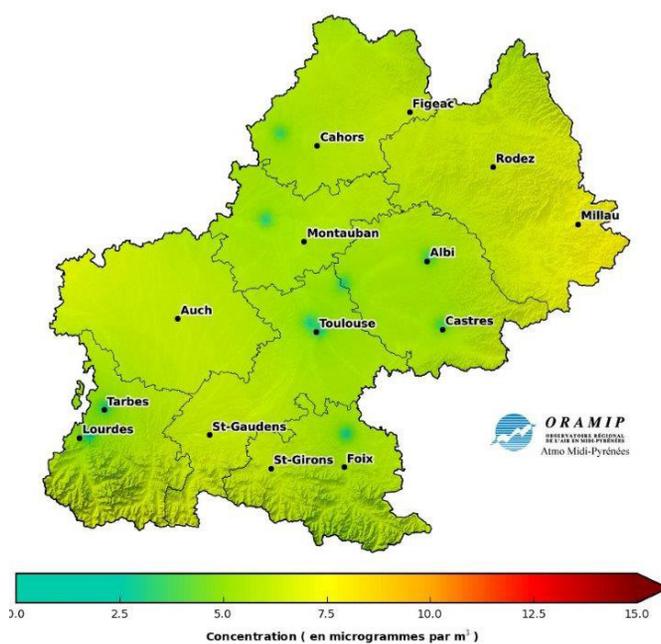




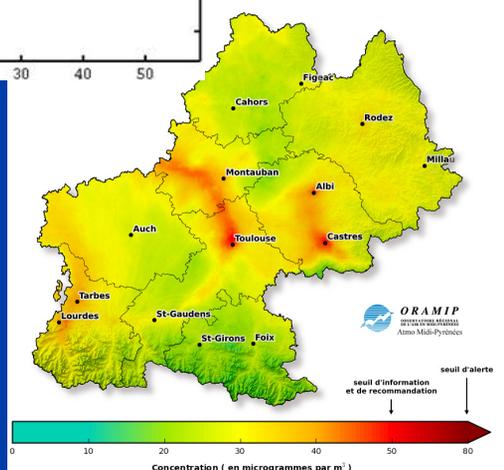
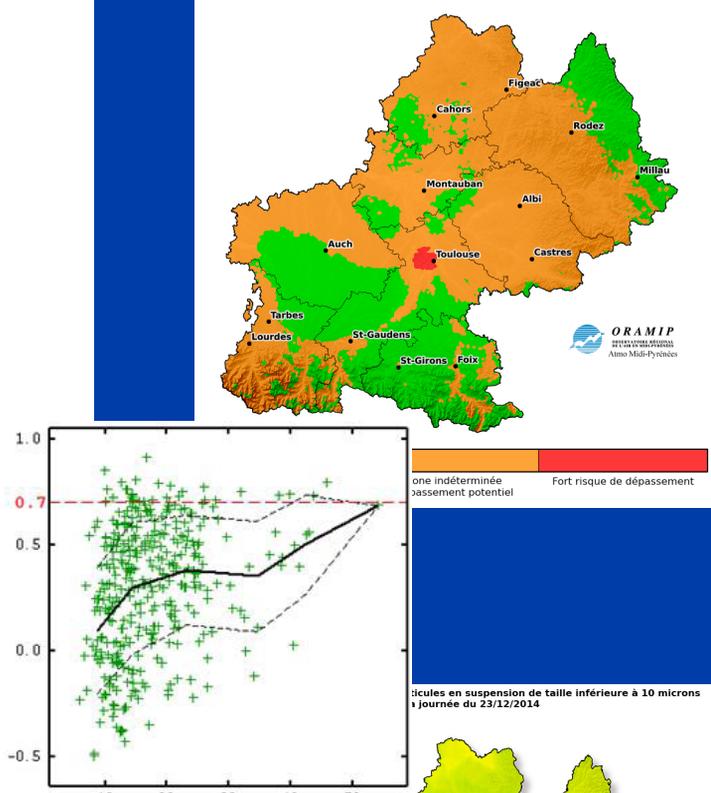
**AMÉLIORATION DE LA CARTOGRAPHIE RÉGIONALE  
DE LA POLLUTION DE L'AIR AUX PARTICULES EN SUSPENSION  
ET MISE EN PLACE DE L'ÉVALUATION  
DE L'INCERTITUDE ASSOCIÉE À CES CARTES**

**CONTRAT PLAN ÉTAT RÉGION 2014**

Incertitudes associées à la prévision de PM10 avec adaptation statistique pour la journée du 10/12/2014



Evaluation de dépassement du seuil d'information et de recommandation à la population pour la journée du 09/12/2013



## CONDITIONS DE DIFFUSION

**ORAMIP Atmo - Midi-Pyrénées**, est une association de type loi 1901 agréée par le Ministère de l'Ecologie, du Développement Durable des Transports et du Logement (décret 98-361 du 6 mai 1998) pour assurer la surveillance de la qualité de l'air sur le territoire de Midi-Pyrénées. ORAMIP Atmo-Midi-Pyrénées fait partie de la fédération ATMO France.

Ses missions s'exercent dans le cadre de la loi sur l'air du 30 décembre 1996. La structure agit dans l'esprit de la charte de l'environnement de 2004 adossée à la constitution de l'Etat français et de l'article L.220-1 du Code de l'environnement. Elle gère un observatoire environnemental relatif à l'air et à la pollution atmosphérique au sens de l'article L.220-2 du Code de l'Environnement. L'ensemble des outils de l'ORAMIP permettant l'évaluation de la pollution de l'air répondent aux exigences du Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) qui est le coordinateur technique au niveau national pour l'ensemble des Association Agréée de Surveillance de la Qualité de l'Air. Le LCSQA est composé du Laboratoire National d'Essai (LNE), de l'Institut National de l'Environnement Industriel et des Risques (INERIS) et de l'école des Mines de Douai (EMD).

ORAMIP Atmo-Midi-Pyrénées met à disposition les informations issues de ses différentes études et garantit la transparence de l'information sur le résultat de ses travaux. A ce titre, les rapports d'études sont librement accessibles sur le site [www.oramip.org](http://www.oramip.org).

Les données contenues dans ce document restent la propriété intellectuelle d'ORAMIP Atmo-Midi-Pyrénées. Toute utilisation partielle ou totale de données ou d'un document (extrait de texte, graphiques, tableaux, ...) doit obligatoirement faire référence à ORAMIP Atmo-Midi-Pyrénées. Les données ne sont pas rediffusées en cas de modification ultérieure.

Par ailleurs, ORAMIP Atmo-Midi-Pyrénées n'est en aucune façon responsable des interprétations et travaux intellectuels, publications diverses résultant de ses travaux et pour lesquels aucun accord préalable n'aurait été donné.

En cas de remarques sur les informations ou leurs conditions d'utilisation, prenez contact avec l'ORAMIP :

- depuis le formulaire de contact sur le site [oramip.atmo-midipyrenees.org](http://oramip.atmo-midipyrenees.org)
- par mail : [contact@oramip.org](mailto:contact@oramip.org)
- par téléphone : 05.61.15.42.46

# SOMMAIRE

## CONDITIONS DE DIFFUSION

SOMMAIRE .....	3
CONTEXTE.....	4
OBJECTIFS DE L'ETUDE .....	5
ANALYSE EXPLORATOIRE DES DONNEES ET INTEGRATION DES CAMPAGNES DE MESURES .....	6
Données exploitées jusqu'alors par la plateforme de modélisation.....	6
Reconstruction statistique des données pour Ax-les-Thermes.....	7
Reconstruction statistique des données pour Cahors .....	8
Reconstruction statistique des données pour Luchon .....	8
Bilan de la methode de reconstruction statistique .....	9
RECHERCHE DU MODELE VARIOGRAPHIQUE ADEQUATE POUR SPATIALISER LES CONCENTRATIONS.....	10
Critères de choix de la méthode.....	10
Tests sur la méthodologie de krigage .....	10
Validation des données.....	10
Recherche des covariables liées aux concentrations de PM10 dans l'air .....	11
Sélection de la méthodologie de krigeage des données sur le territoire régional .....	11
Bilan des tests sur la méthodologie de krigeage des données .....	14
Comparaison des résultats avec la méthode actuelle.....	15
CARTOGRAPHIE DE L'INCERTITUDE ASSOCIEE DES PREVISIONS DE PM10.....	17
PROBABILITE DE RISQUE D'EPISODE DE POLLUTION AUX PARTICULES PM10 .....	19
AMELIORATION ET HARMONISATION DU RENDU CARTOGRAPHIQUE.....	21
CONCLUSION .....	22
ANNEXE : NOTIONS DE GEOSTATISTIQUES .....	23

## CONTEXTE

La plateforme régionale de modélisation de l'ORAMIP fait partie des outils d'évaluation de la qualité de l'air à l'échelle régionale complémentaire aux stations de mesures présente sur le territoire. Elle permet notamment :

- de réaliser des cartes de situation de la qualité de l'air au regard de la réglementation en vigueur et d'évaluer l'exposition des populations sur le territoire régional,
- de prévoir les épisodes de pollution pour l'ozone et les particules en suspension afin d'informer les populations susceptibles d'être exposées.

Depuis la mise en place de la première plateforme régionale en 2002, de nombreuses améliorations ont été apportées : extension de la modélisation à la totalité de la région Midi Pyrénées, mise à jour des données d'émissions de polluants, intégration des particules en suspension pour la prévision des épisodes de pollution, mises à disposition des données à l'échelle communale, mis en place d'une adaptation statistique des données à partir des mesures sur le terrain...

Afin de poursuivre le développement de la plateforme de modélisation, l'ORAMIP a souhaité améliorer la résolution et la précision des cartographies produites prendre en compte l'incertitude associée à une prévision d'un événement de pollution.

## OBJECTIFS DE L'ETUDE

Cette étude s'articule autour de trois axes principaux :

- l'amélioration de la résolution de modélisation affinée en passant de 3 à 1 km de résolution
- Mettre en place un traitement statistique des données pour améliorer la précision du modèle sur les territoires ne disposant pas de mesure en continu en intégrant les données des campagnes de mesures réalisées antérieurement et à venir.
- élaborer des cartes de probabilité associées aux prévisions d'épisodes de pollution de l'air.

Pour se faire des méthodes géostatistiques ont été employées avec de prendre en compte les données géographiques à disposition (inventaire des émissions de polluants, topographie, prévision météorologiques...) et disposer de concentrations ajustées en particules en suspension (PM10) en tout point de la région Midi-Pyrénées.

Une meilleure estimation des concentrations régionales en particules en suspension PM10 permettra de réduire l'incertitude sur les prévisions et d'avoir une meilleure estimation de la probabilité de dépassement de seuils réglementaires

## **ANALYSE EXPLORATOIRE DES DONNEES ET INTEGRATION DES CAMPAGNES DE MESURES**

### **Données exploitées jusqu'alors par la plateforme de modélisation.**

La méthodologie employée par la géostatistique repose sur une couverture de mesures homogène sur l'ensemble de la région et un nombre de stations de mesures suffisant pour que la spatialisation fonctionne correctement.

La répartition des stations de mesures des particules PM10 de l'ORAMIP est principalement liée à la densité de population afin d'évaluer l'exposition de la population sur ces territoires. En parallèle des partenariats spécifiques ont été mis en place pour l'évaluation de site industriel.

Ainsi, sur la région Midi-Pyrénées, les stations de mesures des particules PM10 se répartissent sur les agglomérations de Toulouse, Albi, Castres, Tarbes, et Lourdes, ainsi que sur un site rural implanté à Peyrusse-Vieille dans le Gers. Le site de mesures sur la commune de Bessieres installé dans la cadre d'un partenariat avec un industriel complète le dispositif.

En tout, une dizaine de stations mesurent en continu les particules PM10 sur la région. Certains départements ne comportent pas de moyens de mesures pérennes des particules PM10 tels que les départements du Lot, de l'Aveyron, de l'Ariège et enfin du Tarn-et-Garonne. Cette hétérogénéité de couverture du territoire contraint la méthode d'adaptation statistique.

### **Une stratégie de couverture du territoire à maintenir.**

Afin d'améliorer les connaissances sur les niveaux de polluants sur ces territoires non suivis en continu, l'ORAMIP, en partenariat avec l'Etat et le conseil régional, a mis en place une stratégie de couverture du territoire régional à l'aide de dispositifs de mesures mobiles qui effectuent des campagnes « semi-fixes » de mesures tous les ans pour y évaluer la qualité de l'air.

La fin d'une campagne de mesure temporaire entraîne la perte des données à disposition pour le recalage des cartes de prévisions.

L'analyse géostatistique des données met en évidence l'importance des campagnes semi-fixes pour s'assurer une robustesse méthodologique.

Pour compenser la perte d'une information lors de l'arrêt d'une étude semi-fixe, il est possible de recréer une station dite « virtuelle », en traduisant sous forme d'équations les relations qui liaient la station mobile avec une ou plusieurs stations existantes.

Les données semi-fixe ont ainsi été exploitées afin d'évaluer les éventuelles corrélations de leur mesure avec les stations permanentes du dispositif de l'ORAMIP, et reconstruire une information sur ces sites de mesure temporaire à partir des stations pérennes. L'objectif de cette approche est de compléter les sites de mesures existants sur le territoire régional pour disposer d'une technique de spatialisation plus robuste.

La poursuite de la stratégie de couverture du territoire régional par des campagnes semi-fixes sur de nouvelles zones, permettra une amélioration continue du rendu cartographique de la spatialisation et diminuera l'incertitude.

### **Tests sur les données de l'année 2014.**

Trois campagnes de mesures d'une durée d'une année ont été exploitées sur les communes de Cahors, Ax-les-Thermes et Luchon.

La première phase du projet de spatialisation a donc consisté à calculer des modèles de reconstruction statistique pour ces trois sites.

Les niveaux de pollution aux particules PM10 sont principalement liés à des émissions locales et à un apport extrarégional pouvant venir de région ou de pays voisins.

En partant de ce constat, les stations utilisées dans le cadre de cette étude sont uniquement des sites de fond urbain ou rural pour éviter d'être influencé pour un impact direct des sources d'émissions lié au trafic routier par exemple.

Une reconstruction statistique de données a été réalisée en considérant les concentrations de particules PM10 sont liées régionalement avec

une interdépendance marquée suivant la typologie de stations.

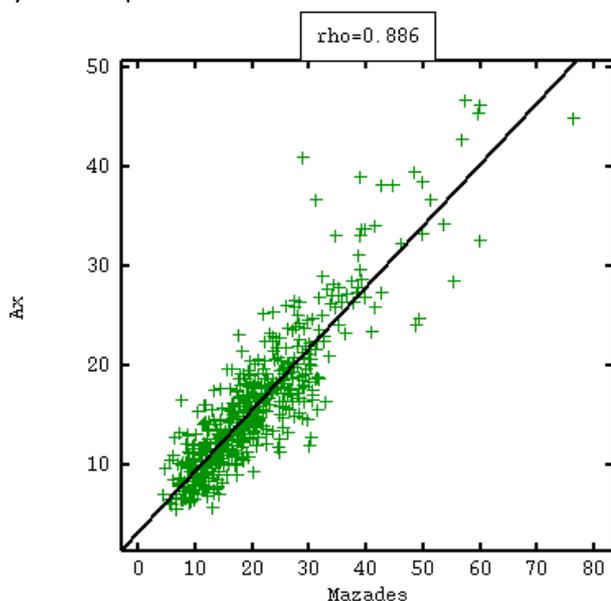
Pour valider, la méthodologie de reconstruction de données d'une station, le modèle statistique a été construit sur un échantillon aléatoire de

données de 75%. Puis sur les 25% de données non utilisées, les données reconstruites à partir du modèle statistiques sont comparées avec les données mesurées.

## Reconstruction statistique des données pour Ax-les-Thermes

Parmi les stations du réseau de l'ORAMIP, la station de Toulouse Mazades donne la meilleure corrélation avec les mesures d'Ax-les-Thermes (coefficient de corrélation = 0.89).

Le graphique de corrélation montre un nuage de point s'alignant selon une droite et avec un biais systématique assez limité.



Nuage de corrélation entre les données de la station Mazades et 75% des données de la station semi-fixe de Ax-les-Thermes.

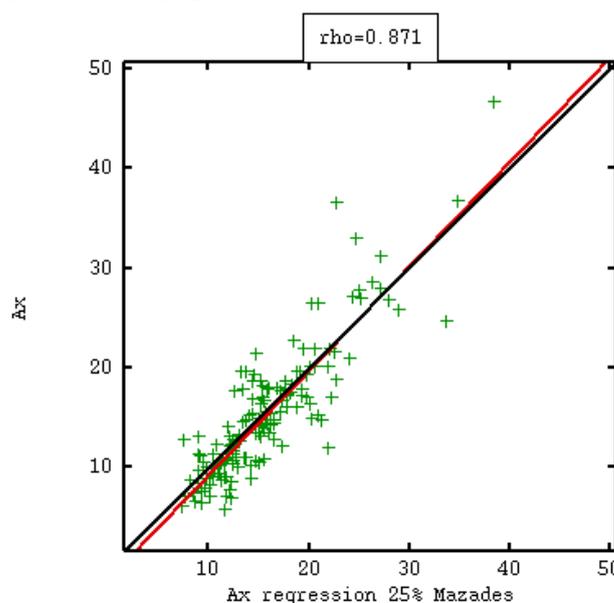
Le modèle statistique de reconstitution des données obtenu est le suivant :

$$[PM10]_{Ax} = 3.362 + 0.616 \times [PM10]_{Mazades}$$

Avec ce modèle, on peut estimer approximativement les concentrations au niveau de la commune d'Ax-les-Thermes à environ 61% des niveaux de concentrations de Toulouse Mazades.

L'utilisation du modèle statistique sur les 25% de données restantes montre une bonne corrélation (coefficient de 0.87) ce qui tend à démontrer que le modèle fonctionne assez correctement avec les mesures de la station de Mazades. Nous pouvons observer que par moment, des points peuvent s'écarter de la droite de régression. Ceci signifie

que même si ce modèle statistique fonctionne dans la majorité des cas, il peut y avoir épisodiquement des différences de concentrations en PM10 entre la zone toulousaine et la vallée d'Ax-les-Thermes. Les différences locales d'émissions et de conditions météorologiques sont à l'origine de ces concentrations différentes en PM10. Certaines journées ne sont pas à l'abri d'un épiphénomène très localisé lié à une activité de voisinage et qui ne serait pas représentatif de la qualité de l'air de la commune d'Ax.



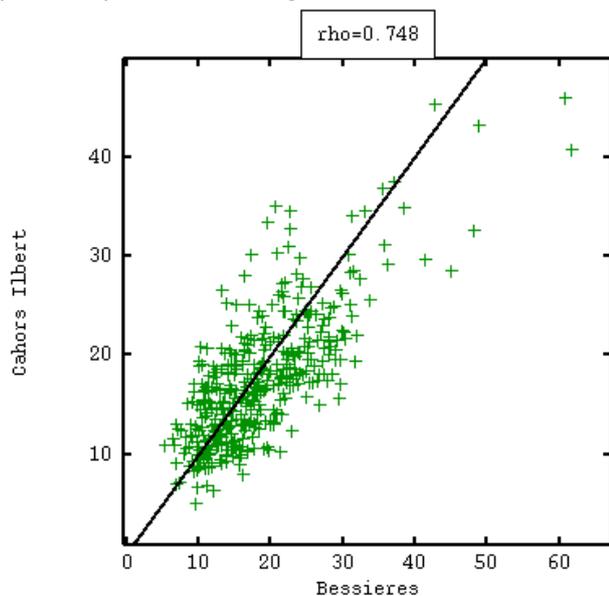
Nuage de corrélation entre les données modélisées statistiquement et les données réelles de la station semi-fixe d'Ax-les-Thermes.

Il est à noter que lorsque la station d'Ax était encore en fonctionnement, une campagne de mesures a eu lieu sur la commune de Pamiers avec une période de recoupement de 15 jours. La comparaison de ces 15 jours ont montré une excellente corrélation entre Ax et Pamiers ce qui paraît cohérent avec la proximité entre ces 2 sites. Néanmoins, le site de Pamiers n'a pas à vocation à devenir un site de mesures pérenne de PM10 et ne peut servir à utiliser un modèle de reconstruction dans la durée.

## Reconstruction statistique des données pour Cahors

L'étude comparative des mesures de Cahors avec les stations pérennes du dispositif de mesures de l'ORAMIP met en évidence une assez bonne corrélation avec la station de Bessières.

Le graphique de corrélation montre un nuage de point dispersé mais s'alignant selon une droite



*Nuage de corrélation entre les données de la station Bessières et 75% des données de la station semi-fixe de Cahors.*

La corrélation avec les mesures de Cahors s'établit avec un coefficient de 0,75. Le modèle obtenu est le suivant :

$$[\text{PM10}]_{\text{Cahors}} = 6.263 + 0.618 \times [\text{PM10}]_{\text{Bessières}}$$

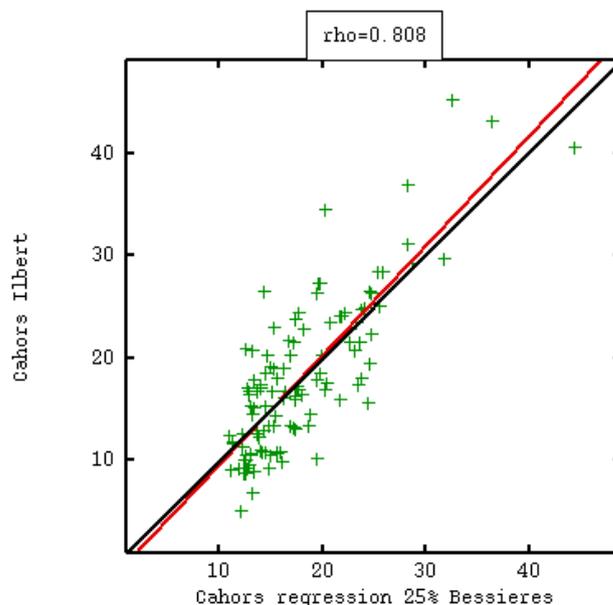
La validation de ce modèle statistique sur les 25% de données restantes montre toujours une dispersion des points autour de la droite du modèle statistique. Ainsi, malgré une estimation correcte de la concentration de Cahors à partir de la station de Bessières, la différence entre valeur réelle et valeur mesurée peut parfois s'avérer importante (à la hausse comme à la baisse).

## Reconstruction statistique des données pour Luchon

L'étude comparative des mesures de Luchon avec les stations fixes du dispositif de l'ORAMIP ne montre pas de corrélations satisfaisantes. Le nuage de point s'articule selon une droite mais est très dispersé. Parmi toutes les stations du réseau, aucune ne semble être en mesure de reproduire les données de Luchon.

L'incertitude sur la valeur estimée sera donc plus élevée pour cette station.

Cette analyse comparative de la station de Cahors permet également de mettre en évidence que les concentrations de PM10 relevées sur la partie nord de la région peuvent par moment être assez différentes des stations situées sur la Haute-Garonne. Ceci confirme qu'une mesure des particules PM10 sur le nord de la région est importante à la fois pour avoir un point de comparaison pour l'évaluation de la prévision mais aussi pour permettre d'ajuster la prévision au quotidien en estimant au mieux les apports extrarégionaux de polluants venant majoritairement du nord de la France.



*Nuage de corrélation entre les données modélisées statistiquement et les données réelles de la station semi-fixe de Cahors.*

Les mesures de Luchon ressortent comme « atypiques » car probablement liées à une problématique plus locale que régionale. L'implantation de la ville de Luchon en fond de vallée fait que les nappes de pollution ont plutôt tendance à se former sous l'influence des

émissions locales plutôt que d'apports venus du reste de la région. De même certaines journées semblent être sous l'influence de masse d'air plus polluée venue d'Espagne.

Déterminer une corrélation et ainsi un modèle robuste semble ainsi risqués avec les influences diverses des sources de pollution de la ville de Luchon.

Ce constat permet de mettre en évidence la problématique particulière de la pollution atmosphérique dans les vallées pyrénéennes. Le climat est singulier avec le reste de la région et chaque vallée en fonction des répercussions de son activité, sa population, son trafic routier et son encaissement connaîtra des niveaux de pollution de l'air qui lui sont propres.

Dans ce travail de reconstruction de données, le meilleur modèle est basé sur la combinaison des données de trois stations situées à Peyrusse-Vieille (Gers), Lourdes et Tarbes.

Le modèle obtenu est :

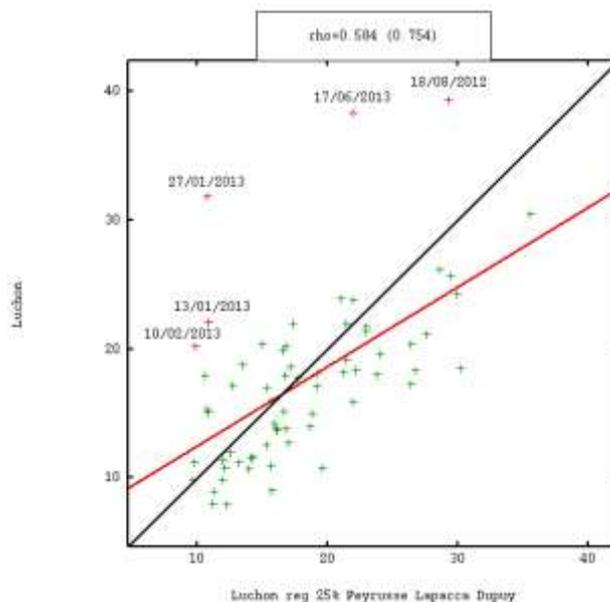
$$[PM10]_{Luchon} = 6.132 + 1.025x [PM10]_{Lapacca} - 0.131x[PM10]_{Dupuy} + 7.521.10^{-2}x[PM10]_{Peyrusse}$$

Le modèle testé sur 25% de données montre que les points ne sont parfois pas alignés avec la droite de régression. L'utilisation d'un tel modèle impliquerait une reconstruction trop aléatoire des données.

## Bilan de la méthode de reconstruction statistique

La création de stations virtuelles permet de compléter le dispositif de stations existantes à partir des campagnes de mesures effectuées par le passé. Les modèles de reconstruction de données s'appuient sur les mesures produites par les stations fixes de l'ORAMIP.

Pour mieux évaluer les concentrations en PM10 dans les vallées pyrénéennes et améliorer la justesse de la prévision, le dispositif de mesures nécessiterait l'installation d'une station en bord de piémont sur une zone influencée par la même masse d'air que les vallées de Haute-Garonne et du Couserans.



*Nuage de corrélation entre les données modélisées statistiquement et les données réelles de la station semi-fixe de Luchon*

Cependant il faut noter qu'il sera toujours difficile d'anticiper les échanges entre l'Espagne et la France ainsi que les pollutions notamment liées à de l'écobuage très fréquent dans cette zone.

En complétant la couverture régionale par des dispositifs de mesures mobiles, le modèle statistique de spatialisation sera régulièrement amélioré et permettra une meilleure évaluation des concentrations en particules PM10 sur prévision.

## RECHERCHE DU MODELE VARIOGRAPHIQUE ADEQUATE POUR SPATIALISER LES CONCENTRATIONS

### Critères de choix de la méthode

L'objectif d'ajuster la prévision issues de la plateforme de modélisation déterministe à l'aide du dispositif de mesures doit conserver une variabilité spatiale propre au milieu avec si possible une descente d'échelle pour délimiter plus précisément les zones impactées.

Plusieurs critères de validation de la méthode de spatialisation des données ont été définis :

- La concentration à l'emplacement de la station de mesure ne doit pas être modifiée par la spatialisation.
- La carte produite doit conserver une cohérence des concentrations par rapport à la connaissance des milieux : les concentrations doivent être plus importantes en milieu urbain que rural, faire ressortir le relief avec l'effet de confinement des vallées et éventuellement faire ressortir les grands axes de circulation.
- Le rendu cartographique doit limiter au maximum les effets de spot des concentrations aux emplacements des stations de mesures. La méthode de spatialisation des données ne doit donc entraîner une décroissance trop rapide des concentrations avec l'éloignement de la station.
- Une amélioration du rendu cartographique avec une résolution plus fine que les données de prévision brute en s'appuyant sur des covariables (altitude, occupation du sol...). Une résolution de 1km est l'objectif visé.
- Les concentrations de PM10 produites doivent être cohérentes avec la saisonnalité des niveaux de PM10 observés par les mesures.

### Tests sur la méthodologie de krigage

Le krigage est une méthode d'interpolation spatiale géostatistique qui consiste à définir une variabilité géographique en tenant compte de la relation qu'il existe entre les distances et les données. Dans notre cas, il s'agit des mesures de PM10 effectuées par les stations de surveillance de la qualité de l'air

La méthodologie de krigage a un impact sur la précision de la donnée de concentration modélisée et sur la qualité de la représentation cartographique de la pollution

Pour spatialiser les concentrations mesurées par les stations de mesures et étendre cette information aux données issues du modèle de prévision Chimere, plusieurs techniques de krigage ont été testées afin de rechercher la plus appropriée à la spécificité des territoires de Midi-Pyrénées.

Dans un premier temps, une étude de cas sur des épisodes de pollution aux particules PM10 a été réalisée pour vérifier la cohérence et le détail visuel de l'information produite.

Plusieurs journées ayant mis en évidence des niveaux de concentration ont été retenus pour évaluer l'impact des tests réalisés :

- 03/03/2011,
- 11/02/2012,
- 03/04/2011,
- 10/12/2012
- 03/03/2013.

### Validation des données

Les méthodes statistiques retenues ont ensuite été confirmées par validation croisée puis tester sur l'ensemble des journées sur les années 2010 et 2013.

La validation de la méthodologie de spatialisation doit prendre en compte deux aspects :

- garantir un bon fonctionnement lors des épisodes de pollution,
- ne pas entraîner une surestimation des concentrations lorsque les niveaux sont plus bas (ce qui représente la majorité des journées de l'année).

## Recherche des covariables liées aux concentrations de PM10 dans l'air

La recherche des covariables est l'étape préalable pour déterminer les données géographiques qui possèdent une relation statistique et spatiale avec les concentrations de particules PM10 sur la région.

Plusieurs types de données ont ainsi été testées afin d'identifier lesquelles pourraient être retenues comme co-variables.

- Les prévisions météorologiques comprenant :
  - Atténuation des nuages (max./min./moy.)
  - Hauteur de couche limite (max./min./moy.)
  - Humidité relative (max./min./moy.)
  - Cumul de pluie
  - Cumul de rayonnement
  - Température (max./min./moy.)
  - Vent méridional (max./min./moy.)
  - Vent zonal (max./min./moy.)
  - Vitesse de vent (max./min./moy.)
- Les prévisions de concentrations journalières en particules PM10 issues de la plateforme de modélisation Chimere selon une grille de 3 km.
- Les quantités d'émissions dans l'air de particules PM10 à une résolution kilométrique
- Les données d'altitude
- La densité de population
- L'occupation du sol

Les corrélations ont été calculées pour l'ensemble des journées des 4 années de données disponibles.

A l'issue des tests réalisées les variables qui ont montré les meilleures corrélations avec les mesures de particules PM10 et qui ont été retenues sont :

- l'altitude,
- le logarithme des émissions de PM10,
- la température max journalière,
- les prévisions de concentrations journalières en PM10 Chimere.

## Sélection de la méthodologie de krigeage des données sur le territoire régional

Afin d'évaluer les différentes méthodologies de krigeage des données, des tests ont été réalisés sur les journées ayant connu des épisodes de pollution aux particules en suspension. La période ainsi considérée couvrait les années 2010 à 2013.

Les méthodologies de krigeage qui ont été testées sont les suivantes :

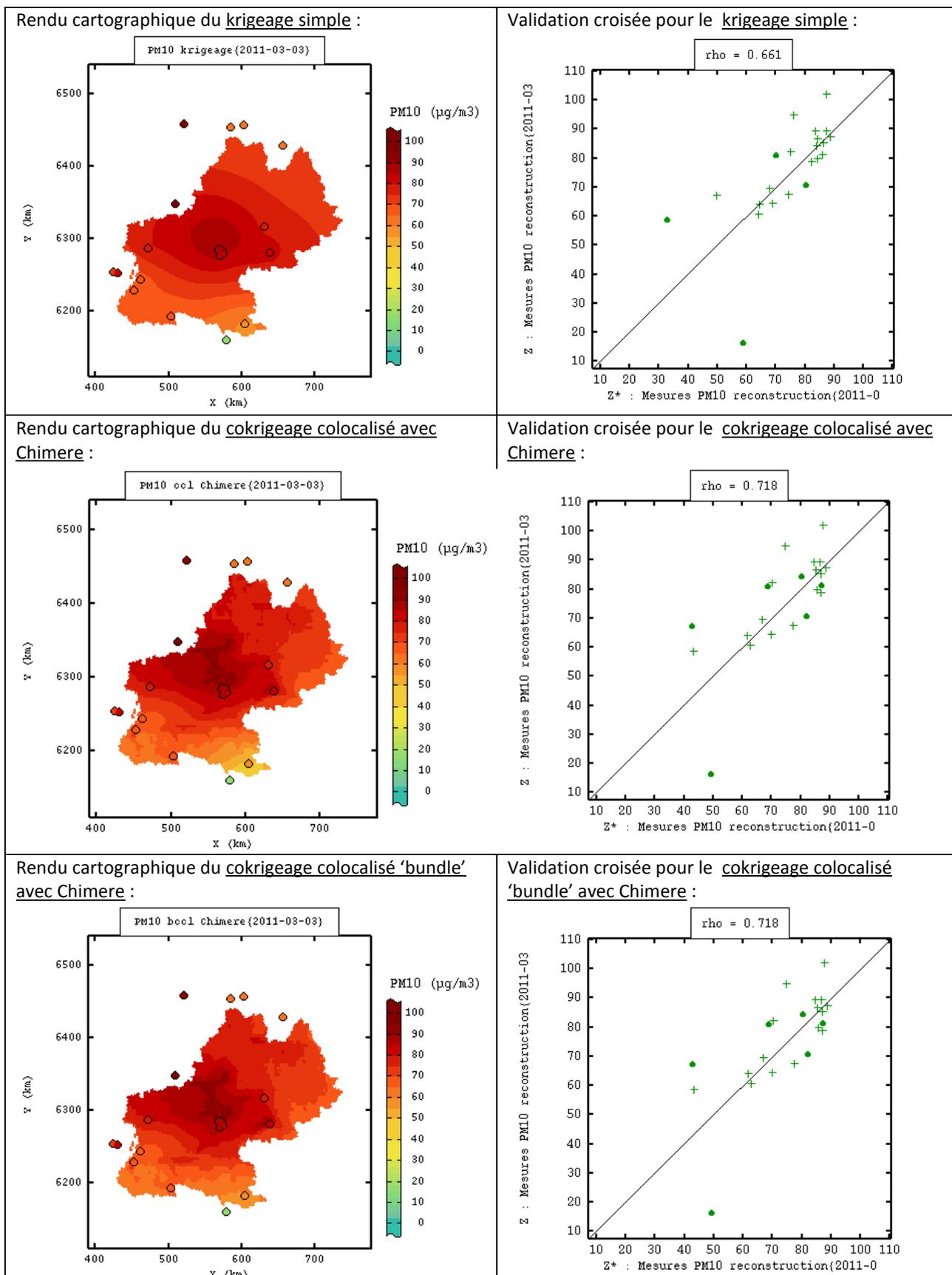
- Krigeage des mesures
- Cokrigeage colocalisé sans lissage
- Cokrigeage colocalisé 'bundle' sans lissage
- Cokrigeage colocalisé avec lissage
- Cokrigeage colocalisé 'bundle' avec lissage

- Cokrigeage colocalisé avec variogramme moyen

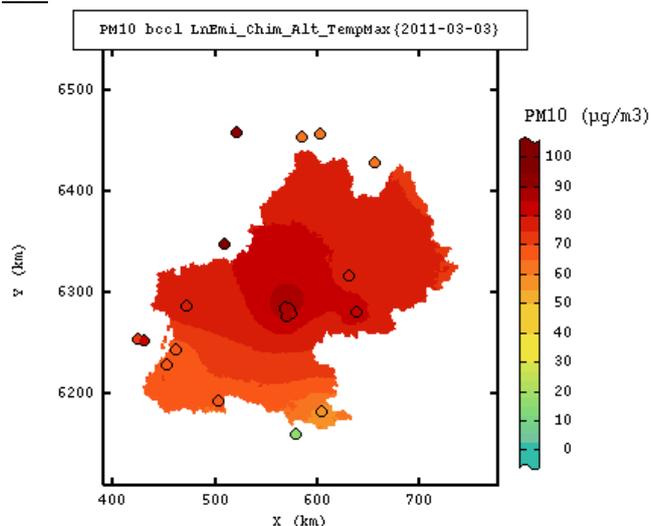
Une comparaison des cartes résultantes de ces tests de spatialisation des données a été effectuée avec l'ensemble des techniques de krigeage.

Cette comparaison permet de visualiser les différents rendus cartographiques. D'autre part une validation croisée a été réalisée pour évaluer la cohérence des résultats obtenue par krigeage avec les données réellement mesurées au niveau des stations de mesures.

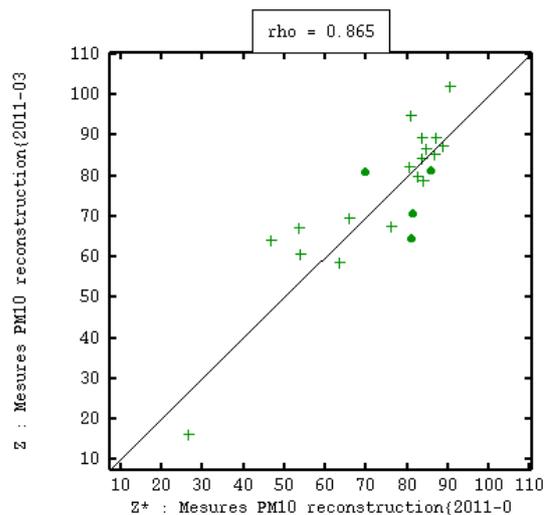
Pour la journée du 3 mars 2011, les résultats des tests des différentes techniques de spatialisation sont présentés à titre d'illustration de la méthodologie de sélection de la méthode de krigage :



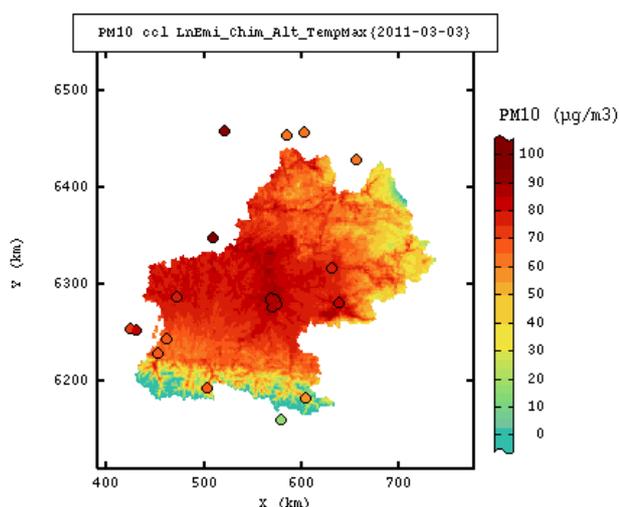
Rendu cartographique du cokrigage colocalisé 'bundle' avec log des émissions, Chimere, altitude et température max.



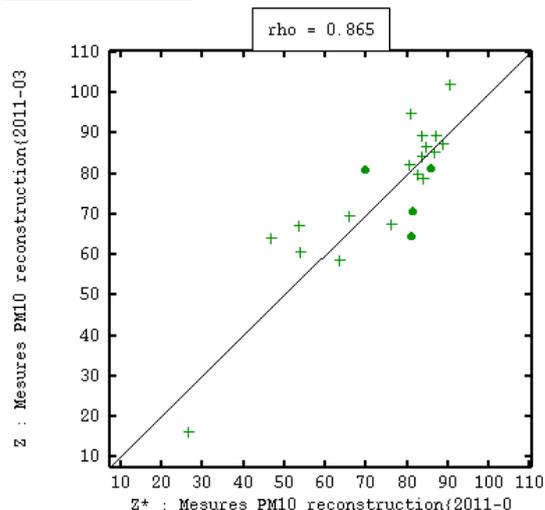
Validation croisée pour le cokrigage colocalisé 'bundle' avec log des émissions, Chimere, altitude et température max.



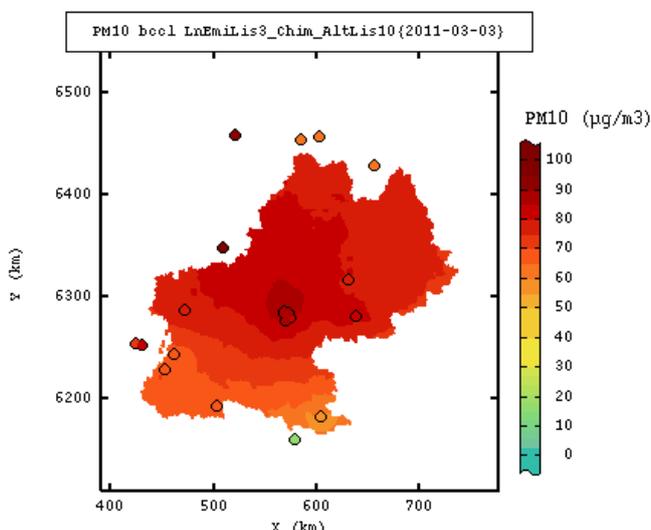
Rendu cartographique du cokrigage colocalisé avec log des émissions, Chimere, altitude et température max.



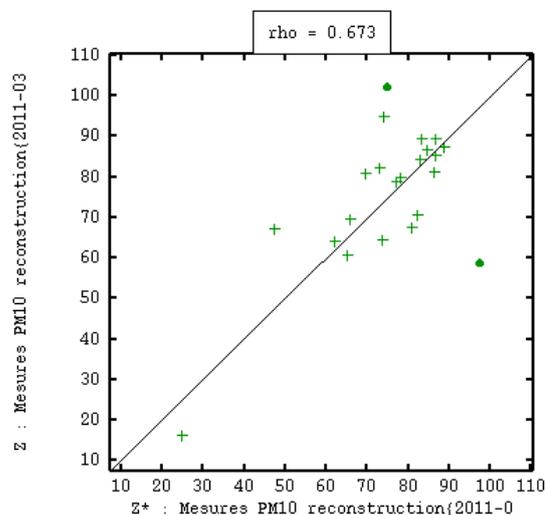
Validation croisée des résultats par cokrigage colocalisé avec log des émissions, Chimere, altitude et température max



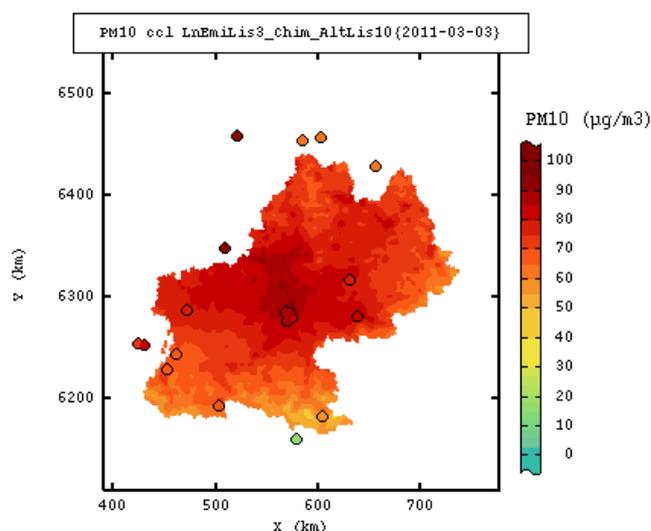
Rendu cartographique du cokrigage colocalisé 'bundle' avec log des émissions lissé à 3km, Chimere, altitude lissée à 10km et température max.



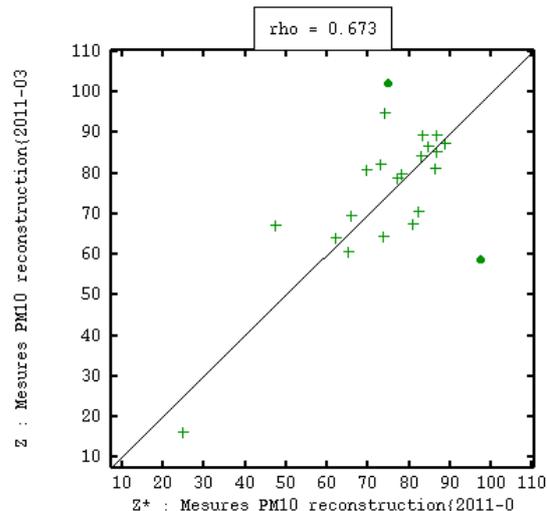
Validation croisée des résultats par cokrigage colocalisé 'bundle' avec log des émissions lissé à 3km, Chimere, altitude lissée à 10km et température max.



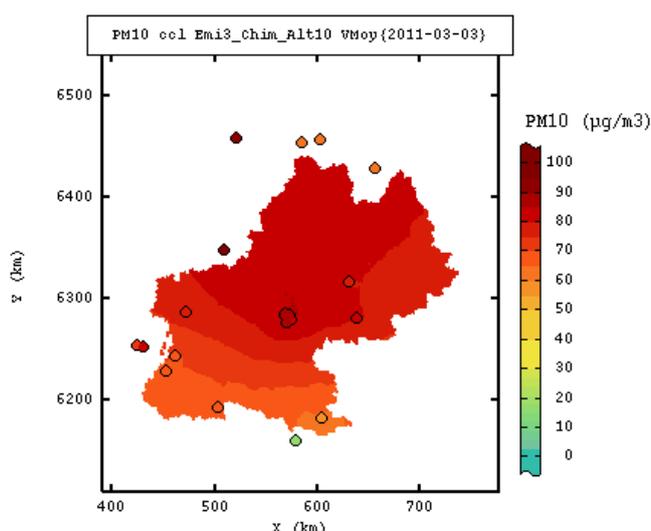
Rendu cartographique du cokrigage colocalisé avec log des émissions lissé à 3km, Chimere, altitude lissée à 10km et température max.



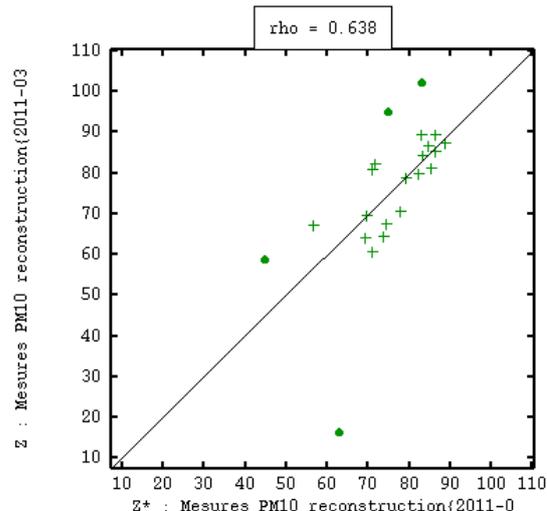
Validation croisée pour le cokrigage colocalisé avec log des émissions lissé à 3km, Chimere, altitude lissée à 10km et température max.



Rendu cartographique du cokrigage colocalisé avec log des émissions lissé à 3km, Chimere, altitude lissée à 10km et température max et l'utilisation d'un variogramme moyen.



Validation croisée pour le cokrigage colocalisé avec log des émissions lissé à 3km, Chimere, altitude lissée à 10km et température max et l'utilisation d'un variogramme moyen .



## Bilan des tests sur la méthodologie de krigage des données

Le bilan des tests sur les épisodes de pollution montre que les méthodologies donnant les meilleurs résultats sur les validations croisées sont le cokrigage colocalisé et colocalisé « bundle » avec l'utilisation des covariables : logarithme des émissions, les prévisions chimere, l'altitude et les températures maxima sans lissage.

Les autres méthodologies donnent des corrélations sensiblement identiques. Au niveau de l'analyse visuelle, le cokrigage colocalisé (sans bundle) permet un rendu cartographique

un niveau de détail assez performant et cohérent avec les typologies régionales.

Pour confirmer le choix définitif, cette méthodologie a été employée sur l'ensemble des journées de la période test soit de 2010 à 2013.

Au cours de ces tests, certains dysfonctionnements ont été identifiés dans la méthodologie.

Il y avait notamment, une augmentation des concentrations en particules avec l'altitude, qui a été invalidée du fait de notre connaissance de la situation sur le terrain et les études portant sur

les niveaux verticaux de particules en fonction de l'altitude.

D'autre part, sur certaines journées, les variogrammes avaient tendance à donner trop de poids aux émissions de polluants ce qui entraînait une trop grande variabilité spatiale sur les cartes et des effets de spots de concentrations.

Les données de la station d'Andorre-la-Vieille paraissent pour certaines journées très éloignées des autres stations du domaine. La représentativité de cette station est peut être une cause de cet écart ou l'influence d'une source locale peut compromettre la concentration de fond que l'on cherche à évaluer. Compte tenu de cette situation et de notre manque de connaissance de la représentativité de cette station, elle a été

retirée des stations de référence pour le redressement des données régionales.

Enfin, le fait de calculer un variogramme spécifique à chaque journée de prévision a montré que cette technique manquait de robustesse et donnait des résultats trop aléatoires en raison de faible nombre de stations utilisé pour le calcul du variogramme.

Le choix définitif s'est donc porté sur l'utilisation d'un variogramme moyen sur une période d'étude qui pourra être recalculé périodiquement pour le réajuster.

Le fait d'utiliser un variogramme moyen a permis de pallier le problème lié à l'augmentation des niveaux de particules avec l'altitude.

## Comparaison des résultats avec la méthode actuelle

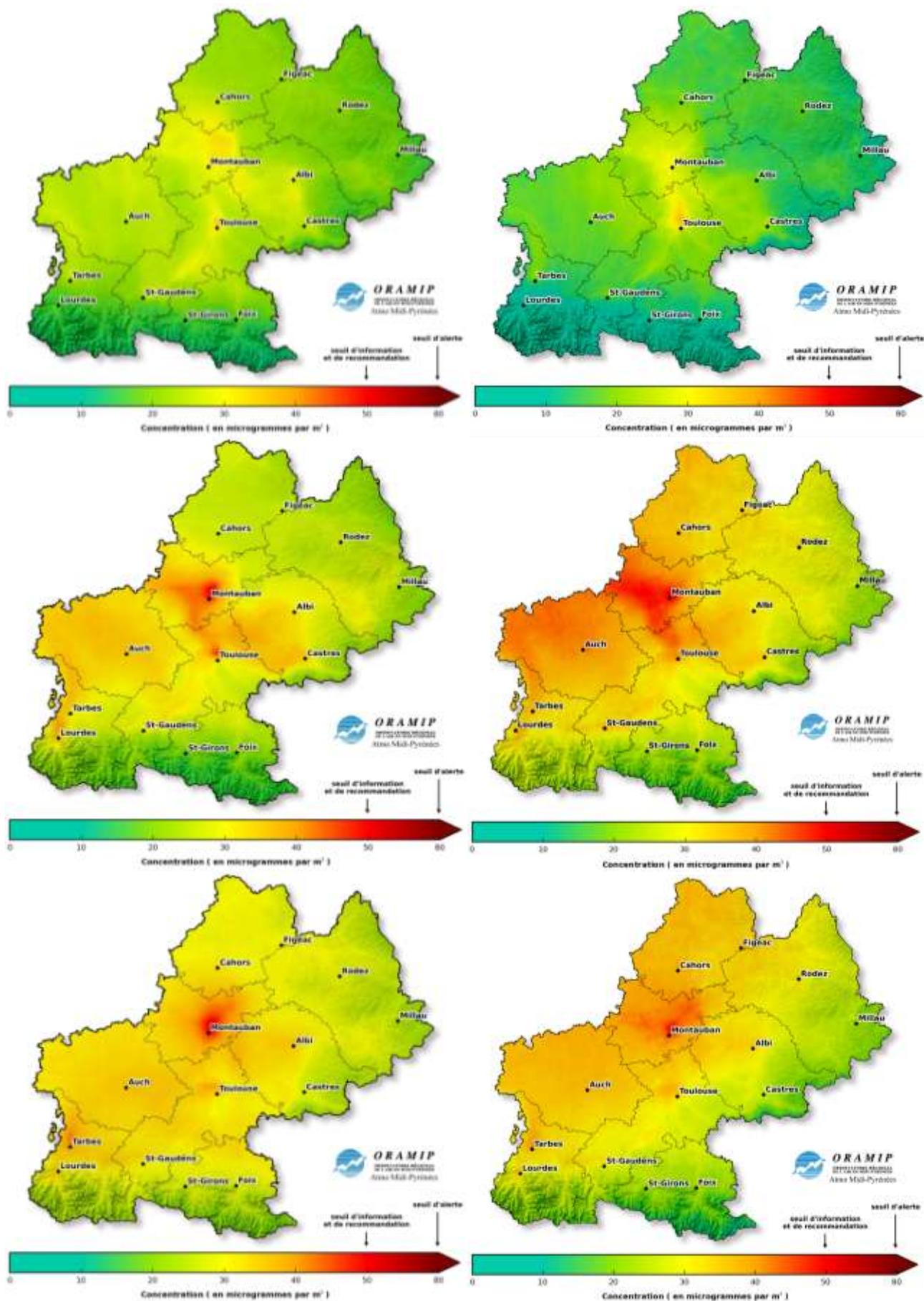
Les résultats des validations croisées avec les données des stations de mesures réalisées sur la méthode de spatialisation utilisée jusqu'alors par l'ORAMIP ont été comparés avec ceux du cokrigage colocalisé.

Les résultats donnent des corrélations nettement supérieures avec la nouvelle méthodologie avec des écart-types plus faibles et des biais moyen proche de 0. Ces résultats confirment le choix de la méthodologie retenue

Méthode	Corrélation	Min.	Max.	Moy.	Ecart-type
Cokrigage colocalisé avec variogrammes moyens	0.722	-3.81	6.09	-0.05	1.95
Spatialisation des écarts par inverse distance (ORAMIP)	0.584	-25.01	31.72	0.17	11.59

Tableau de comparaison des statistiques des validations croisées des deux méthodologies.

Des exemples de représentations cartographiques selon l'ancienne méthodologie de spatialisation des données par « inverse carré distance » et selon la nouvelle méthodologie de spatialisation par « Cokrigage colocalisé avec variogrammes moyens » sont présentés ci-après.



Carte de prévision des concentrations en particules PM10 réalisée selon la méthode « Inverse carré distance » (à gauche) et selon la nouvelle méthode de spatialisaton (à droite)

## CARTOGRAPHIE DE L'INCERTITUDE ASSOCIEE DES PREVISIONS DE PM10

Les cartes de prévision d'épisode de pollution sont disponibles quotidiennement tant en interne pour le prévisionniste que sur le site internet de l'ORAMIP.

Ces cartes sont font partie des outils permettant d'anticiper un épisode de pollution et de communiquer sur cet évènement notamment auprès des populations les plus sensibles à la pollution de l'air.

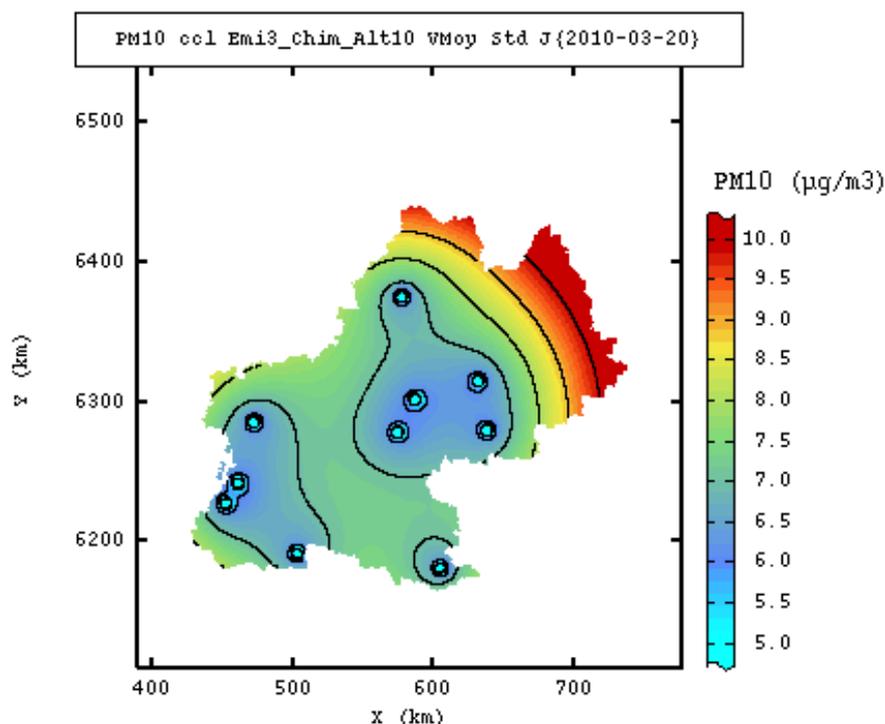
Afin de fournir un critère de confiance aux cartes de prévision des moyennes journalières en particules PM10, un travail de cartographie de l'incertitude est venu compléter le projet de spatialisation.

Une première approche est d'utiliser la cartographie de l'écart-type de cokrigage de la méthode de spatialisation. Cependant cette approche ne constitue qu'une appréciation qualitative de l'incertitude car elle ne tient pas compte de la variabilité journalière des concentrations de particules.

Pour quantifier, l'incertitude de manière plus exhaustive, il est nécessaire de considérer cette variabilité temporelle entre les différentes prévisions journalières et en tenant compte de la variabilité spatiale des concentrations. Pour se faire, l'écart-type journalier des données est pondéré avec la distance moyenne entre les stations de référence utilisées dans la cadre de la méthodologie de spatialisation.

Pour compléter le calcul d'incertitudes, l'incertitude de la mesure au niveau des stations fixes de référence ou bien celle de la donnée calculée par reconstruction statistique.

La mise en œuvre opérationnel de la réalisation des cartes d'incertitude a nécessité d'intégrer à la plateforme de modélisation un fichier d'entrée contenant les valeurs d'incertitude aux stations qui alimente le post-traitement des données.



Exemple de carte d'incertitude associée à la carte de prévision du 20 mars 2010

Sur les cartes produites, l'incertitude est minimum aux emplacements des stations et

correspond à la valeur de l'incertitude associée à l'analyseur.

L'incertitude augmente avec l'éloignement des stations et lié à l'écart-type de cokrigage.

L'incertitude augmente lorsque l'amplitude et les concentrations augmentent en raison de la pondération avec l'écart-type des données journalières et la distance moyenne entre les stations.

En compilant un historique de données d'incertitudes, il est envisagé d'identifier les

zones où les incertitudes sont régulièrement élevées. Ces zones traduisant une estimation moins fiable des prévisions en PM10. Elles pourront ainsi faire l'objet de campagne de mesures pour réduire, à terme, l'incertitude des prévisions régionales des concentrations en PM10 sur ces zones.

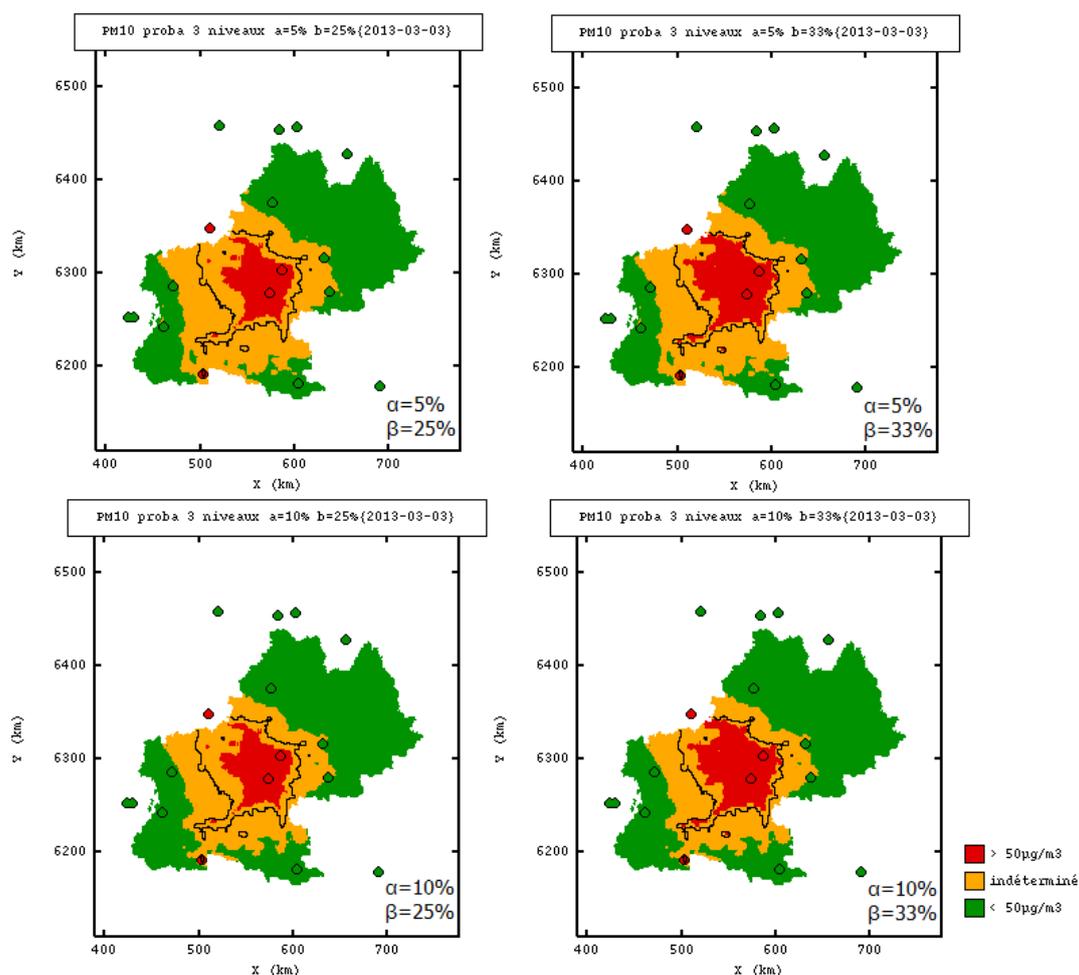
## PROBABILITE DE RISQUE D'EPISODE DE POLLUTION AUX PARTICULES PM10

En complément de la cartographie de l'incertitude associée à une prévision régionale il a également été mené une réflexion sur la mise à disposition au sein de l'ORAMIP d'une carte de probabilité de survenue d'un épisode de pollution aux particules PM10.

Pour évaluer la survenue et l'étendue d'un épisode de pollution au regard de la réglementation, le prévisionniste étudiait les cartographies de concentration afin d'identifier d'éventuelles zones en dépassement des seuils de concentrations fixés à 50 et 80  $\mu\text{g}/\text{m}^3$  de particules PM10. Ces deux seuils correspondant respectivement au niveau d'information et au niveau d'alerte. Cette approche ne tenait pas compte de l'incertitude associée à la prévision de concentrations en particules PM10.

Pour estimer une probabilité de risque de dépassement des seuils réglementaires en tenant compte dorénavant de l'incertitude, la méthodologie employée a été revue en utilisant une méthode dite d'espérance conditionnelle.

En modifiant des coefficients d'ajustement, cette technique permet de délimiter une zone de probabilité de non détection de dépassement (coefficient  $\alpha$ ) et une zone de probabilité de détection à tort du seuil (coefficient  $\beta$ ) à partir des incertitudes associées aux données de la carte de prévision.



Cartes représentant différentes valeurs des coefficients d'ajustement ( $\alpha$  et  $\beta$ ) pour la probabilité de risque de dépassement (trait noir : isoconcentration de 50  $\mu\text{g}.\text{m}^{-3}$  de PM10, zone rouge : fort risque de dépassement, zone orange : risque de dépassement modéré, zone verte : risque de dépassement faible)

En modifiant les coefficients d'ajustement  $\alpha$  et  $\beta$ , il est ainsi possible d'étendre ou de réduire les zones de probabilité de dépassement d'un seuil réglementaire définissant un épisode de pollution aux particules en suspension. Ainsi, en diminuant la valeur  $\alpha$ , on réduit le nombre de non détection de dépassement. En augmentant le coefficient  $\beta$ , on augmente le risque de détection de dépassement à tort.

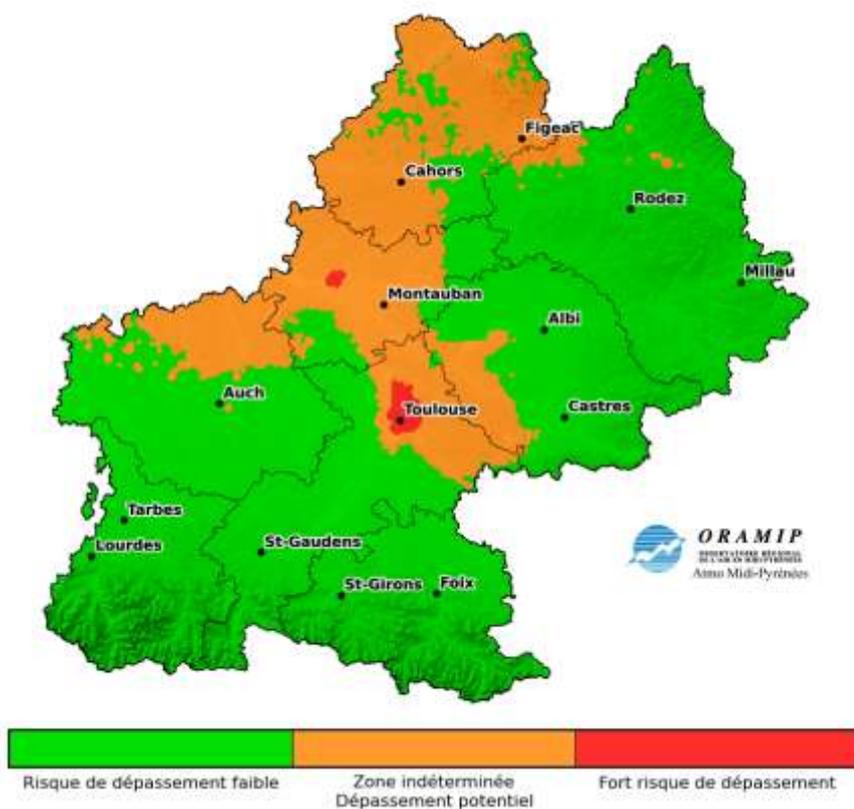
L'objectif de cette méthodologie est de trouver un équilibre entre ces deux coefficients pour éviter toute erreur d'anticipation d'un épisode de pollution.

Une validation croisée de cette méthodologie appliquée sur des journées ayant connu un épisode de pollution aux particules PM10 permettra d'affiner ces coefficients.

Les cartes de probabilité d'épisodes de pollution ne sont pas destinées à un usage de communication auprès du grand public. Elle reste un outil technique qui permet avant tout de faciliter l'évaluation des zones potentiellement sujettes à un épisode de pollution par les personnels de l'ORAMIP.

Ces cartes ont fait l'objet d'un traitement spécifique pour améliorer la lecture du niveau de risque de dépassement de seuil de pollution

Evaluation de dépassement du seuil d'information et de recommandation à la population pour la journée du 03/01/2015



Exemple de carte de risque de dépassement de seuil d'information et de recommandation.

## AMELIORATION ET HARMONISATION DU RENDU CARTOGRAPHIQUE

Dans le but de faciliter la compréhension de l'information sur les concentrations de polluants atmosphériques par le grand public, et afin d'harmoniser les gammes colorimétriques utilisées par les différents acteurs français de la qualité de l'air (AASQA et Ineris), un travail de refonte des cartes a été mené par l'ORAMIP dans le cadre de ce projet pour valoriser les résultats de spatialisation.

En 2013, le Laboratoire Central de Surveillance de la Qualité de l'Air (LCSQA) a établi une nouvelle palette de couleur permettant de faire ressortir les épisodes de pollution au regard de la réglementation en vigueur.

A la suite de ces travaux, en aout 2014, le Ministère en charge de l'Ecologie a demandé la mise en place de cette palette de couleur pour les cartographies de pollution dans toutes les structures en charge de la surveillance de la qualité de l'air. L'utilisation de ces couleurs par tous les systèmes de prévision de qualité

de l'air devant simplifier l'intercomparaison entre les différentes cartes régionales et nationales.

Pour la prévision de pollution quotidienne, un dépassement de seuil réglementaire est représenté par la couleur rouge. Les deux premiers paliers de concentrations utilisent la même couleur pour ne pas faire ressortir d'informations visuelles lorsque les niveaux de pollution sont faibles.

Ensuite, les couleurs des paliers sont liées à des concentrations basées sur la valeur du seuil d'information et de recommandation. Ainsi, chaque palier de concentration est calculé en diminuant de 20% la concentration du seuil d'information. Le dernier étage correspond au seuil et toutes concentrations prévues supérieures à cette valeur seront en rouge foncé 'carmin'.

Toutes les couleurs sont ensuite interpolées entre les paliers de concentrations pour avoir un dégradé sur l'ensemble de la gamme et avoir des variations de concentrations plus fines.

Couleur	Code couleur RGB	Concentrations moyenne journalière en particules PM10	Concentrations maximum horaire journalier d'ozone
	(0,204,170)	0 µg/m <sup>3</sup>	0 µg/m <sup>3</sup>
	(0,204,170)	10 µg/m <sup>3</sup>	36 µg/m <sup>3</sup>
	(153,230,0)	20 µg/m <sup>3</sup>	72 µg/m <sup>3</sup>
	(255,255,0)	30 µg/m <sup>3</sup>	108 µg/m <sup>3</sup>
	(255,170,0)	40 µg/m <sup>3</sup>	144 µg/m <sup>3</sup>
	(255,0,0)	50 µg/m <sup>3</sup>	180 µg/m <sup>3</sup>
	(128,0,0)	80 µg/m <sup>3</sup>	240 µg/m <sup>3</sup>

Seuil d'information et de recommandation →

Seuil d'alerte →

Tableau de concordance des couleurs avec les paliers de concentrations.

Pour améliorer le rendu cartographique un ombrage du relief régional a été ajouté au fond de carte de prévision de la pollution.

Ainsi, le relief permet par exemple de faire ressortir les niveaux de pollution dans les vallées pyrénéennes.

Les concentrations sont détournées pour ne conserver que l'information régionale et se focaliser uniquement sur le territoire d'intérêt.

Enfin, les seuils réglementaires sont indiqués sur l'échelle de concentrations pour faciliter l'interprétation de la carte par le public. Ceci permet

de mieux cerner les zones sujettes à un événement de pollution au regard des seuils réglementaires.

Pour les particules en suspension, un travail spécifique a été menée dans le cadre du post traitement des données afin de récupérer l'information spatialisée et pouvoir redessiner la carte avec une meilleure bibliothèque graphique que celle fournie par le logiciel de traitement géostatistique.

Suite à cette étude, toutes les cartes sont disponibles en ligne sur le site internet de l'ORAMIP à l'adresse : <http://oramip.atmo-midipyrenees.org/l-air-de-ma-region/les-chiffres-du-jour/prevision>

## CONCLUSION

Les cartes de prévisions de la pollution de l'air issues de la plateforme de modélisation régionale de l'ORAMIP ont été fortement améliorées en termes de résolution, de précision et de rendu cartographique.

Ces cartes sont désormais associées à une évaluation de la probabilité de survenue d'un épisode de pollution intégrant l'incertitude de la méthode cartographique mais également l'incertitude des dispositifs de mesures en station. La probabilité de dépassement facilite l'aide à la décision en exploitant l'incertitude associée de la prévision pour fournir des zones sujettes à un risque d'épisode de pollution aux particules PM10. Le calcul d'incertitude fournit un critère qualitatif de la spatialisation qui permet d'évaluer la fiabilité d'une prévision. Ce nouvel outil autorise la mise en place d'une stratégie d'implantation des futures campagnes de mesures pour améliorer la spatialisation et réduire l'incertitude de la prévision avec une meilleure connaissance terrain.

D'autre part, les campagnes de mesures semi-fixes sur les territoires ne bénéficiant pas d'un suivi en continu sont désormais intégrées dans les données d'entrée de la plateforme de modélisation et cela permet d'améliorer la justesse des cartes de prévision de la pollution de l'air. La reconstruction statistique des données a permis de valoriser les campagnes de mesures effectuées sur des durées d'environ un an, en créant des stations virtuelles, complémentaires au dispositif fixe de mesure, et qui permettent de réduire les incertitudes de prévisions.

L'intérêt de la stratégie de couverture du territoire à l'aide de campagnes de mesures semi-fixes est donc confirmé par cette étude. Elle devra être poursuivie tant sur les territoires n'ayant pas encore fait l'objet d'un suivi que sur les sites ayant fait l'objet d'une campagne de mesures il y a plusieurs années.

## ANNEXE : Notions de géostatistiques

Le terme **covariable** désigne une variable géographique qui a une corrélation spatiale avec la variable étudiée, dans notre cas les concentrations en PM10. Une covariable est donc une variable auxiliaire qui a un rôle explicatif en rapport avec la variable étudiée.

Le **krigeage** est une méthode d'interpolation spatiale géostatistique qui consiste à définir une variabilité géographique en tenant compte de la relation qu'il existe entre les distances et les données. Dans notre cas, il s'agit des mesures de PM10 effectuées par les stations de surveillance de la qualité de l'air.

Le **cokrigeage colocalisé** est un krigeage qui utilise une ou plusieurs covariables pour affiner une distribution spatiale de la variable d'intérêt.

Un **variogramme** est une description de la variabilité moyenne d'un phénomène dans l'espace. Le variogramme est construit à partir des relations qu'il existe entre la distance entre les stations et les mesures associées. Il fournit une estimation de l'influence, ou portée, d'une mesure dans l'espace.

Le **modèle variographique** est défini et ajusté à partir du variogramme. Il reproduit le plus fidèlement possible la forme du variogramme et sera utilisé comme modèle théorique pour spatialiser les mesures dans l'espace.

La **validation croisée** consiste à enlever alternativement une station de mesures parmi le réseau, puis de comparer l'information estimée par la spatialisation avec la mesure effectuée. Cette validation croisée donne une estimation de l'erreur de la méthode statistique pour l'ensemble des stations du réseau de mesures, donc pour des implantations de typologie différente.



**ORAMIP**  
OBSERVATOIRE RÉGIONAL  
DE L'AIR EN MIDI-PYRÉNÉES  
Atmo Midi-Pyrénées

# Surveillance de la qualité de l'air en Midi-Pyrénées

24 heures/24 • 7 jours/7

• • prévisions • •

• • mesures • •



L'information  
sur la qualité de l'air  
en Midi-Pyrénées :

<http://oramip.atmo-midipyrenees.org>

ÉTUDE RÉALISÉE PAR ATMO MIDI-PYRÉNÉES ORAMIP  
FINANCEMENT DANS LE CADRE  
DU CONTRAT DE PROJETS ÉTAT-RÉGION MIDI-PYRÉNÉES

