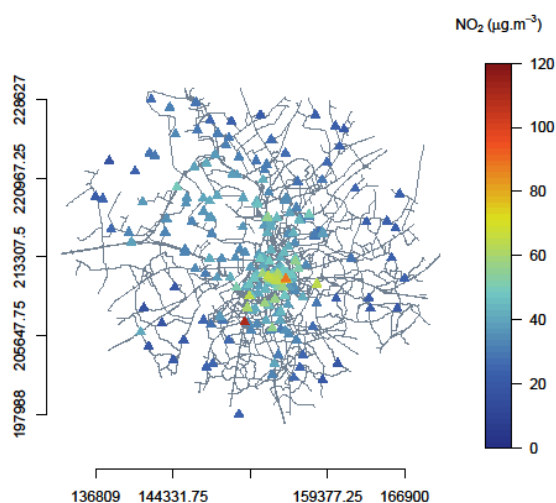


Figure 12 : Campagne fictive par tubes à diffusion passive NO_2

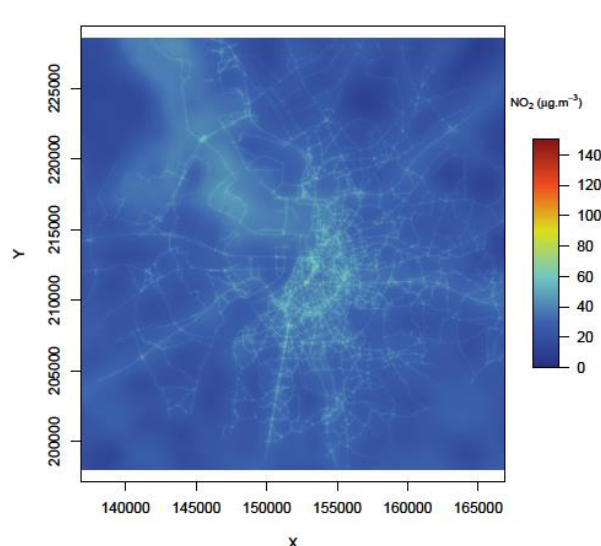


Figure 13 : Cartographie de la moyenne annuelle de NO_2 sur l'agglomération d'Anvers (année 2012). Spatialisation sur une grille de 20m de résolution des moyennes issues des tubes NO_2 fictifs avec la nouvelle approche par krigeage, qui introduit une dérive polynomiales

4. CONCLUSION

Une nouvelle méthodologie a été développée pour réaliser des cartographies à partir de données de modélisation urbaine. Fondée sur le krigeage, elle permet de prendre en compte la non-stationnarité des concentrations au voisinage des routes sous la forme d'une dérive polynomiale fonction des émissions routières. Palliant ainsi les défauts du krigeage ordinaire, elle se révèle sensiblement plus performante que les deux interpolateurs recommandés jusqu'alors (linéaire, Delaunay). La précision du résultat final n'en reste pas moins liée à la répartition spatiale des points de simulation qui doivent couvrir les différentes situations de pollution rencontrées sur le domaine. Plus complexe à mettre en œuvre que les deux interpolateurs précités, cette nouvelle approche s'ajoute à la liste des méthodes d'interpolation recommandées.

Le krigeage avec dérive polynomiale peut aussi s'appliquer à des données d'échantillonnage issues de campagnes par tubes. Il permet de traiter en une seule étape d'estimation l'ensemble des points de mesure. Là encore, la bonne répartition de ces derniers (avec un nombre suffisant de points disposés perpendiculairement aux routes) joue un rôle fondamental dans la qualité du résultat.

5. REFERENCES

ASPA. Evaluation de la qualité de l'air sur l'agglomération mulhousienne. Rapport technique, ASPA, 2014. URL www.atmo-alsace.net/medias/produits/Evaluation_de_la_qualite5.pdf.

N. Baldwin, O. Gilani, S. Raja, S. Batterman, R. Ganguly, P. Hopke, V. Berrocal, T. Robins, and S. Hoogterp. Factors Affecting Pollutant Concentrations in the Near-Road Environment. *Atmospheric Environment*, May 2015.

M. Beauchamp, L. Malherbe et L. Létinois. Estimation de l'exposition des populations aux dépassements de seuils réglementaires. Rapport technique, LCSQA, 2014. URL <http://www.lcsqa.fr>.

M. Beauchamp, L. Malherbe et L. Létinois. Application de méthodes géostatistiques pour la détermination de zones de représentativité en concentration et la cartographie des dépassements de seuils. Rapport technique, LCSQA, 2011. URL <http://www.lcsqa.fr>.

J.N. Cape, Y.S. Tang, N. van Dijk, L. Love, M.A. Sutton, and S.C.F. Palmer. Concentrations of ammonia and nitrogen dioxide at roadside verges, and their contribution to nitrogen deposition. *Environmental Pollution*, 1320 (3):0 469 – 478, 2004. ISSN 0269-7491. doi: <http://dx.doi.org/10.1016/j.envpol.2004.05.009>. URL <http://www.sciencedirect.com/science/article/pii/S0269749104001952>.

D.J. Carruthers, H. E. C. M., and R. Singles. Development of ADMS-Urban and comparison with data for urban areas in the UK. Proc. of Air Pollution Modelling and its Application xii. Technical report, CERC, 1998.

J.P. Chiles and P. Delfiner. *Geostatistics : modelling spatial uncertainty*. Wiley, New-York, second edition, 2012.

Directive 2008/50/EC. *Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe*, 2008.

T. Faus-Kessler, M. Kirchner, and G. Jakobi. Modelling the decay of concentrations of nitrogenous compounds with distance from roads. *Atmospheric Environment*, 420 (19):0 4589 – 4600, 2008. ISSN 1352-2310. doi: <http://dx.doi.org/10.1016/j.atmosenv.2008.01.073>. URL <http://www.sciencedirect.com/science/article/pii/S1352231008001167>.

A. Font, T. Baker, I.S. Mudway, E. Purdie, C. Dunster, and G.W. Fuller. Degradation in urban air quality from construction activity and increased traffic arising from a road widening scheme. *Science of The Total Environment*, 497-498:0 123–132, 2014. ISSN 0048-9697. doi: <http://dx.doi.org/10.1016/j.scitotenv.2014.07.060>.

URL <http://www.sciencedirect.com/science/article/pii/S0048969714010900> .

N. Fortin, E. Bocher, J. Picaut, G. Petit, and G. Dutilleux. An opensource tool to build urban noise maps in a GIS. YVERDON-LES-BAINS, Switzerland, October 2012.

N. L. Gilbert, S. Woodhouse, D.M. Stieb, and J.R. Brook. Ambient nitrogen dioxide and distance from a major highway. *Science of The Total Environment*, 3120 (1-3):0 43–46, 2003. ISSN 0048-9697. doi: [http://dx.doi.org/10.1016/S0048-9697\(03\)00228-6](http://dx.doi.org/10.1016/S0048-9697(03)00228-6).

URL <http://www.sciencedirect.com/science/article/pii/S0048969703002286> .

L. Malherbe et G. Cárdenas. Application des méthodes géostatistiques pour l'exploitation conjointe des mesures de fond et de proximité. Rapport technique, LCSQA, 2005. URL <http://www.lcsqa.fr>.

CH. Monn, V. Carabias, M. Junker, R. Waeber, and H.U Wanner. Small-scale spatial variability of particulate matter <10m (PM₁₀) and nitrogen dioxide. *Atmospheric Environment*, 31(15):0 2243–2247, 1997.

T.M. Naser, I. Kanda, T. Ohara, K. Sakamoto, S. Kobayashi, H. Nitta, and T. Nataami. Analysis of traffic-related {NO_x} and {EC} concentrations at various distances from major roads in japan. *Atmospheric Environment*, 430 (15):0 2379 – 2390, 2009. ISSN 1352-2310. doi: <http://dx.doi.org/10.1016/j.atmosenv.2009.02.002>.

URL <http://www.sciencedirect.com/science/article/pii/S1352231009001216> .

H. Omre. Bayesian kriging - Merging observations and qualified guesses in kriging. *Mathematical Geology*, 190 (1):0 25–39, 1987. ISSN 0882-8121. doi: 10.1007/BF01275432 . URL <http://dx.doi.org/10.1007/BF01275432> .

H. Pleijel, G. Pihl Karlsson, and E. Binsell Gerdin. On the logarithmic relationship between NO₂ concentration and the distance from a highroad. *Science of The Total Environment*, 3320 (1-3):0 261–264, 2004. ISSN 0048-9697. doi: <http://dx.doi.org/10.1016/j.scitotenv.2004.03.020>.

URL <http://www.sciencedirect.com/science/article/pii/S0048969704001986> .

M. C. Roorde-Knape, N.A.H. Janssen, J. de Hartog, P.H.N. Van Vliet, H. Harssema, and B. Brunekreef. Traffic related air pollution in city districts near motorways. *Science of The Total Environment*, 2350 (1-3):0 339–341, 1999. ISSN 0048-9697. doi: [http://dx.doi.org/10.1016/S0048-9697\(99\)00217-X](http://dx.doi.org/10.1016/S0048-9697(99)00217-X).

URL <http://www.sciencedirect.com/science/article/pii/S004896979900217X> .

S. Seguret, L. Malherbe et G. Perron. Estimation du dioxyde d'azote dans la vallée de la Thur. Rapport technique, LCSQA, 2003. URL <http://www.lcsqa.fr>.

M.P Taylor, L.J Kristensen, P.J Davies, J. Csavina, A.K. Mackay, N.C Munksgaard, and K.A. Hudson-Edwards. Omissions about the sources of contaminant emissions and depositions - A reply to comments on Taylor, M.P., Davies, P.J., Kristensen, L.J., Csavina, J., 2014. Licenced to pollute but not to poison: The ineffectiveness of regulatory authorities at protecting public health from atmospheric arsenic, lead and other contaminants resulting from mining and smelting operations. *Aeolian Research*, 17:0 205–213, June 2015.

X. Zou, Z. Shen, T. Yuan, S. Yin, J. Cai, L. Chen, and W. Wang. Shifted power-law relationship between NO₂ concentration and the distance from a highway: A new dispersion model based on the wind profile model. *Atmospheric Environment*, 400 (40):0 8068 – 8073, 2006. ISSN 1352-2310. doi: <http://dx.doi.org/10.1016/j.atmosenv.2006.09.012>. URL <http://www.sciencedirect.com/science/article/pii/S1352231006009496> .