

Note technique

VOISINAGE DE KRIGEAGE ET CARTES ANALYSEES

Maxime Beauchamp, Laure Malherbe (Ineris)

SYNTHESE

La réalisation de cartes analysées est un outil important dans le système de prévision national de la qualité de l'air PREV'AIR. Les observations du réseau de mesures et les sorties du modèle de chimie-transport CHIMERE sont intégrés par des méthodes géostatistiques afin de fournir les cartographies les plus vraisemblables possible de la qualité de l'air. Ces méthodes géostatistiques, couramment appelées krigeage, font intervenir des paramètres de voisinage qui peuvent modifier fortement l'aspect final des cartes. Dans cette courte note, nous discuterons des performances obtenues avec les différents voisinages de krigeage testés et des problèmes qui sont susceptibles d'en découler

1. LE KRIGEAGE DANS LE SYSTEME PREV'AIR

La réalisation de cartes analysées vise à offrir une représentation aussi vraisemblable que possible de la qualité de l'air. Dans le système PREV'AIR, ces cartes sont construites à partir d'un krigeage des observations avec les sorties du modèle de chimie-transport CHIMERE en dérive externe¹.

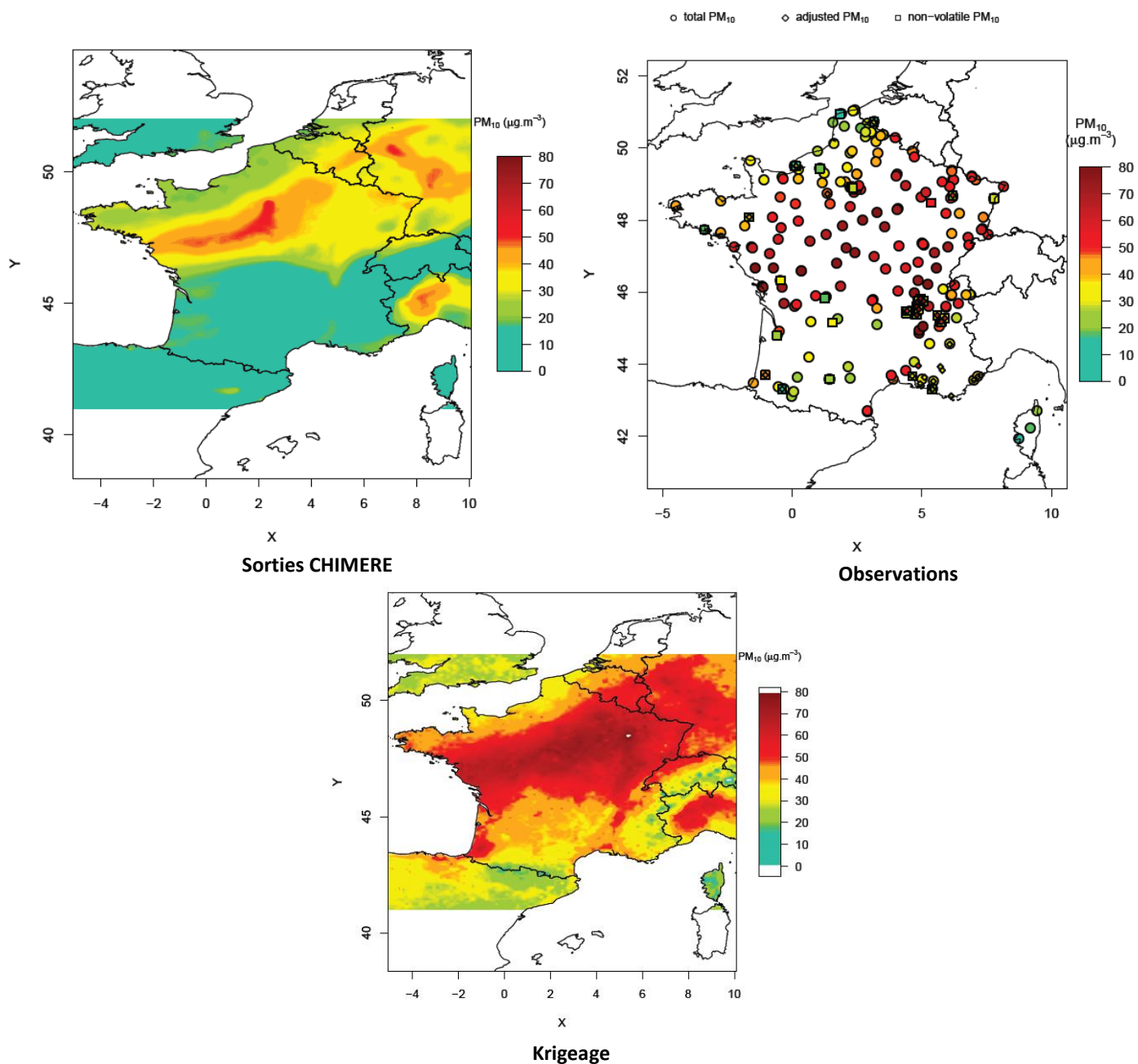


Figure 1 : Krigeage en dérive externe des concentrations de PM₁₀ le 21 mars 2015

¹ Cf. Rapport LCSQA (Malherbe et Ung, 2009). Travaux relatifs à la plate-forme nationale de modélisation PREV'AIR : Réalisation de cartes analysées d'ozone (2/2).

Le système de krigeage en dérive externe est défini comme suit :

Krigeage en Dérive externe (KDE)

$$\begin{bmatrix} \gamma_{\alpha\beta} & V_{\alpha} \\ V_{\beta}^t & 0 \end{bmatrix} \begin{bmatrix} \lambda_{\beta} \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma_{\alpha 0} \\ V_0 \end{bmatrix} \quad (1)$$

où $\gamma_{\alpha\beta}$ est la valeur du variogramme des résidus en validation croisée de la régression entre les observations (Z_{α}) et le modèle CHIMERE (V_{α}). μ est le multiplicateur de Lagrange dans la procédure d'optimisation.

L'estimation Z_0^* au point x_0 et son écart-type d'estimation sont obtenus après résolution du système linéaire (1) et l'obtention des poids de krigeage λ_{β} .

$$\begin{aligned} Z_0^* &= \begin{bmatrix} Z_{\alpha} \\ 0 \end{bmatrix}^t \times \begin{bmatrix} \lambda_{\alpha} \\ \mu \end{bmatrix} \\ \sigma_K^2 &= \gamma_{00} + \begin{bmatrix} \lambda_{\alpha} \\ \mu \end{bmatrix}^t \times \begin{bmatrix} \gamma_{\alpha 0} \\ V_0 \end{bmatrix} \end{aligned} \quad (2)$$

2. VOISINAGE DE KRIGEAGE

Le système de krigeage peut être résolu en voisinage unique (on prend alors toutes les stations du réseau disponibles pour estimer la valeur du polluant en x_0) ou en voisinage glissant (on ne prend que les stations les plus proches de x_0 , dans un voisinage défini en distance ou en nombre). Le voisinage glissant, bien qu'impliquant moins de stations de mesure à chaque estimation, est plus coûteux en temps de calcul que le voisinage unique car la matrice de variance-covariance définie en (1) doit alors être inversée à chaque nouvelle estimation.

Le système PREV'AIR utilise actuellement un voisinage glissant en nombre fixé à 80 stations. Nous avons cherché à déterminer s'il ne serait pas préférable de mettre en place un voisinage unique, afin de limiter les temps de calculs lors de la prochaine mise à jour du système PREV'AIR (production d'une analyse journalière calculée comme la moyenne de 24 analyses horaires).

Si le gain en temps de calcul s'est révélé important avec un facteur de réduction du temps CPU de l'ordre de 4, les scores de performance du krigeage en voisinage unique se sont révélés légèrement moins bons qu'en voisinage glissant, tant en corrélation qu'en erreur quadratique moyenne (voir Figure 2). Le voisinage unique peut donc potentiellement être à l'origine d'un taux de détection des dépassements plus faible qu'en voisinage glissant ainsi que d'un taux de fausses alertes plus élevé.

Les tests réalisés sur les voisinages en distance ont révélés qu'il était difficile de trouver une distance qui optimiserait à la fois les temps de calcul et les scores de performance, qui sont d'une journée à une autre tantôt meilleurs que les scores obtenus avec le voisinage en nombre, tantôt moins satisfaisants. De plus, le voisinage en distance ne permet pas de fournir une estimation dans les zones où il n'y a aucune station de mesure, ce qui est problématique en termes de communication.

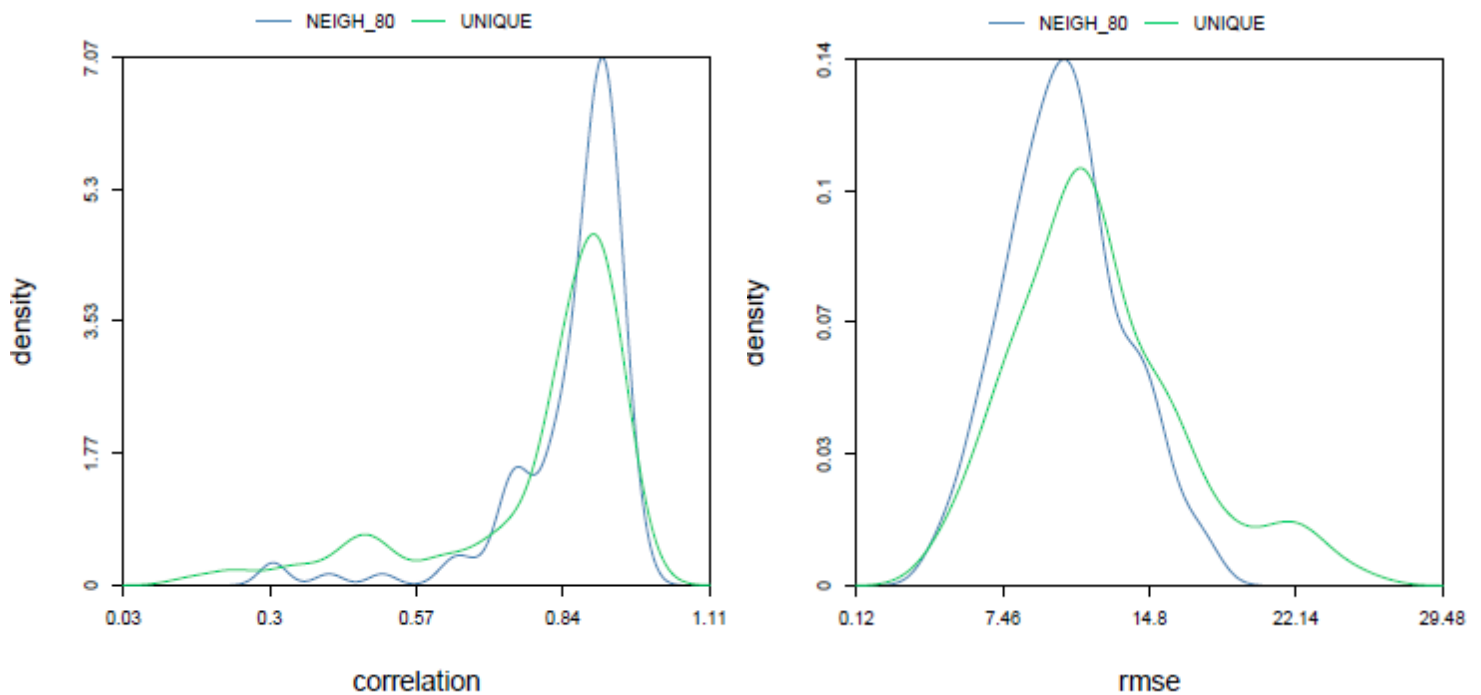


Figure 2 : Distribution (ajustée) des corrélations et des RMSE2 journalières sur l'année 2013 pour les PM10

3. SUR LE PROBLEME DES DISCONTINUITES EN VOISINAGE GLISSANT

Le voisinage glissant en nombre peut faire apparaître des discontinuités dans l'estimation liées à la sélection des stations qui diffère d'un nœud de grille à un autre. Ce problème ne peut être véritablement résolu qu'en utilisant un voisinage en distance et en intervenant directement dans le système de krigeage. Cette intervention consiste à ajouter un effet de pépité résiduel sur la diagonale de la matrice de variance-covariance, qui augmente avec la distance entre la station concernée et le nœud de grille.

Les travaux conduits par Gribov (2004) ou Rivoirard (2011) pourront être testés lors de travaux à venir pour prévenir ce problème.

² Root Mean Square Error : racine carrée de la moyenne des écarts quadratiques