



Qualité de l'air

Études

Juillet 2009



CART O3 : Surveillance et prévision de la pollution par l'ozone : Cartographies des concentrations d'ozone dans la région de l'étang de Berre

En collaboration avec :
NUMTECH



Association pour la Surveillance de la Qualité de l'Air de la Région de l'Etang de Berre et de l'Ouest des Bouches-du-Rhône
Route de la Vierge - 13 500 Martigues - Tel. 04 42 13 01 20 - Fax. 04 42 13 01 29
Site internet: www.airfobep.org - e-mail : airfobep@airfobep.org
Serveur vocal 04 42 49 35 35 (selon tarification téléphonique en vigueur)



Cartographie de la pollution par l'ozone dans La région de l'Etang de Berre Résumé

La région de l'Etang de Berre connaît des niveaux de pollution par l'ozone très élevés. Les maximums de concentration d'ozone en France, voir en Europe y sont observés.

Dans cette région, AIRFOBEP assure la surveillance de la pollution par l'ozone grâce à un réseau dense de capteurs et aux outils de modélisation mis en place au niveau national et régional .

En plus de cette surveillance, AIRFOBEP entreprend des études pour mieux comprendre cette pollution et mieux représenter son évolution spatio-temporelle dans la région de l'Etang de Berre.

AIRFOBEP a développé CARTO3, un outil de modélisation afin de représenter la spatialisation des épisodes de pollution par l'ozone dans la région de l'Etang de Berre.

Cet outil de modélisation a deux types de données d'entrée :

- les mesures du réseau de capteurs,
- les sorties des modèles de simulation.

La modélisation photochimique :

Utilisation des données de la plate forme nationale PREV'AIR

Pour réaliser les cartographies de pollution par l'ozone, AIRFOBEP utilise des données de modélisation photochimique. Ces données sont issues de la plate forme de modélisation PREV'AIR.

Le système PREV'AIR a été mis en place en 2003 à l'initiative du Ministère en charge de l'environnement afin de générer et de diffuser quotidiennement des prévisions et des cartographies de qualité de l'air, issues de simulations numériques.

Les modélisations réalisées par PREV'AIR sont le résultat de simulations numériques effectuées à l'aide de modèles eulériens déterministes tridimensionnels (dits "de chimie-transport"). Ces outils permettent notamment de calculer l'évolution de la pollution photo-oxydante et particulaire dans les basses couches de l'atmosphère.

Les modèles utilisés dans le système PREV'AIR sont :

- MOCAGE pour la météorologie,
- CHIMERE pour la photochimie.

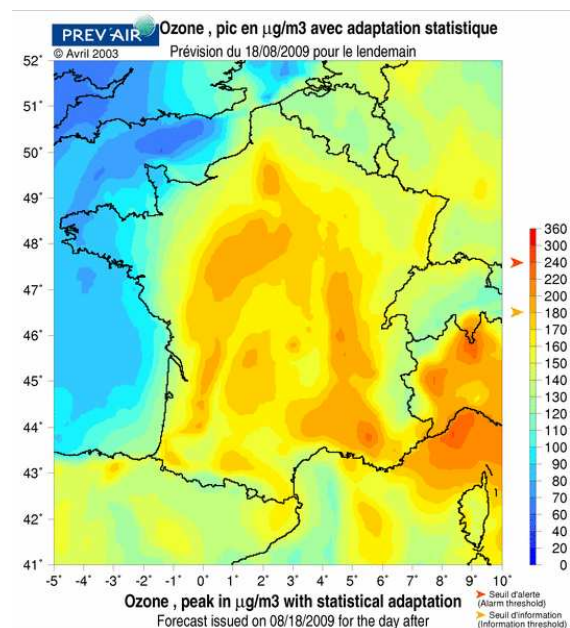


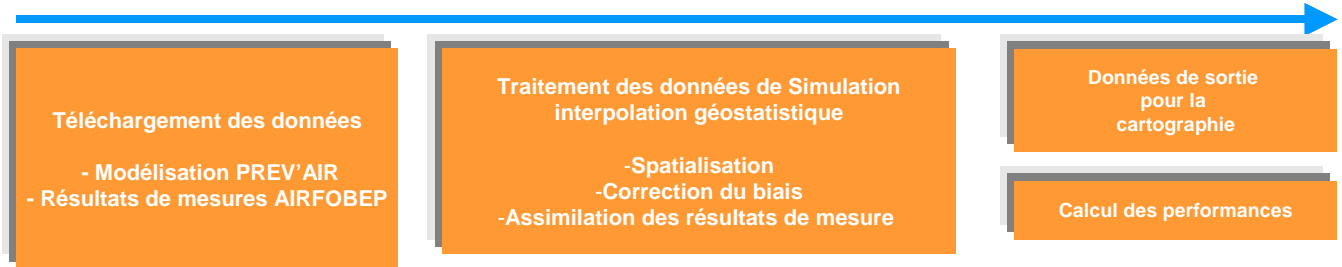
Illustration d'une simulation réalisée par PREV'AIR :
Prévision du 18 août 2900 des niveaux d'ozone pour le lendemain.

Traitement des données de modélisation : l'interpolation géostatistique

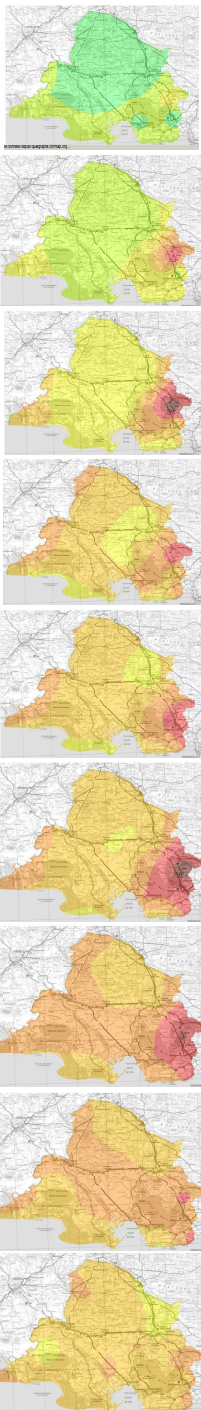
Les données de modélisation issues de PREV'AIR sont traitées à l'aide de méthodes de statistiques et d'interpolation complexes. Une méthode d'interpolation géostatistique, **le krigeage des innovations**, est notamment utilisée pour :

- **spatialiser les informations dans le but de les cartographier,**
- **corriger le biais évalué à partir de l'historique de fonctionnement de PREV'AIR (2004 – 2008),**
- **« assimiler » les résultats de mesures : c'est à dire corriger les données de modélisation avec les résultats de mesures obtenus sur la terrain.**

Les cartes de pollution par l'ozone



Cycle de production des cartographies de pollution par l'ozone



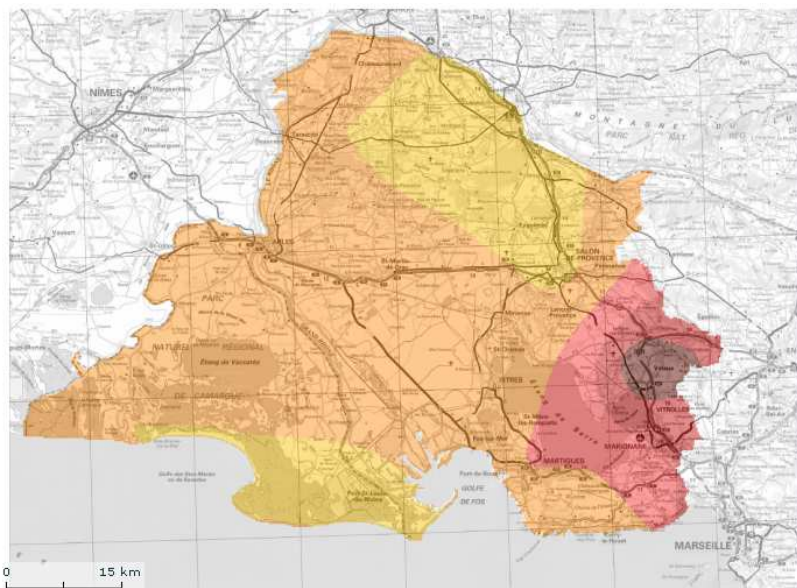
Une cartographie de la pollution par l'ozone est réalisée pour tout le territoire de surveillance d'AIRFOBEP.

Cette cartographie fournit les niveaux de pollution par l'ozone :

- constatés pour la veille,
- en prévision pour le jour même et pour les deux jours qui suivent.

Pour toutes ces échéances, sont fournies :

- des cartographies des maximums journaliers,
- les scénarii, heure par heure, de l'évolution spatio-temporelle de la pollution par l'ozone.



Cartographies de la pollution par l'ozone : Echelle de l'indice ATMO.
Journée du 18 août 2009 : maximum journalier des concentrations d'ozone

Ci-contre : évolution, heure par heure, des niveaux de concentration de l'ozone au cours de la journée du 18 août 2009 (10h – 18h)

Plus d'infos :

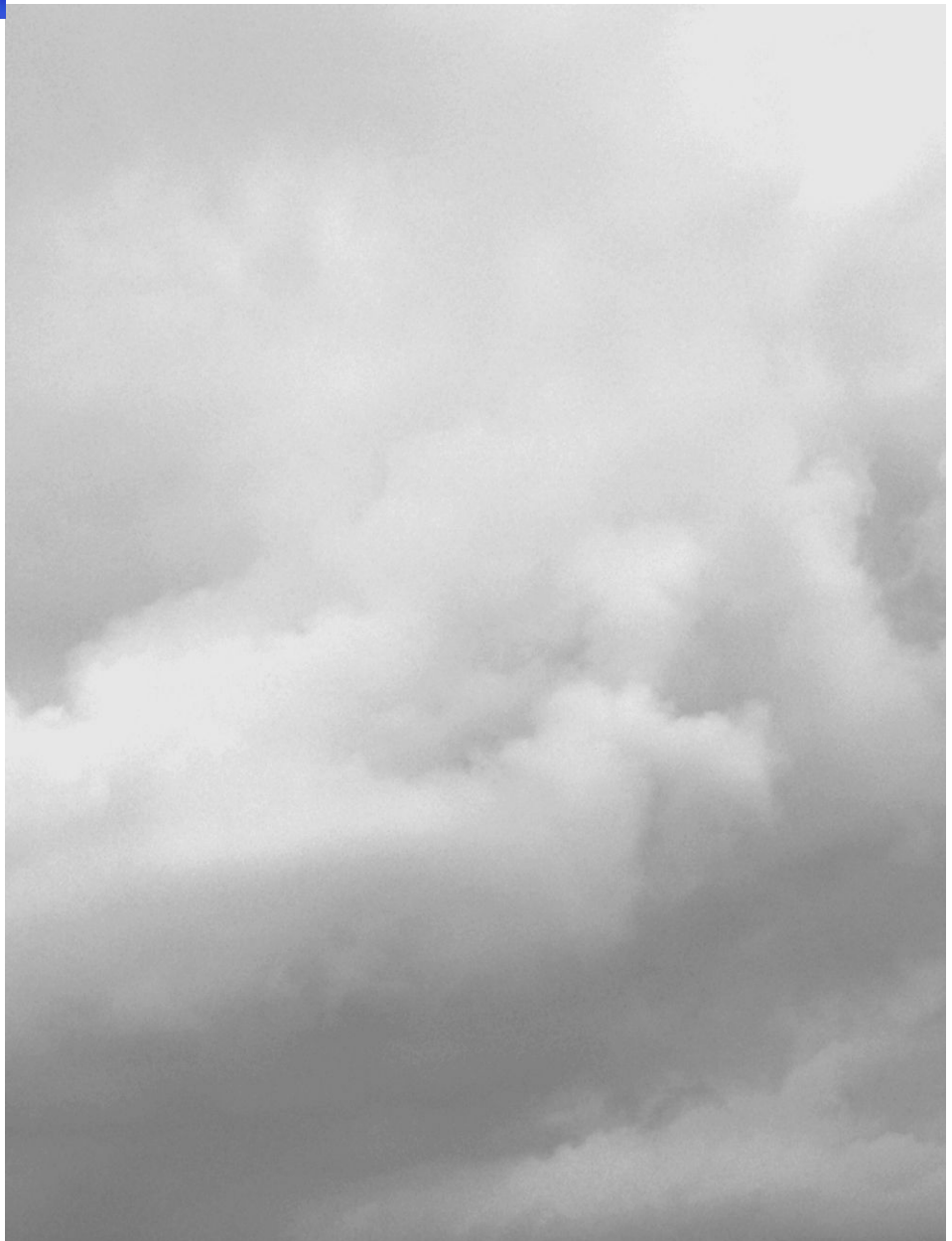
Surveillance et prévision des concentrations d'ozone sur le pourtour de l'étang de Berre : rapport AIRFOBEP 2009
Etude chronologique des pics d'ozone sur le pourtour de l'étang de Berre : Rapport AIRFOBEP 2007



Table des matières

1. Objet de l'étude	5
2. Description de la plate-forme opérationnelle et de son fonctionnement.....	7
2.1. Architecture du système	7
2.2. Principe de fonctionnement.....	8
2.3. Interface de contrôle et lancement des simulations.....	14
2.4. Interface web de suivi et calcul statistique.....	17
2.5. Interpolation et assimilation des données.....	21
2.6. Prise en compte du biais.....	24
<i>Table des figures.....</i>	26
<i>Table des tableaux.....</i>	27

Objet de l'étude



1. Objet de l'étude

AIRFOBEP souhaitait mettre en place une cartographie opérationnelle des niveaux d'ozone à l'échelle de son territoire d'intervention. Le présent document décrit la mise en place de la plate-forme CartO3, développée par NUMTECH, qui s'appuie sur les données disponibles dans la base BADOS (données d'observation et données de simulation PREV'AIR) pour l'élaboration des cartographies locales à l'échelle du domaine d'intervention d'AIRFOBEP. Il s'agit ainsi d'une application informatique qui réalise l'interpolation de données de concentration d'ozone disponible dans BADOS et de générer des fichiers qui pourront être utilisés pour visualiser des cartographies de concentration d'ozone. Le but est de disposer chaque matin de cartographies de concentrations horaires et maximales journalières, pour les jours J, J+1 et J+2 (cartes prévues) et pour le jour J-1 (cartes analysées).

Ce présent document décrit le fonctionnement de la plate-forme opérationnelle.



Description de la plate-forme opérationnelle et de son fonctionnement

2.1. Architecture du système

2.2. Principe de fonctionnement

2.3. Interface de contrôle et lancement

2.4. Interface web de suivi et calcul statistique

2.5. Interpolation et assimilation des données

2.6. Prise en compte du biais

2. Description de la plate-forme opérationnelle et de son fonctionnement

2.1. Architecture du système

L'outil CartO3 est lié au système opérationnel BADOS, base de données chimiques, qui lui fournit les données d'entrée qui sont nécessaires à son fonctionnement :

- des mesures chimiques observées par les stations du réseau AIRFOBEP ;
- des prévisions chimiques issues des simulations PREV'AIR.

A l'issue de ses calculs, l'outil CartO3 livre différents types de résultats :

- des grilles de concentrations en O3 au format binaire **.PGM** à destination du serveur FTP de **SIMALIS** pour l'affichage des **cartographies sur le site Internet (cartographies horaires et maximaux horaires sur la journée)** ;
- des **images JPG** permettant de valider les résultats obtenus (interface web de suivi NUMTECH) ;
- les **résultats intermédiaires et finaux** dans un format binaire compressé, qui seront stockés par l'**unité de téléchargement de BADOS** pour analyse ultérieure ;
- un historique par jour des observations horaires et des simulations pour les calculs statistiques.
- un **journal des opérations réalisées**, qui est transmis par **e-mail** (à ce jour, uniquement à NUMTECH) ;
- des **fichiers statistiques**, utilisés sur le site internet de suivi NUMTECH.

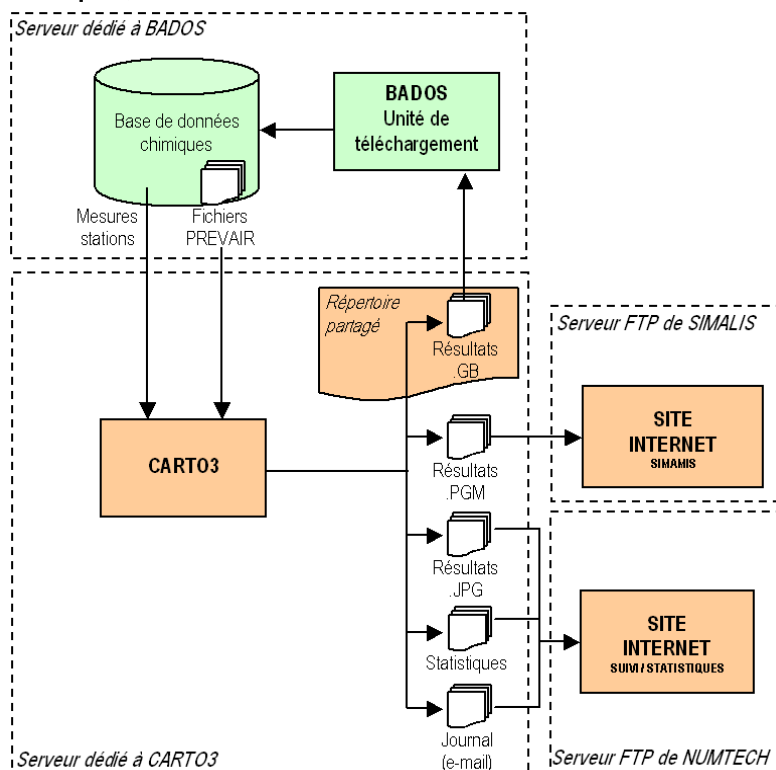


Figure 1 - Architecture du système

2.2. Principe de fonctionnement

La plate-forme CartO3 suit en opérationnel le principe de fonctionnement décrit ci-dessous. Les étapes d'interpolation et de prise en compte du biais sont décrites respectivement dans les sections 2.5 et 2.6. A noter que les cartographies horaires affichées sur le site SIMALIS sont en heures TU et non en heures locales.

Etape 1 : CartO3 récupère l'archive des prévisions PREVAIR du jour, la décompresse, et génère un fichier grille pour chaque heure de prévision disponible.

Pour cela, CartO3 se connecte au serveur partagé de BADOS, et cherche le fichier PREVAIR requis dans le répertoire de la base de données. Ce fichier devra correspondre aux prévisions en ozone pour le jour courant, et pour la grille PREVAIR définie dans le fichier de paramètres de CartO3. La grille PREVAIR à utiliser est spécifiée dans le fichier de configuration de CartO3, ainsi que le chemin d'accès à BADOS.

Une fois le fichier copié, celui-ci sera décompressé. Chaque archive est en effet compressé au format ZIP, et contient un fichier texte (ASCII) contenant chacune des heures de prévision, de J-1 (analysé) à J+2. CartO3 générera un fichier texte pour chaque heure de prévision, en spécifiant les coordonnées de chacun des points de la grille PREVAIR.

Par exemple, le 21/10/2008, le fichier PREVAIR correspondant aux prévisions en ozone sur la grille AWM se nommera « AWM_O3.20081021.zip » et contiendra le fichier texte « AWM_O3.20081021.txt » au format suivant :

2008102000	66.7	64.5	64.5	63.5	63.0	59.5	55.0	53.8	38.5	40.9	54.4	...
2008102001	64.6	62.6	64.0	64.2	63.3	57.8	54.5	55.5	48.0	43.6	52.1	...
2008102002	62.1	60.6	63.2	65.0	64.2	56.0	52.9	55.5	51.4	45.1	50.8	...
...												
2008102100	62.1	56.8	51.9	52.0	60.2	59.5	52.5	47.4	41.8	40.6	57.5	...
2008102101	62.0	56.4	51.3	52.4	60.1	59.6	52.0	46.6	41.6	40.1	57.6	...
2008102102	61.9	56.4	50.8	52.7	59.8	59.7	51.9	46.0	41.6	39.8	57.6	...
...												
2008102200	67.9	62.0	60.5	59.6	60.6	62.1	55.4	55.1	55.0	59.9	62.3	...
2008102201	67.7	60.9	59.9	59.6	60.7	60.6	53.3	52.7	54.1	59.8	58.0	...
2008102202	67.0	59.6	59.0	59.7	60.6	57.9	50.4	49.7	53.6	59.5	51.5	...
...												
2008112301	61.6	56.9	53.0	54.8	59.9	46.6	37.9	34.9	44.2	58.7	44.7	...
2008112302	60.8	57.0	53.2	54.1	58.1	45.8	36.3	32.3	39.2	54.0	45.0	...
2008112303	60.1	57.4	54.1	54.3	56.7	44.7	34.5	29.8	35.2	50.0	44.9	...
...												

Les fichiers générés par CartO3 se nommeront « 2008102000.txt », « 2008102001.txt », ...« 2008102023.txt » pour les fichiers analysés de J-1 entre 0 et 23 heures, « 2008102100.txt », « 2008102101.txt », ...« 2008102123.txt » pour les fichiers de prévisions sur J entre 0 et 23 heures, « 2008102201.txt », ...« 2008102223.txt » pour les fichiers de prévisions sur J+1 entre 0 et 23 heures, et « 2008102301.txt », ...« 2008102323.txt » pour les fichiers de prévisions sur J+2 entre 0 et 23 heures. Les fichiers auront le format ci-dessous :

579	4772	48.170577693043
579.5	4772	48.128617856927
580	4772	48.086204526735
580.5	4772	48.043332611837
581	4772	47.999999949762
581.5	4772	47.995577983064
582	4772	47.990700079485
582.5	4772	47.985373146223
583	4772	47.979606890902
583.5	4772	47.973413568014
...		
751.5	4888	63.308126804937
752	4888	63.352592933488
752.5	4888	63.396399652566
753	4888	63.439676910233

Les deux premières colonnes correspondent aux coordonnées X et Y de chacune des mailles PREVAIR.

Dans le fichier journal de CartO3, ces étapes correspondent aux éléments suivants :

- « *Lecture de la configuration des grilles PREVAIR sous BADOS...* »
- « *Création du répertoire de sortie...* »
- « *Décompression du fichier PREVAIR...* »
- « *Décomposition en fichiers horaires...* »

Etape 2 : CartO3 réalise une interpolation sur la grille fine de l'ensemble des fichiers horaires entre J-1 et J+2 à l'aide de Surfer.

La grille fine est spécifiée dans le fichier de configuration de CartO3, ainsi que la résolution de sortie. La méthode d'interpolation utilisée sous Surfer est le kriging.

Dans le fichier journal de CartO3, cette étape est indiquée par la ligne suivante : « *Krigeage en grille fine des fichiers horaires...* ».

Etape 3 : Pour chacune des cartes horaires entre J-1 et J+2, CartO3 introduit une correction des valeurs simulées au niveau et autour des stations d'observation au moyen d'un krigeage des innovations.

CartO3 réalise les opérations suivantes :

- extraction de la valeur simulée pour chaque station dont on possède des valeurs de biais (recherche des coordonnées des points les plus proches sur la grille correspondant à chacune des stations) à partir des grilles de calcul,
- extraction du biais au niveau de chaque station en fonction des valeurs tabulées par classes de concentrations,
- krigeage sur la grille fine des valeurs de biais,
- pour chaque point de la grille fine, addition de la valeur obtenue par le modèle avec la valeur du biais krigé.

Le biais est stocké dans un fichier « *biais.ini* » décrit au chapitre suivant ; il contient la liste des stations à considérer pour le calcul du biais, et les valeurs de biais associées, par classes de concentrations.

Dans le fichier journal de CartO3, ces étapes correspondent aux éléments suivants :

- « *Calcul des mailles les plus proches de chaque station sur la grille fine...* »
- Pour chaque heure de J-1 à J+2 :
 - o « *Jx 0h : lecture de la grille fine issue du modèle...* »
 - o « *Jx 0h : calcul du biais...* »
 - o « *Jx 0h : krigeage en grille fine de la correction du biais...* »
 - o « *Jx 0h : calcul sur la grille fine de modèle+krigeage...* »

Etape 4 : Pour chacune des heures de J-1 uniquement, les cartes analysées sont calculées à partir des mesures chimiques issues de BADOS.

CartO3 réalise les opérations suivantes pour chacune des heures de J-1 :

- connexion au serveur de partage de la base de données chimiques BADOS, et extraction des valeurs mesurées aux stations ;
- recherche des coordonnées des points les plus proches sur la grille correspondant à chacune des stations ;
- calcul de la différence entre modèle et mesures ;
- krigeage sur la grille fine des différences mesures - modèle ;
- pour chaque point de la grille fine, addition de la valeur obtenue par le modèle avec la différence mesures - modèle calculée précédemment.

Les stations considérées sont issues du fichier « *stations.ini* », décrit au chapitre suivant.

Dans le fichier journal de CartO3, ces étapes correspondent aux éléments suivants :

- « *Lecture des données mesurées à J-1 par les stations...* »
- « *Calcul des mailles les plus proches de chaque station sur la grille fine...* »
- Pour chaque heure de J-1 :
 - o « *J-1 analysé 0h : lecture de la grille fine issue du modèle...* »
 - o « *J-1 analysé 0h : calcul des différences mesures-modèle...* »
 - o « *J-1 analysé 0h : krigeage en grille fine des différences mesures-modèle...* »
 - o « *J-1 analysé 0h : calcul sur la grille fine de modèle+différence modèle/mesures...* »

Etape 5 : Les grilles (cartes analysées) obtenues pour chacune des heures de J-1 sont transformées en fichiers PGM, et des cartographies JPG sont tracées avec Surfer.

Ce format de fichier binaire est le format d'entrée nécessaire à SIMALIS pour le tracé des cartographies sur le site internet.

Ces fichiers PGM ont la notation suivante : « [date au format yyyy_mm_dd]_AJm1_[heure]_O3.pgm ». Par exemple, le fichier de la carte analysée de J-1 13h fourni lors du calcul du 09/11/2008 se nommera : « 2008_11_09_AJm1_13_O3.pgm »

Dans le fichier journal de CartO3, cette étape est spécifiée par la ligne suivante : « *J-1 analysé 0h : conversion de la grille fine en PGM...* »

L'ensemble des données sont également tracées sous forme de cartographies (avec fusion avec un fond de plan en transparence) afin d'être affichées sur le site internet de suivi NUMTECH. Les images sont stockées sous le nom : « [date au format yyyy_mm_dd]_AJm1_[heure]_O3.jpg ».

Etape 6: Les maxima journaliers, pour chaque jour entre J-1 et J+2, sont calculés à partir des grilles fines calculées à l'étape 2 et à l'étape 4 (carte analysée pour J-1).

Le site internet affiche en effet les valeurs maximales journalières en O3 (et non les moyennes journalières souhaitées dans l'application de calcul PM).

Dans le fichier journal de CartO3, cette étape est spécifiée par les lignes suivantes :

- « *J-1 : calcul du maximum journalier...* »
- « *J : calcul du maximum journalier...* »
- « *J+1 : calcul du maximum journalier...* »
- « *J+2 : calcul du maximum journalier...* »

Etape 7 : Les grilles de maxima journaliers, pour chaque jour entre J-1 et J+2, sont converties au format PGM, et des cartographies JPG sont tracées avec Surfer.

Ce format de fichier binaire est le format d'entrée nécessaire à SIMALIS pour le tracé des cartographies sur le site internet.

Le fichier PGM du maxima journalier de J-1 analysé a la notation suivante : « [date au format yyyy_mm_dd]_AJm1_O3_max.pgm ». Par exemple, le fichier des maxima de la carte analysée de J-1 fourni lors du calcul du 09/11/2008 se nommera : « 2008_11_09_AJm1_O3_max.pgm ».

Le fichier PGM du maxima journalier des prévisions pour J, J+1 et J+2 ont la notation suivante : « [date au format yyyy_mm_dd]_PJP[prévision]_O3_max.pgm ». Par exemple, le fichier des maxima pour J fourni lors du calcul du

09/11/2008 se nommera « 2008_11_09_PJ0_O3_max.pgm » ; celui pour J+1 « 2008_11_09_PJ1_O3_max.pgm » ; celui pour J+2 : « 2008_11_09_PJ2_O3_max.pgm »

Dans le fichier journal de CartO3, cette étape est spécifiée par les lignes suivantes :

- « J-1 : conversion du maximum journalier en PGM... »
- « J : conversion du maximum journalier en PGM... »
- « J+1 : conversion du maximum journalier en PGM... »
- « J+2 : conversion du maximum journalier en PGM... »

L'ensemble des données sont également tracées sous forme de cartographies (avec fusion avec un fond de plan en transparence) afin d'être affichées sur le site internet de suivi NUMTECH. Les images sont stockées sous les noms : « [date au format yyyy_mm_dd]_AJm1_O3.jpg » et « [date au format yyyy_mm_dd]_PJp[prévision]_O3_max.jpg ».

Etape 8 : Les grilles de chacune des heures entre J et J+2, sont converties au format PGM (les grilles de chacune des heures de J-1 – carte analysée – ont été converties à l'étape 4) , et des cartographies JPG sont tracées avec Surfer.

Ce format de fichier binaire est le format d'entrée nécessaire à SIMALIS pour le tracé des cartographies sur le site internet.

Ces fichiers PGM ont la notation suivante : « [date au format yyyy_mm_dd]_PJp[prévision]_[heure]_O3.pgm ». Par exemple, le fichier de la carte analysée de J 8h fourni lors du calcul du 09/11/2008 se nommera : « 2008_11_09_PJp0_8_O3.pgm »

Dans le fichier journal de CartO3, cette étape est spécifiée par les lignes suivantes :

- « J : conversion en PGM des valeurs horaires... »
- « J+1 : conversion en PGM des valeurs horaires... »
- « J+2 : conversion en PGM des valeurs horaires... »

L'ensemble des données sont également tracées sous forme de cartographies (avec fusion avec un fond de plan en transparence) afin d'être affichées sur le site internet de suivi NUMTECH. Les images sont stockées sous le nom : « [date au format yyyy_mm_dd]_AJm1_PJp[prévision]_[heure]_O3.jpg ».

Etape 9 : Les données calculées par CartO3 sont stockées dans un fichier binaire (compressé) destiné à être transmis à BADOS.

Les différentes valeurs stockées dans les fichiers .gb (format binaire propre aux applications développées pour AIRFOBEP) par CartO3 sont les suivantes :

- valeurs horaires obtenues sur la grille fine à partir des extractions PREVAIR, pour J-1 (issues de l'étape 2)
- valeurs horaires calculées sur la grille fine à partir des extractions PREVAIR additionnées au biais krigé, pour J-1 (issues de l'étape 3)
- valeurs horaires calculées sur la grille fine à partir des extractions PREVAIR additionnées au biais krigé et analysées (issues de l'étape 4)
- valeurs horaires obtenues sur la grille fine à partir des extractions PREVAIR, pour J (issues de l'étape 2)
- valeurs horaires calculées sur la grille fine à partir des extractions PREVAIR additionnées au biais krigé, pour J (issues de l'étape 3)
- valeurs horaires obtenues sur la grille fine à partir des extractions PREVAIR, pour J+1 (issues de l'étape 2)
- valeurs horaires calculées sur la grille fine à partir des extractions PREVAIR additionnées au biais krigé, pour J+1 (issues de l'étape 3)
- valeurs horaires obtenues sur la grille fine à partir des extractions PREVAIR, pour J+2 (issues de l'étape 2)
- valeurs horaires calculées sur la grille fine à partir des extractions PREVAIR additionnées au biais krigé, pour J+2 (issues de l'étape 3)

Le fichier .gb généré se nomme : « [date au format yyyyymmdd].gb ».

Etape 10 : Les statistiques sont calculées pour J-1.

Pour cela, CartO3 réalise les opérations suivantes :

- connexion au serveur de la base de données chimiques BADOS et extractions des mesures aux stations pour J-1 ;
- lecture dans le fichier binaire .gb du calcul courant, des valeurs calculées aux stations sur la grille fine à partir des extractions au biais krigé, et analysées, pour J-1
- lecture dans le fichier binaire .gb du calcul de la veille, des valeurs calculées aux stations sur la grille fine à partir des extractions au biais krigé pour J
- lecture dans le fichier binaire .gb du calcul de l'avant-veille, des valeurs calculées aux stations sur la grille fine à partir des extractions au biais krigé pour J+1
- lecture dans le fichier binaire .gb du calcul de l'avant-avant-veille, des valeurs calculées aux stations sur la grille fine à partir des extractions au biais krigé pour J+2

Par exemple, lors du calcul du 07/05/2008, CartO3 extraira :

- les valeurs mesurées aux stations le 06/05/2008 (J-1)
- les données calculées sur la grille fine + biais + analyse le 07/05/2008 pour J-1
- les données calculées sur la grille fine + biais le 06/05/2008 pour J
- les données calculées sur la grille fine + biais le 05/05/2008 pour J+1
- les données calculées sur la grille fine + biais le 04/05/2008 pour J+2

Le fichier de statistiques généré aura le format ci-dessous :

```

dates;stations;mesures;J-1   PREVAIR;J-1   PREVAIR+biais;J-1   PREVAIR+biais   analysé;J   PREVAIR;J
PREVAIR+biais;J+1 PREVAIR;J+1 PREVAIR+biais;J+2 PREVAIR;J+2 PREVAIR+biais
07/04/2009 0h;ACGM;25;48.69;30.88;24.99;48.85;30.8;68.91;49.26;66.45;46.68
07/04/2009 0h;AIXA;27;74.3;44.19;27.98;74.45;44.3;84.49;54.11;88.42;55.89
07/04/2009 0h;AIXP;34;70.03;39.43;34.28;70.18;39.56;83.37;52.46;87.82;54.73
07/04/2009 0h;BETG;19;72.33;48.7;19.51;72.52;48.89;84.95;59.33;95.37;69.13
07/04/2009 0h;CRAU;16;71.6;51.05;16.19;71.81;49.69;81.17;58.43;91.23;67.3
07/04/2009 0h;FSCB;31;79.89;57.01;31.23;80.12;54.57;88.22;62.66;95.23;67.88
...
07/04/2009 1h;ACGM;20;44.05;26.24;20.02;44.1;26.35;66.28;46.61;66.17;47.09
07/04/2009 1h;AIXA;27;69.78;39.43;26.99;76.26;46.27;82.27;52.19;84.65;53.52
07/04/2009 1h;AIXP;13;65.48;34.64;13.51;73.32;42.9;81.05;50.5;83.31;51.8
07/04/2009 1h;BETG;19;65.4;42.07;19.54;73.53;49.9;80.39;54.76;90.58;64.35
07/04/2009 1h;CRAU;27;66.27;45.66;27.09;70.69;50.03;81.15;58.37;90.67;66.82
07/04/2009 1h;FSCB;26;75.38;52.51;26.23;78.86;55.99;85.2;59.63;93.75;66.4
...
    
```

Le fichier se nomme simplement « stat.txt » et sera transmis au site internet NUMTECH de suivi.

La liste des stations prises en compte est stocké dans le fichier « stations.ini », décrit au chapitre suivant.

Dans le fichier journal de CartO3, cette étape est spécifiée par la ligne suivante : « Calcul des statistiques sur J-1 ».

Etape 11 : L'ensemble des fichiers PGM créés aux étapes 4, 6 et 7, sont transmis au serveur FTP de SIMALIS afin de tracer les cartographies sur le site internet.

Dans le fichier journal de CartO3, cette étape est spécifiée par la ligne suivante : « Transfert des fichiers PGM sur le serveur FTP... ».

Etape 12 : Le fichier binaire .gb, qui contient les résultats de CartO3, est transmis au serveur FTP dédié au partage de fichiers BADOS avec AIRFOBEP.

Dans le fichier journal de CartO3, cette étape est spécifiée par la ligne suivante : « Envoi sur le serveur FTP de BADOS du fichier .GB... ».

Etape 13 : Un e-mail, contenant le journal d'exécution, est automatiquement envoyé à CartO3 afin de permettre le contrôle de l'état des calculs.

Un exemple de journal est présenté ci-dessous :

```

04/09/2008 19:26:30  DEBUT DU TRAITEMENT DU 29/07/2008...
04/09/2008 19:26:30  Lecture de la configuration des grilles PREVAIR sous BADOS...
04/09/2008 19:26:30  Création du répertoire de sortie (D:\CartO3\sorties\20080729\)...
04/09/2008 19:26:31  Décompression du fichier PREVAIR (\\bados\2008\out-AFM_03.20080729.zip)...
04/09/2008 19:26:31  Décomposition en fichiers horaires...
04/09/2008 19:26:31  Krigeage en grille fine des fichiers horaires...
04/09/2008 19:28:15  Lecture des données mesurées à J-1 par les stations...
04/09/2008 19:28:15  Calcul des mailles les plus proches de chaque station sur la grille fine...
04/09/2008 19:28:16  J-1 analysé (28/07/2008 0h) : lecture de la grille fine issue du modèle...
04/09/2008 19:28:17  J-1 analysé (28/07/2008 0h) : calcul des différences mesures - modèle...
04/09/2008 19:28:17  J-1 analysé (28/07/2008 0h) : krigeage en grille fine des différences mesures
- modèle...
04/09/2008 19:28:22  J-1 analysé (28/07/2008 0h) : calcul sur la grille fine de modèle + différence
modèle/mesures...
04/09/2008 19:28:24  J-1 analysé (28/07/2008 0h) : conversion de la grille fine en PGM...
04/09/2008 19:28:24  J-1 analysé (28/07/2008 1h) : lecture de la grille fine issue du modèle...
04/09/2008 19:28:24  J-1 analysé (28/07/2008 1h) : calcul des différences mesures - modèle...
04/09/2008 19:28:24  J-1 analysé (28/07/2008 1h) : krigeage en grille fine des différences mesures
- modèle...
04/09/2008 19:28:30  J-1 analysé (28/07/2008 1h) : calcul sur la grille fine de modèle + différence
modèle/mesures...
04/09/2008 19:28:32  J-1 analysé (28/07/2008 1h) : conversion de la grille fine en PGM...
...
04/09/2008 19:31:25  J-1 (28/07/2008) : calcul du maximum journalier...
04/09/2008 19:31:32  J (29/07/2008) : calcul du maximum journalier...
04/09/2008 19:31:39  J+1 (30/07/2008) : calcul du maximum journalier...
04/09/2008 19:31:47  J+2 (31/07/2008) : calcul du maximum journalier...
04/09/2008 19:31:54  J-1 analysé (28/07/2008) : conversion du maximum journalier en PGM...
04/09/2008 19:31:55  J (29/07/2008) : conversion du maximum journalier en PGM...
04/09/2008 19:31:55  J+1 (30/07/2008) : conversion du maximum journalier en PGM...
04/09/2008 19:31:55  J+2 (31/07/2008) : conversion du maximum journalier en PGM...
04/09/2008 19:31:55  J (29/07/2008 0h) : conversion en PGM des valeurs horaires...
04/09/2008 19:31:56  J (29/07/2008 1h) : conversion en PGM des valeurs horaires...
04/09/2008 19:31:56  J (29/07/2008 2h) : conversion en PGM des valeurs horaires...
...
04/09/2008 19:32:02  J+1 (30/07/2008 0h) : conversion en PGM des valeurs horaires...
04/09/2008 19:32:03  J+1 (30/07/2008 1h) : conversion en PGM des valeurs horaires...
...
04/09/2008 19:32:09  J+2 (31/07/2008 0h) : conversion en PGM des valeurs horaires...
04/09/2008 19:32:09  J+2 (31/07/2008 1h) : conversion en PGM des valeurs horaires...
...
04/09/2008 19:32:16  Suppression des fichiers TXT temporaires...
04/07/2009 19:32:19  Transfert des fichiers PGM sur le serveur FTP...
04/07/2009 19:32:27  Envoi sur le serveur FTP de BADOS du fichier .GB...
04/09/2008 19:34:50  TRAITEMENT DU 29/07/2008 TERMINE AVEC SUCCES.
    
```

Ce journal est un fichier nommé « [date au format yyyy_mm_dd].log », et sera transmis au site internet NUMTECH de suivi.

Etape 14 : Les images JPG, le fichier de statistiques, et le fichier de journal sont transmis par FTP au site internet NUMTECH de suivi.

2.3. Interface de contrôle et lancement des simulations

2.3.1 Interface de contrôle

Un applicatif spécifique a été conçu par NUMTECH afin d'automatiser toutes les opérations de mise en fonctionnement de la plate-forme, allant du rapatriement des données d'observations chimiques et des prévisions chimiques à partir de la base de données BADOS d'AIRFOBEP, la production de cartographies en intégrant les étapes d'interpolation et calcul du biais, l'envoi sur le site ftp de SIMALIS pour affichage sur le site internet dédié à AIRFOBEP et au final la sauvegarde dans BADOS des champs de concentrations simulés.

Cet applicatif permet notamment l'envoi d'e-mails pour le suivi du bon fonctionnement et déroulement des étapes de production, ainsi que les relances manuelles de l'application (voir section suivante).

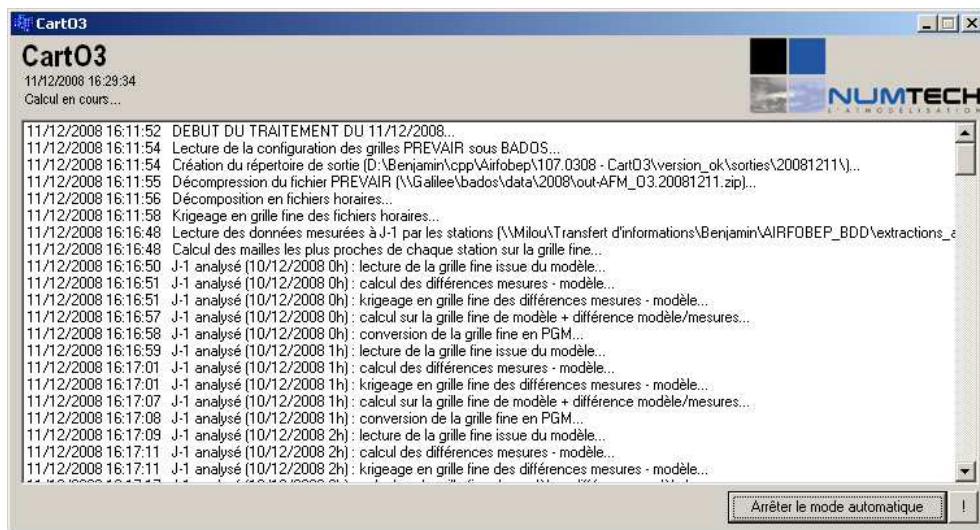


Figure 2 – Interface principale de CartO3.

L'interface de CartO3 affiche le journal du dernier calcul réalisé ; la couleur de fond de la fenêtre dépend de l'état du calcul :

- blanc lorsqu'un calcul est en cours,
- vert lorsque toutes les opérations ont pu être réalisées,
- orange lorsque des données étaient manquantes, ou lorsqu'un transfert n'a pu être effectué,
- rouge en cas d'erreur bloquante qui a empêché CartO3 de fournir des résultats en sortie.

Le bouton en bas à droite permet d'activer ou de désactiver le mode automatique de CartO3. Lorsque le mode automatique est activé, les opérations sont réalisées automatiquement à l'heure planifiée dans le fichier de configuration « options.ini » (situé dans le répertoire de CartO3).

Le bouton « ! » en bas à droite permet de forcer les calculs, pour une période donnée, lorsque ceux-ci n'ont pu être réalisés. La fenêtre suivante apparaît alors, et permet de choisir la période à calculer :

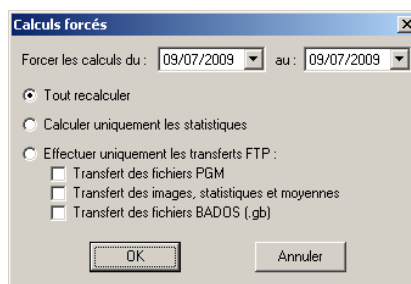


Figure 3 – Fenêtre de forçage de CartO3.

La fenêtre de forçage permet de configurer les calculs souhaités :

- effectuer un calcul complet (calculs puis transferts FTP aux différents serveurs), grâce à l'option « Tout recalculer » ;
- calculer uniquement les fichiers statistiques (« stat.txt »), grâce à l'option « Calculer uniquement les statistiques » (attention, ces fichiers ne sont pas automatiquement envoyés au site internet NUMTECH de suivi ;
- effectuer uniquement les transferts FTP, grâce à l'option du même nom, en choisissant le type de fichiers à transférer :
 - o fichiers PGM pour SIMALIS,
 - o fichiers images et statistiques pour le site internet NUMTECH de suivi
 - o fichiers .GB pour BADOS

2.3.2 Fichier de configuration générale de CartO3

Le fichier de configuration de CartO3, nommé « options.ini », et situé dans le répertoire de CartO3, contient les éléments suivants :

[Répertoires]	
Configuration BADOS=	répertoire contenant les fichiers de configuration de BADOS
Prévisions O3 PREVAIR (BADOS)=	répertoire contenant la base de données BADOS
Sorties=	répertoire de stockage des sorties de CartO3
Surfer=	répertoire contenant l'exécutable de Surfer (version 8 ou 9)
bados_donnees=	répertoire contenant la base de données chimiques BADOS
[Grille PREVAIR]	
Grille=	grille PREVAIR à utiliser
Facteur_GrilleSortie=	facteur avec la grille de sortie ; la grille PREVAIR étant définie en km, et la grille de sortie étant définie en m, le facteur à spécifier est 1000
[Grille de sortie]	
x_min=	coordonnée x min de la grille fine en sortie de CartO3 (en m)
x_max=	coordonnée x max de la grille fine en sortie de CartO3 (en m)
y_min=	coordonnée y min de la grille fine en sortie de CartO3 (en m)
y_max=	coordonnée y max de la grille fine en sortie de CartO3 (en m)
resolution=	résolution de la grille fine en sortie de CartO3 (en m)
[Mode automatique]	
heure=	heure d'exécution du calcul de CartO3, en mode automatique
mail_serveur=	adresse du serveur de mail à utiliser pour l'envoi des e-mails
mail_port=	port du serveur de mail à utiliser pour l'envoi des e-mails
mail_login=	login sur le serveur de mail à utiliser pour l'envoi des e-mails
mail_destinataires=	liste des adresses e-mail destinataires des e-mails de CartO3
[Envoi FTP PGM]	
serveur=	paramètres du serveur FTP de SIMALIS recevant les PGM de sortie
port=	
login=	
password=	
repertoire=	
[Envoi FTP web_stat]	
serveur=	paramètres du serveur FTP de NUMTECH recevant les jpg et stat
port=	
login=	
password=	
repertoire=	
[Envoi FTP BADOS]	
serveur=	paramètres du serveur FTP de AIRFOBEP pour BADOS
port=	


```
login=
password=
repertoire=
```

2.3.3 Fichier de configuration des stations dans CartO3

La liste des stations à prendre en compte pour le calcul des cartes analysées sur J-1 et des statistiques est stockée dans le fichier « stations.ini ». Les coordonnées des stations spécifiées dans ce fichier sont également utilisées pour le calcul du biais krigé ; cela signifie que toutes les stations spécifiées dans le fichier « biais.ini » doivent être également définies dans le fichier « stations.ini ». En outre, les codes des stations doivent être identiques à ceux spécifiés dans la configuration de la base de données BADOS.

Le format du fichier est : code station BADOS ; coordonnée x (en m) ; coordonnée y (en m)

```
ACGM;619621;4824728
AIXA;683459;4798273
AIXP;684327;4800205
BETG;675633;4817059
CRAU;649853;4818103
FSCB;656474;4813592
ISTR;660228;4819300
LPNT;622090;4825634
MNDM;665398;4809113
MRMV;661979;4827737
RBRT;679423;4819906
SAZE;613396;4824346
SLPV;668693;4833800
SMMR;615420;4812578
SRMY;647508;4849996
SSLP;671688;4799793
VTRL;681570;4814036
```

2.3.4 Fichier de configuration du calcul de biais dans CartO3

La liste des stations à prendre en compte pour le calcul du biais krigé est spécifiée dans le fichier « biais.ini ». Les codes des stations doivent être identiques à ceux spécifiés dans la configuration de la base de données BADOS, et les coordonnées des stations doivent être définies dans le fichier « stations.ini » de CartO3.

Le biais est défini pour chaque station, et chaque classe ; le format du fichier est : classe ; biais station 1 ; biais station 2 ; biais station 3 ; etc...

```
classes;BETG;FSCB;ISTR;MNDM;MRMV;RBRT;SLPV;SMMR;SRMY;SSLP;VTRL
10;-6;-5.2;-1;9.1;0.4;0;47.1;0;-2.1;0;-3.1
20;-11.7;-6;-8.7;1.2;-0.4;0.7;-12.6;1.9;-8.5;3.6;1
30;-10.2;-8.7;-3.3;-6.6;-1.7;-4.4;-7.9;-4.2;-18.1;-5.3;-13.5
40;-17.8;-12.7;-10.1;-13.3;-9;-9.8;-14;-6.9;-20.1;-17;-20.5
50;-20.2;-16;-13;-14.7;-10.6;-17;-18.2;-8.5;-20;-22.3;-23.8
60;-21.1;-18.9;-15;-15.4;-12.4;-19.6;-20.4;-13.9;-20;-20.8;-23.7
70;-23.3;-20.7;-17.1;-17.6;-13.9;-19.9;-22.1;-16.4;-19.6;-21;-23.5
80;-23.6;-22.9;-19.1;-18.5;-14.8;-20.2;-22.4;-17.2;-19.9;-21.8;-22.6
90;-25.6;-25.6;-21.4;-20.2;-17.1;-21.3;-22.4;-18.4;-20.8;-24.4;-23.8
100;-26.2;-27.4;-22.2;-21.8;-16.9;-21.7;-22.2;-19.5;-25.1;-26.1;-25.7
110;-25.2;-28.3;-23.7;-23.2;-16;-20.2;-21.7;-19.8;-24.5;-28.6;-25.1
...
```

Dans le fichier de configuration ci-dessus, cela signifie que le biais ajouté à la valeur de la station ISTR sera de -3.3 si la valeur calculée à ISTR est supérieure ou égale à 30 et inférieure à 40.

2.3.5 Relance automatique et manuelle

La procédure qualité mis en place à AIRFOBEP conduit à une ré-actualisation ou correction des données de mesures chimiques de la base BADOS. Une première ré-actualisation est généralement effectuée les 2-3 jours après la mise à

disposition des données brutes d'observations dans BADOS. Ensuite, une mise à jour mensuelle, puis annuelle, est éventuellement réalisée.

Vis-à-vis des corrections hebdomadaires, il a été décidé que chaque jour, la plate-forme CartO3 vérifie si, sur les dix derniers jours, les données chimiques de BADOS ont été mises à jour par rapport aux données chimiques qui ont été extraites lors du fonctionnement quotidien de la plate-forme. Si cela est le cas, les cartographies analysées pour le jour correspondant sont refaites pour affichage sur le site SIMALIS, ainsi que les statistiques aux stations affichées sur l'interface web NUMTECH.

Vis-à-vis des corrections mensuelles ou annuelles, il n'est pas possible d'envisager une relance automatique de la plate-forme CartO3. Il a été convenu avec AIRFOBEP de pouvoir relancer manuellement la plate-forme sur les dates considérés. Cette relance se fera par AIRFOBEP à partir de l'interface de contrôle, puisqu'à terme l'ensemble de l'application sera transféré à AIRFOBEP. La procédure de relance manuelle est décrite dans la guide utilisateur associée à l'application.

2.4. Interface web de suivi et calcul statistique

Dans le but de suivre le bon fonctionnement de la plate-forme en terme de performance de modélisation, un module de statistiques a été mis en place. Celui-ci permet de comparer les mesures, faites aux stations AIRFOBEP, aux simulations effectuées quotidiennement., c'est-à-dire qu'il permet de suivre au jour-le-jour, via une interface web, les performances de la plate-forme. De plus, cette interface permet de calculer des statistiques de performances globales de la plate-forme en pouvant sélectionner des périodes de calcul.

Cette interface web a donc les fonctions suivantes :

- Afficher les cartes produites par NUMTECH avec SURFER. Ceci permet de contrôler la cohérence des cartes affichées sur le site SIMALIS et ainsi de pointer éventuellement un problème de transfert ftp entre NUMTECH et SIMALIS.
- Afficher les statistiques journalières pour un suivi quotidien des performances.
- Calculer et afficher des statistiques de performance sur une période donnée.

Cette interface est actuellement en construction et cette note sera mise à jour lors de la mise en opérationnelle de cette interface. Néanmoins, les sections suivantes décrivent les statistiques envisagées à ce jour sur la page journalière et la page de statistique globale.

En effet, en complément de la sauvegarde des concentrations sur une grille 2D permettant la création de cartographies, un fichier texte est créé (et mis à jour) chaque jour pour le jour J-1. Ce fichier texte contient pour chaque heure de la journée et pour chaque station la valeur mesurée, ainsi que les valeurs simulées avec la simulation basée sur les observations de ce jour (calcul J-1 du jour J), avec la prévision J du jour précédent (J-1), avec la prévision J+1 du jour J-2 et avec la prévision J+2 du jour J-3. Ceci permet ainsi de pouvoir comparer d'une part les observations et d'autre part les résultats de simulation de ce jour à partir des calculs J-1, J, J+1 et J+2.

A partir de ces fichiers textes, deux types d'affichage statistique sont alors réalisés :

- **Un affichage des statistiques au jour le jour :**

Dans l'interface correspondante, le jour J, les mesures effectuées pour J-1 sont comparées à :

- la simulation de J-1 obtenue le jour J,
- la simulation de J effectuée le jour J-1,
- la simulation de J+1 effectuée le jour J-2,
- la simulation de J+2 effectuée le jour J-3.

Les statistiques qui sont alors calculées et affichées concernent,

- d'une part, les valeurs horaires :
 1. les valeurs horaires observées,
 2. les valeurs horaires simulées,
 3. le biais entre les valeurs horaires simulées et les valeurs horaires observées,
 4. Le sous-indice ATMO issu de la valeur horaire observée,
 5. Le sous-indice ATMO issu de la valeur horaire simulée,
 6. Le biais en terme de sous-indice ATMO entre observation et simulation.
- d'autre part, les valeurs journalières :
 7. les valeurs journalières observées,
 8. les valeurs journalières simulées,
 9. le biais entre les valeurs journalières simulées et les valeurs journalières observées.
 10. le maximum horaire observé sur la journée,
 11. le maximum horaire simulé sur la journée,
 12. le biais entre le maximum horaire simulé et observé sur la journée,
 13. Le sous-indice ATMO issu du maximum horaire observé,
 14. Le sous-indice ATMO issu du maximum horaire simulé,
 15. Le biais en terme de sous-indice ATMO entre observation et simulation.

Ces statistiques concernent une station donnée ou toutes les stations, un cycle de simulation donné (J-1¹, J, J+1 ou J+2) et une date donnée. L'affichage par défaut de l'interface est la date du jour courant. Cependant, il est possible de choisir une date passée afin de visualiser les statistiques journalières de toute date.

- **Un calcul et affichage des statistiques de performance sur une période donnée :**

Cette interface permet de calculer et afficher les performances de la plate-forme CartO3 sur une période donnée.

Ces calculs et affichages sont réalisés :

- pour une station donnée,
- ou pour l'ensemble des stations,
- ou par type de stations (rurale, semi-urbaine, ...),
- et pour un cycle de simulation donné (J-1², J, J+1 et J+2).

De plus, ils concernent :

- les sorties PREV'AIR uniquement,
- les sorties de la plate-forme CartO3 avant analyse (PREV'AIR + biais),
- et les sorties de la plate-forme CartO3 après analyse (PREV'AIR + biais analysé).

Plusieurs calculs de performances sont alors calculés :

- des statistiques sur les concentrations horaires :
 - ✓ Par classe de concentration, le nombre d'éléments observés, le nombre d'éléments simulés, par rapport aux observations le biais moyen simulé, la NMSE, le quantile 2.5% des erreurs, le quantile 97.5% des erreurs et l'intervalle à 95%. Les classes de concentration retenues sont : 0-29, 30-54, 55-79, 80-104, 105-129, 130-149, 150-179, 180-209, 210-239 et > 240 µg/m³.
 - ✓ Toute concentration confondue, la moyenne des observations, la moyenne des concentrations simulées, le biais entre observation et simulation, la NMSE, le quantile 2.5% des erreurs, le quantile 97.5% des erreurs, l'intervalle à 95% des erreurs et le paramètre E-20% qui désigne le pourcentage d'heures pour lesquelles les concentrations simulées et observées diffèrent de moins de 20%.
- des statistiques sur les maximums horaires journaliers :

¹ Ceci correspond bien évidemment aux valeurs J-1 avant création de la carte analysée.

² Ceci correspond bien évidemment aux valeurs J-1 avant création de la carte analysée.

- ✓ Par classe de concentration, le nombre d'éléments observés, le nombre d'éléments simulés, par rapport aux observations le biais moyen simulé. Les classes de concentration retenues sont : 0-29, 30-54, 55-79, 80-104, 105-129, 130-149, 150-179, 180-209, 210-239 et > 240 $\mu\text{g}/\text{m}^3$.
 - ✓ La moyenne des concentrations maximales horaires journalières mesurées.
 - ✓ La moyenne des concentrations maximales horaires journalières simulées.
 - ✓ Le biais moyen entre ces moyennes.
 - ✓ Le pourcentage de bonne prévision (sous-indice ATMO).
 - ✓ Le pourcentage de prévision (sous-indice ATMO) sous-estimante.
 - ✓ Le pourcentage de prévision (sous-indice ATMO) sur-estimante.
 - ✓ Le biais moyen sur le sous-indice ATMO.
- des statistiques sur les moyennes journalières :
- ✓ Par classe de concentration, le nombre d'éléments observés, le nombre d'éléments simulés, par rapport aux observations le biais moyen simulé, le quantile 2.5% des erreurs, le quantile 97.5% des erreurs et l'intervalle à 95%. Les classes de concentration retenues sont : 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99, 100-119 et > 120 $\mu\text{g}/\text{m}^3$.
 - ✓ Toute concentration confondue, la NMSE journalière (erreur quadratique moyenne normalisée).
 - ✓ Toute concentration confondue, l'écart type des biais journaliers.
 - ✓ Toute concentration confondue, l'écart type élargie à 2 des biais journaliers.
 - ✓ Toute concentration confondue, le coefficient de corrélation R^2 journalier.
 - ✓ Toute concentration confondue, la moyenne des moyennes journalières mesurées.
 - ✓ Toute concentration confondue, la moyenne des moyennes journalières simulée.
 - ✓ Toute concentration confondue, le biais moyen entre ces moyennes.
 - ✓ Toute concentration confondue, le paramètre E-20% qui désigne le pourcentage de jours pour lesquels les concentrations simulées et observées diffèrent de moins de 20%
- des statistiques sur les dépassements de la valeur seuil de 180 $\mu\text{g}/\text{m}^3$ pour les valeurs horaires (vis-à-vis des observations d'une part, et vis-à-vis des simulations d'autre part) :
- ✓ le nombre et la liste des journées concernées sur la base des observations : la valeur observée, la valeur simulée et le biais sont alors affichés.
 - ✓ le nombre et la liste des journées concernées sur la base des simulations : la valeur observée, la valeur simulée et le biais sont alors affichés
- des statistiques sur des seuils ou paramètres réglementaires :
- ✓ seuil de 120 $\mu\text{g}/\text{m}^3$ sur une heure :
 - un tableau de contingence qui présente le nombre d'heures observées pour la période sélectionnée (paramètre X), le nombre d'heures observées qui dépassent ce seuil (Y), le nombre d'heures observées qui ne dépassent pas ce seuil (Z), le nombre d'alertes correctes (A), le nombre de fausses alertes (B), le nombre d'évènements manqués (C) et le nombre de non-alertes correctes (D).

Nb échéances horaires : X	Obs > seuil : Y	Obs < seuil : Z
Prévision > seuil	A	B
Prévision < seuil	C	D
Taux de bonne prévision		
		%
Taux de prévision fausse		
		%
Paramètre Alpha		
		%
Paramètre Béta		
		%

Associé à ce tableau sera fournis également :

- Le taux de bonne prévision (alertes et non alertes) : $(A+D)/(A+B+C+D)$.
- Le taux de prévision fausse (fausses alertes ou alertes manquées) : $(B+C)/(A+B+C+D)$.
- La probabilité de détection (POD ou paramètre alpha, valeur optimale = 1) : $A / (A+C)$.
- Le taux de fausse alerte (FAR ou paramètre bêta, valeur optimale = 0) : $B/(A+B)$

- La moyenne des concentrations mesurées dépassant ce seuil.
- La moyenne des concentrations simulées dépassant ce seuil en ne considérant que les échéances pour lesquelles une observation existe.
- Le biais moyen entre ces moyennes observées et simulées.
- La moyenne des concentrations simulées dépassant ce seuil en considérant toutes les échéances, y compris celles pour lesquelles une observation n'est pas présente.
- ✓ seuil de 180 $\mu\text{g}/\text{m}^3$ sur une heure : Les mêmes paramètres que pour le seuil de 120 $\mu\text{g}/\text{m}^3$ sur une heure.
- ✓ seuil de 240 $\mu\text{g}/\text{m}^3$ sur une heure : Les mêmes paramètres que pour le seuil de 120 $\mu\text{g}/\text{m}^3$ sur une heure.
- ✓ seuil de 110 $\mu\text{g}/\text{m}^3$ sur 8 heures (moyenne glissante) : Les mêmes paramètres que pour le seuil de 120 $\mu\text{g}/\text{m}^3$ sur une heure.
- ✓ AOT40 (cumul des écarts positifs de concentration au-dessus de 40 ppb pour les valeurs entre 8 et 20h) :
 - La valeur observée sur la période.
 - La valeur simulée sur la période en ne considérant que les échéances pour lesquelles une observation existe.
 - Le biais moyen entre observation et simulation.
 - La valeur simulée sur la période en considérant toutes les échéances, y compris celles où aucune observation existe.
- ✓ AOT60 (cumul des écarts positifs de concentration au-dessus de 60 ppb pour les valeurs entre 8 et 20h) : Les mêmes paramètres que pour l'AOT40.
- ✓ AOT90 (cumul des écarts positifs de concentration au-dessus de 90 ppb pour les valeurs entre 8 et 20h) : Les mêmes paramètres que pour l'AOT40.

Les stations concernées par le calcul des statistiques sont les suivantes (ceux sont également les stations qui sont utilisées pour la création des cartes analysées) :

- ACGM : pas de type défini
- AIXA : pas de type défini
- AIXP : pas de type défini
- BETG : station urbaine
- CRAU : pas de type défini
- FSCB : pas de type défini
- ISTR : station urbaine
- LPNT : pas de type défini
- MNDM : station urbaine
- MRMV : station urbaine
- RBRT : station industrielle
- SAZE : pas de type défini
- SLPV : station urbaine
- SMMR : station rurale
- SRMY : station rurale
- SSLP : station industrielle
- VTRL : station urbaine

2.5. Interpolation et assimilation des données

Pour rappel, la plate-forme opérationnelle a deux objectifs principaux :

- ✓ Fournir chaque matin une prévision pour le jour même (J), le lendemain (J+1) et le sur-lendemain (J+2) de la pollution en ozone sur la zone d'intervention d'AIRFOBEP. Ces cartes complètent les outils de surveillance et de prévision de la pollution en ozone opérés par ailleurs par AIRFOBEP, tel que VIGIPOL.
- ✓ Fournir chaque matin une carte analysée de la pollution en ozone de la veille (J-1) sur la zone d'intervention d'AIRFOBEP. Cette carte analysée complète ainsi les outils de surveillance et suivi de la pollution opérés par ailleurs par AIRFOBEP.

Les résultats des simulations PREV'AIR sont obtenus en des points localisés explicitement (x,y,z) avec un espacement de 10 km dans les deux directions X et Y (grille haute résolution FRAN01, voir figure ci-dessous). Or, les cartes qui seront présentées sur le site Internet d'AIRFOBEP sont réalisées par la société SIMALIS. La production de ces cartes nécessite de fournir à SIMALIS des données sur une grille régulière. Afin d'avoir un rendu visuel de qualité (notamment suite à l'étape d'introduction d'un biais) et afin de faciliter l'étape suivante d'assimilation des données d'observation pour me cycle J-1, il a été retenu de réaliser une première étape d'interpolation spatiale des sorties brutes PREV'AIR. Ceci est présenté dans la section 2.5.1.

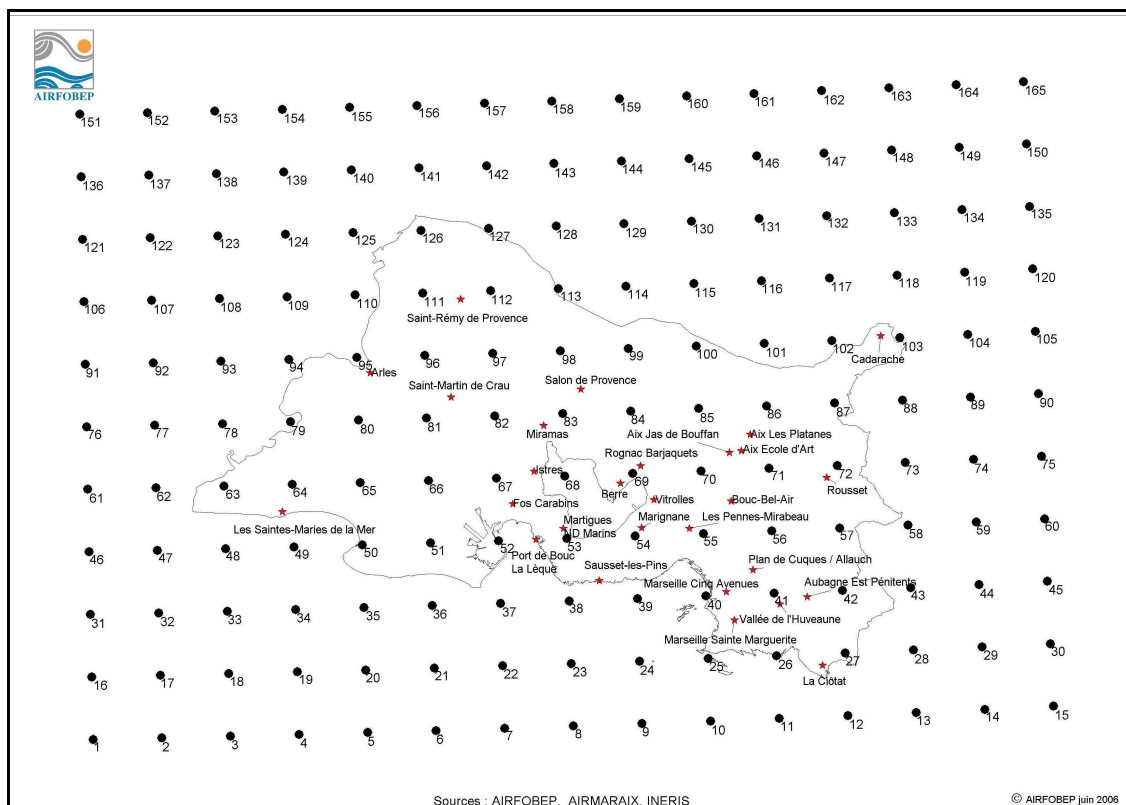


Figure 4 – Illustration de la grille de sortie haute résolution des simulations PREV'AIR (les points rouges correspondent à la localisation des stations d'observations AIRFOBEP).

Comme indiqué ci-dessus, la production de la carte analysée pour le jour J-1 consiste à d'intégrer au champ 2D simulés les données d'observation afin d'obtenir une cartographie en ozone la plus proche possible de la réalité, l'objectif étant de disposer d'un moyen de surveillance de cette pollution. Cette étape est écrite dans la section 2.5.2.

2.5.1 Interpolation des champs de concentration sur une grille régulière

Le choix de la résolution de la grille régulière doit satisfaire à deux critères :

- ✓ ne pas introduire d'artefact d'interpolation en certaines zones du domaine par rapport aux sorties brutes PREV'AIR,

- ✓ limiter la résolution afin de ne pas créer des fichiers trop importants qui rendraient leurs traitements et leurs mises à disposition difficiles,
- ✓ permettre d'appliquer l'étape d'assimilation des données d'observations en n'impactant pas des zones trop importantes (ce qui serait le cas en gardant une résolution spatiale de 10 km).

Le choix de NUMTECH s'est porté sur une résolution de 500 mètres qui permet de satisfaire aux critères précédemment cités. Cette interpolation des données est effectuée grâce au logiciel Golden Software SURFER en utilisant une méthode d'interpolation qui préserve le champ initial en évitant de créer des valeurs interpolées éloignées des valeurs initiales (sorties brutes PREV'AIR).

2.5.2. Assimilation des données de mesures pour le cycle J-1

Pour la production de la carte analysée pour le jour J-1, il a été retenu de faire une interpolation géostatistique en utilisant des méthodes de krigeage. Le krigeage est une méthode d'interpolation spatiale qui permet une estimation linéaire basée sur l'espérance mathématique et aussi sur la variance de la donnée spatialisée. A ce titre le krigeage se base sur le calcul, l'interprétation et la modélisation du variogramme qui est une appréciation de la variance en fonction de la distance entre données.

Il existe plusieurs types de krigeage, seuls certains d'entre eux sont applicables au cas qui nous intéressent ici :

- ✓ krigeage des observations,
- ✓ krigeage des innovations,
- ✓ krigeage avec le modèle en dérive externe.

Chacune de ces méthodes a ses spécificités en terme de temps de calcul et de difficulté de mise en œuvre. Dans le cadre de ce projet, c'est une méthode de krigeage des innovations qui a été retenue. Cette méthode est celle utilisée par l'INERIS dans le cadre de la plate-forme PREV'AIR pour l'assimilation des données d'observations d'ozone. Cette méthode est également celle reprise par différentes AASQA pour « régionaliser » les prévisions PREV'AIR sur leur domaine d'intervention en intégrant leurs mesures locales. Contrairement au krigeage des observations qui ne s'intéressent qu'aux observations, le krigeage des innovations consiste à ajouter aux résultats du modèle une combinaison linéaire des erreurs (différences modèle-mesure) calculés aux stations de mesure. Autrement dit, cette méthode est une adaptation statistique qui permet de corriger par les observations les simulations de concentration. La méthode utilisée est décrite dans les deux rapports suivants : « Honoré C., Malherbe L., 2003, Application de modèles grande échelle à la problématique régionale : cas de l'ozone, Rapport LCSQA » et « Honoré C., Malherbe L., 2004, Application de modèles grande échelle à la problématique régionale : cartographie de l'ozone et du dioxyde d'azote, Rapport LCSQA ».

Dans la méthode de krigeage des innovations, l'hypothèse formulée est que la concentration du polluant considéré en un point p est la somme de la concentration obtenue par modélisation en ce point et d'un résidu aléatoire :

$$Z(p) = \text{mod}(p) + \tilde{Z}(p)$$

Aux points p_s des stations de mesure, ce résidu est la différence entre la concentration mesurée et concentration simulée. On désigne cette différence sous le terme d'innovations.

Le krigeage des innovations consiste à interpoler par krigeage ordinaire la variable \tilde{Z} aux points de la grille. En un point p de cette grille, l'estimation de Z s'obtient donc en ajoutant au modèle la valeur interpolée de \tilde{Z} :

$$\tilde{Z}(p) = \text{mod}(p) + \sum_i \lambda_i \tilde{Z}(p_i)$$

- $\text{mod}(p)$ est la valeur du modèle au point p .
- $\tilde{Z}(p_s)$ est l'innovation à la station de mesure p_s , calculée par la relation :

$$\tilde{Z}(p_s) = Z(p_s) - \text{mod}(p_s)$$

- $Z(p_s)$: observation au point p_s
- $\text{mod}(p_s)$: valeur du modèle interpolée au point p_s

Les poids λ_i sont alors solutions du système de krigeage associé à la variable \tilde{Z}

Les deux figures suivantes illustrent l'impact du krigeage sur les sorties PREV'AIR. Notons toutefois que l'impact est d'une part de corriger localement les concentrations au niveau des données d'observations et d'autre part de corriger les niveaux de concentrations autour de ces points. Cependant, cette méthode n'a pas pour objet de modifier fortement le champ de concentration initial. L'interaction de données d'observation avec des valeurs différentes couplés au champ initial peut conduire, comme cela est le cas sur les figures présentées, à observer la création de zones très locales qui se distinguent du champ général (donnée locale différente des autres données locales autour de ce point et très différent du champ initial).

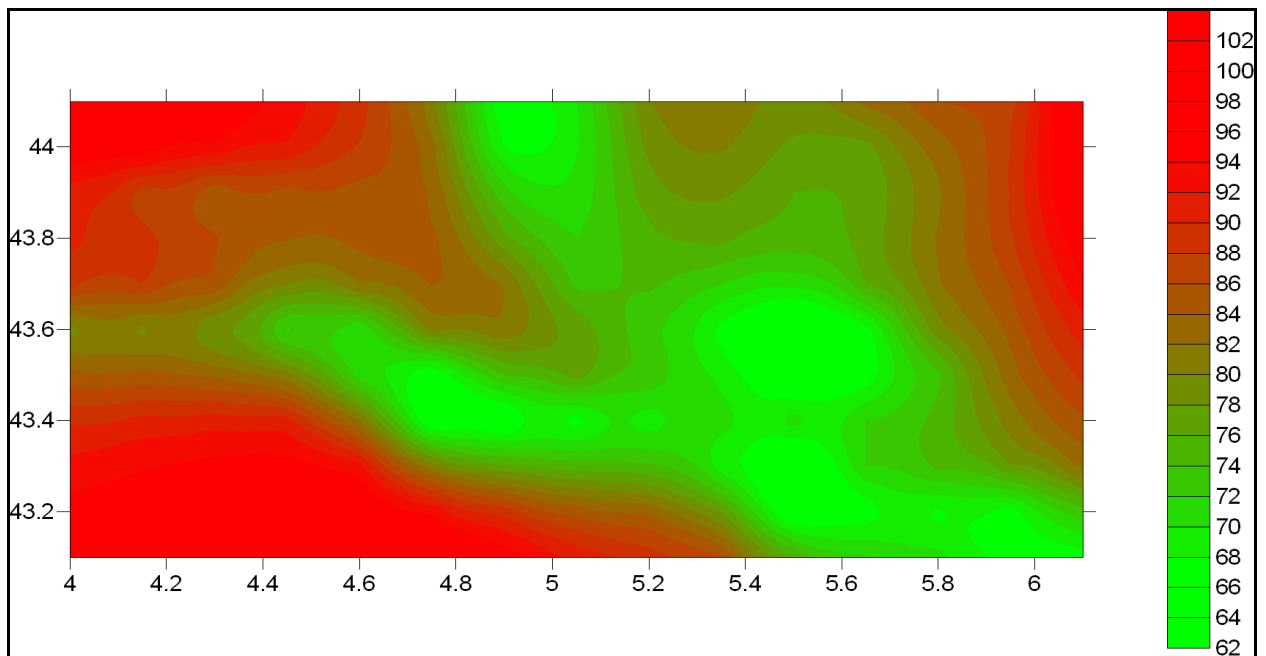


Figure 5 – Exemple de cartographie de concentrations horaires en ozone issus directement des sorties PREV'AIR

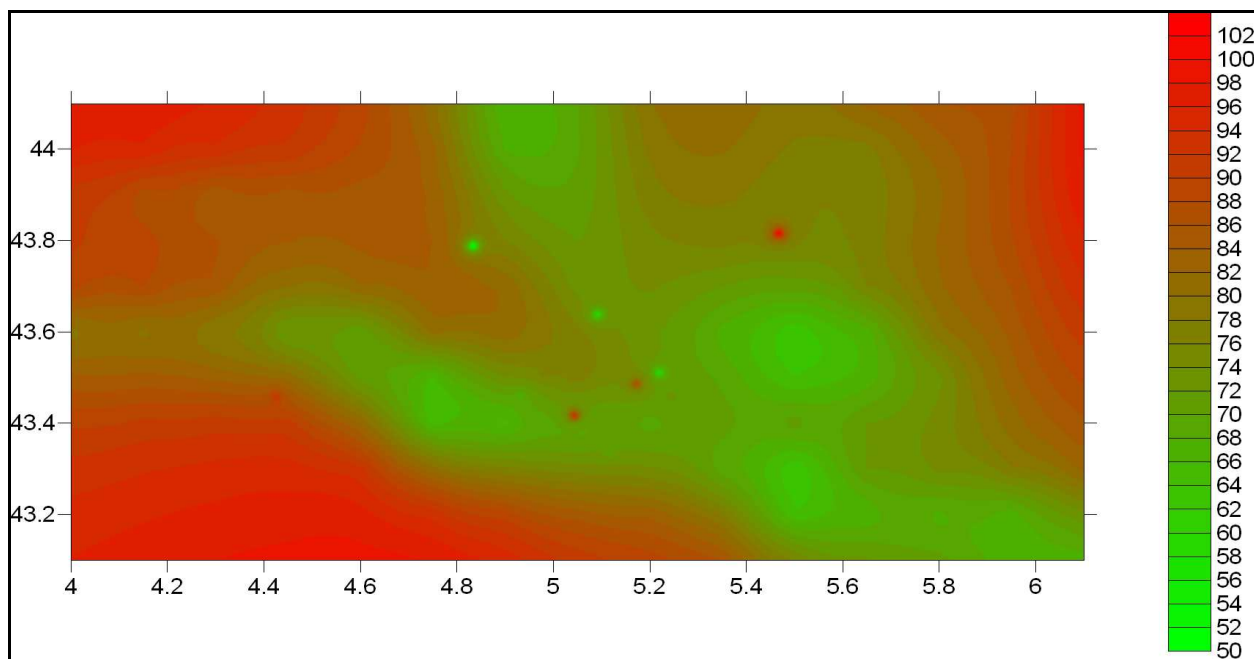


Figure 6 – Exemple de carte analysée qui intègre les observations locales par une interpolation par krigeage

Pour rappel, les stations utilisées pour la création des cartes analysées sont indiquées au niveau du calcul statistique.

2.6. Prise en compte du biais

Après l'étape d'interpolation (pour toutes les échéances) et avant l'étape d'assimilation pour le jour J-1, les sorties PREV'AIR au niveau de chaque station sont corrigées au moyen d'un biais.

En effet, le système PREVAIR a été mis en place en 2003 sur l'initiative du Ministère de l'écologie et du développement durable (MEDD) afin de générer et de diffuser quotidiennement des prévisions et des cartographies de qualité de l'air, issues de simulations numériques, à différentes échelles spatiales. Les prévisions sont effectuées quotidiennement jusqu'à trois jours d'échéance : J, J+1, J+2. Des simulations pour J-1 sont également réalisées. Celles-ci sont de deux types : des sorties brutes du modèle, et des sorties issues d'un krigeage utilisant les simulations numériques comme variable secondaire. En 2006, AIRFOBEP a réalisé une analyse des performances du modèle PREVAIR par comparaison modèle-mesure. Le rapport AIRFOBEP présente alors une évaluation des performances sur les prévisions de concentrations horaires, pour l'ensemble des stations de la zone AIRFOBEP, pour la grille haute résolution (période du 01/05 au 30/09 sur les étés de 2004 et 2005), et pour chaque échéance de prévisions. Pour les simulations à J-1 seules les sorties brutes du modèle ont été étudiées.

En résumé, il a été observé pour les simulations horaires estimées aux points des stations de mesure (la concentration est estimée au niveau de chaque station de mesures par une moyenne pondérée en fonction de la distance aux quatre nœuds de la maille contenant la station), des performances proche les unes des autres pour les différentes échéances de prévision :

- biais moyen de + 13 $\mu\text{g}/\text{m}^3$,
- moyenne des erreurs en valeur absolue de 20 $\mu\text{g}/\text{m}^3$,
- sous estimation de 20% des concentrations supérieures à 180 $\mu\text{g}/\text{m}^3$.

Compte tenu des évolutions de la plate-forme PREV'AIR depuis 2005, il est apparu important de mettre à jour les comparaisons d'AIRFOBEP sur la base des simulations 2006 – 2007. L'annexe 1 de cette note présente ainsi des résultats (graphiques et tableaux) sur ces deux années, tels qu'ils sont disponibles dans le rapport AIRFOBEP. Notons toutefois que cette analyse ne porte que sur les concentrations horaires, que sur les simulations J-1 (seules simulations disponibles) et que sur les simulations haute résolution de PREV'AIR.

A l'issu de cette analyse, il a été retenu avec AIRFOPEB de considérer une correction des simulations PREV'AIR au niveau de chaque station par ajout d'un biais. Ce biais a été construit selon la méthodologie suivante :

- Pour les années 2006 et 2007 entière et pour chaque station, un biais a été calculé entre les observations et les simulations pour différentes classes de concentration. Seul le biais sur le cycle J-1 a pu être calculé.
- Pour les étés 2004 et 2005 (seules données disponibles par AIRFOBEP, la période estivale couvrant les mois d'avril à septembre), un biais a également été calculé pour chaque station pour différentes classes de concentration. Seul le biais pour le cycle J-1 a été calculé.
- Enfin, pour chaque station et chaque classe de concentration, un biais moyen a été calculé à partir du biais sur [2004-2005] et du biais sur [2006-2007]. Notons que ce biais correspond à la concentration d'ozone qu'il faut théoriquement ajouter aux sorties de simulation PREV'AIR pour atteindre les concentrations observées.

Le biais est calculé uniquement sur le cycle J-1, mais il sera également appliqué sur les cycles de prévisions (J, J+1 et J+2). Ce biais dépend donc de la classe de concentration simulée pour chaque station (voir tableau ci-dessous). Plutôt que de corriger les simulations qu'au niveau des points d'observations, il est apparu plus judicieux d'appliquer ce biais correction dans une zone d'influence autour de la station. La méthode retenue est alors celle utilisé lors de l'étape d'assimilation des observations pour le jour J-1, à savoir une méthode de krigeage des innovations. Contrairement à l'étape d'assimilation pour le jour J-1 qui utilise les observations pour corriger le champ simulé, au niveau du biais sont alors utilisées comme « innovation » les valeurs simulées au niveau des stations plus le biais correctif. Cette méthode permet de prendre en compte ainsi le biais autour des stations tout en intégrant un lissage spatial.

classe de concentrations simulées ($\mu\text{g}/\text{m}^3$)	BETG	FSCB	ISTR	MNDM	MRMV	RBRT	SLPV	SMMR	SRMY	SSLP	VTRL
0 ≤ [C] < 10	-6,0	-5,2	-1,0	9,1	0,4	0	47,1	0	-2,1	0	-3,1
10 ≤ [C] < 20	-11,7	-6,0	-8,7	1,2	-0,4	0,7	-12,6	1,9	-8,5	3,6	1,0
20 ≤ [C] < 30	-10,2	-8,7	-3,3	-6,6	-1,7	-4,4	-7,9	-4,2	-18,1	-5,3	-13,5
30 ≤ [C] < 40	-17,8	-12,7	-10,1	-13,3	-9,0	-9,8	-14,0	-6,9	-20,1	-17,0	-20,5
40 ≤ [C] < 50	-20,2	-16,0	-13,0	-14,7	-10,6	-17,0	-18,2	-8,5	-20,0	-22,3	-23,8
50 ≤ [C] < 60	-21,1	-18,9	-15,0	-15,4	-12,4	-19,6	-20,4	-13,9	-20,0	-20,8	-23,7
60 ≤ [C] < 70	-23,3	-20,7	-17,1	-17,6	-13,9	-19,9	-22,1	-16,4	-19,6	-21,0	-23,5
70 ≤ [C] < 80	-23,6	-22,9	-19,1	-18,5	-14,8	-20,2	-22,4	-17,2	-19,9	-21,8	-22,6
80 ≤ [C] < 90	-25,6	-25,6	-21,4	-20,2	-17,1	-21,3	-22,4	-18,4	-20,8	-24,4	-23,8
90 ≤ [C] < 100	-26,2	-27,4	-22,2	-21,8	-16,9	-21,7	-22,2	-19,5	-25,1	-26,1	-25,7
100 ≤ [C] < 110	-25,2	-28,3	-23,7	-23,2	-16,0	-20,2	-21,7	-19,8	-24,5	-28,6	-25,1
110 ≤ [C] < 120	-24,0	-28,4	-24,7	-24,3	-15,1	-16,5	-22,0	-19,6	-26,1	-27,3	-23,3
120 ≤ [C] < 130	-24,0	-29,7	-24,8	-24,8	-14,3	-16,4	-21,2	-18,5	-25,7	-31,0	-25,4
130 ≤ [C] < 140	-22,2	-31,4	-26,6	-25,0	-11,8	-15,5	-21,4	-19,5	-28,6	-32,7	-23,8
140 ≤ [C] < 150	-21,0	-33,1	-26,0	-27,3	-13,6	-14,3	-23,8	-21,9	-30,9	-33,7	-23,5
150 ≤ [C] < 160	-27,4	-37,1	-34,5	-32,7	-19,4	-19,3	-30,2	-24,8	-39,1	-40,4	-29,3
160 ≤ [C] < 170	-33,1	-49,5	-41,8	-44,8	-25,1	-20,0	-35,2	-31,4	-50,1	-45,6	-30,2
170 ≤ [C] < 180	-36,2	-47,2	-49,1	-37,7	-30,2	-19,2	-38,9	-39,5	-45,7	-45,4	-26,8
180 ≤ [C] < 190	-22,3	-44,8	-40,8	-29,9	-28,1	-11,6	-38,6	-47,3	-48,7	-38,0	-20,3
190 ≤ [C] < 200	-36,9	-47,9	-42,9	-38,6	-22,1	-11,0	-39,0	-41,2	-63,1	-36,4	-17,1
200 ≤ [C] < 210	-37,1	-55,9	-56,4	-28,6	-41,0	-7,4	-10,4	-46,6	-51,4	-64,4	-17,5
220 ≤ [C]	-55,4	-76,3	-81,5	-59,9	-51,0	-55,1	-65,7	-73,2	-69,2	-64,4	-79,4

Tableau 1 – Biais correctif utilisé pour corriger les simulations PREV'AIR au niveau des points d'observations

■ Table des figures

Figure 1 - Architecture du système.....	7
Figure 2 – Interface principale de CartO3.....	14
Figure 3 – Interface principale de CartO3.....	14
Figure 2 – Illustration de la grille de sortie haute résolution des simulations PREV’AIR (les points rouges correspondent à la localisation des stations d’observations AIRFOBEP).....	21
Figure 3 – Exemple de cartographie de concentrations horaires en ozone issus directement des sorties PREV’AIR 23	
Figure 4 – Exemple de carte analysée qui intègre les observations locales par une interpolation par krigeage.....	24

■ Table des tableaux

TABLEAU 1 – BIAIS CORRECTIF UTILISE POUR CORRIGER LES SIMULATIONS PREV’AIR AU NIVEAU DES POINTS D’OBSERVATIONS.....	25
--	-----------