

CS MODÉLISATION

Calcul d'exposition : rappel des travaux LCSQA 2014-2015 et nouveaux développements

*Maxime Beauchamp,
Laure Malherbe*

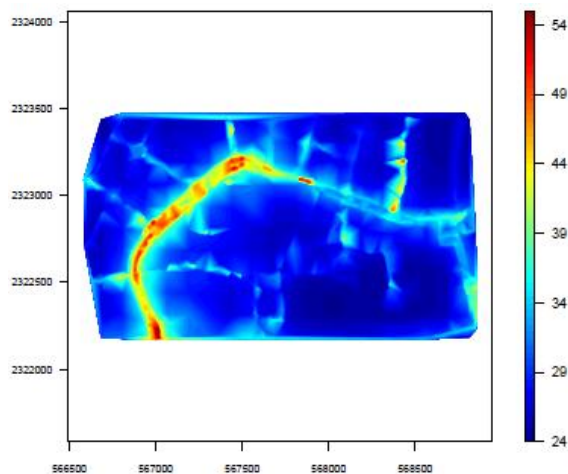
À l'INERIS, site d'Hauteville.
Jeudi 7 juillet 2016

- **Octobre 2012**: séminaire LCSQA sur la répartition spatiale de la population
- **Septembre 2013** : séminaire LCSQA sur l'exposition des populations aux dépassements de valeurs réglementaires
 - ↳ Besoin d'harmonisation
 - ↳ Développement d'une méthodologie nationale confiée par le MEEM au LCSQA
- **2014: Travaux LCSQA sur la répartition spatiale de la population**
 - **Publication du rapport** : Létinois L. Méthodologie de répartition spatiale de la population. Rapport LCSQA 2014 disponible sur le site du LCSQA : <http://www.lcsqa.org/rapport/2014/ineris/methodologie-repartition-spatiale-population>
- **2014** : Travaux LCSQA sur l'interpolation des sorties de modèles urbains et les calculs d'exposition
 - **Publication du rapport** : Beauchamp M., Malherbe L. et Létinois L. *Estimation de l'exposition des populations aux dépassements de seuils réglementaires*. Rapport LCSQA 2014 disponible sur le site du LCSQA : <http://www.lcsqa.org/rapport/2014/ineris/estimation-populations-exposees-aux-depassements-seuils-reglementaires-1-echelle>
- **2015** : Note de synthèse méthodologique reprenant les travaux de 2014 étendus aux calculs d'exposition à l'échelle nationale/régionale
 - **Publication de la note de synthèse** : Beauchamp M., Malherbe L. et Létinois L., Drevet J. *Synthèse méthodologique sur la caractérisation des situations de dépassement de seuil : délimitation des zones de dépassement et estimation des populations et écosystèmes exposés*. Rapport LCSQA 2015 disponible sur le site du LCSQA : <http://www.lcsqa.org/rapport/2015/ineris/synthese-methodologique-caracterisation-situations-depassement-seuil-delimitatio>

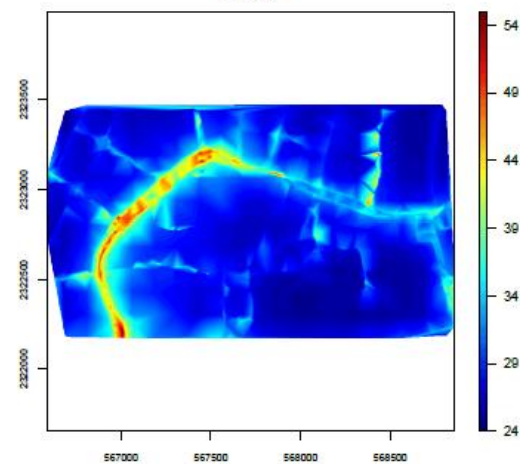
Rappels sur les travaux 2014

- **Objectif** : Evaluer différentes méthodes d'interpolation appliquées sur des sorties de modèles urbains
- **6 jeux de données fournis par les AASQA** : Bourges, Nantes, Niort, Orléans, Reims, Troyes.
- **3 interpolateurs testés** : linéaire, Delaunay et krigeage (ponctuel et de blocs avec des voisinages allant de 2 à 10 voisins) :
 - **Linéaire** : la valeur estimée est obtenue à partir d'une combinaison linéaire des observations aux sommets du triangle dans lequel le point est inclus.
 - **Delaunay** : Une triangulation de Delaunay est effectuée sur les observations, puis les valeurs de concentrations sont interpolées linéairement le long des arêtes des triangles pour définir des polygones où les concentrations sont comprises entre deux isovaleurs.
 - **krigeage** : la valeur estimée en un point est la moyenne pondérée des observations dans un voisinage restreint ou non, après ajustement de la fonction de covariance spatiale par un modèle admissible.

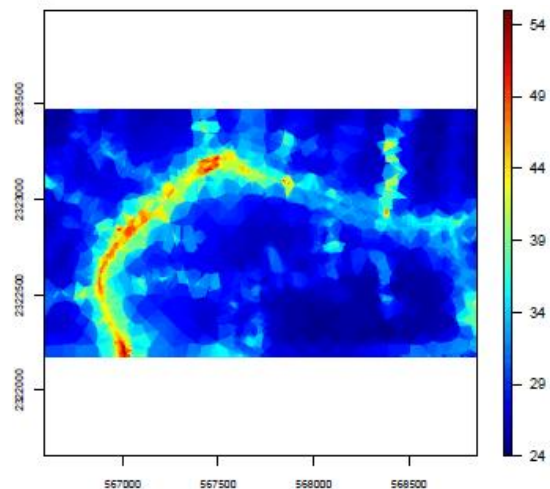
Delaunay



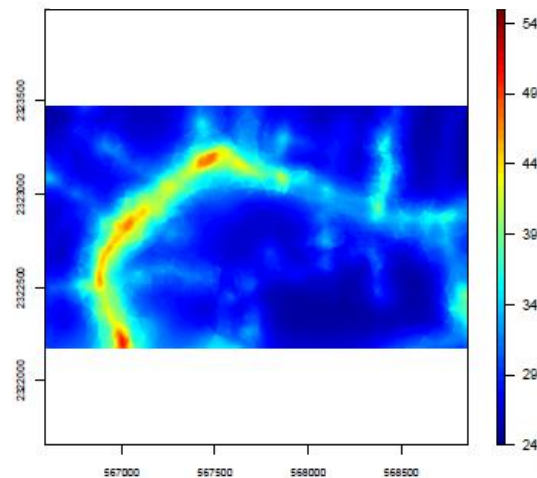
linéaire



krigeage à 2 voisins

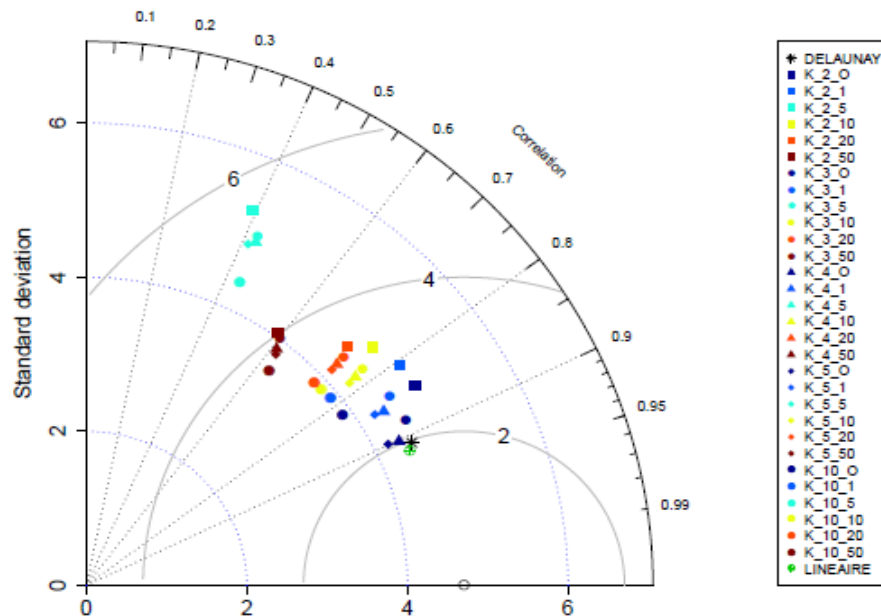


krigeage à 10 voisins



Cartographies obtenues après interpolation sur l'agglomération d'Orléans

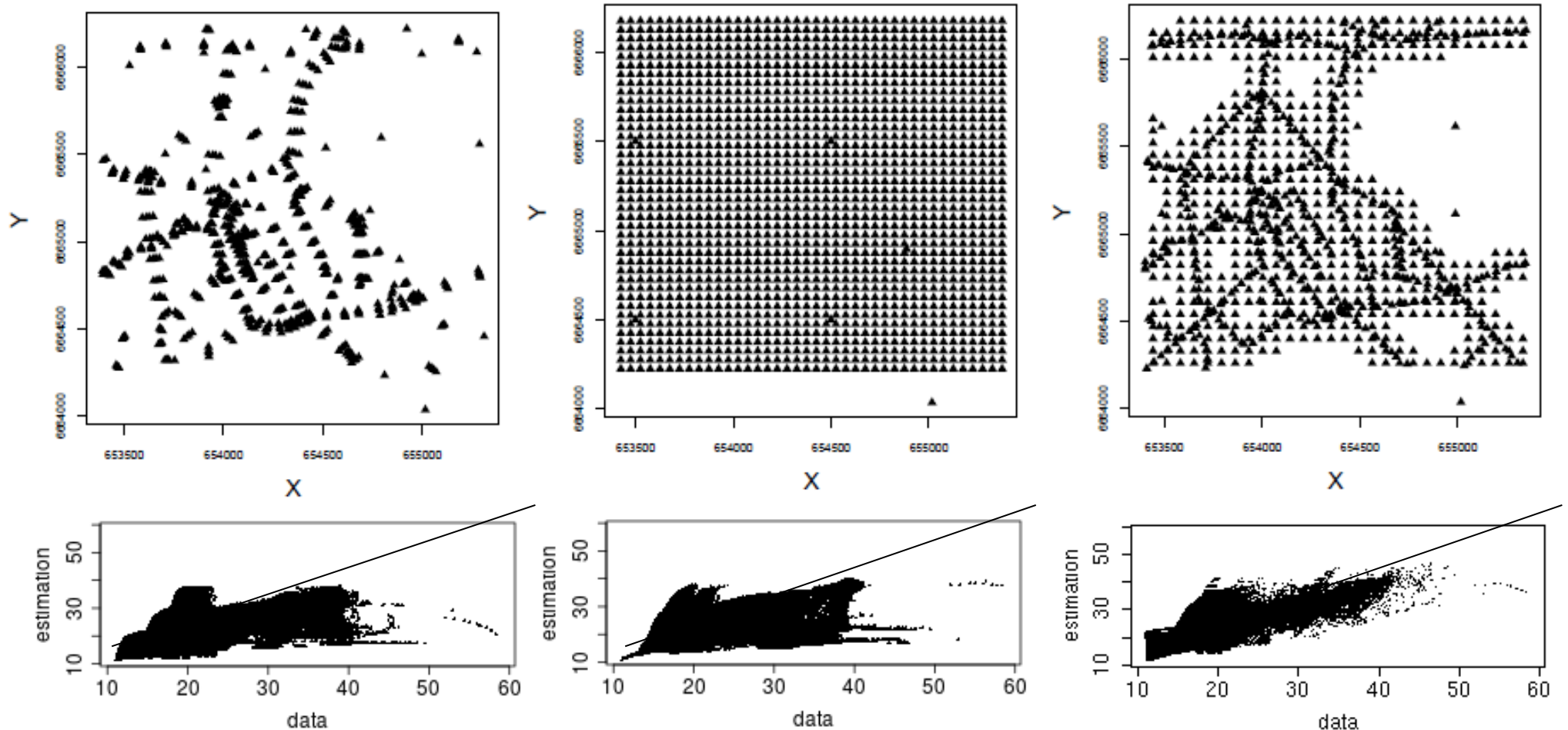
- 2 méthodes de validation (réalisées sur les zones où les concentrations sont les plus fortes) :
 - **Validation croisée** : on retire un point du jeu de données et on le ré-estime.
 - **Validation sur grille de référence** : le modèle fournit les concentrations sur une grille haute résolution et basse définition. La grille basse définition est utilisée pour interpoler les concentrations sur la grille haute définition.



- Les résultats de validation montrent que le krigeage n'est pas (encore) adapté à l'interpolation sur des données de typologies différentes (urbain / proximité automobile).
- Les scores en validation croisée montrent que l'interpolateur linéaire et l'interpolateur de Delaunay sont quasi-équivalents, ce qui est confirmé par la validation sur grille de référence. On notera que les scores de l'interpolateur linéaire sont tout de même très légèrement meilleurs que ceux de l'interpolateur de Delaunay.
- Les scores obtenus en validation croisée peuvent être très bons mais la validation sur grille de référence montrent une tendance générale à la sous-estimation des valeurs fortes.

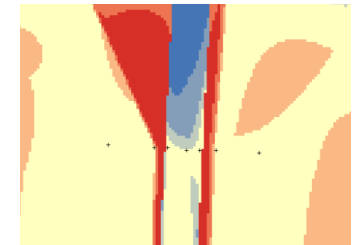
Du fait d'une différence très faible entre les méthodes de Delaunay et l'interpolateur linéaire, la suite de l'étude utilise uniquement l'interpolateur linéaire

- Cette tendance est plus liée au maillage utilisé qu'à l'interpolateur.
- Illustration avec les maillages utilisés dans PREVISION'AIR : meilleure est l'organisation du maillage et plus faible sera la tendance à la sous-estimation.



Maillages et Nuage de points en validation sur grille de référence (interpolateur linéaire) pour les maillages 1, 3 et 6 utilisés dans PREVISION'AIR

- **Tests d'optimisation du maillage** : Une attention particulière doit être portée à l'interpolation en proximité des axes routiers : il ne faut pas surestimer les concentrations en s'éloignant des axes et ainsi inclure dans le calcul d'exposition des bâtiments qui ne devraient pas l'être.
 - > placer des points récepteurs en bordure d'axe pour limiter la surestimation et privilégier des points récepteurs à faible distance de l'axe où la décroissance est la moins linéaire.
 - > mais il faut malgré tout estimer la zone en dépassement sur l'axe en vue du rapportage (même si aucune population n'est rattachée à cette zone)



- **La définition du maillage intelligent est dépendante du contexte local.**
 - > la procédure utilisée dans l'étude sur Bourges (cf. Rapport 2014) consistant à améliorer le maillage intelligent à partir d'une sortie de modèle haute résolution semble être une solution simple et générale applicable par tous. Afin de d'alléger les temps de calcul sur la sortie haute définition, seules quelques journées représentatives pourraient être simulées.

- **Croisement avec la population**

- **Quelle méthode ? Plusieurs méthodes étudiées:**

- Population projetée sur le même maillage que les données de modélisation (après interpolation) et croisement maille à maille
 - Intersection entre les mailles (ou polygones) en dépassement et les bâtiments habités. Comptage au prorata de la surface de bâtiment intersectée.

Cette méthode équivaut à la précédente puisque dans cette dernière, le nombre d'habitants par maille est évalué au prorata des surfaces de bâtiments intersectés (pour éviter les doubles comptes).

- Intersection entre les mailles (ou polygones) en dépassement et les bâtiments habités. La population totale de chaque bâtiment recoupé est prise en compte.

- **Méthode retenue** (cf. réunion de la CS du 25/11/2015):

- Population projetée sur le même maillage que les données de modélisation (après interpolation) et **croisement maille à maille**

équivalente à

Intersection entre les mailles (ou polygones) en dépassement et les bâtiments habités. Comptage au prorata de la surface de bâtiment intersectée.

Exception possible: dans des zones de taille réduite, avec un habitat de type maison, la population totale des maisons peut être comptabilisée.

- Pas de buffer appliqué à la zone de dépassement

Rappels sur les travaux 2015

- Synthèse des travaux déjà réalisés par le LCSQA sur la définition d'une méthodologie pour l'estimation des populations exposées aux dépassements de seuils réglementaires
- Ajout d'orientations méthodologiques pour l'estimation de la végétation et des écosystèmes exposés.
- Les recommandations du LCSQA sont présentées :
 - par échelle spatiale (régionale, urbaine)
- et - par type de seuil (protection de la santé, protection de la végétation et des écosystèmes).

L'ozone, les PM_{10} et le NO_2 , polluants pour lesquels des dépassements de valeurs cibles et limites sont observés, sont principalement concernés.

Exemple de recommandations pour les seuils de protection de la végétation

Type de données	Origine possible	Description	Echelle/résolution
Carte spatialisée de l'AOT ₄₀ en moyenne sur les cinq dernières années	-Plateformes de modélisation régionales <u>OU</u> - Plateforme de modélisation nationale PREV'AIR.	Carte de l'AOT ₄₀ , en moyenne sur 5 ans ou pour une année civile, construite à partir des sorties horaires de modélisation régionale corrigées à l'aide des données d'observation.	Ordre de grandeur de la résolution spatiale: ≤ 10 km
Carte des zones de végétation et écosystèmes	Base de données Corine Land Cover d'occupation des sols <u>ET</u> Zones naturelles protégées identifiées par le LCSQA dans la méthodologie Zones sensibles (2010)	Sont incluses dans les zones de végétation les classes Corine Land Cover 2 (toutes les sous-classes de niveau 2, 3 (toutes les sous-classes de niveau 3), 4 (toutes les sous-classes de niveau 3) et 5 jusqu'à la sous-classe 5.2.2 incluse (classe 5.2.3 exclue). A ces classes, sont ajoutées les zones naturelles protégées identifiées dans la méthodologie Zones sensibles (Malherbe, 2010) et appartenant aux catégories suivantes : protection de biotope, réserve naturelle nationale, réserve naturelle de Corse, parc national (zone de cœur et d'adhésion), parc naturel régional.	—

Recommandations additionnelles pour :

- *Cas des dépassements liés à des sources industrielles*
- *Calcul des statistiques et Interpolation*

Approche 1 (conseillée) : spatialiser les données d'entrée horaires ou journalières nécessaires au calcul de la statistique, calcul de la statistique, calcul d'exposition

Approche 2 (déconseillée) : calcul de la statistique sur les sorties de modèles, spatialisation de la statistique, calcul d'exposition

→ Test à effectuer par des AASQA volontaires si besoin car l'approche 1 est plus coûteuse que l'approche 2

- *Format de sortie :*

Pour chaque situation de dépassement identifiée par la mesure fixe, quelle que soit l'échelle spatiale, les AASQA devront fournir :

- une cartographie décrivant la zone de dépassement (contours et surface) ;
- si le seuil dépassé est un objectif environnemental pour la protection de la santé humaine, une estimation du nombre d'habitants inclus dans la zone de dépassement ;
- si le seuil dépassé est un objectif environnemental pour la protection de la végétation, une estimation de la surface de végétation/d'écosystème incluse dans la zone de dépassement ;

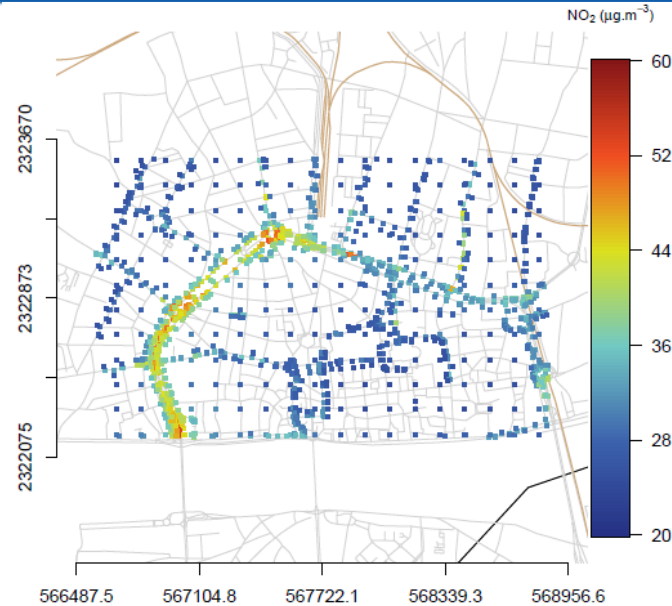
Le format de sortie (ASCII/géoréférencés) sera défini après discussion en CS EMTD.

Travaux 2015-2016

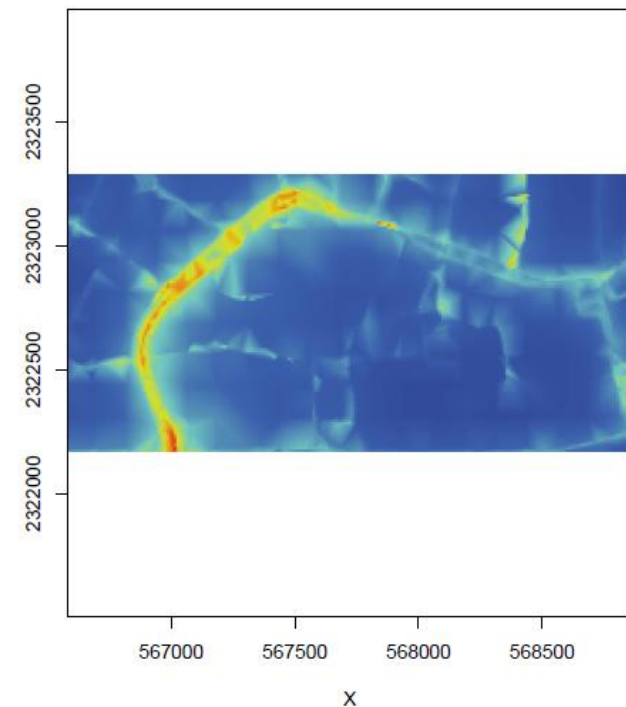
Amélioration du krigage pour l'interpolation

LCSQA Amélioration du krigeage pour l'interpolation

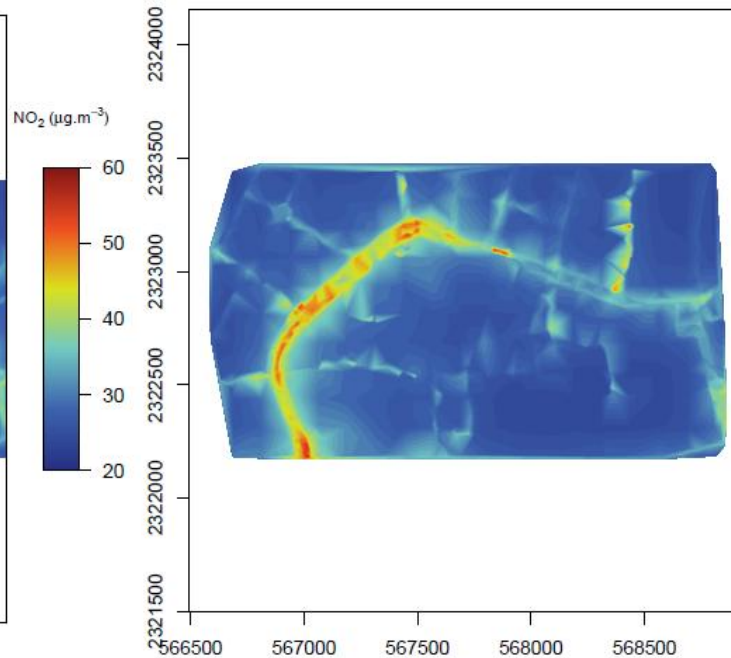
Précédemment : Utilisation de l'interpolation linéaire ou Delaunay pour interpoler les sorties de modèles urbains. Le krigeage, dans ses hypothèses classiques, n'est pas satisfaisant et fait apparaître un lissage trop important



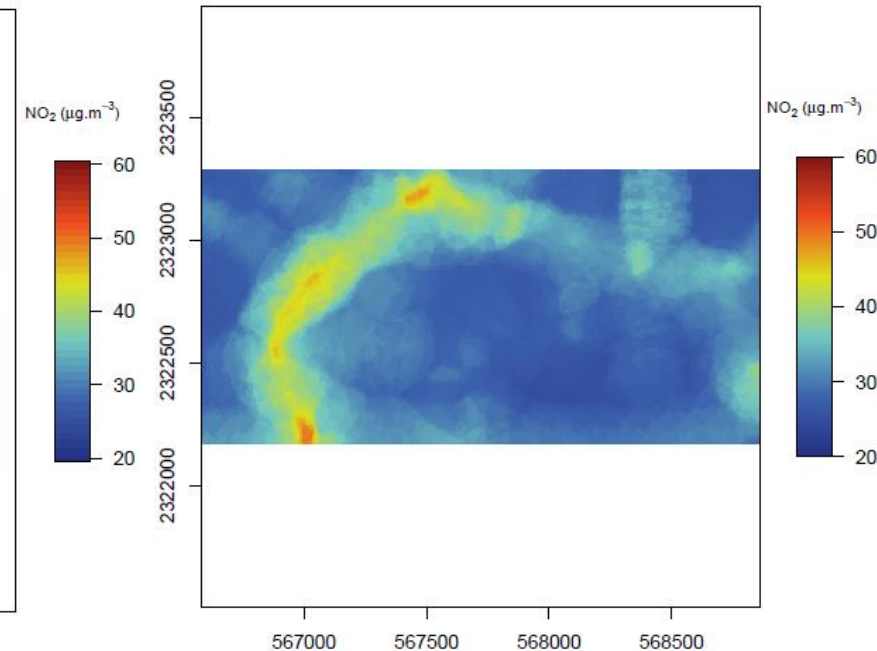
Sorties ADMS



Interpolation linéaire



Interpolation de Delaunay



Krigeage Ordinaire

Idée : utiliser le cadastre des émissions routières et construire un modèle statistique adapté en entrée du krigage

- Un point s situé en (x,y) peut être influencé par k routes, notées R^k , qui amène un surplus S^k au point s

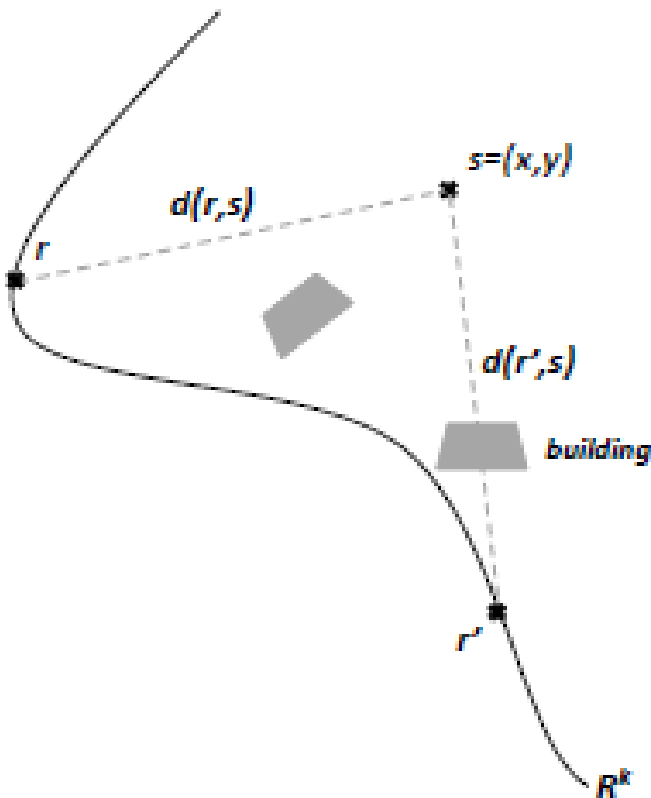
- Chaque route R^k est décomposée en un terme intégral, où chaque point r de la route amène un surplus de pollution au point s

$$S^k(x) = \int_{R^k} f(S(r), d(r, x)) dr$$

- Ce surplus dépend de la distance à la route (hypothèse simplifiée qui ne prend pas en compte les obstacles que sont les bâtiments et la chimie) et des émissions au point r de la route

- On introduit une hypothèse exponentielle de décroissance à la route :

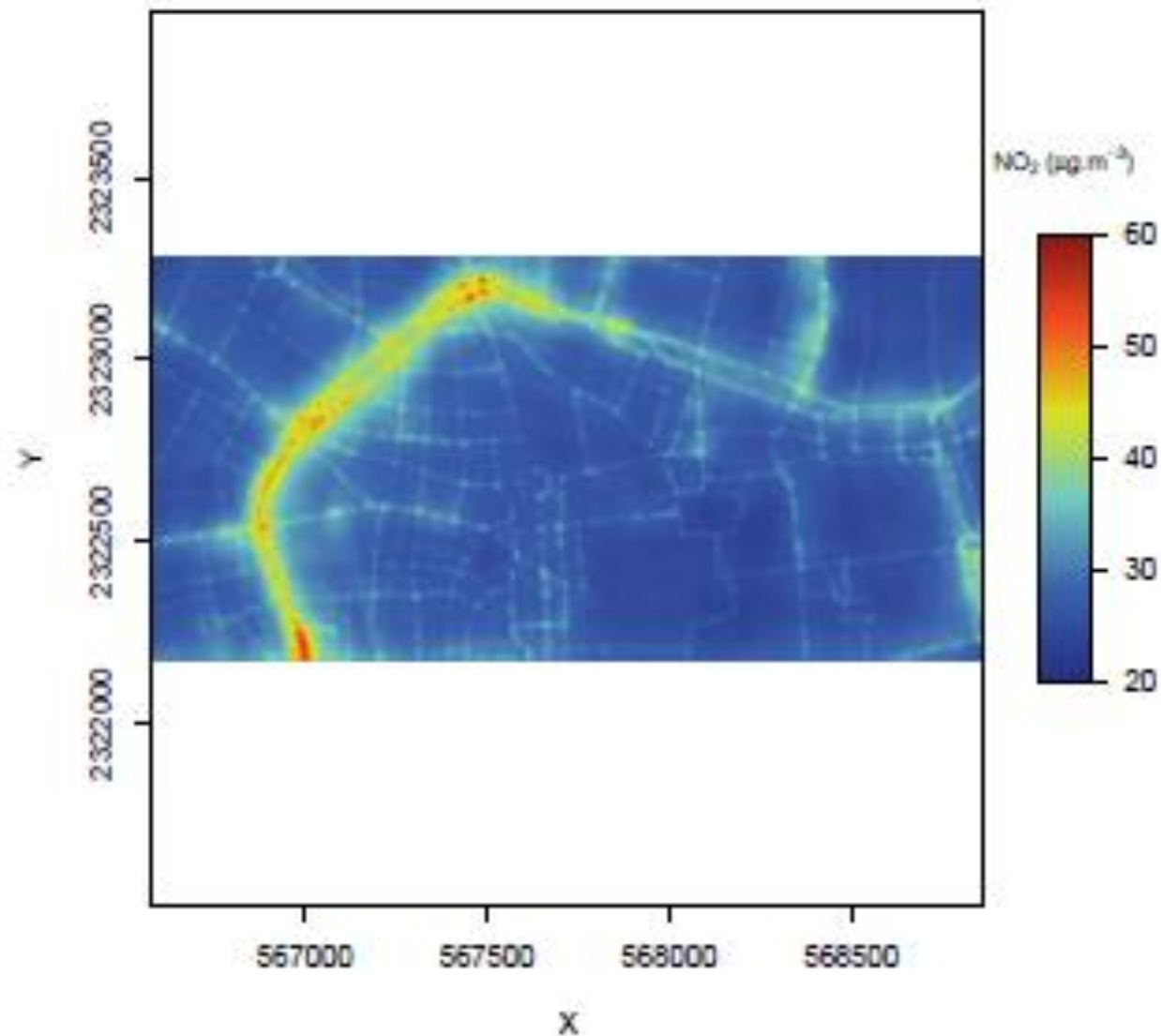
$$f(x) = \frac{AE_k}{|R^k|} e^{-\alpha d(r,x)}$$



- On cherche à ne pas fixer le paramètre α dans la fonction exponentielle afin qu'il s'ajuste localement en fonction de la fenêtre de voisinage utilisé à l'estimation
- Pour cela, on utilise le développement en série de la fonction exponentielle qui la linéarise et autorise l'utilisation d'une dérive externe :

$$S(x) = \sum_{k=1}^N \int_{R^k} \sum_{u=0}^{+\infty} \frac{AE_k}{|R^k|} \frac{(-\alpha d(r, x))^u}{u!} dr$$

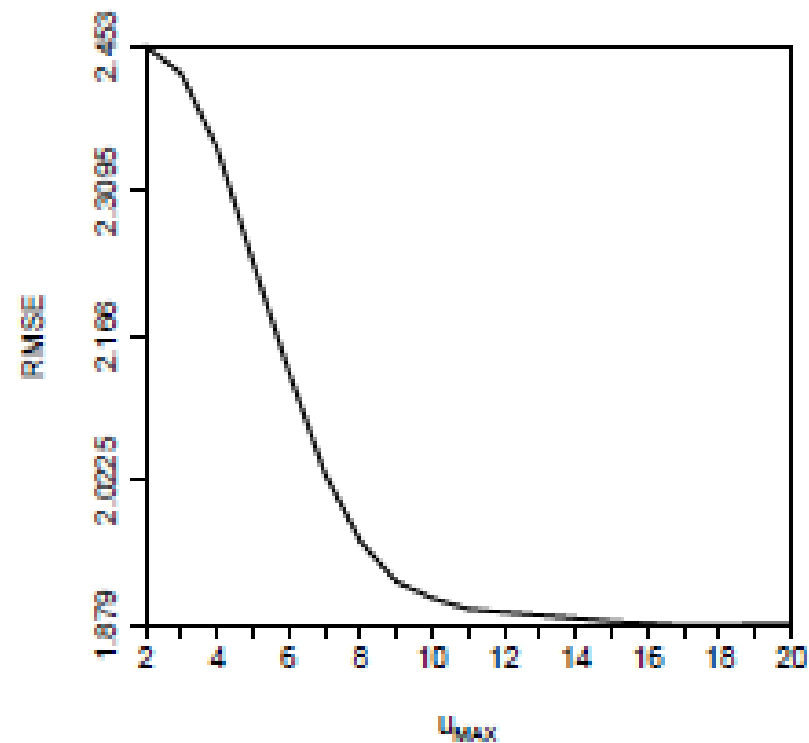
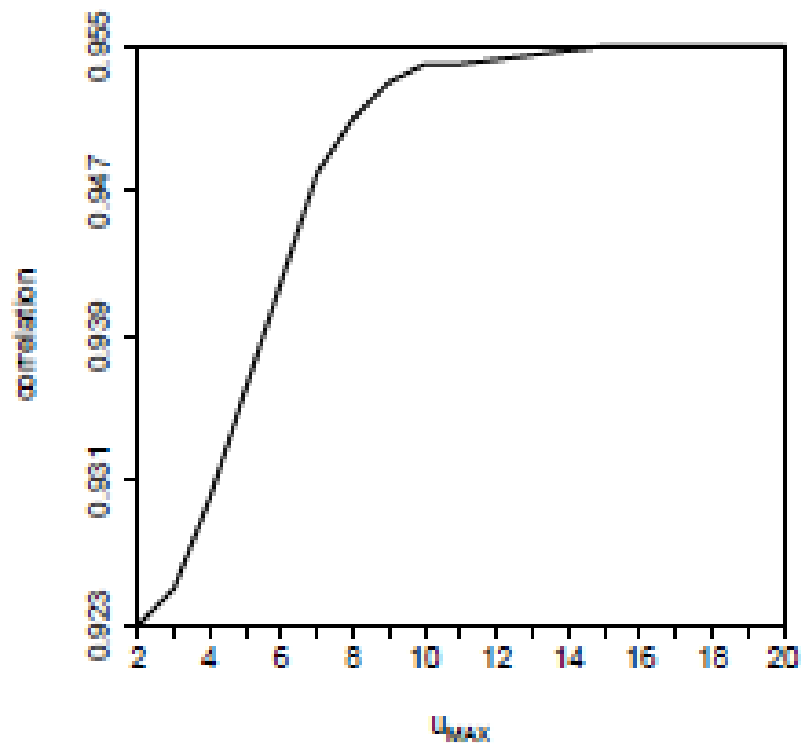
- Pour calculer l'estimateur sur toute la grille (centre ville d'Orléans à 5m de résolution), un couplage R/C++/GDAL a été utilisé (temps CPU très efficace) :
 - Le programme C++ importe le cadastre des émissions via la librairie GDAL et procède à la discrétisation des routes et au calcul des distances
 - R est utilisé pour l'ajustement du modèle géostatistique (variogramme des résidus)
 - Le programme C++ reprend la main pour résoudre les équations du krigeage



Krigeage avec dérive externe exponentielle

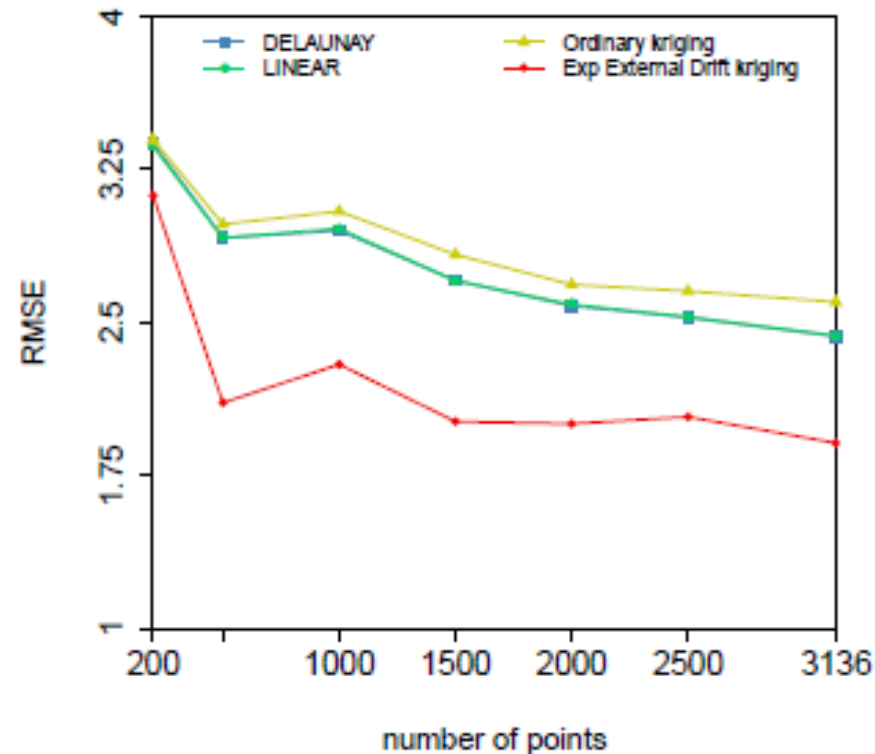
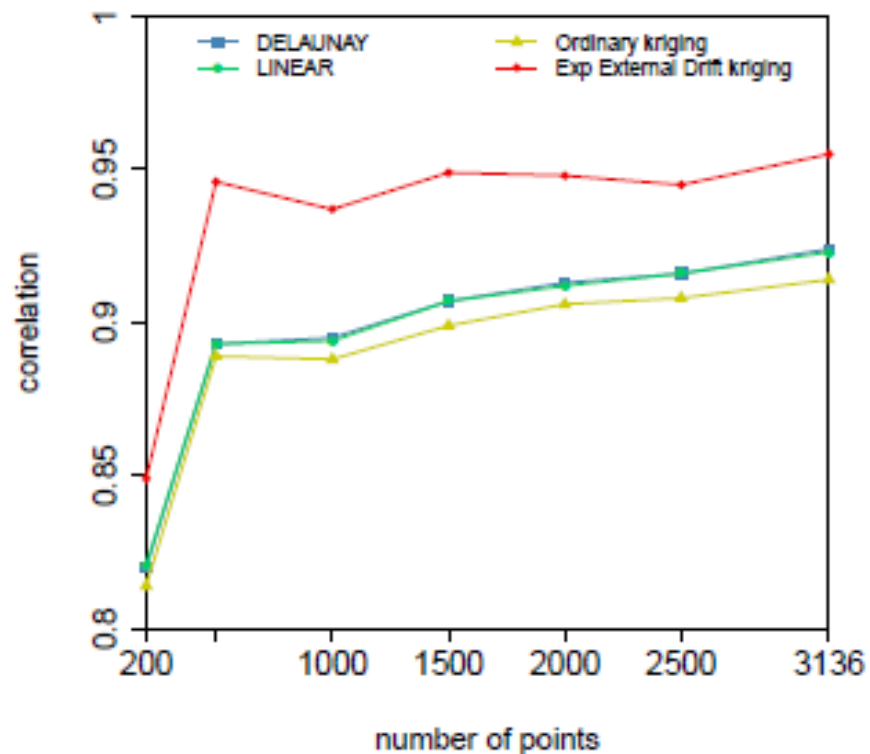
- Résultats de validation croisée

- les scores (corrélation / RMSE) s'améliorent lorsque l'ordre de la troncature de l'exponentielle augmente
- la limite de performance est atteinte pour $u_{\text{MAX}}=15$



- Résultats de validation croisée

- les scores (corrélation / RMSE) sont bien meilleurs avec le krigeage qui introduit une dérive exponentielle
- si on diminue le nombre de points, les scores de la nouvelle méthode sont stables jusqu'à N=500 points récepteurs (3136 dans la simulation ADMS originale)



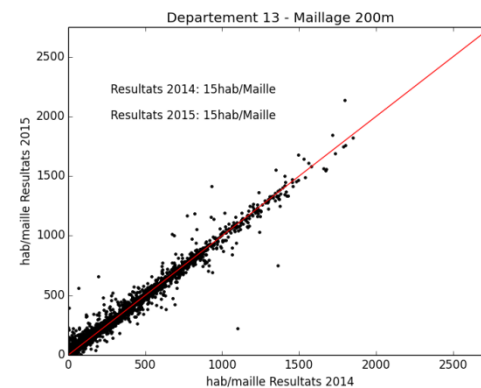
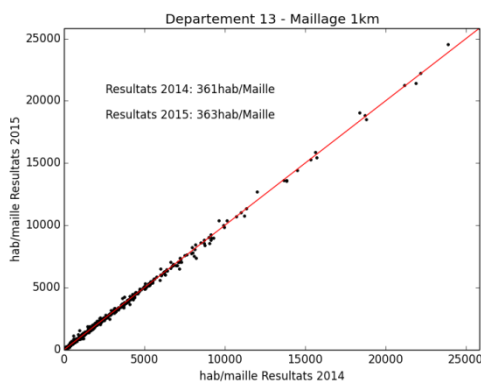
Comparaison des chiffres d'exposition LCSQA/AASQA

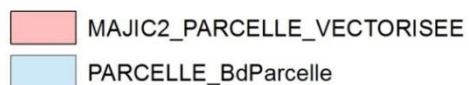
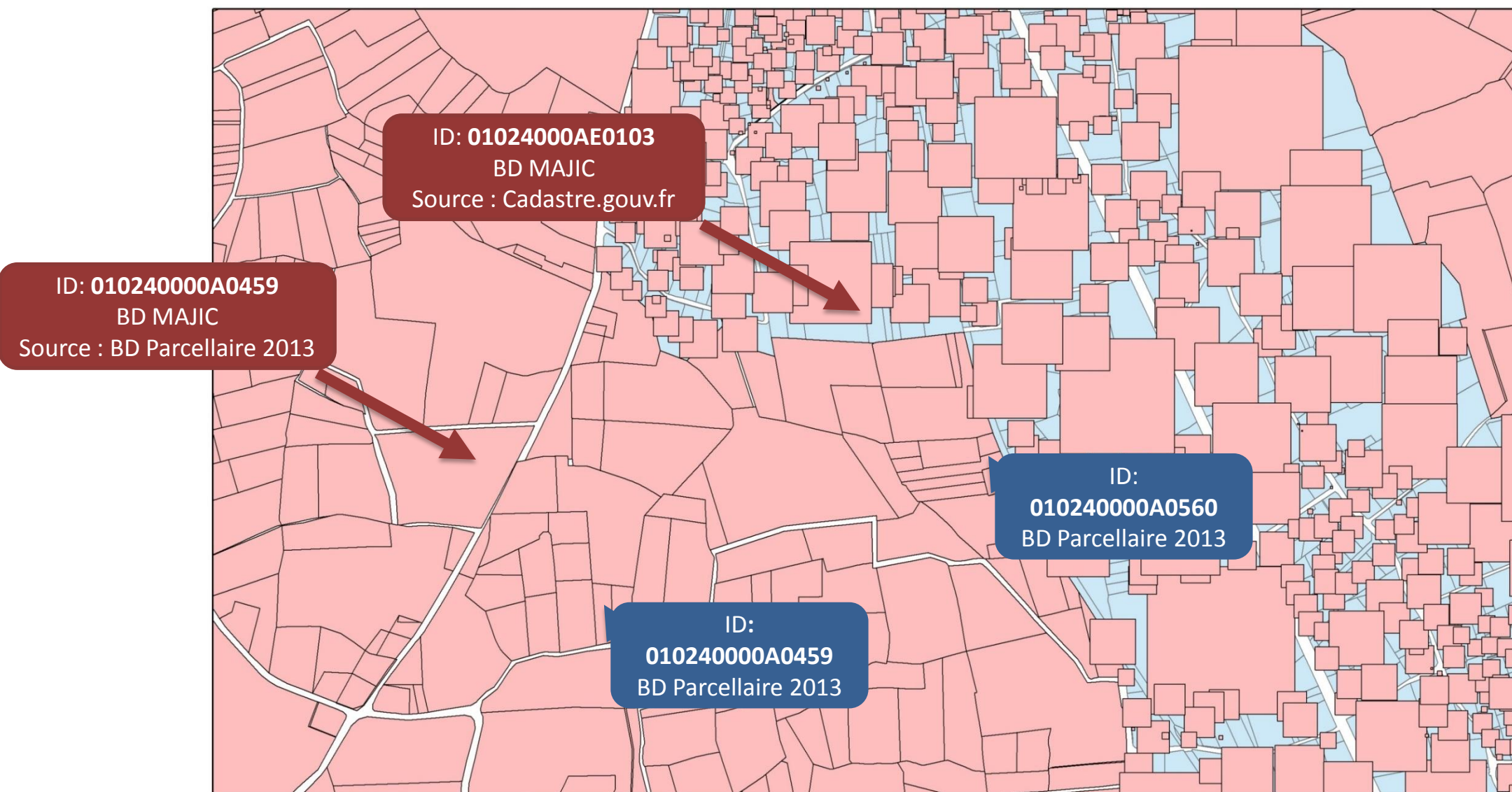
- Comparaison des chiffres d'exposition aux dépassements de valeurs limites, d'objectifs long terme et valeurs cibles calculés par le LCSQA pour l'échelle nationale et par les AASQA pour l'échelle régionale.
- Les valeurs cibles et objectifs long terme O_3 ainsi que la valeur limite annuelle PM_{10} sont considérés dans cette note. Pour des raisons pratiques, les données de rapportage 2014 des AASQA n'étant pas encore toutes disponibles au moment de la rédaction de cette note, la période d'étude porte sur les statistiques calculées pour l'année 2013.
- Structure de la note :
 - Les données utilisées pour calculer les chiffres d'exposition par le LCSQA et les AASQA sont rappelées.
 - Les chiffres d'exposition sont ensuite présentés et comparés pour chacun des polluants
 - Des explications possibles des différences sont fournies (construction des cartes analysées, prise en compte des composantes locales, données de population / zones de végétation et écosystèmes, etc.)

Mise à jour des données de population

- Début 2015: Fourniture des données de population spatialisées selon la méthodologie nationale (MAJIC) pour l'année de référence 2011, avec une note d'utilisation:
<http://www.lcsqa.org/rapport/2014/ineris/fourniture-validation-donnees-population-spatialisees-selon-methodologie-nationa>
- Avril 2016: Fourniture des données de population spatialisées, réactualisées pour l'année de référence 2012, avec une note présentant les mises à jour et comparant les données ancienne et nouvelle version.
<http://www.lcsqa.org/rapport/2015/ineris/fourniture-donnees-population-spatialisees-selon-methodologie-nationale-methode->

A l'échelle kilométrique, bonne cohérence entre les données 2011 et 2012 mais détection ponctuelle d'écarts sensibles du fait de parcelles manquantes → investigations complémentaires pour en rechercher les causes. Méthode de correction à définir.





Cas des grandes agglomérations

- **Discussions...**

Merci de votre attention

