

ACHILLE Newsletter

Lettre semestrielle à l'attention de la communauté ADOMOCA

Numéro 11, Octobre 2013

Edito

Les programmes ADOMOCA-I et sa suite ADOMOCA-II ont été soutenus pendant près d'une dizaine d'années par le PNCA de l'INSU puis par les LEFEs CHAT et ASSIM/MANU, sur des financements de l'ADEME. Ils ont permis à la communauté de chimie atmosphérique française de s'insérer et de coordonner des projets scientifiques basés sur l'assimilation de mesures spatiales. Initialement orienté autour du développement d'un outil générique adaptable à tout type de modèle via le coupleur PALM, le programme s'est ré-orienté autour d'une approche pragmatique fédérant 7 laboratoires (GAME, LA, CERFACS-SUC, LATMOS, CERE, LSCE, LISA) en s'appuyant sur 1) le développement d'outils d'assimilation pour 2) répondre à des questions scientifiques. Dans sa dernière mouture, ADOMOCA-II a permis la publication de 19 articles de rang A et de 4 thèses. Une réunion annuelle a systématiquement été organisée aux Lecques (83). Une lettre ACHILLE, téléchargeable depuis le site ETHER, a été envoyée à la communauté élargie 2 à 3 fois par an.

La communauté ADOMOCA a souhaité soumettre une nouvelle proposition (ADOMOCA-III) au LEFE sur les trois années à venir en mettant en avant l'importance des réunions annuelles permettant de présenter et de discuter les avancées opérées dans chaque laboratoire tant sur les plans technique que scientifique ainsi que la lettre trimestrielle ACHILLE. Nous élargissons la communauté à d'autres équipes au sein de laboratoires identifiés précédemment, soit une trentaine de chercheurs. Sur la période des trois années à venir, nous souhaitons définir des axes de recherche prioritaire qui vont nous permettre de bâtir une proposition d'envergure nationale (e.g. ANR) et/ou internationale (e.g. FP7) afin de financer prioritairement des ressources humaines (thèses et post-docs). Le programme ADOMOCA-III est ventilé en quatre thèmes : 1) aérosols, 2) composition gazeuse, 3) inversion de sources et estimations des paramètres et 4) méthodologie. Parmi ces 4 thématiques, certaines

actions sont déjà structurées comme des projets collaboratifs autour de questions bien identifiées. D'autres, au contraire, sont présentées sous la forme d'actions propres à chaque laboratoire et, dans ce cas, ADOMOCA-III doit servir à enclencher la synergie entre les différentes équipes.

Nous attendons avec impatience les résultats du Conseil Scientifique du LEFE et vous tiendrons informés.

Philippe Ricaud, coordinateur du programme ADOMOCA-III

Modélisation de la qualité de l'air

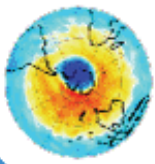
Voir la pollution à l'ozone depuis l'espace (suite)

G. Foret¹ et al.

L'observation/la surveillance de l'ozone troposphérique revêt une importance particulière du fait du caractère nocif de l'ozone pour son environnement et de son impact sur le bilan radiatif terrestre (3^{ème} gaz à effet de serre). Une partie importante de cet ozone est produite par les activités anthropiques via l'émission de polluants primaires NOX et COV (aussi produits de façon naturelle par la végétation et les sols) qui sous certaines conditions (notamment en présence de rayonnement solaire) vont produire de l'ozone. Cette production d'ozone photochimique qui mène à des pics de pollution à l'ozone en situation anti-cyclonique est surveillée par les réseaux de mesures de surface mais aussi les modèles de chimie-transport régionaux qui sont utilisées comme base des systèmes de prévision/surveillance de la pollution (Prevoir, MACC-II).

Au cours de la dernière décennie, de nouveaux instruments satellitaires ont ouvert de nouvelles perspectives pour compléter les systèmes de surveillance des concentrations troposphériques d'ozone. Ces instruments opérant soit dans le domaine UV (GOME2, OMI) soit dans le domaine IR (IASI, TES) ont déjà démontré leur capacité à observer l'ozone dans la troposphère avec une assez bonne précision et une fréquence journalière. Plus spécifiquement, plusieurs études ont montré le

¹ Laboratoire Inter-universitaire des Systèmes atmosphériques, LISA/IPSL, Université Paris-Est Créteil et Université Paris Diderot, CNRS/INSU UMR 7583, Créteil, France



potentiel de telles observations pour détecter, au moins qualitativement, des panaches de pollution à l'ozone (Eremenko et al, 2008 ; Dufour et al, 2010 ; Liu et al, 2010 ; Kar et al, 2010).

Pour essayer de répondre plus spécifiquement à la question de la détection des panaches de pollution d'ozone par les satellites, il a été proposé dans le cadre du projet ADOMOCA d'étudier spécifiquement la capacité des observations satellitaires à « voir » mais aussi à quantifier cette pollution à l'ozone. Nous avons donc choisi de nous placer dans un cas de fortes concentrations, c'est-à-dire une situation anti-cyclonique marquée au cours de laquelle de fortes concentrations d'ozone ont été observées notamment en surface par les réseaux d'observation régionaux (Figure 1). Dans notre cas, cet anticyclone est chassé par une dépression ce qui va créer l'apparition d'un front chaud selon un axe Sud-Nord. Ce front chaud est associé à un soulèvement vertical des masses d'air qui permet de transférer une partie de l'ozone produit dans la couche limite vers la troposphère libre.

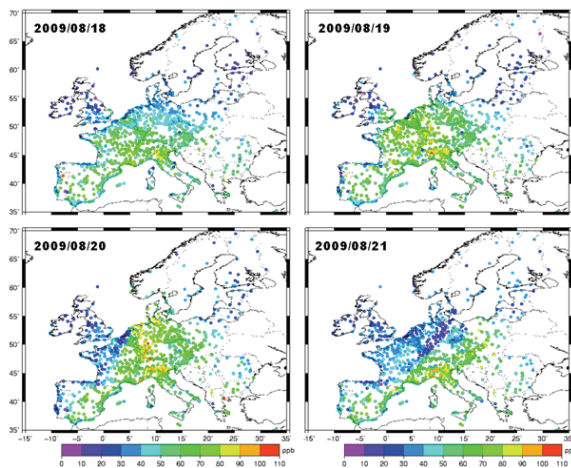


Figure 1. Surface ozone concentrations (ppb) at 3pm at European (background urban, suburban and rural) stations (<http://acm.eionet.europa.eu/databases/airbase/>).

L'inspection des différentes observations satellitaires montrent que les observations de l'instrument IASI permettent de reproduire la structure de l'événement de transport notamment en comparaison de simulations réalisées par des modèles de chimie-transport (CHIMERE et MOCAGE). Les observations de GOME2 et OMI ne permettent pas en revanche d'observer cet événement alors que l'échantillonnage de TES reste trop faible. L'inspection de coupes latitudinales transverses confirme la capacité de l'instrument IASI à représenter le panache d'ozone. En revanche, on remarque que l'ozone dans la couche limite n'est pas vu par l'instrument du fait de son manque de sensibilité dans la couche limite (Figure 2).

Cette étude a pu montrer la capacité de l'instrument IASI à détecter un panache de pollution transporté dans la troposphère libre mais aussi la difficulté à observer ce même panache dans la couche limite atmosphérique. Une étude plus systématique de ces situations de transport pourrait permettre de mieux caractériser l'apport de l'ozone produit dans la couche limite atmosphérique et ainsi de mieux comprendre le bilan d'ozone troposphérique. La production d'analyses (après assimilation des observations IASI dans un modèle de chimie transport) pourrait permettre d'évaluer plus précisément les flux d'ozone vers la troposphère.

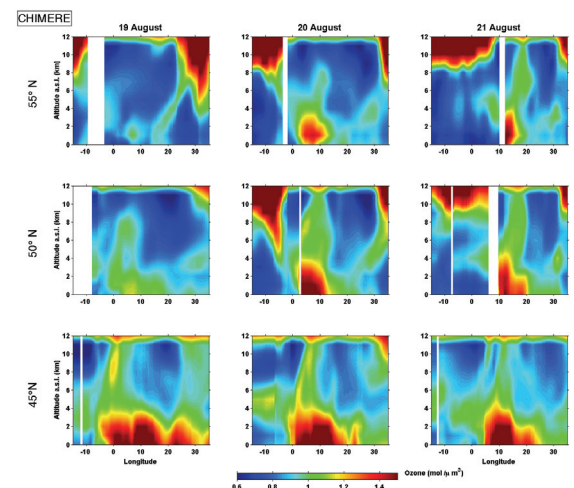


Figure 2. Vertical cross sections of ozone concentration simulated by the CHIMERE model along tree latitudes (45°N, 50°N and 55°N) and for three days (2009/08/19-21) at 10 p.m

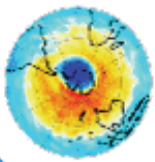
Assimilation de l'ozone à la surface

La chaîne d'assimilation CHIMERE-EnKF, évaluation de l'analyse de l'ozone sur l'été 2009

B. Gaubert¹, A. Coman¹, G. Foret¹, F. Meleux², A. Ung², L. Rouil², M. Beekmann¹

Dans le cadre du projet MACC I et II (Monitoring Atmosphere and Climate Change), un système d'assimilation a été développé au LISA (Coman et al. 2012, Gaubert et al. 2013) à partir du modèle de Chimie-Transport CHIMERE (Menut et al. 2013) couplé à un filtre de Kalman d'ensemble (Evensen et

² Institut National de l'Environnement Industriel et des Risques, INERIS, Parc Technologique ALATA-B.P. No. 2, 60550 Verneuil en Halatte, France



al. 2004). Cette méthode séquentielle repose d'une part sur un ensemble de simulations fournies par le modèle de chimie-transport CHIMERE et d'autre part d'une étape d'analyse des concentrations d'ozone de surface à l'échelle horaire. Les matrices de covariances d'erreurs d'ébauches et d'analyses peuvent alors être directement approximées à partir de la covariance de l'ensemble. Cependant, l'échantillonnage par un ensemble de taille finie implique d'une part la localisation de l'analyse sur une fenêtre spatiale donnée et d'autre part d'employer un facteur (additif) d'inflation de la dispersion de l'ensemble afin d'éviter la sous-estimation de ces erreurs. Dans un premier temps, une démarche objective permettant de qualifier et classer les stations de mesures en fonction de leur degré de pollution a été mise en place à partir de la classification de Flemming et al. (2005). Celle-ci a permis de sélectionner un nombre important de stations pour l'assimilation et l'évaluation. Les stations qualifiées d'urbaines présentant une faible représentativité spatiale, notamment la nuit, et pouvant introduire ainsi un biais important ont été écartées de la procédure d'assimilation. Différents tests de sensibilité du système d'assimilation à la formulation de la covariance d'erreur d'ébauche via la modification du facteur d'inflation et à la formulation de la variance des erreurs d'observations ont été réalisés à partir des diagnostics de Desroziers (Desroziers et al. 2005). On veut maintenant se placer dans un contexte d'utilisation plus systématique de ce système par exemple dans un contexte opérationnel comme celui du projet MACC-II. Ainsi une analyse des mois de Juin-Juillet-Aout 2009 (JJA, 01/06/2009 au 29/08/2009, soit 90 jours) a été réalisée. Afin de quantifier le biais et l'estimation des variances d'erreurs à chaque pas d'assimilation, nous avons utilisé un indicateur appelé variable aléatoire centrée réduite (Candille et al., 2007), que nous noterons RCRV pour Reduced Centered Random Variable (RCRV). Cette variable est définie comme le rapport de la différence entre la valeur observée (pour chaque station i) et la valeur simulée (ou analysée) à l'endroit de l'observation, divisée par la racine carrée de la somme des variances prescrites dans l'algorithme d'assimilation c'est-à-dire la variance de l'ensemble prévue $(\sigma_b)^2$ ou analysée $(\sigma_a)^2$ et de la variance des erreurs d'observations $(\sigma_o)^2$. Le calcul est donc réalisé dans l'espace des observations, en utilisant la relation mathématique suivante :

$$RCRV(\text{ébauche}) = \frac{y_i^o - H_i y^b}{\sqrt{\sigma_o^2 + \sigma_b^2}} \quad (1)$$

$$\text{et } RCRV(\text{analyse}) = \frac{y_i^o - H_i y^a}{\sqrt{\sigma_o^2 + \sigma_a^2}} \quad (2)$$

Les statistiques des distributions de RCRV (figure 3) permettent d'estimer d'une part le biais pondéré (moyenne) et d'autre part de vérifier les erreurs

prescrites (écart-type). La nuit, on observe un biais systématique de prévision et d'analyse accru dans la mesure où les concentrations observées sont surestimées. Entre 9 h et 18 h, le biais de prévision devient positif, mais reste généralement inférieur à 20 %. Globalement, le champ analysé obtenu n'est pas biaisé (faiblement la nuit) et les erreurs estimées par l'ensemble sont représentatives. Dans cette configuration d'analyse le facteur d'inflation est estimé en ligne à partir des valeurs des erreurs d'ébauche diagnostiquées de la journée précédente. Dans ce cas, les erreurs sont correctement estimées pour l'analyse et la prévision pendant la journée, mais sont globalement sous-estimées la nuit, en raison de la sous-estimation des erreurs de représentativité.

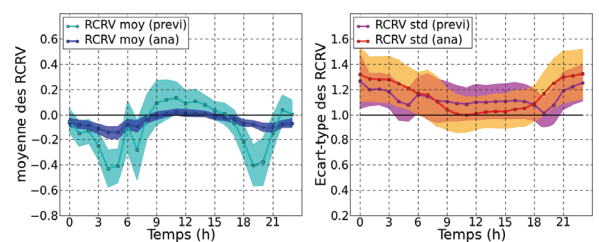
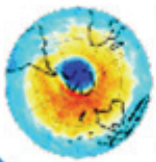


Figure 3 : Moyenne (à gauche) et écart-type (à droite) des distributions de RCRV en moyenne (ligne) \pm l'écart-type (bande) pour la période de JJA 2009 pour la prévision (en bleu clair pour la moyenne et en violet pour l'écart-type) et l'analyse (en bleu foncé pour la moyenne et en orange pour l'écart-type).

La comparaison de la simulation de référence et de l'analyse avec l'ensemble des stations d'évaluation à été réalisée sur la base des maxima journaliers des moyennes sur 8 heures (figure 4). On note une amélioration sur le nuage de points, plus recentré sur l'axe $y = x$, le coefficient de corrélation entre les concentrations observées et simulées est de 0,74 pour CHIMERE et de 0,85 pour l'analyse. On distingue également quatre régions sur le scatter-plot, dont deux indiquent une bonne estimation, correspondant aux dépassements (B) et aux non-dépassements (C) du seuil de 60 ppb à la fois observés et simulés. Les deux autres régions indiquent les dépassements de seuils prévus, mais non observés (A) et les dépassements de seuils observés mais non prévus (D). On note une nette amélioration du nombre de dépassements observés et prévus tout en réduisant le nombre de dépassements observés mais non prévus.

Finalement, l'emploi d'un filtre de Kalman d'Ensemble avec un schéma d'analyse de type racine carrée assure une erreur faible pour un ensemble de taille raisonnable (20 membres) au regard du temps de calcul associé (5 heures / jour sur 12 processeurs). Ainsi, l'algorithme d'analyse peut être employé dans un cadre opérationnel. Pour confirmer la robustesse de ce système d'assimilation nous avons testé la reproductibilité des résultats en réalisant une



assimilation sur un été. L'analyse montre de bonnes performances avec une réduction des RMSE (de 30 % en moyenne) pour 94 % des stations de validation. L'assimilation des observations du dioxyde d'azote est une piste envisagée, mais elle nécessite de prendre en compte correctement les erreurs instrumentales et de représentativité. L'assimilation conjointe de colonnes partielles IASI (Coman et al. 2012) où IASI+GOME2 (Cuesta et al. 2013) et d'observations de surfaces permettrait aussi une meilleure contrainte sur la colonne troposphérique d'ozone.

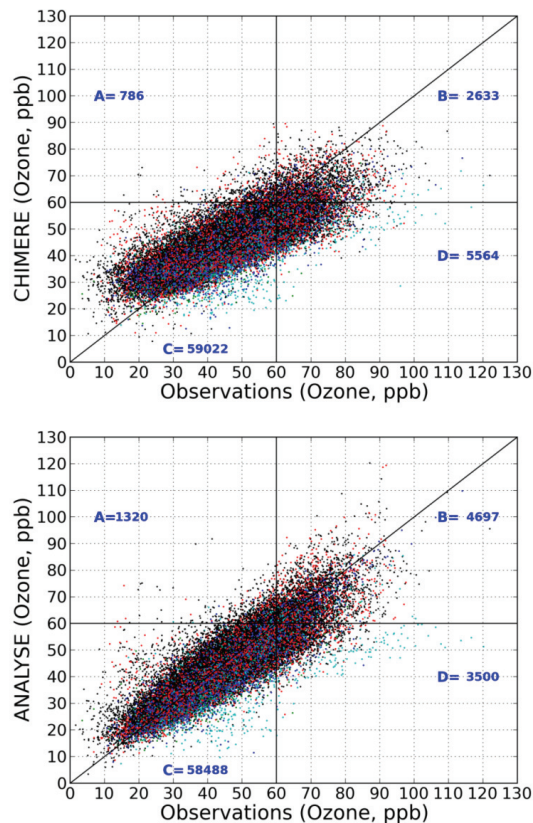


Figure 4 : Scatter-plots des maxima journaliers des moyennes sur huit heures d'ozone estimées par CHIMERE (en haut) et par l'analyse (en bas) par rapport à la totalité des observations (issues de stations d'évaluation) pour chaque jour de la période de JJA 2009. On dispose donc de 68005 observations réparties sur 359 stations urbaines (point noirs), 18 stations de fond (en vert), 112 stations rurales (en bleu), 187 stations suburbaines (en rouge) et 92 stations d'altitude (en bleu clair) entre 300 et 800 m (dont 27 stations de fond, 32 stations rurales et 33 stations suburbaines).

Références :

G. Candille, C. Côté, P. Houtekamer, et G. Pellerin. Verification of an ensemble prediction system against observations. Monthly Weather Review, 135(7) :2688–2699, 2007.
A. Coman, G. Foret, M. Beekmann, M. Eremenko, G. Dufour, B. Gaubert, A. Ung, C. Schmechtig, J. M. Flaud, et G. Bergametti. Assimilation of iasi partial tropospheric

columns with an ensemble kalman filter over europe. Atmos. Chem. Phys., 12(5) :2513–2532, 2012.

J. Cuesta, M. Eremenko, X. Liu, G. Dufour, Z. Cai, M. Höpfner, T. von Clarmann, P. Sellitto, G. Foret, B. Gaubert, M. Beekmann, J. Orphal, K. Chance, R. Spurr, et J.-M. Flaud. Satellite observation of lowermost tropospheric ozone by multispectral synergism of IASI thermal infrared and GOME-2 ultraviolet measurements over Europe, Atmos. Chem. Phys., 13, 9675-9693, doi:10.5194/acp-13-9675-2013, 2013.

G. Desroziers, L. Berre, B. Chapnik, et P. Poli. Diagnosis of observation, background and analysis error statistics in observation space. Quarterly Journal of the Royal Meteorological Society, 131 (613) :3385–3396, 2005.

G. Evensen. Sampling strategies and square root analysis schemes for the enfk. Ocean Dynamics, 54(6) :539–560, 2004.

J. Flemming, R. Stern, et R. J. Yamartino. A new air quality regime classification scheme for O₃, NO₂, SO₂ and PM₁₀ observations sites. Atmospheric Environment, 39(33) :6121–6129, 2005.

B. Gaubert, A. Coman, G. Foret, F. Meleux, A. Ung, L. Rouil, A. Ionescu, Y. Candau et M. Beekmann. Regional scale ozone data assimilation using an Ensemble Kalman Filter and the CHIMERE Chemical-Transport Model. Geosci. Model Dev. Discuss., 6, 3033-3083, 2013.

L. Menut, B. Bessagnet, D. Khvorostyanov, M. Beekmann, N. Blond, A. Colette, I. Coll, G. Curci, G. Foret, A. Hodzic, S. Mailler, F. Meleux, J.-L. Monge, I. Pison, G. Siour, S. Turquety, M. Valari, R. Vautard, and M. G. Vivanco. CHIMERE 2013: a model for regional atmospheric composition modelling. Geosci. Model Dev., 6, 981-1028, 2013

Modélisation dynamique de la matrice des covariances d'erreur d'ébauche avec VALENTINA-Ensemble

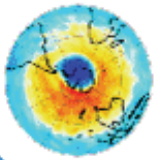
E. Jaumouillé³, E. Emili³, O. Pannekoucke^{3,4}, S. Massart⁵, A. Piacentini³

La qualité des analyses et des prévisions fournies par les modèles de qualité de l'air dépend fortement des erreurs d'ébauche qui sont difficiles à estimer dans la pratique. Dans les systèmes d'assimilation de données de type variationnel les erreurs d'ébauche sont souvent considérées uniformes en espace et en temps. Dans cet article, on s'intéresse à l'estimation et à la modélisation des erreurs d'ébauche dans un modèle de chimie de l'atmosphère à l'échelle régionale. L'objectif étant d'obtenir une estimation de la dynamique spatio-temporelle de la matrice des

³ CERFACS/SUC CNRS-URA 1875, Toulouse, France

⁴ CNRM-Game, Météo-France/CNRS-UMR 3589, Toulouse, France

⁵ ECMWF, Reading, UK



covariances d'erreur d'ébauche (appelée par la suite matrice B) afin de prendre en compte notamment sa structure spatiale et sa dépendance de la situation météorologique.

Dans cette étude, on propose une méthode d'assimilation d'ensemble pour estimer les erreurs d'ébauche. Cela consiste à simuler les erreurs du système à l'aide de perturbations aléatoires ainsi que leur évolution temporelle au cours des étapes successives d'analyse et de prévision. Le modèle de chimie-transport MOCAGE, en configuration de grille régionale sur l'Europe à une résolution de $0.2^\circ \times 0.2^\circ$, 47 niveaux verticaux (du sol à 35km environ), avec le schéma chimique RACMOBUS (118 espèces modélisées, 91 transportées) est utilisé. Les conditions aux bords du domaine sont imposées par un couplage avec les champs chimiques calculés sur une grille globale à $2^\circ \times 2^\circ$ et les mêmes niveaux verticaux. Le système d'assimilation VALENTINA est utilisé avec une méthode variationnelle 3D-Var et des fenêtres d'assimilation d'une heure. On a considéré l'assimilation des observations d'ozone de surface, disponibles dans la base de données Airbase et les simulations ont été faites pour la journée du 7 juillet 2010 (voir Fig. 5 pour la position des observations assimilées). Sur cette journée, une large zone de haute pression est basée sur l'Europe centrale, avec des vents faibles et des conditions stables alors qu'une zone de basse pression se situe au sud de l'Islande accompagnée de vents modérés qui soufflent en direction O-E sur l'Irlande et l'Angleterre (Figure 6).

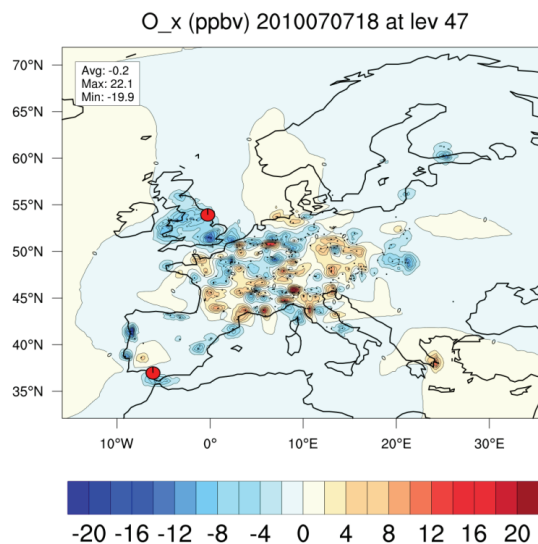


Figure 5 : Incréments d'analyse obtenus après l'assimilation des stations de surface pour l'ozone pour le 7 juillet 2010 à 18h. Les points noirs montrent la position géographique des observations assimilées. Les points rouges indiquent les villes de Hull en Angleterre et Hueva en Espagne.

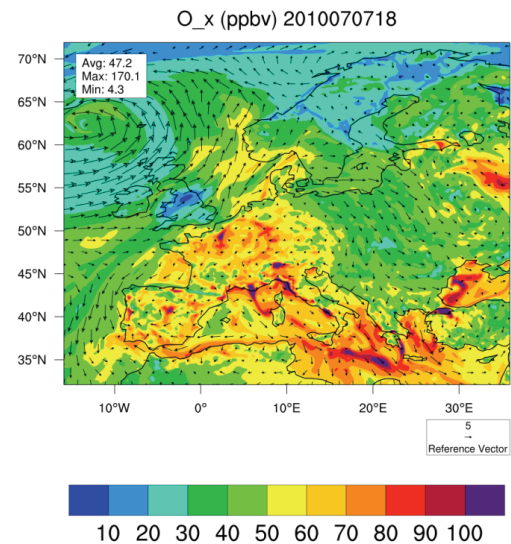
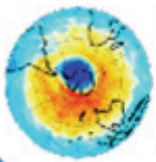


Figure 6 : Champ d'ozone et vent moyen à la surface pour la journée du 7 juillet 2010 à 18h

L'assimilation d'ensemble s'appuie sur l'utilisation d'un gestionnaire de tâches (ECFLOW, <https://software.ecmwf.int/wiki/display/ECFLOW/Home>) qui permet de distribuer plusieurs instances de la chaîne d'assimilation VALENTINA sur les machines disponibles. Un ensemble de N analyses d'une durée de 24h est calculé d'abord, en utilisant des statistiques d'erreur d'ébauche 'climatologiques' et uniformes (voir Tab. 1). L'analyse des $N \times 24$ prévisions à une heure ainsi obtenues permet l'estimation des structures d'erreur d'ébauche avec une résolution horaire. Cette information est utilisée pour calculer une analyse finale qui prend en compte la dynamique du modèle et les observations assimilées dans la matrice B. Ce type d'implémentation est une variante des méthodes hybrides Ensemble-3D-Var (Buehner, M., 2005), avec des cycles prévision-analyse de 24h dans notre cas. Les perturbations stochastiques implémentées dans l'ensemble sont de type gaussien et, dans le cas de champs de type géophysique, sont corrélées en espace par l'application de l'opérateur $B^{1/2}$ de VALENTINA. Les perturbations qui ont été retenues dans cette étude sont les suivantes (voir Tab. 1 pour plus de détails): sur la condition initiale du champ analysé (O_3), sur le forçage météorologique (vent, pression atmosphérique, température, humidité), sur les émissions des espèces primaires, les taux de dépôt sec à la surface et sur les observations assimilées. Une taille de l'ensemble N égale à 10 a été utilisée dans cette étude, en considérant que le temps de calcul pour une journée d'assimilation sur le domaine européen est d'environ 8 heures avec la configuration utilisée (Workstation avec 16 processeurs Intel Sandy Bridge). Un travail de parallélisation de la chaîne sur machine de type scalaire est en cours et permettra de réduire les temps de calculs de plusieurs facteurs dans le futur. Le choix des taux de perturbation (l'écart type des perturbations gaussiennes) et de leur structure



spatiale (paramétrée en terme de longueurs de corrélation zonale, longitudinale et verticale) a été décidé sur la base des études précédentes menées avec le système VALENTINA (Massart et al. 2011, Jaumouillé et al. 2012), qui avaient évalué les caractéristiques climatologiques de l'erreur d'ébauche. Une étude de sensibilité aux choix des perturbations et une vérification de leur validité avec des observations permettront un réglage plus fin de l'ensemble dans le futur. Il faudra notamment considérer la présence de corrélation temporelle pour certains types de perturbations (e.g. les émissions).

Perturbations	Ecart type	Longueur de corrélation horizontale	Longueur de corrélation verticale
Champs O ₃ à 00 UTC	25% du champ	4.0 deg (Global) 0.4 deg (Europe)	5 points grille modèle
Champs météo	Ensemble CEP	Ensemble CEP	Ensemble CEP
Emissions et dépôt sec	25% pour chaque champ	2.0 deg (Global) 0.2 deg (Europe)	/
Observations	10%	no	/
Matrice B 'climatologique'	25%	0.4 deg (Europe)	1 point de grille

Tableau 1: Paramètres utilisés pour perturber les membres de l'ensemble. Dans le cas des champs 2D ou 3D les champs de perturbation sont obtenus avec l'application de l'opérateur $B^{1/2}$ de VALENTINA sur un champ de bruit blanc à moyenne zéro et écart type indiqué dans la deuxième colonne. Les valeurs négatives présentes dans les champs chimiques après les perturbations sont mises à zéro. L'ensemble météo du Centre Européen de Prévisions (CEP) est formé par les 10 analyses opérationnelles du modèle météorologique IFS.

Dans un premier temps, on s'est intéressés au diagnostic et à la prise en compte des structures d'erreur d'ébauche horizontales (le long de la latitude et de la longitude), notamment au niveau surface du modèle, où se situent aussi les observations assimilées. Les structures verticales et la variabilité verticale, ayant un impact plus significatif dans le cas d'assimilation d'observations satellitaires, seront détaillées dans une étude future.

Une première évaluation de l'ensemble est faite en regardant la dispersion horaire des concentrations d'ozone à la surface issues des 10 membres de l'ensemble sur la journée du 7 juillet 2010. Nous présentons les résultats sur deux localités: Hull, à l'Est de l'Angleterre (53.75 N, 0.25 O, Figure 7) se

trouvant dans une zone de vents modérés en direction O-E dominants (Figure 6) et Huelva, au Sud de l'Espagne (36.93 N, 6.4 O, Figure 7) se situant dans une zone de vents plus faibles et températures plus élevées (>30 °C).

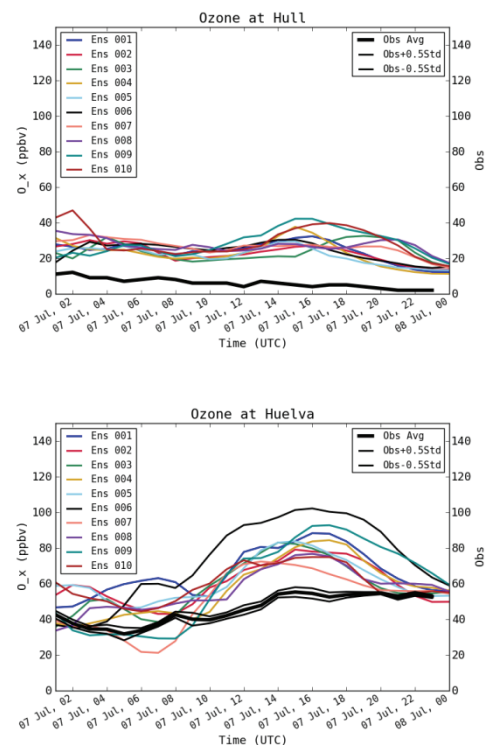
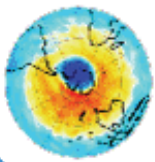


Figure 7 : Profil temporel de l'ozone de surface sur les villes de Hull en Angleterre (en haut) et de Huelva (en bas) pour les 10 ensembles générés (en couleur) et les observations (en noir) pour la journée du 7 juillet 2010.

Les trajectoires d'ozone sur les deux sites montrent que la dispersion de l'ensemble diffère significativement: à Hull en Angleterre les 10 membres de l'ensemble sont peu dispersés (différence entre 10 et 20 ppbv) alors qu'à Huelva en Espagne la dispersion est plus importante allant jusqu'à 40 ppbv de différence entre les membres. Le différent comportement de l'ensemble sur ces deux sites ne peut pas s'expliquer en prenant seulement en compte la concentration moyenne d'ozone, plus élevée en Espagne, mais c'est la combinaison des observations assimilées et du transport qui joue un rôle significatif. Dans le cas du site en Angleterre les observations en amont du couloir de vent et la forte directionnalité du transport réduisent l'incertitude de l'ensemble sur le site examiné au bout de quelques heures d'assimilation. En Espagne les faibles vents peuvent par contre contribuer à augmenter la dispersion de l'ensemble des trajectoires d'O₃ à cause des changements de direction du vent de l'ensemble météo. La comparaison avec les observations non assimilées dans un cercle fixé à 40 km de rayon autour des deux villes (courbe noire pour la moyenne des observations) permet d'évaluer les biais du champ d'ozone et, en regroupant plusieurs points



observation-modèle, les propriétés statistiques de l'ensemble (e.g. écart type). Ce type de validation sera faite sur tous les sites européens et permettra un réglage plus fin des perturbations utilisées pour générer l'ensemble.

La variabilité de l'ensemble (Figure 8) dépend fortement de l'endroit géographique où l'on se trouve, son écart-type étant compris entre 1 et 40 ppbv avec une moyenne de 5 ppbv sur le domaine. Les écarts sont en général plus élevés sur terre et correspondent aux régions continentales où il n'y a pas d'observations assimilées (par exemple la partie Est du domaine). Mais on retrouve aussi une forte variabilité sur la Méditerranée et sur certaines régions comme le Portugal, la côte Italienne ou la Grèce où la perturbation des émissions et la photochimie plus intense de l'ozone peuvent jouer un rôle significatif.

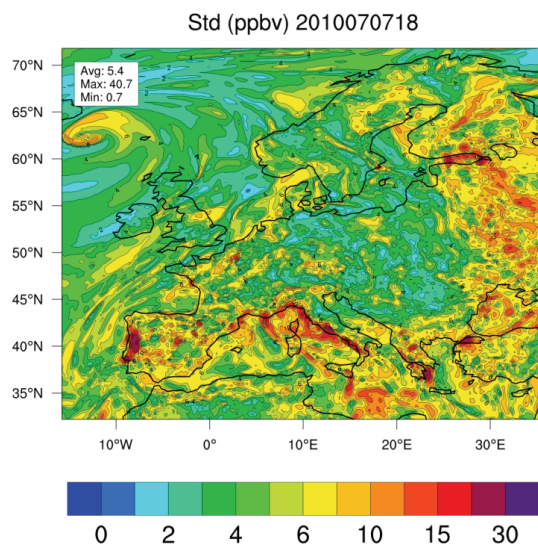


Figure 8 : Ecart-type diagnostiqué sur les 10 ensembles générés pour l'ozone à la surface du 18 juillet 2010 à 18h

Nous voulons utiliser une matrice B dynamique horaire pour l'analyse finale et évaluer son impact sur les concentrations d'ozone à la surface. On va alors modéliser les structures spatiales de l'erreur d'ébauche avec une approche consistant à calculer localement les longueurs caractéristiques des corrélations des erreurs (longueurs de portées, Pannekoucke et al. 2008). Pour cela, nous avons mis au point des diagnostics horaires pouvant être réalisés sur les membres de l'ensemble généré. La corrélation des erreurs est calculée entre deux points distants d'un kilomètre fixé (i.e. 50km dans cette étude). Les longueurs de portées horizontales sont diagnostiquées en supposant que la fonction de corrélation est parabolique et elles sont calculées dans les 8 directions autour d'un point ce qui permet de tenir compte de leur anisotropie. On propose de regarder des statistiques pour le 7 juillet 2010 à 18h pour évaluer les effets de la dynamique du modèle et de l'assimilation des observations sur les structures d'erreur imposées à l'heure 00 UTC (Tab. 1). On

s'attend à un impact significatif sur les structures d'erreur initialement homogènes et isotropes au bout de plusieurs heures de simulation et que le transport puisse redistribuer l'information sur plusieurs dizaines/centaines de kilomètres selon l'intensité locale des vents. Les portées horizontales (L_x le long de la longitude) diagnostiquées à partir de l'ensemble sont représentées sur la Figure 9. Leur répartition spatiale n'est pas homogène sur l'Europe variant de 30km à 500-1000km avec une moyenne de 80 km. On remarque principalement : des structures à grande échelle (500 km) qui rentrent à l'Ouest du domaine à cause de l'advection de l'erreur du champ d'ozone global, des structures à courte échelle (30-100 km) sur le continent européen où les observations assimilées sont assez denses (en accord avec Jaumouillé et al. 2012), des structures avec une échelle de 150-200 km sur la mer Méditerranée, au nord de l'Algérie et du Maroc, à cause des vents zonaux intenses qui ont pour effet d'allonger les corrélations d'erreur initiales (Pannekoucke et al. 2013).

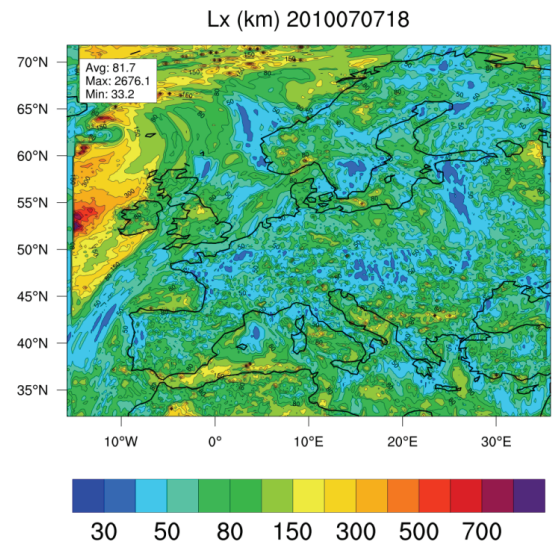
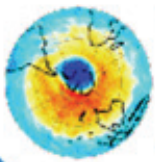


Figure 9 : Portées horizontales (L_x :Est-Ouest) diagnostiquées à partir des 10 ensembles pour l'ozone à la surface du 18 juillet 2010 à 18h

Avec les longueurs de portées horaires pour l'ozone de surface diagnostiquées à partir des 10 membres de l'ensemble généré, on définit une matrice B hétérogène en espace et dynamique. Cette matrice B dynamique est utilisée pour calculer une analyse finale de la journée du 7 juillet 2010. Pour comprendre l'impact de l'utilisation des longueurs de corrélation dynamiques sur les concentrations d'ozone, on va comparer cette analyse avec l'analyse obtenue avec une matrice B climatologique, analyse dans laquelle les longueurs de portées sont fixées à 0.4 degrés sur l'Europe et ne varient pas sur la journée. On représente sur la Figure 10 l'évolution temporelle des concentrations d'ozone à la surface sur les deux villes étudiées pour la journée du 7 juillet



2010 pour le run direct MOCAGE sans assimilation (en bleu), l'analyse obtenue avec la matrice B climatologique (en rouge), l'analyse obtenue avec la matrice B dynamique (en vert) et issues des observations non assimilées (en noir). L'impact de l'assimilation des données de surface sur le modèle se voit au bout de quelques heures : 6h à Hull et 12h à Huelva. Ensuite, l'analyse obtenue avec la matrice B dynamique est meilleure que celle obtenue avec la matrice B climatologique, se rapprochant des concentrations d'ozone issues des observations. L'écart avec les observations passe alors de 20 ppbv (écart avec l'analyse issue de la matrice B climatologique) à 10 ppbv (écart avec l'analyse issue de la matrice B dynamique) à certains instants de la journée.

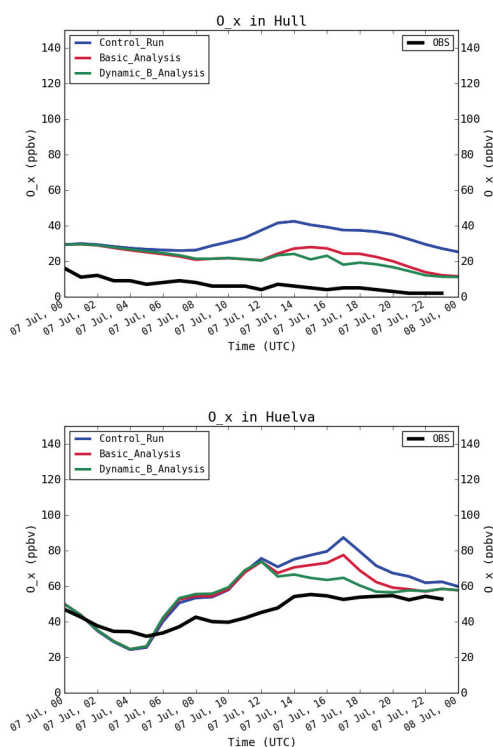


Figure 10: Profil temporel de l'ozone de surface sur les villes de Hull en Angleterre (en haut) et de Huelva (en bas) pour le run direct (bleu), l'analyse obtenue avec une matrice B classique (rouge), l'analyse obtenue avec une matrice B dynamique diagnostiquées à partir des 10 ensembles (vert) et les observations (noir) pour la journée du 7 juillet 2010.

On peut conclure que l'ensemble généré pour la journée du 7 Juillet 2010 contient une dynamique des covariances d'erreur d'ébauche assez riche et peut fournir des informations additionnelles pour améliorer l'analyse sur certains sites. Ce type d'étude a pour objectif non seulement la possible amélioration des analyses des champs chimiques mais l'évolution du système d'assimilation VALENTINA vers des prévisions/analyses de type probabiliste, avec un écart type associé qui peut fournir une information additionnelle assez précieuse

aux utilisateurs des données. Pour poursuivre la validation de la méthode d'assimilation d'ensemble présentée, nous envisageons une validation sur la totalité des sites disponibles, de prendre en compte la corrélation verticale et étudier son impact sur l'analyse et de prendre en compte l'anisotropie horizontale des portées (pour l'instant nous savons modéliser B dans le seul cas d'ellipses dont l'axe principal est N-S ou E-O mais l'élongation des structures d'erreur diagnostiquées est souvent dans la direction de l'advection). Ensuite, il serait intéressant de considérer une période d'étude plus longue et d'étudier une autre saison. Pour compléter l'étude, nous souhaitons comparer la méthode d'assimilation d'ensemble à une méthode NMC qui consiste à estimer les erreurs entre plusieurs prévisions produites consécutivement pour une même période donnée.

Références:

Buehner, M. (2005). Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Quarterly Journal of the Royal Meteorological Society*, 131(607), 1013–1043. doi:10.1256/qj.04.15

Jaumouillé, E., Massart, S. and Piacentini, A. and Cariolle, D. and Peuch, V.-H., Impact of a time-dependent background error covariance matrix on air quality analysis, *Geosci. Model Dev.*, 5, 1075-1090, 2012, doi:10.5194/gmd-5-1075-2012

Pannekoucke, O., Massart, S. (2008). Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quarterly Journal of the Royal Meteorological Society*, 14, 1–14. doi:10.1002/qj.288

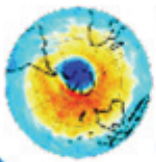
Pannekoucke, O., Emili, E., & Thual, O. (2013). Modeling of local length-scale dynamics and isotropizing deformations. *Quarterly Journal of the Royal Meteorological Society*, n/a–n/a. doi:10.1002/qj.2204

Algorithme d'assimilation

Modélisation des fonctions de covariances : isotropisation locale

O. Pannekoucke^{3,4}, E. Emili³, O. Thual³

Dans les algorithmes actuellement utilisés en assimilation de données, la modélisation de matrice de covariances en grande dimension joue un rôle essentiel. Ces matrices représentent les fonctions de covariances (colonnes de la matrice) qui décrivent les liens statistiques entre les points du domaine d'intérêt. Par exemple, dans le formalisme du filtre de Kalman,



la matrice de covariances d'erreur de prévision, la matrice B, contribue à filtrer l'erreur d'observation et à propager, dans l'ébauche, l'information apportée par les observations. Ainsi il faut être capable de produire ce type de matrice de covariances, en représentant au mieux leur caractéristique essentielle. Autre exemple d'utilisation : dans le formalisme du filtre de Kalman d'ensemble, la modélisation de matrice de grande taille peut avoir un intérêt pour la localisation des fonctions de structures, polluées par le bruit d'échantillonnage associée à la taille finie des ensembles utilisés.

Afin de modéliser correctement ces matrices de covariances de grande taille, il est nécessaire de les caractériser. Une première difficulté est de chercher à extraire des informations synthétisant une part de leur propriétés. Souvent, on introduit en assimilation de données la notion de longueur de portée, qui correspond à une longueur caractéristique de corrélation significative. Pour donner une expression analytique à cette grandeur, on fait souvent référence à la définition donnée par Daley (1991), et qui correspond à la longueur pour laquelle la parabole osculatrice à la fonction de corrélation, au point origine, atteint la valeur 0.5 (Fig. 11). Différentes formules pratiques pour l'estimations ont été introduites dans la littérature (voir les références dans Pannekoucke et al. 2008), elles font de ce diagnostic un outils simple et informatif de l'étalement des fonctions de corrélations, mais aussi de leur forme en dimension deux ou trois. En effet, l'extension aux dimensions supérieures de la longueur de portée, en dimension un, est donnée par le tenseur diffusion locale qui caractérise des directions de corrélations préférentielles. Ceci se représente sous la forme de cercle ou d'ellipse dont l'orientation n'est pas nécessairement celle des axes du repères (Fig. 12).

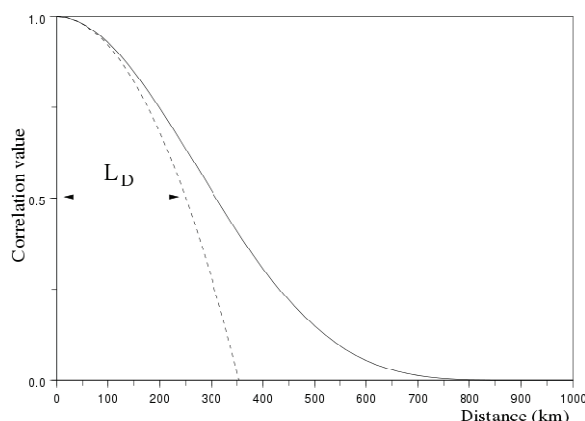


Fig. 11 : Illustration de ce que représente une longueur de portée : la longueur de portée L_D correspond à la longueur pour laquelle l'approximation quadratique au voisinage de l'origine (courbe tirété) de la fonction de corrélation (courbe continue) atteint la valeur 0.5.

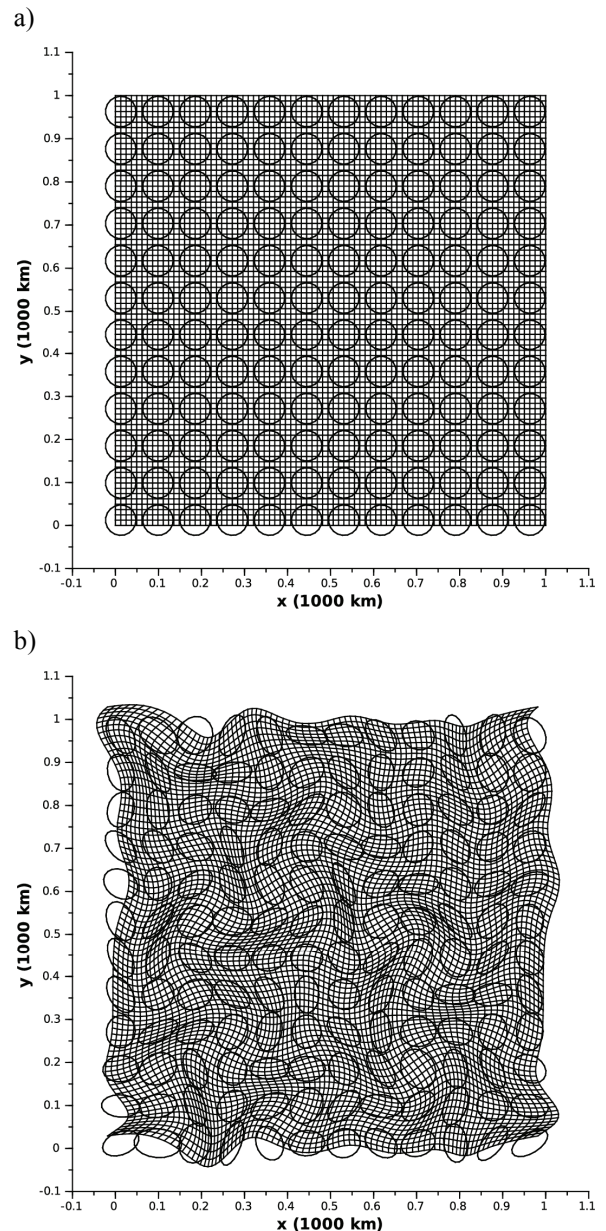
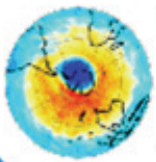


Fig. 12 : Illustration du changement des formes des fonctions de corrélation par changement de coordonnées : partant de fonctions de corrélations ayant toute la même isotropie (a), un changement de coordonnées conduit à des fonctions de corrélations localement anisotrope (b).

Différents modèles pour les matrices de covariances ont été introduits en assimilation de données (approche spectrale, ondelette, filtre récursif,...) et parmi eux l'approche par opérateur de diffusion ; c'est cette technique qui est implémenté dans la suite VALENTINA. L'idée est de construire des fonctions de corrélations proche de Gaussiennes en résolvant une équation de diffusion (dont le noyau est la Gaussienne) plutôt qu'en stockant explicitement ces fonctions dans une matrice dont la taille dépasserait les capacités de nos super-calculateurs. Cette formulation a l'avantage de construire directement des fonctions de covariances à partir des longueurs



portées diagnostiquées (Pannekoucke et Massart, 2008). Dans le cas particulier où les tenseurs de diffusion locale sont isotropes (cercle sur la Figure 12.(a)) il est même possible de construire un algorithme numériquement très efficace où la diffusion se ramène au calcul, en parallèle, de diffusion 1D. Cet algorithme est également disponible dans le cas où les tenseurs de diffusion sont localement isotropes : quand le rayon des cercles varie de façon régulière d'un point à l'autre du domaine (ce cas sera illustré dans la suite).

Naturellement, les fonctions de corrélations localement isotropes ne représentent qu'un cas particulier et quand on les diagnostiques, les fonctions de covariances qui nous intéressent sont principalement anisotropes. Une technique possible pour construire de telles fonctions anisotropes est d'utiliser un changement de coordonnées, comme introduit par Desroziers (1997). Comme le montre la figure 12, si l'on considère un système de coordonnées rectiligne et des fonctions isotropes schématisées par les cercles (panel a), un changement de coordonnées conduit à des coordonnées curvilignes et les fonctions de covariances deviennent anisotropes telle que schématisées par les ellipses (panel b). Ainsi, on comprend qu'à partir de la construction de fonctions de covariances isotropes (par exemple à l'aide d'une diffusion isotrope) on parvient par un changement de coordonnées adapté à reproduire des fonctions anisotropes telles que diagnostiquées dans nos systèmes.

Aussi, se pose la question de comment déterminer un tel changement de coordonnées. Une première tentative d'estimation objective de ce changement de coordonnées a été introduite par Michel (2012) qui utilise les outils d'analyse en ondelette développés en traitement d'images pour déduire le changement de coordonnées. Nous avons généralisé cette approche à l'aide des outils de la géométrie Riemannienne (Pannekoucke et al. 2013). Dans ce cadre, le tenseur de diffusion locale (donc les longueurs de portées) est vu comme une métrique locale, et on calcul le changement de coordonnée, qui conduit à une métrique localement isotrope, à l'aide d'une résolution itérative d'une équation de Poisson non-linéaire. Ainsi, comme le montre la figure 13, à partir du diagnostic de la métrique locale (tenseur de diffusion représenté par les ellipses au panel (a)) on obtient par résolution de l'équation de Poisson non-linéaire des coordonnées curvilignes dans lesquelles la métrique est isotrope pour l'ensemble des points du domaine (panel (b)) : ce qui est appelé globalement isotrope. Ce jeu d'équations donne la forme générale des transformations possibles, puisqu'il permet de relaxer cette hypothèse d'isotropie globale pour décrire le cas de l'isotropie locale. La figure 14 montre le cas où d'un champ de tenseur de diffusion anisotrope (a) (dont on montre qu'il ne pourrait être ramené à champ isotrope), l'algorithme permet de

trouver un système de coordonnées dans lequel les tenseurs de diffusion sont localement isotropes (b).

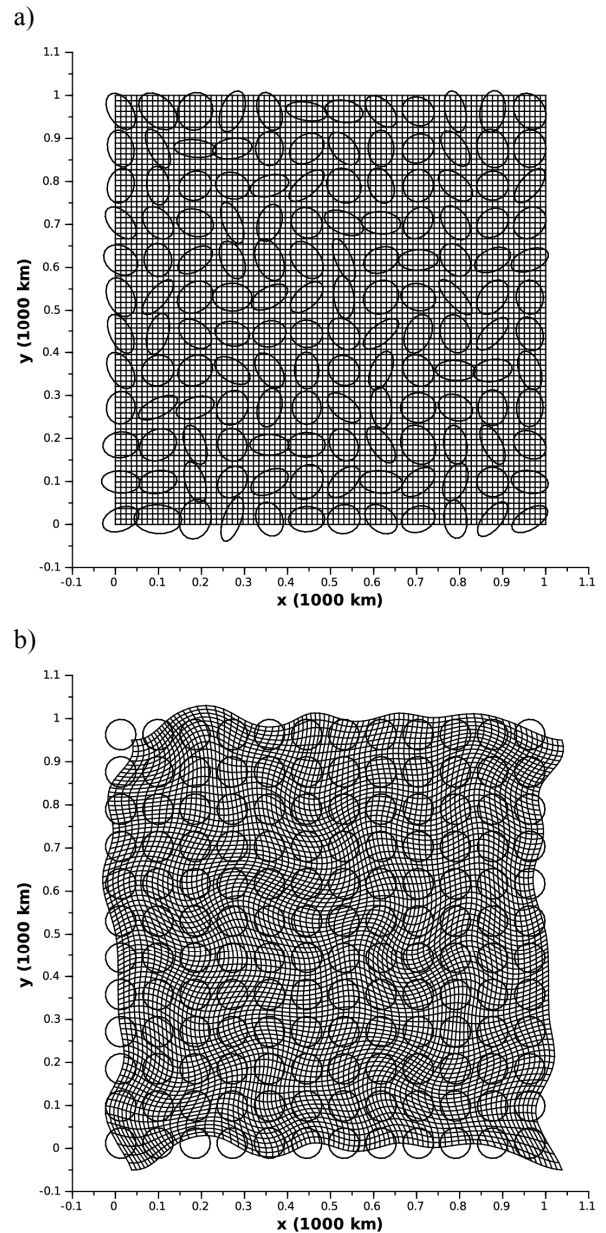


Fig.13 : Procédure d'isotropisation : en diagnostiquant les propriétés locales des longueurs de portées (a), on parvient à trouver un changement de coordonnées pour lequel les fonctions de corrélations sont isotropes (b).

Ces outils théoriques vont être maintenant utilisés et mis en œuvre pour la modélisation des fonctions de structures dans la suite VALENTINA où l'on utilisera les diagnostics ensemblistes pour une modélisation de la matrice B dépendante à l'écoulement (voir l'article de lettre : Jaumouillé et al.). Un autre aspect non développé ici est l'utilisation de cette approche pour décrire l'évolution temporelle des longueurs de portées, en particulier leur transport et leur dilatation/contraction par l'écoulement atmosphérique.

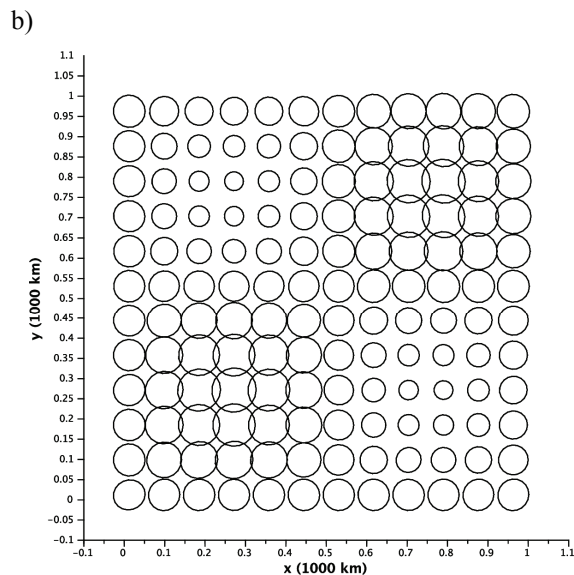
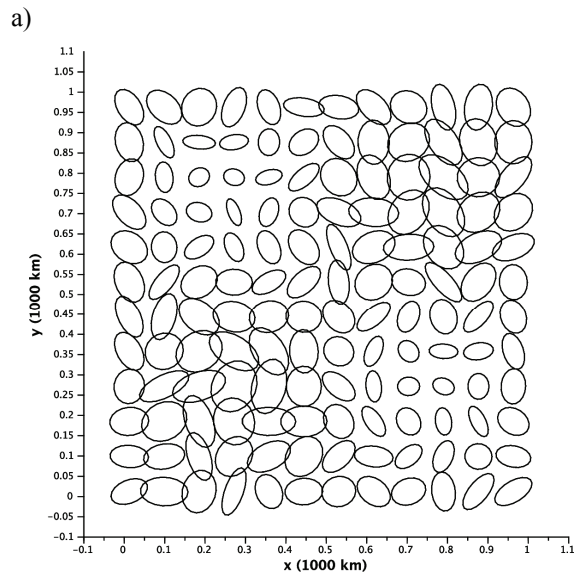
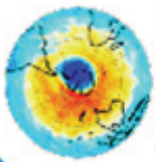


Fig.14 : Illustration de l'isotropisation d'un champ de tenseurs de diffusion anisotropes (a) vers un système de coordonnées pour lequel le champ de tenseur devient localement isotrope (b). Pour ce cas il n'existe pas de changement de coordonnées conduisant à une isotropie globale.

Références :

Daley, R. Atmospheric Data Analysis Cambridge University Press, 1991, p472.

Desroziers, G. A Coordinate Change for Data Assimilation in Spherical Geometry of Frontal Structures Monthly Weather Review, 1997, 125, 3030-3038.

Michel, Y. Estimating deformations of random processes for correlation modelling: methodology and the one-dimensional case Quarterly Journal Royal Meteorological Society, 2012.

Pannekoucke, O.; Berre, L. & Desroziers, G. Background error correlation length-scale estimates and their sampling statistics Quarterly Journal Royal Meteorological Society, 2008, 134, 497-508.

Pannekoucke, O. & Massart, S. Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. Quarterly Journal Royal Meteorological Society, 2008, 134, 1425-1438

Pannekoucke, O.; Emili, E. & Thual, O. Modeling of local length-scale dynamics and isotropizing deformations Quarterly Journal Royal Meteorological Society, 2013, doi :10.1002/qj.2204

Lien ADOMOCA sur ETHER :

<http://ether.ipsl.jussieu.fr>

Contact

elodie.jaumouille@cerfacs.fr

Prochaine lettre : Avril 2014