



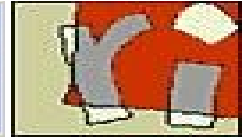
UNIVERSITE
PARIS-SUD XI



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE



INRIA



Master Recherche Informatique

Bâtiment 490 – 91 405 – ORSAY CEDEX

DIRECTION : Daniel ETIEMBLE

Responsable des Stages : Anne VILNAT

Année 2005-2006

Date de pré soutenance : 28 juin 2006

Nom de l'étudiant : Boisgontier Hervé

Organisme d'accueil : INRIA Rocquencourt – équipe CLIME

Nom de l'encadrant : J.-P. Berroir

Titre du mémoire : Faisabilité d'assimilation de mesures satellite de la chimie atmosphérique pour la prévision de la qualité de l'air.

Thème du stage : Assimilation de données, Simulation et analyse numérique, Validation et modélisation de système complexes.

Tuteur : R. Guerbi

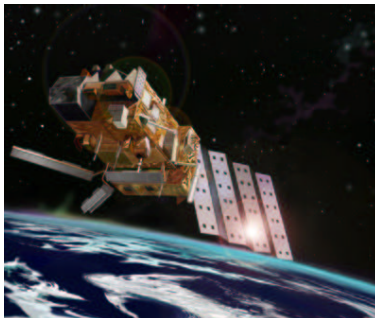
Rapporteur : C. Jacquemin

Secrétariat – tél : 01 69 15 75 18 Fax : 01 69 15 42 72

e-mail : Martine.Croissant@lri.fr

Faisabilité d'assimilation de mesures satellite de la chimie atmosphérique pour la prévision de la qualité de l'air.

Hervé Boisgontier
INRIA - Clime
78153 Le Chesnay, France
herve.boisgontier@laposte.net



IASI, embarqué sur le satellite MetOp ©ESA

Résumé

L'objectif de ce travail est d'améliorer la prévision de la qualité de l'air grâce aux données acquises par un capteur satellite : IASI. Ce capteur étant en phase de pré-lancement, nous réalisons une étude, en amont, pour caractériser le potentiel de ces données.

Ce rapport présente la méthodologie proposée pour étudier l'assimilation de données satellite de la basse troposphère dans un modèle de chimie-transport à l'échelle continentale et expose les résultats obtenus dans le cadre de ce protocole expérimental.

Dans une première partie de stage, nous montrons l'intérêt théorique de l'assimilation de colonnes d'ozone à travers deux études préliminaires. Dans un second temps, nous testons, sur des expériences numériques jumelles, l'apport de cette assimilation. Dans un troisième temps, nous réalisons des simulations de mesures qu'aurait obtenues IASI. Enfin dans un dernier temps, nous testons l'assimilation de colonnes d'ozone obtenues par simulation du capteur IASI.

Mots-Clefs

Assimilation de données, Qualité de l'air, Satellite, MetOp, IASI, Simulation et analyse numérique, Validation et modélisation de systèmes complexes.

1. INTRODUCTION.

1.1 Contexte et objectifs

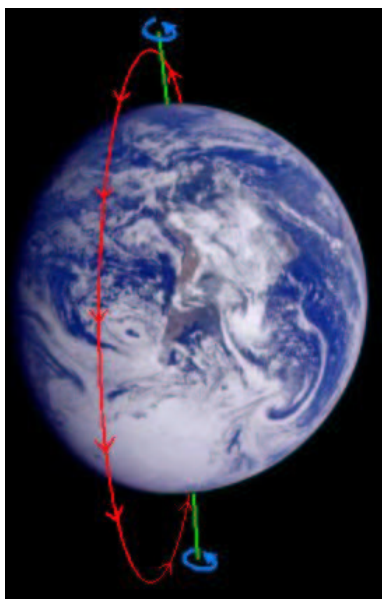
L'équipe CLIME (tous les sigles sont explicités en annexe page 19) est un projet commun entre l'INRIA Rocquencourt et le CERA (ENPC Marne la vallée). Dans le cadre du contrat "EPS MetOp RAO", l'ESA et l'EUMETSAT mettront à disposition des données réelles acquises par le capteur IASI embarqué sur le satellite EPS MetOp (le lancement prévu le 7 octobre 2006, après plusieurs reports).

Le but de ce stage est d'exploiter les données IASI pour améliorer la prévision de la qualité de l'air. Nous nous appuierons sur une technique d'assimilation de données : l'Interpolation Optimale. Notre objectif est de réaliser une étude de faisabilité de l'assimilation d'observations satellitaires, provenant d'IASI, dans un modèle de chimie-transport.

Le capteur IASI [2, 3, 13] mesure le spectre infrarouge émis par la terre. Pour notre étude, après une modélisation inverse, ces données renseignent sur la colonne d'ozone 0-6 km ; c'est-à-dire la quantité totale d'ozone comprise entre le sol et 6000 mètres d'altitude.

Le satellite MetOP (contenant IASI) est un satellite polaire, c'est-à-dire que son orbite passe au dessus des pôles (cf. figure 1). La terre tournant sur elle-même d'ouest en est, le satellite peut donc prendre la totalité de la terre en image après un certain nombre de révolutions. MetOp est synchrone par rapport au soleil, ce qui nous permet d'obtenir des captures d'information toujours à la même heure. Toutes les acquisitions de données IASI sont donc effectuées

à 9h. L'orbite de MetOp est à une altitude de 800 km, ce qui permet d'obtenir des acquisitions plus précises par rapport aux satellites géostationnaires (36000 km d'altitude).



Rouge : La trajectoire de MetOp ;
Vert : Axe de rotation de la terre ;

Figure 1 : L'orbite polaire de MetOp.

Les expériences d'assimilation sont menées en utilisant la plate-forme Polyphemus [10], développée par le CEREA. Cette plate-forme est un système, écrit en C++, permettant de piloter un modèle de prévision de la qualité de l'air appelé Polair3D [11] : fourniture des données d'entrée, mise en oeuvre d'applications telles que prévision et prévision d'ensemble, assimilation de données, visualisation... Polair3D est un programme écrit en Fortran résolvant un système d'équations différentielles modélisant le transport et les réactions chimiques des constituants de l'atmosphère.

Les phénomènes à prendre en compte dans la prévision de la qualité de l'air sont nombreux [12] : advection par le vent, diffusion turbulente, réactions chimiques pour les espèces réactives, effets dynamiques (condensation/évaporation et coagulation) pour les particules (aérosols), transferts de masse multiphasiques (entre les gaz, les particules, les gouttes de pluie et les gouttes de nuage), émissions et dépôts au niveau du sol. L'ensemble de ces processus doit être modélisé puis paramétré au sein de ce qui est communément appelé un modèle de Chimie-Transport. L'implémentation de ce modèle de Chimie-Transport, développée par le CEREA, est Polair3D.

Le document décrit tout d'abord les études préliminaires sur l'intérêt théorique de l'assimilation de colonnes d'ozone dans un modèle de Chimie-Transport (cf. section 2). L'assimilation de colonnes d'ozone, mise en place dans Polyphemus, est expliquée dans la section 3. Les expériences jumelles,

comparant des simulations avec ou sans assimilation de colonnes artificielles, sont présentées en section 4. Les données du capteur IASI n'étant pas encore disponibles, nous devons travailler sur des données simulées (cf. chapitre 5). La section 6 expose les résultats des tests d'assimilation des colonnes simulées. Nous effectuons une assimilation quotidienne, c'est-à-dire à la fréquence d'acquisition d'IASI, pour contraindre le modèle de chimie-transport Polair3D.

L'étude porte sur une période d'un mois allant du 1^{er} au 31 juillet 2001, au niveau de l'Europe. Ce mois particulier a été retenu, car les chercheurs du CEREA ont validé le modèle que nous utilisons (Polair3d) sur cette période. De plus, les émissions de polluants sont bien connues sur la même période, en raison d'études antérieures.

1.2 Motivations

Les modèles sont imprécis : ils souffrent d'incertitudes sur les conditions initiales, sur les données d'entrée, ainsi que sur la paramétrisation.

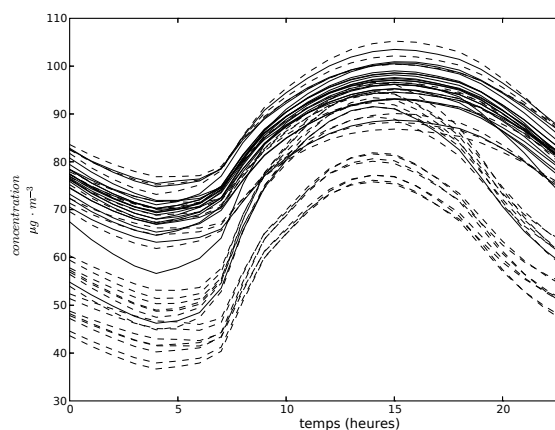


Figure 2 : Prévisions issues du modèle Polair3D en faisant varier les paramètres physiques et les données d'entrée.

La figure 2 représente des concentrations d'ozone au niveau du sol, en moyenne sur l'Europe, en fonction du temps pour une journée donnée. Chacune des différentes courbes a été obtenue par le même modèle (Polair3D) en appliquant de légères perturbations sur les paramètres physiques et les données d'entrée. Les différences de concentrations d'ozone entre les différentes prévisions peuvent atteindre près de $40\mu\text{g}\cdot\text{m}^{-3}$. Cet exemple illustre le fait que le modèle est sous-contraint : il a besoin d'informations supplémentaires. Actuellement, il est possible d'assimiler des mesures d'ozone effectuées au niveau du sol. Les mesures au sol permettent de contraindre le modèle, mais elles sont mal réparties géographiquement et n'apportent pas d'information sur l'ozone de plus haute altitude. C'est pourquoi nous souhaitons assimiler des colonnes d'ozone. Ainsi, l'incertitude sera réduite et le modèle sera mieux contraint.

1.3 Travaux antérieurs

L'assimilation de données [1] est une méthodologie qui utilise les outils de l'optimisation des Équations aux Dérivées Partielles, déterministes et stochastiques, pour trouver l'état le plus probable d'un système, compte-tenu d'une « ébauche *a priori* » fournie par une simulation numérique et des mesures expérimentales hétérogènes. Cette méthodologie permet de tenir compte des erreurs affectant aussi bien le modèle numérique que l'acquisition des observations.

L'assimilation de données satellite de la basse troposphère n'a encore jamais été réalisée en raison du manque de données appropriées. Toutefois, des travaux concernant l'assimilation de données satellite mesurant la colonne totale d'ozone (du sol jusqu'à la limite de l'atmosphère) ont été réalisés, avec succès, par Eskes et al. [6] pour étudier le trou dans la couche d'ozone. Nous avons donc dû mettre en place, pour cette étude, un protocole expérimental afin de tester l'intérêt de l'assimilation de colonnes d'ozone de basse altitude, en vue d'étudier la qualité de l'air et sa prévision.

2. ÉTUDES PRÉLIMINAIRES

2.1 Introduction

L'atmosphère terrestre est divisée en plusieurs couches. Chaque couche est définie par des caractéristiques physiques différentes. Nous nous intéressons aux altitudes comprises entre 0 et 5 km. Il y a deux couches présentes dans cette zone : la couche limite aux basses altitudes et la troposphère libre pour les altitudes élevées. Cette dernière s'élève en fait au delà de 5 km, mais nous ne nous intéressons qu'à sa partie inférieure. La limite entre ces deux couches n'est pas fixe et varie suivant l'heure de la journée.

La couche limite est définie comme la partie de la troposphère directement soumise à l'influence de la surface terrestre [15]. Les polluants présents dans la couche limite se dispersent et se mélangent rapidement en raison de la turbulence de cette couche ; ceux qui atteignent la troposphère libre se dispersent plus lentement en raison de la plus grande stabilité verticale et de la plus faible intensité de la turbulence [8].

La motivation principale de ce travail est une meilleure connaissance de la qualité de l'air de la couche limite, c'est-à-dire l'air que nous respirons. Le capteur IASI, embarqué sur satellite, est plus sensible aux couches de haute altitude qu'aux couches de basse altitude. Il sera donc intéressant d'assimiler les colonnes uniquement si l'ozone de la couche limite représente une part significative de la colonne mesurée (section 2.2), ou bien si une meilleure connaissance des hautes altitudes (en connaissant mieux les concentrations) permet de mieux connaître les concentrations aux basses altitudes (section 2.3).

Pour pouvoir effectuer des tests quantitatifs, nous créons tout d'abord une référence en exécutant une simulation de Polyphemus avec les meilleures données d'entrée possibles

du modèle sur la période choisie. Polyphemus prend en entrée une masse de données (2,5 Go), répartie dans plus de 200 fichiers, représentant les différents paramètres physiques du modèle (cf. figure 3).

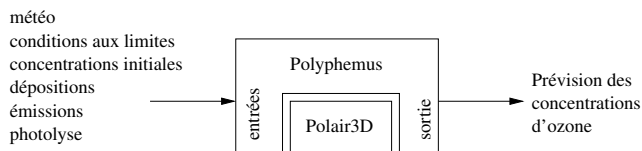


Figure 3 : L'architecture de Polyphemus.

Le temps de calcul pour une simulation d'un mois telle que cette référence est sur la machine utilisée (Pentium 4 cadencé à 3,2 GHz) est de 8 heures.

2.2 Calcul de la contribution de la couche limite dans une colonne d'ozone entre 0 et 5 km.

2.2.1 Introduction

Nous désirons connaître la contribution de l'ozone de la couche limite à la colonne d'ozone 0-5 km, c'est-à-dire la part que représente la quantité d'ozone contenue dans la couche limite par rapport à la quantité totale d'ozone comprise entre le sol et 5 km. Pour avoir une idée de cette valeur, nous avons effectué une étude sur la période de référence (mois de juillet 2001) en Europe. Nous avons calculé, toutes les 3 heures, la quantité d'ozone contenue dans la couche limite et nous l'avons comparée avec la colonne d'ozone entre 0 et 5 km.

2.2.2 Intérêt du calcul

Le modèle de chimie-transport n'est validé complètement que dans la seule couche limite. Il repose sur des observations au sol, alors que les données satellite intègrent des informations sur la troposphère libre et la couche limite. Nous cherchons à assimiler des colonnes d'ozone issues de mesures satellite, donc plus sensibles aux couches hautes de l'atmosphère. Nous désirons savoir ce que représente la mesure du satellite et en quoi elle nous renseigne sur la couche limite, car la principale finalité de l'étude de la qualité de l'air est de mieux connaître la pollution au niveau du sol.

2.2.3 Comment est effectué le calcul ?

D'une part, parmi les données d'entrée de Polyphemus, une matrice à trois dimensions donne l'altitude de la couche limite par rapport au sol, pour toute l'Europe, toutes les trois heures, pour le mois de juillet 2001. Cette donnée est issue des modèles météo de l'ECMWF.

D'autre part, notre simulation de référence est une matrice à quatre dimensions contenant pour 10 intervalles d'altitude (cf. équation 1), pour toute l'Europe, toutes les heures, la concentration d'ozone.

$$z = \left\{ \begin{array}{l} [0; 50] \\ [50; 300] \\ [300; 600] \\ [600; 1000] \\ [1000; 1500] \\ [1500; 2100] \\ [2100; 2700] \\ [2700; 3400] \\ [3400; 4200] \\ [4200; 5000] \end{array} \right\} \text{ (en mètres)} \quad (1)$$

Pour calculer une colonne d'ozone, entre le sol et 5 km, en un point de coordonnées (i, j) , à un instant t , nous utilisons la formule 2 :

$$Colonne_{t,i,j} = \sum_{h=0}^{10} C_{O_3}(t,h,i,j) \cdot (z_{h_{max}} - z_{h_{min}}) \quad (2)$$

Si la frontière de la couche limite se situe à 700 mètres d'altitude, l'ozone comprise entre 0 et 600 mètres est comptabilisée comme contribution à la couche limite alors que celle comprise entre 1000 et 5000 mètres est comptabilisée comme contribution à la troposphère libre. La quantité d'ozone comprise entre 600 et 1000 mètres est répartie entre la couche limite ($C_{O_3}(t,h=4,i,j) \cdot (700 - 600)$) et la troposphère libre ($C_{O_3}(t,h=4,i,j) \cdot (1000 - 700)$). Nous avons donc créé un programme qui lit le fichier contenant la hauteur de la couche limite et celui contenant les concentrations d'ozone, puis calcule la contribution de la couche limite à la colonne d'ozone.

2.2.4 Résultats

En moyenne sur la période du mois de juillet 2001, nous trouvons que la couche limite contribue à hauteur d'environ 15% à la colonne d'ozone entre le sol et 5 km.

La contribution est plus importante au niveau des terres qu'au niveau des mers (cf. figure 4).

Nous avons étudié plus précisément les périodes de forte activité d'ozone (vers 15h) et de faible activité (la nuit). Nous avons écrit un programme calculant la contribution moyenne, en chaque point du domaine, pour un horaire donné. Par exemple, nous nous intéressons à la contribution à 15h, pour cela, nous effectuons la moyenne de toutes les cartes quotidienne de contribution à 15h. Ainsi, nous obtenons une carte de la contribution à 15h, en moyenne sur le mois, de la couche limite à la colonne d'ozone (cf. figure 5).

La contribution de la couche limite est plus importante à 15h (au niveau des terres, elle représente entre 20 et 60% de la colonne d'ozone). Par contre, la contribution est moins importante la nuit (elle représente moins de 10% au niveau des terres, cf. figure 7).

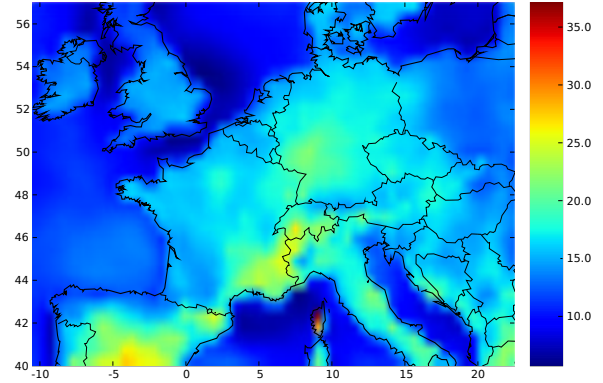


Figure 4 : Contribution (en %) de la couche limite à la colonne d'ozone en moyenne sur le mois de juillet 2001.

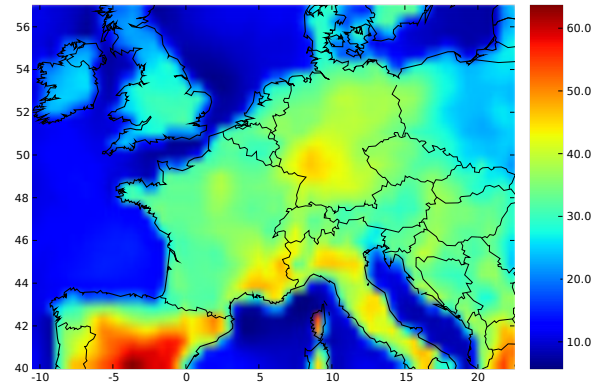


Figure 5 : Contribution (en %) de la couche limite moyenne sur le mois de juillet à 15h.

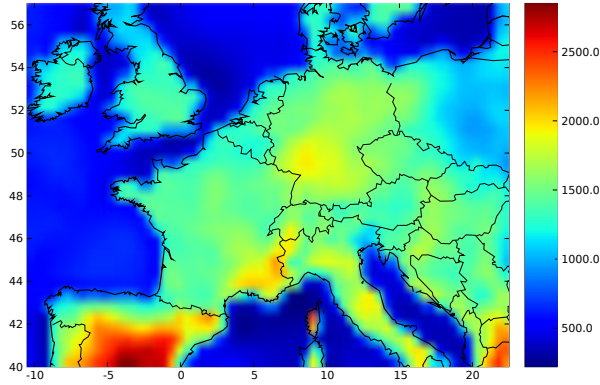


Figure 6 : Hauteur de la frontière (en mètres) entre la couche limite et la troposphère libre à 15h en moyenne sur le mois de juillet 2001.

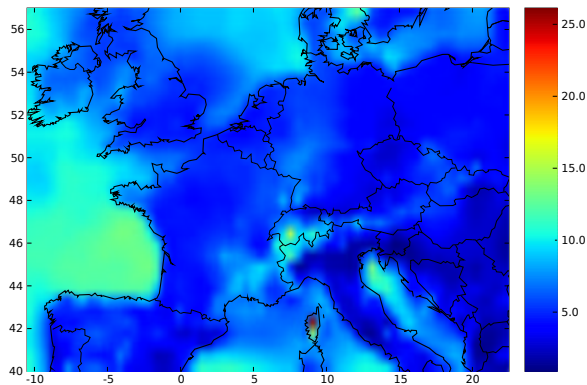


Figure 7 : Contribution (en %) de la couche limite en pourcentage moyen à minuit sur le mois de juillet 2001.

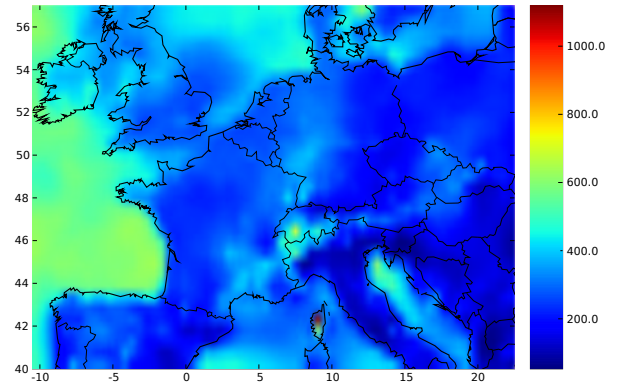


Figure 8 : Hauteur de la frontière (en mètres) entre la couche limite et la troposphère libre à minuit en moyenne sur le mois de juillet 2001.

La figure 9 nous indique que la contribution de la couche limite à la colonne d’ozone entre le sol et 5 km est cyclique, importante en journée et plus faible la nuit.

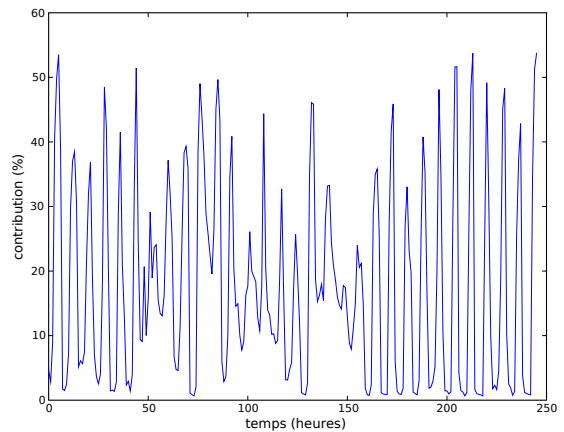
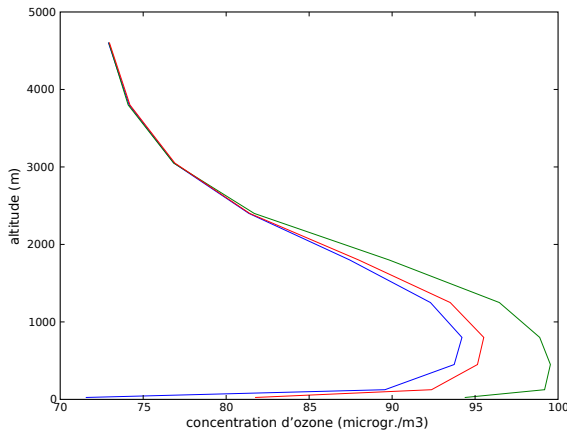


Figure 9 : Contribution de la couche limite en pourcentage pour un point situé au milieu de la carte ($48^{\circ}30'N$, $06^{\circ}E$ - près de Nancy).

Nous avons cherché à expliquer la différence entre le jour et la nuit. Nous avons calculé le coefficient de corrélation entre la hauteur de la frontière entre les deux couches et la contribution de la couche limite à la colonne d’ozone. Nous avons obtenu une corrélation de 96,53 % (cf. figures 6 et 8 à comparer respectivement aux figures 5 et 7). Nous avons écrit un script en Python pour visualiser la répartition verticale de l’ozone, à différents horaires de la journée (cf. figure 10). Ce profil indique qu’il y a plus d’ozone présent au niveau du sol en journée que la nuit alors que la quantité d’ozone est identique pour les hautes altitudes quelle que soit l’heure de la journée.

Nous pouvons donc expliquer la différence entre le jour et la nuit par deux phénomènes : principalement, la hauteur de la couche limite et d'autre part, le profil vertical de l'ozone.



Rouge : moyenne.
Bleu : moyenne à minuit.
Vert : moyenne à 15h.

Figure 10 : Profil vertical de la quantité d'ozone en moyenne sur tout le domaine.

2.2.5 Conclusion de cette étude en vue de l'assimilation de colonnes d'ozone

L'étude présentée ci-dessus nous a montré qu'au niveau des terres, à 15h, nous avons une contribution de la couche limite à la colonne d'ozone 0-5 km de l'ordre de 33%. À l'heure de l'acquisition (9h) par IASI des données, à partir desquelles seront calculées les colonnes d'ozone, cette contribution est d'environ 25%. Le résultat est donc intéressant et encourageant pour assimiler des colonnes d'ozone.

2.3 Test de sensibilité

2.3.1 Introduction

Comme nous venons de le voir, les données satellite que nous cherchons à assimiler représentent majoritairement la troposphère libre. Une amélioration de la prévision de la qualité de l'air (dans la couche limite) sera effective si les estimations des concentrations dans la couche limite sont sensibles à des corrections (dûes à une meilleure connaissance) effectuées dans la troposphère libre. À quel point une perturbation des quantités d'ozone dans la troposphère libre modifie-t-elle les quantités d'ozone de la couche limite ?

Pour tester cette sensibilité, nous changeons artificiellement dans Polyphemus les concentrations d'ozone dans la troposphère libre, puis une simulation est effectuée. Nous comparons alors les résultats obtenus dans la couche limite avec ceux de la simulation de référence.

2.3.2 Modification des conditions initiales

Dans une première expérience, les concentrations initiales d'ozone au dessus de la couche limite sont multipliées par une constante α , puis la simulation avec Polyphemus est effectuée. Pour cela, a été conçu un programme créant un fichier de concentrations initiales d'ozone modifiées, à partir du fichier des concentrations initiales et du fichier de hauteur de la frontière entre les deux couches. La figure 11 montre l'erreur relative en fonction du temps, pour une perturbation initiale de $\alpha = -50\%$, sur tout le domaine, entre la simulation perturbée et la référence.

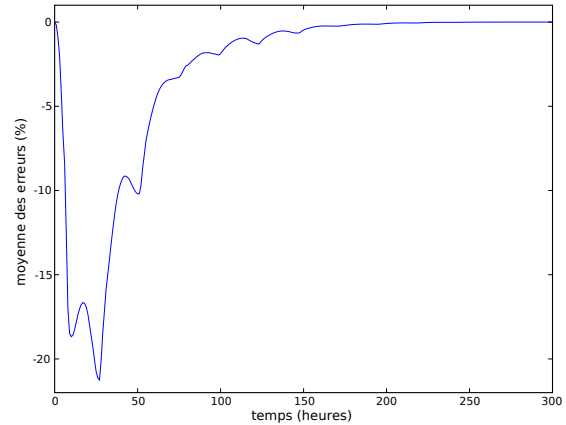


Figure 11 : Erreur relative moyenne entre la simulation test et la référence.

Le pic (impact maximal au niveau du sol) est atteint à $t_{max} = 27$ heures après le début de la simulation et la différence de quantité d'ozone (noté $effet_{t=t_{max}}$), à ce moment là, est en moyenne de $-21,26\%$, mais en certain point il peut y avoir des différences de l'ordre de -90% (cf. figure 12).

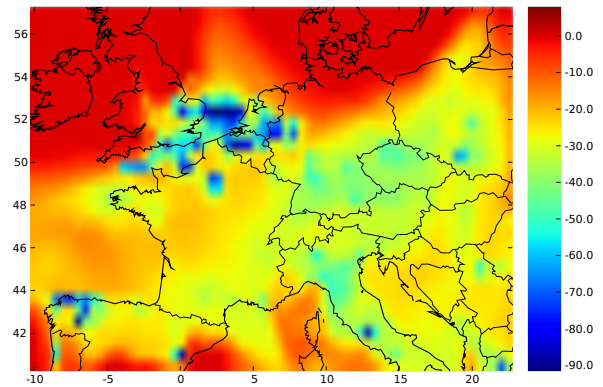


Figure 12 : Erreur relative moyenne entre la simulation test et la référence.

Table 1 : Résultats des expériences de modification des concentrations initiales d’ozone dans la troposphère libre.

α	t_{max}	$effet_{t=t_{max}}$	t_{fin}
+100%	1 jour et 3 h	+67,34 %	6 jours et 6 h
+20%	1 jour et 3 h	+9.93 %	6 jours et 6 h
-20%	1 jour et 3 h	-8.96 %	6 jours et 6 h
-50%	1 jour et 3 h	-21,26 %	6 jours et 6 h
-100%	1 jour et 3 h	-38,94 %	6 jours et 6 h

Les effets de la perturbation sont complètement atténués après 6 jours et 6 heures (inférieur à 1 % de la perturbation, ce laps de temps est noté t_{fin}).

Les résultats des expériences suivantes (avec des coefficients multiplicateurs différents) sont présentés dans le tableau 1.

Ceux-ci sont similaires : Le pic est toujours 27 heures après la perturbation. Les effets de la perturbation sont tangibles jusqu’à 6 jours et 6 heures.

Les perturbations +100 % et -100 % donnent des résultats exceptionnels (en certains points la quantité d’ozone a pu être modifiée de +3500 %), Nous choisissons donc de les écarter. Ces résultats sont sûrement dus à une discontinuité de la concentration d’ozone entre une zone modifiée (au dessus de la couche limite) et une zone non-modifiée (ce qui provoque des effets de bords).

Dans ces simulations, nous obtenons, en moyenne et au plus fort de l’effet ressenti, au niveau du sol, un impact de l’ordre de 46 % de la perturbation initiale dans la troposphère libre.

Le capteur IASI sera en capacité de nous donner des colonnes d’ozone toutes les 24 heures. Étant donné qu’une donnée différente dans la troposphère libre a un impact au niveau du sol effective pendant environ 48 heures, nous pouvons affirmer qu’assimiler ces colonnes d’ozone doit permettre d’améliorer la qualité de la prévision si les données assimilées sont de bonne qualité.

2.3.3 Simulation de modification cyclique de la quantité d’ozone pour les altitudes comprises entre 1500 et 5000 mètres

Nous souhaitons assimiler des colonnes issues de données satellite, acquises chaque jour par le capteur IASI. Pour estimer l’impact de l’assimilation de ces données, nous testons la sensibilité de Polyphemus à des perturbations quotidiennes. Chaque jour à 14h, dans le processus de simulation, nous multiplions, par un coefficient α , les concentrations d’ozone entre 1500 et 5000 mètres. Pour cela, nous remplaçons le pilote de base (une classe C++) par un nouveau pilote que nous avons écrit afin de modifier les concentrations d’ozone

au cours de la simulation.

9h est l’horaire d’acquisition de données par le capteur satellitaire IASI. C’est pourquoi, nous perturbons les concentrations d’ozone dans la partie haute du domaine étudié, à cette heure là.

Nous avons choisi de ne pas modifier toutes les concentrations de la troposphère libre car, à 9h (horaire de activité d’ozone modérée), la hauteur de la limite entre les deux couches est très variable suivant le lieu (différence terre et mer cf. section 2.2).

La figure 13 représente la différence relative entre la simulation perturbée avec $\alpha = -50 %$ et la référence, au niveau du sol.

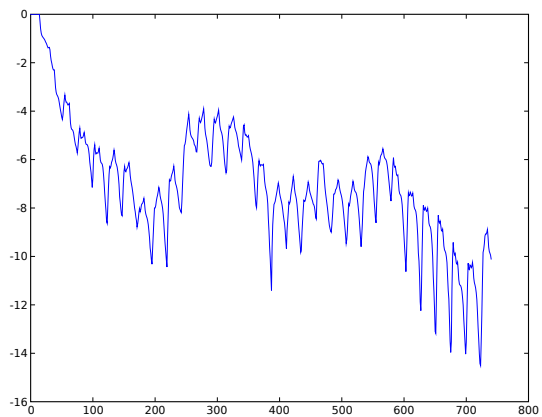


Figure 13 : Erreur relative moyenne entre la simulation et la référence.

Toutes les 24 heures, la simulation est perturbée. Au bout de 24 heures, à la première perturbation qui n’est pas terminée, s’ajoute la seconde perturbation. Après 5 jours environ (120 heures), le nombre de perturbations ressenties reste constant. Ainsi, l’erreur moyenne effectuée par cette simulation reste comprise entre -4 et -15 %.

Le fait de modifier les concentrations d’ozone de -50 % a donc un impact, au niveau du sol, sur les concentrations entre -4 % et -15 %, ce qui correspond environ au quart de la perturbation.

Nous avons effectué d’autres simulations avec pour valeur de α +100 %, +20 % et -20 %. Pour les mêmes raisons, nous écartons le +100 %. Nous avons obtenu des résultats qui sont semblables, pour +20 % et -20 %. L’impact au niveau du sol est un peu moins important : de l’ordre du cinquième de la perturbation effectuée sur les hautes altitudes.

2.3.4 Conclusion

Au vu de ces résultats, nous pouvons conclure que les concentrations d'ozone dans la couche limite sont sensibles via Polyphemus à des variations de concentrations d'ozone dans la troposphère libre. En supposant que le modèle Polyphemus représente fidèlement la réalité de la relation entre les concentrations d'ozone de la troposphère libre et de la couche limite, nous en concluons que le fait de mieux connaître les concentrations dans la troposphère libre influe sur la qualité de la prévision dans la couche limite. Ainsi, l'assimilation des colonnes d'ozone issues d'IASI doit permettre d'améliorer les prévisions des concentrations d'ozone au niveau du sol.

2.4 Bilan des études préliminaires

Ces deux résultats sont concluants : effectuer des corrections des concentrations dans la troposphère libre a un impact significatif (1/5) sur les concentrations dans la couche limite ; cet impact est maximum 27 heures après la perturbation et dure suffisamment longtemps (50 heures) au regard de la fréquence d'acquisitions des données IASI. L'assimilation de colonnes d'ozone est susceptible d'améliorer la prévision de la qualité de l'air.

3. L'ASSIMILATION DE DONNÉES AVEC LA PLATE-FORME POLYPHEMUS.

3.1 Introduction

Les études préliminaires ont montré l'intérêt théorique d'assimiler des colonnes d'ozone pour améliorer la qualité de la prévision au niveau du sol. Nous mettons en place les éléments permettant cette assimilation. La méthode d'assimilation de données utilisée est une de celles disponibles dans Polyphemus : l'Interpolation Optimale [1].

3.2 L'assimilation de données par la méthode de l'Interpolation Optimale

Nous disposons de deux informations : une prévision, notée X_b , effectuée avec un modèle, et des observations, notées Y_o . Les fiabilités sont données sous forme de matrice de covariance B et R . B correspond à l'erreur *a priori* du modèle, alors que R correspond aux erreurs de mesure effectuées par l'instrument. Le but de l'assimilation de données est de trouver une meilleure prévision, ou « analyse », notée X_a , qui réalise un compromis entre la prévision issue du modèle et les observations. Pour pouvoir travailler sur des données de même nature, nous utilisons un opérateur d'observation, noté H , qui fait passer de l'espace des prévisions à l'espace des observations.

La méthode de l'Interpolation Optimale cherche à minimiser une fonction J (cf. équation 3) qui évalue le coût associé à une prévision X (cf. [5, 9, 7]) :

$$\begin{aligned} J(X) &= J_b(X) + J_o(X) \\ &= \underbrace{(X - X_b)^T B^{-1} (X - X_b)}_{\text{écart à la prévision du modèle}} \\ &\quad + \underbrace{(Y_o - H(X))^T B^{-1} (Y_o - H(X))}_{\text{écart aux observations}} \end{aligned} \quad (3)$$

Elle est constituée de deux parties : un terme caractérisant l'écart à la prévision du modèle, noté J_b , et un autre l'écart aux observations, noté J_o (cf. figure 14).

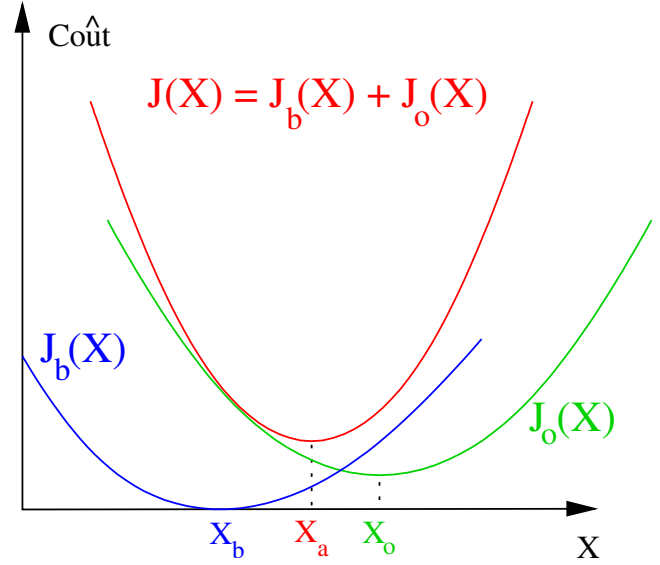


Figure 14 : Fonction de coût J .

X_a est la prévision pour laquelle la fonction de coût J est minimale (cf. équation 4), c'est-à-dire que cette prévision réalise le compromis optimal entre la prévision issue du modèle et les observations.

$$X_a = \text{Arg min } J \quad (4)$$

Après résolution de l'équation 4 nous obtenons l'équation suivante :

$$\begin{aligned} X_a &= X_b + K \underbrace{(Y_o - H(X_b))}_{\text{Innovation}} \\ \text{avec } K &= \underbrace{B^T (HBH^T + R)^{-1}}_{\text{Matrice de gain}} \end{aligned} \quad (5)$$

La démonstration est donnée en annexe (page 18).

3.3 L'assimilation de colonnes d'ozone avec la méthode de l'Interpolation Optimale avec Polyphemus

Nous utilisons le pilote d'Interpolation Optimale déjà existant dans Polyphemus. La matrice de covariance B est définie dans le fichier de configuration *polair3d-oi.cfg*. Dans le fichier de configuration *observation.cfg*, nous conservons la rubrique qui définit la matrice de covariance R relative aux observations, et nous ajoutons une nouvelle rubrique pour indiquer le chemin d'accès aux colonnes observées.

Ensuite, j'ai écrit, sous forme d'une classe C++, un « manager » qui décrit l'opérateur d'observation H , définit les matrices de covariance B et R (décrites dans les fichiers de configuration) et lit les données observées Y_o (les colonnes) (cf. figure 15). À chaque instant, pour lequel des observations sont disponibles (acquisition d'une image satellite), une assimilation de données est effectuée.

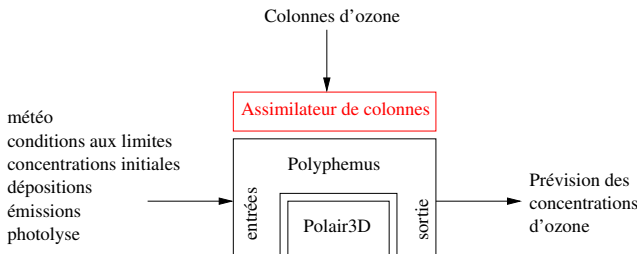


Figure 15 : Ajout du module d'assimilation de colonnes à Polyphemus contenant le manager et le modèle dérivé.

Le module d'assimilation de Polyphemus permet de choisir les altitudes d'intérêt, pour lesquelles les concentrations d'ozone doivent être prises en compte dans le vecteur d'état X_b . Dans un premier temps, nous considérons un vecteur d'état contenant l'ensemble des concentrations sur le domaine 3D, soit une dimension de 21 450. L'opérateur d'observation réalise alors l'intégration verticale des concentrations pour obtenir un champ 2D de colonnes.

En raison de la taille du vecteur d'état, le temps de calcul associé est prohibitif. Polyphemus devait inverser des matrices carrées de 21 450 * 21 450 éléments. Dans un second temps, j'ai donc cherché à diminuer la dimension de l'état : nous ne considérons que les colonnes (taille du vecteur d'état divisée par 10 et la taille des matrices à inverser a été divisée par 100). Nous avons donc dû écrire un nouveau module d'assimilation, dérivé du précédent, appelé *modèle dérivé*. Celui-ci permet d'inclure, dans le vecteur d'état, les colonnes d'ozone. L'assimilation se déroule ainsi en trois temps :

1. Calcul des colonnes à partir des concentrations 3D sorties du modèle ;
2. Assimilation des colonnes avec H devenu opérateur identité ;

3. Correction des concentrations à l'aide de la colonne obtenue, en utilisant la même forme de profil que celle issue de la prévision du modèle.

Ainsi le temps de calcul de l'assimilation des colonnes a été réduit à environ 2 minutes sur la machine utilisée (Pentium 4 cadencé à 3,2 GHz). Le module d'assimilation de colonnes d'ozone est donc constitué de deux entités : le manager et le *modèle dérivé*.

4. ASSIMILATION DE COLONNES SIMULÉES À PARTIR DE LA RÉFÉRENCE.

4.1 Introduction

Nous cherchons, par ces tests, à valider l'écriture du manager d'une part, et à tester l'utilité d'assimiler les colonnes d'ozone d'autre part.

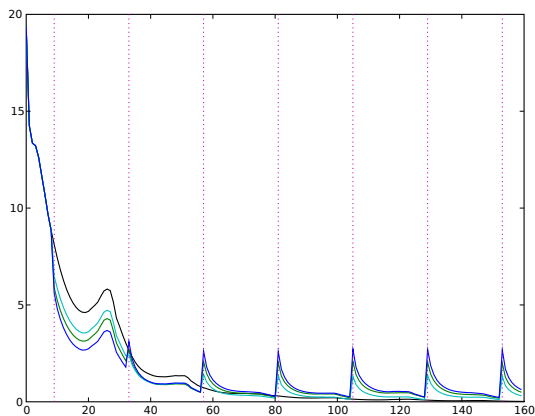
Pour cela, nous créons un programme calculant des colonnes d'ozone à partir des concentrations d'ozone correspondant à la simulation de référence. Grâce à un autre programme, nous perturbons ces colonnes par un bruit blanc, afin de se rapprocher de l'assimilation de colonnes réelles. Nous modifions également un des paramètres d'entrée de façon à ce que le modèle devienne incorrect. Des expériences jumelles sont effectuées ensuite : l'une avec et l'autre sans l'assimilation des colonnes perturbées dans le modèle erroné.

4.2 Première expérience : concentrations initiales d'ozone perturbées

Cette première expérience est le test permettant de valider les travaux de la section 3. Les concentrations d'ozone initiales sont perturbées par un bruit blanc de variance 20%. Cela correspond au modèle erroné. Nous assimilons les colonnes obtenues en ajoutant un bruit blanc de variance 5%.

Nous avons effectué l'assimilation de ces colonnes avec plusieurs valeurs pour la matrice de covariance R . Les résultats sont présentés sur la figure 16 : nous représentons la moyenne de la valeur absolue de la différence relative entre la simulation et la référence en fonction du temps (en heures) à l'altitude 0. Nous observons deux phases : la première journée, où la simulation est très influencée par les concentrations initiales ; les jours suivants, où les conditions initiales ont été « oubliées ».

Dans la première phase, les simulations avec assimilation des colonnes simulées donnent de meilleurs résultats que la simulation sans assimilation. Plus la matrice de covariance d'erreur R est faible, plus l'assimilation fait confiance aux colonnes (courbe en bleu sur la figure 16), et donc la simulation se rapproche plus rapidement de la référence. Au contraire, la simulation ne faisant que peu confiance aux colonnes (courbe en cyan sur le figure 16) se rapproche plus doucement de la référence.



Noir : Sans assimilation ;
 Bleu : Assimilation avec $R = 0,5.id$;
 Vert : Assimilation avec $R = 1.id$;
 Cyan : Assimilation avec $R = 2.id$;
 Magenta pointillé : Horaire d'assimilation.

Figure 16 : Test d'assimilation de colonnes simulées d'ozone avec les concentrations initiales perturbées.

Dans la seconde phase, les concentrations initiales ne sont plus influentes. Les observations sont, par contre, toujours perturbées. C'est pour cela que les résultats des simulations avec assimilation sont moins bons : à chaque assimilation, l'écart avec la référence augmente, puis s'atténue jusqu'à l'assimilation suivante.

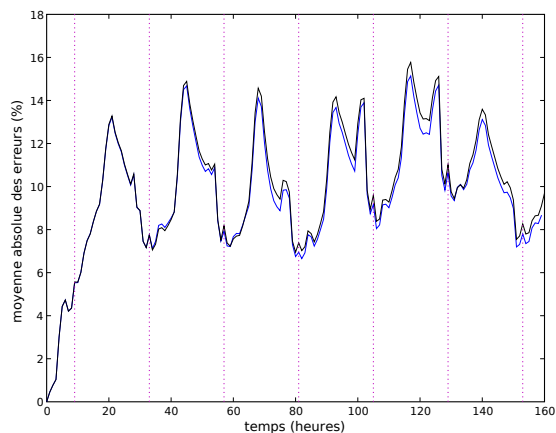
Ce test valide d'une part l'écriture du manager et du modèle dérivé, et d'autre part montre l'intérêt de l'assimilation de colonnes dans le cas de concentrations initiales mal connues. Nous passons ensuite à un test où une des données d'entrée du modèle est erronée.

4.3 Seconde expérience : diffusion verticale inexacte

Un polluant se disperse plus ou moins facilement dans l'air en fonction des conditions météorologiques de l'atmosphère et, notamment, de la température dont le profil permet la diffusion verticale ainsi que la turbulence de la couche. Généralement, dans les basses couches de l'atmosphère, la température décroît régulièrement avec l'altitude, ce qui permet aux polluants émis de s'élever et donc de se mélanger avec l'air ambiant.

Dans une seconde expérience, nous avons perturbé le modèle en modifiant un des paramètres physiques, la diffusion verticale, qui contrôle les échanges verticaux dus à la turbulence. Pour ce faire, nous avons remplacé le fichier de diffusion verticale par un autre correspondant à une diffusion verticale différente.

La figure 17 indique que l'assimilation de colonnes d'ozone



Noir : Sans assimilation ;
 Bleu : Avec assimilation ;
 Magenta pointillé : Horaire d'assimilation.

Figure 17 : Moyenne de la valeur absolue de la différence relative entre la simulation et la référence.

ne permet pas de pallier une mauvaise connaissance de ce paramètre, car nous n'apportons qu'une information sur la quantité totale d'ozone dans une colonne et non sur la répartition verticale à l'intérieur de celle-ci.

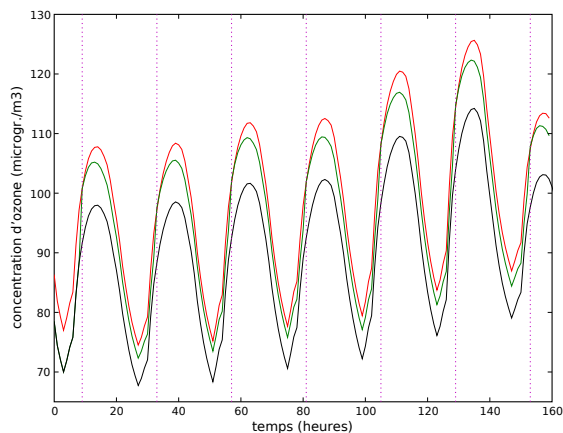
4.4 Troisième expérience : assimilation de colonnes dans un modèle erroné

Cette expérience teste l'assimilation de colonnes dans un modèle erroné. Pour créer ce modèle, nous avons multiplié toutes les concentrations de la référence par un coefficient α . L'assimilation des colonnes permettra de réaliser un compromis entre les observations, générées à partir du modèle de référence, et la simulation issue du modèle erroné.

Dans un premier temps, nous cherchons à connaître l'impact de l'assimilation de colonnes parfaites : nous ne perturbons pas les colonnes à assimiler et nous positionnons la matrice d'erreur R à 0, afin d'indiquer que l'erreur sur ces colonnes est nulle.

Les résultats de cette expérience sont illustrés dans la figure 18 avec $\alpha = 1,1 = +10\%$. Elle présente trois courbes : la courbe rouge représente la prévision issue du modèle de référence, la courbe noire est la prévision obtenue grâce au modèle erroné et la courbe verte est le résultat de la simulation avec assimilation.

Jusqu'à la première assimilation, la simulation et la prévision issue du modèle erroné donnent exactement la même prévision : puisqu'elles partent toutes deux des mêmes concentrations initiales et possèdent les mêmes conditions aux limites. Nous observons que, à la première assimilation, la simulation avec assimilation atteint exactement la prévision issue du bon modèle : ce qui est normal, puisque les co-



Rouge : La prévision issue du bon modèle ;
 Vert : Simulation avec assimilation ;
 Noir : La prévision issue du modèle erroné ;
 Magenta pointillé : Horaire d'assimilation.

Figure 18 : Concentration au niveau du sol pour une assimilation de bonnes colonnes dans un modèle erroné.

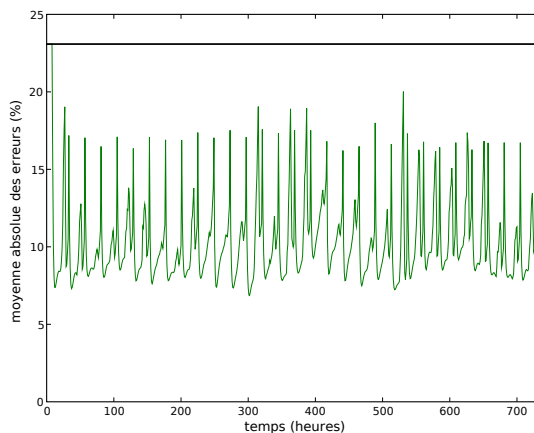
lonnes sont supposées parfaites. Entre deux assimilations de colonnes, la simulation s'éloigne de la référence car elle est obtenue par le modèle erroné.

Au bout de 24 heures (le temps entre deux assimilations) la simulation se trouve en moyenne à mi-distance entre la référence et le modèle. L'assimilation des colonnes est influente assez longtemps pour maintenir la prévision proche de celle obtenue avec le bon modèle.

Dans un second temps, nous nous intéressons à une simulation plus réaliste : avec des erreurs sur les colonnes et un modèle plus fortement erroné. L'incertitude prévue sur les colonnes issues d'IASI étant de l'ordre de 20 %, nous avons choisi cette valeur comme perturbation sur les colonnes de référence. Nous avons utilisé un modèle erroné d'environ 23 % (incertitude probable du modèle). La figure 19 présente deux courbes : les erreurs commises par une simulation sans assimilation (erreurs constantes = 23 %, en noir) et les erreurs d'une simulation avec assimilation des colonnes (en vert).

Avec des colonnes perturbées (20 %), nous obtenons en moyenne des erreurs comprises entre 7 % et 17 %. La prévision obtenue avec assimilation des colonnes d'ozone effectuée au minimum 25 % d'erreur en moins par rapport à la prévision sans assimilation de données.

Selon ces résultats, nous pouvons espérer, en assimilant des colonnes d'ozone issues des données IASI, améliorer la qualité de la prévision au niveau du sol avec un modèle dont les paramètres physiques sont imparfaits.



Vert : Simulation avec assimilation ;
 Noir : Simulation sans assimilation.

Figure 19 : Moyenne de la valeur absolue de la différence relative entre la simulation et la référence.

5. SIMULATION DES DONNÉES PRODUITES PAR LE CAPTEUR SATELLITAIRE IASI.

5.1 Cadre général de la simulation des colonnes simulées

Actuellement, le satellite MetOp, contenant le capteur IASI, n'a pas encore été envoyé en orbite. Pour effectuer l'étude, nous avons donc besoin de simuler les données qu'auraient obtenues IASI pendant la période d'intérêt. Celles-ci doivent être les plus proches possible des données qu'aurait mesurées le capteur IASI dans les mêmes conditions. Ces données subissent un traitement, en chaque point, afin d'obtenir une image où chaque valeur ponctuelle correspond à la colonne d'ozone 0-6 km.

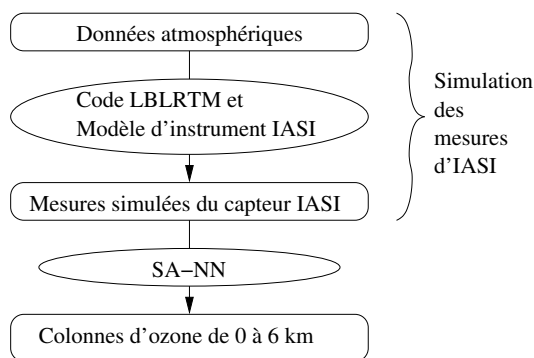


Figure 20 : Chaîne logicielle simplifiée de simulation des données IASI et des colonnes d'ozone.

Le code LBLRTM [4], développé par la société AER, permet d'estimer, en fonction d'un profil d'ozone et de données atmosphériques (cf. section 5.2), les radiances présentes au niveau de l'orbite satellitaire de MetOp (cf. section 5.3.2).

Le modèle d'instrument IASI est appliqué à ces radiances et ainsi est obtenue la mesure simulée du capteur IASI (cf. figure 20).

Le programme SA-NN [16] (cf. section 5.3.4), développé par le Service d'Aéronomie (SA) de l'IPSL, transforme cette radiancie simulée en colonne d'ozone entre le sol et 6 km, grâce à un réseau de neurones (cf. figure 20).

5.2 Données d'entrée atmosphériques

Pour simuler les colonnes d'ozone, nous avons besoin de connaître un ensemble de propriétés de l'atmosphère : la température, la pression, la densité de l'air et les concentrations d'ozone en fonction de l'altitude.

5.2.1 Le profil d'ozone

Pour le profil d'ozone, le point de départ est la simulation de référence de Polyphemus (supposée parfaite). Elle donne un profil d'ozone entre 0 et 4,6 km pour chaque point du domaine, et pour toutes les heures (en particulier pour l'heure pour lequel nous souhaitons simuler la mesure IASI). Pour obtenir un profil entre 0 et 100 km, nous utilisons un profil « *a priori* » entre 5 et 100 km. Ce profil « *a priori* » provient de données d'atmosphère standard pour les latitudes de l'Europe (Sources : AER). Pour chaque point du domaine, nous obtenons ainsi un profil d'ozone 0-100 km en juxtaposant le profil 0-4,6 km de ce point (issu de Polyphemus) et le profil standard 5-100 km pour les altitudes données dans l'équation 6. La jonction est effectuée en reliant les deux parties de profil par un segment de droite (cf. figures 21 et 22).

$$\left. \begin{array}{l}
 alt[0] \rightarrow 0 \\
 alt[1] \rightarrow 0,025 \\
 alt[2] \rightarrow 0,125 \\
 alt[3] \rightarrow 0,45 \\
 alt[4] \rightarrow 0,8 \\
 alt[5] \rightarrow 1,25 \\
 alt[6] \rightarrow 1,8 \\
 alt[7] \rightarrow 2,4 \\
 alt[8] \rightarrow 3,05 \\
 alt[9] \rightarrow 3,8 \\
 alt[10] \rightarrow 4,6 \\
 \vdots \\
 alt[31] \rightarrow 25 \\
 alt[32] \rightarrow 27,5 \\
 alt[33] \rightarrow 30 \\
 \vdots \\
 alt[41] \rightarrow 50 \\
 alt[42] \rightarrow 55 \\
 alt[43] \rightarrow 60 \\
 \vdots \\
 alt[51] \rightarrow 100
 \end{array} \right\} \begin{array}{l}
 \text{Polyphemus} \\
 \\
 \text{Profil_atm6}
 \end{array} \quad (6)$$

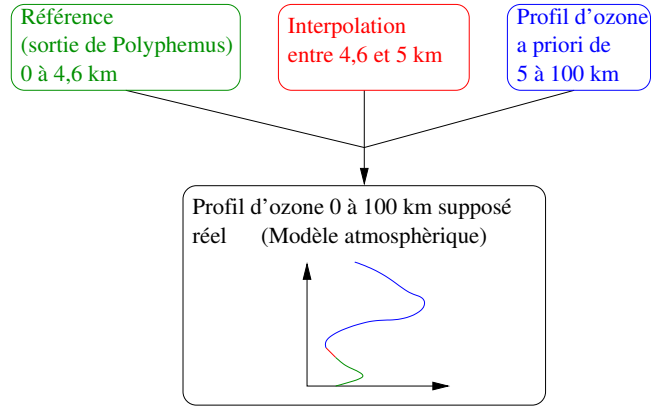
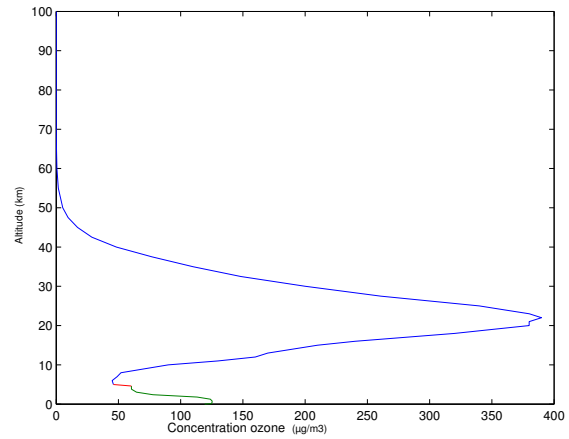


Figure 21 : Création du profil d'ozone en fonction de l'altitude.



Bleu : Profil moyen « *a priori* » 5-100 km ;
Rouge : Interpolation 4,6-5 km ;
Vert : Profil issu de la référence 0-4,6 km.

Figure 22 : Exemple de profil d'ozone supposé réel 0-100 km pour un point donné.

5.2.2 Les profils de température et de pression

De même, nous utilisons des profils standards pour la température et la pression, pour les altitudes comprises entre 5 et 100 km, provenant de la même source. Pour les altitudes comprises entre le sol et 4,6 km, nous utilisons les données provenant de l'ECMWF. Nous avons choisi ces données, car c'est avec elles qu'ont été calculées les simulations effectuées avec Polyphemus. Un exemple de profil est présenté figure 23 pour la température et figure 24 pour la pression.

5.2.3 Le profil de densité de l'air

Le profil de densité de l'air est, quant à lui, calculé à partir de la pression et de la température suivant la formule 7.

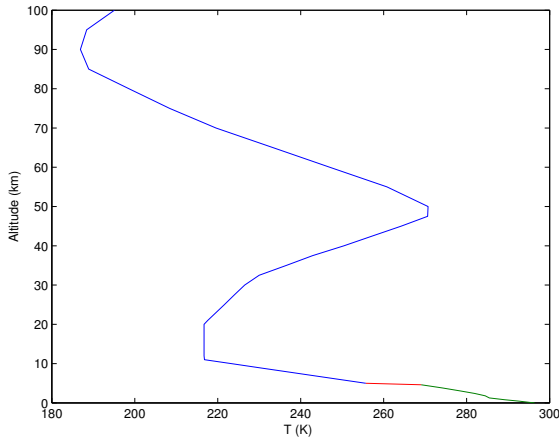
$$\rho = \frac{P}{r * T} \quad (7)$$

avec ρ la densité de l'air ;

P la pression ;

T la température ;

$r = 8.314472$: la constante des gaz parfaits.



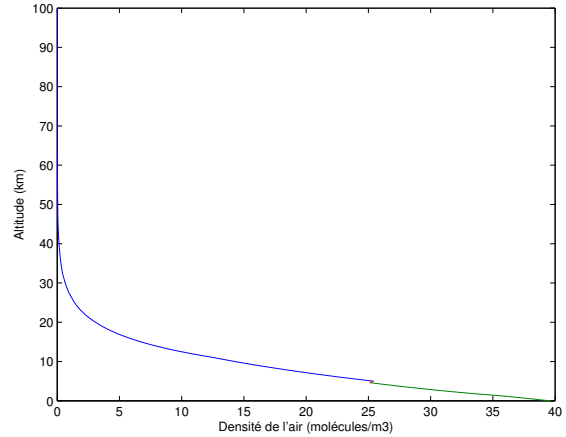
Bleu : Profil moyen « *a priori* » 5-100 km ;

Rouge : Interpolation 4,6-5 km ;

Vert : Profil issu d'ECMWF 0-4,6 km.

Figure 23 : Exemple de profil des températures en fonction de l'altitude.

Nous obtenons ainsi un profil tel que celui présenté figure 25.



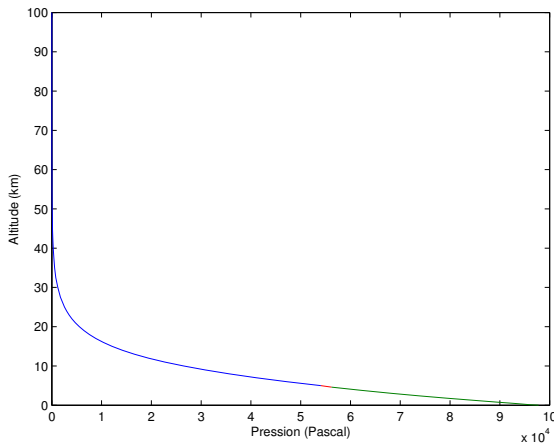
Bleu : Profil moyen « *a priori* » 5-100 km ;

Rouge : Interpolation 4,6-5 km ;

Vert : Profil issu d'ECMWF 0-4,6 km.

Figure 25 : Exemple de profil de densité de l'air en fonction de l'altitude.

Toutes ces données atmosphériques sont enregistrées dans un fichier nommé *TAPE5* qui est une entrée de LBLRTM.



Bleu : Profil moyen « *a priori* » 5-100 km ;

Rouge : Interpolation 4,6-5 km ;

Vert : Profil issu d'ECMWF 0-4,6 km.

Figure 24 : Exemple de profil de pression en fonction de l'altitude.

5.3 Les éléments de la simulation des colonnes d'ozone

5.3.1 Rôle de LNFL et de HITRAN

Le capteur IASI mesure la radiance en sortie d'atmosphère pour un grand nombre de longueurs d'onde dans l'infrarouge, notamment pour les bandes d'absorption de l'ozone.

La base de données HITRAN [14] contient l'ensemble des bandes d'absorption pour un certain nombre de molécules dont l'ozone. Nous utilisons le programme LNFL qui, à partir de cette base de données, crée un fichier de données spectrales *TAPE3* contenant les bandes d'absorption de l'ozone pour le domaine spectral de l'instrument IASI.

5.3.2 Rôle du programme LBLRTM

Le programme LBLRTM calcule la radiance théorique au niveau du satellite en fonction d'un modèle d'atmosphère

(le fichier *TAPE5* de la section 5.2) et de données spectrales (le fichier *TAPE3* produit par LNFL).

5.3.3 Rôle du modèle d'instrument IASI

La sortie de LBLRTM est un spectre de radiances présenté par la figure 26.

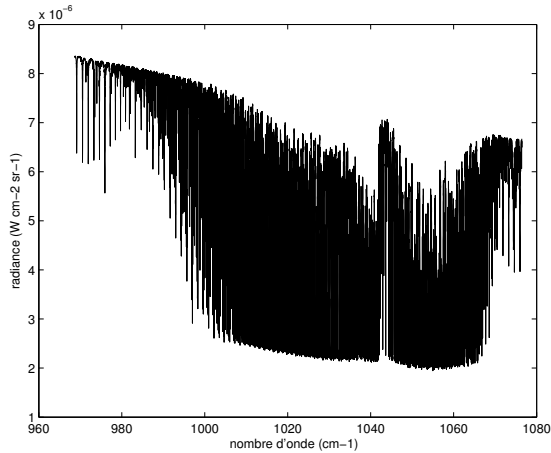


Figure 26 : Radiance simulée.

Le modèle d'instrument IASI représente la fonction de transfert de l'instrument IASI. Une convolution est appliquée à ce spectre, par la fonction d'instrument d'IASI, afin d'obtenir la mesure d'IASI simulée, c'est-à-dire le spectre de radiances présenté sur la figure 27.

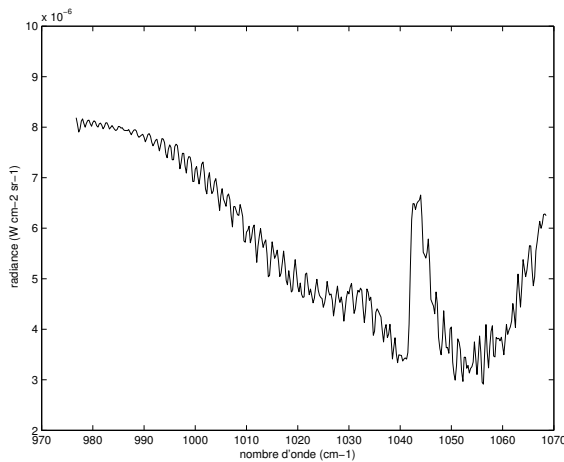


Figure 27 : Mesure IASI simulée.

5.3.4 Rôle de SA-NN

SA-NN est un réseau de neurones, avec deux couches cachées, qui permet d'estimer la valeur de la colonne d'ozone 0-6 km.

Le réseau de neurones (cf. figure 28) prend en entrée les radiances obtenues grâce à LBLRTM, notées y , et les tempé-

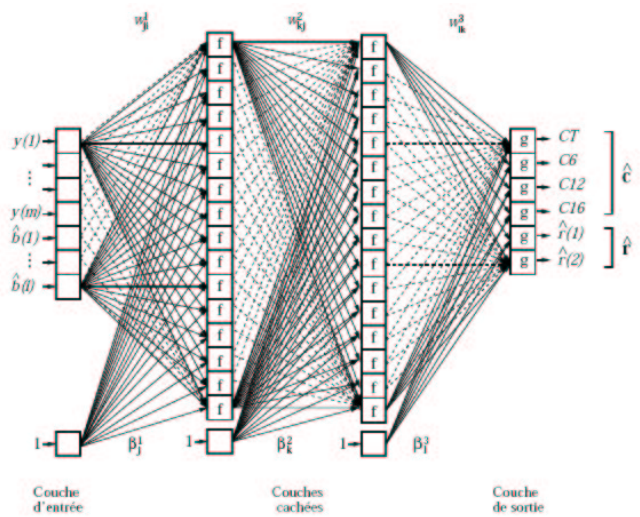


Figure 28 : Architecture du réseau de neurones.

ratures associées à des niveaux de pression fixés, notées \hat{b} . En sortie, le réseau de neurones retourne d'une part quatre colonnes d'ozone, et d'autre part deux radiances $\hat{r}(1)$ et $\hat{r}(2)$. CT , $C6$, $C12$, $C16$, représentent respectivement la colonne d'ozone totale (jusqu'en haut de l'atmosphère), la colonne 0-6 km (celle qui nous intéresse), la colonne 0-12 km et la colonne 0-16 km. Deux radiances témoins ne sont pas données en entrée du réseau de neurones ; le réseau cherche à les retrouver. SA-NN compare donc les résultats obtenus en sortie, $\hat{r}(1)$ et $\hat{r}(2)$, avec les radiances témoins, afin de tester si les sorties du réseau de neurones ne sont pas aberrantes (par exemple si l'atmosphère, sur laquelle nous travaillons, est trop éloignée des exemples de la base d'apprentissage du réseau de neurones).

5.4 Mise en place de la chaîne de simulation

Nous avons créé un programme **Fortran** qui, à partir de l'atmosphère « *a priori* » (*profils_atm6*), des données de l'ECMWF et de la simulation de référence issue de Polyphemus, génère le fichier *TAPE5* contenant les données de l'atmosphère à simuler ainsi qu'un fichier contenant le profil d'ozone *prof_test* (cf. figure 29).

Le logiciel LBLRTM génère, à partir des fichiers *TAPE5* (représentant l'atmosphère) et *TAPE3* (représentant les données spectrales de l'ozone), un fichier *TAPE29* qui contient la radiance estimée à la sortie de l'atmosphère pour le domaine spectral d'IASI (cf. figure 29).

Nous avons écrit un autre programme **Fortran** permettant d'appliquer, par une convolution, le modèle de l'instrument IASI, de lancer le réseau de neurones SA-NN et de calculer l'erreur commise en comparant la colonne obtenue par rapport aux données contenues dans *prof_test* (cf. figure 29).

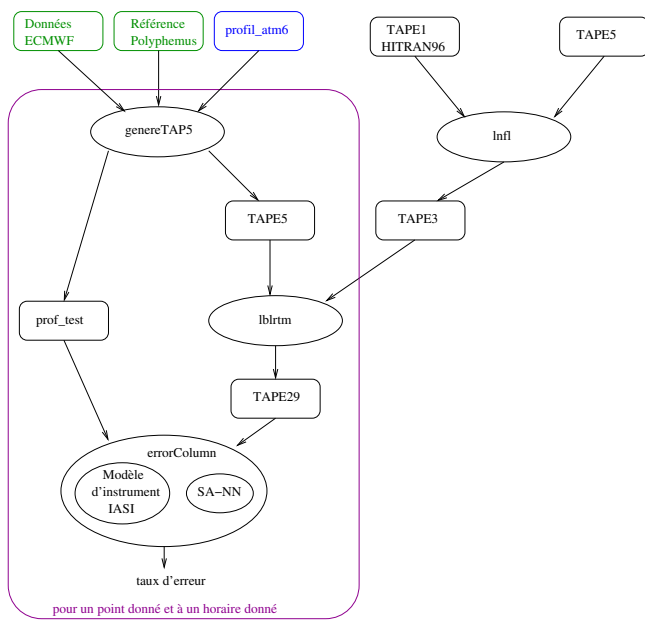


Figure 29 : Chaîne de simulation de la mesure IASI.

5.5 Résultats

Nous obtenons ainsi, pour un point donné et à un horaire donné, la simulation de l'erreur commise par le capteur IASI. Un script `bash` permet de répéter l'ensemble de ces opérations de manière à obtenir une carte des erreurs commises en chaque point, pour un horaire donné (cf. figure 30). Au final, nous simulons, de cette manière, l'ensemble des colonnes d'ozone pour toute la zone d'étude, chaque jour à 9h.

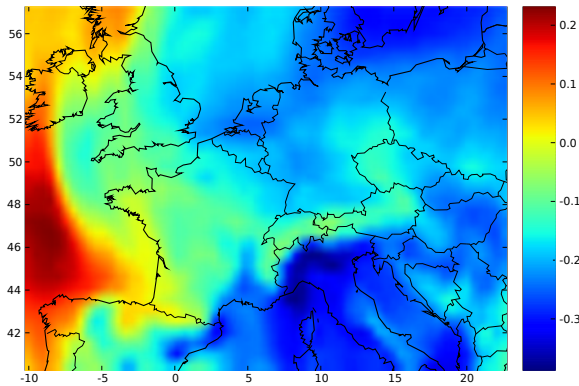
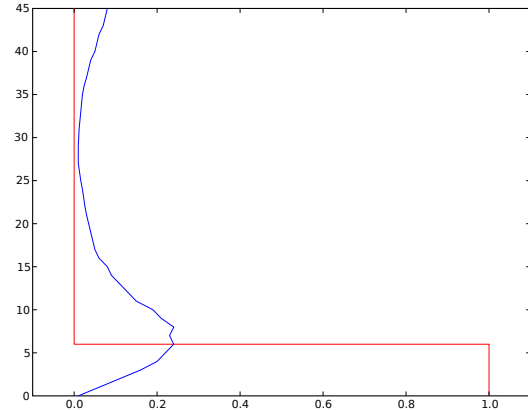


Figure 30 : Exemple de carte d'erreurs commises par le capteur IASI (simulation).

Pour calculer, sur l'ensemble de la période d'étude, la valeur de l'erreur simulée du capteur IASI, il a fallu 10 jours et 19 heures de calculs (sur une machine Pentium 4 cadencée à 3,2 GHz). L'élément le plus onéreux en temps de calcul étant LBLRTM.

En moyenne, la valeur absolue de l'erreur commise est de l'ordre de 27 %, pour la période d'étude. Toujours sur le mois de juillet 2001, l'erreur la plus importante est 53 %.

Pour améliorer ces résultats, il faudrait prendre en compte le noyau caractérisant la sensibilité de la colonne 0-6 km à chaque niveau d'altitude, obtenu par l'algorithme SA-NN. La figure 31 représente ce noyau en bleu et en rouge le noyau idéal.



Bleu : Noyau de SA-NN ;
Rouge : Noyau idéal.

Figure 31 : Noyaux caractérisant la sensibilité de la colonne 0-6km à chaque niveau d'altitude.

Nous calculons une colonne d'ozone selon la formule 2, c'est-à-dire que chaque niveau compris entre le sol et le haut de la colonne est pris en compte à 100 % alors que ceux qui sont au-dessus ne sont pas du tout pris en compte. Nous calculons donc notre colonne d'ozone suivant le noyau idéal, or SA-NN n'est pas sensible uniquement aux altitudes comprises entre le sol et le haut de la colonne. Pour palier ce problème, il faudrait calculer la colonne d'ozone ne se servant du noyau obtenu par SA-NN.

6. ASSIMILATION DE COLONNES SIMULÉES

6.1 Introduction

Dans la section 4.4, nous avons testé l'assimilation de pseudo colonnes dans un modèle erroné. Grâce à la section 5, nous connaissons l'erreur commise en chaque point par le capteur satellitaire IASI. Nous calculons les colonnes d'ozone simulées grâce à l'équation 8 :

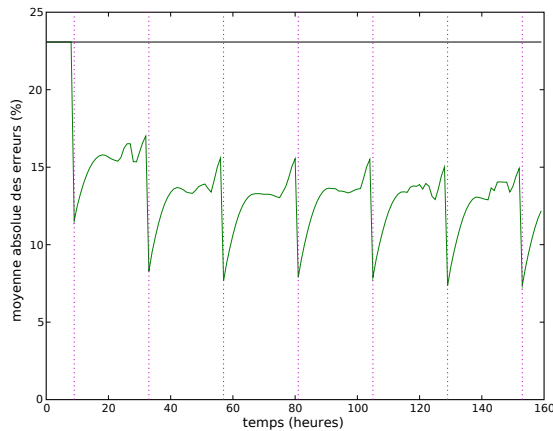
$$\begin{aligned}
 & ColonneSimulée_{t,i,j} \\
 &= ColonneParfaite_{t,i,j} * (1 + Erreur_{t,i,j})
 \end{aligned} \tag{8}$$

ColonneParfaite correspond à la colonne calculée à partir de la référence (correspondant au bon modèle); elle est calculée suivant la formule de l'équation 2. $Erreur_{t,i,j}$ est l'erreur calculée par la simulation de la section 5, pour le point (i, j) à l'instant t .

Nous n'utilisons pas directement la colonne produite par SA-NN, mais l'erreur commise par celle-ci, car Polyphemus ne permet pas d'effectuer des prévisions au-delà de 5 km. Pour poursuivre ce travail, il faudrait étudier la relation entre les colonnes 0-6 km produites par SA-NN et les colonnes 0-5 km assimilables par Polyphemus.

6.2 Assimilation de colonnes simulées dans un modèle erroné

L'expérience de la section 4.4 va être refaite avec les colonnes issues de la simulation. Nous avons précédemment des colonnes erronées de l'ordre de 20 %, or la simulation du capteur IASI, a montré que l'erreur sur les colonnes est de l'ordre de 23 %. Les résultats de l'expérience sont présentés sur la figure 32.



Noir : Sans assimilation ;
Vert : Avec assimilation ;
Magenta pointillé : Horaire d'assimilation.

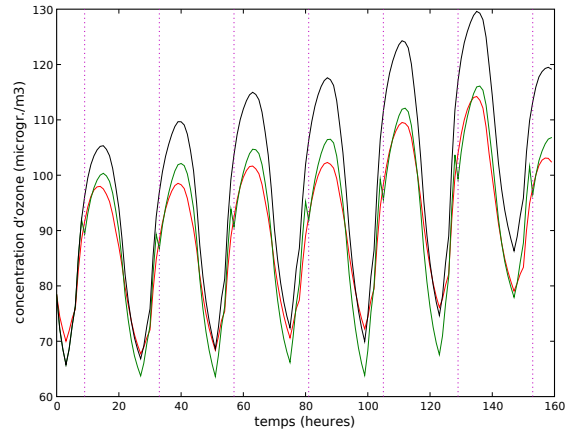
Figure 32 : Moyenne de la valeur absolue de la différence relative entre la simulation et la référence.

Elle représente deux courbes : les erreurs commises par une simulation sans assimilation (erreurs constantes = 23 %, en noir) et les erreurs d'une simulation avec assimilation des colonnes (en vert). En assimilant nos colonnes simulées, la simulation commet une erreur comprise entre 8 et 16 %. Ainsi, l'assimilation des colonnes simulées d'IASI améliore la prévision d'au moins 30 %, pour un modèle erroné à hauteur de 23 %.

6.3 Assimilation de colonnes simulées dans un modèle au données d'entrée perturbées

Dans un soucis de se rapprocher le plus possible de la réalité, nous avons testé l'assimilation des colonnes simulées d'IASI dans un modèle où certaines données d'entrée sont perturbées. Ces perturbations ont été choisies avec soin pour être les plus réalistes possible. Dans le modèle perturbé, les émissions de dioxyde d'azote sont sur-estimées de 30 % ; les dépôts d'ozone sont sous-estimés de 15 % ; et les entrées et sorties du domaine de l'ozone sont sur-estimées de 15 %.

Nous avons effectué deux simulations, l'une avec et l'autre sans assimilation des colonnes d'ozone simulées d'IASI. Les résultats sont présentés sur les figures 33 et 34.



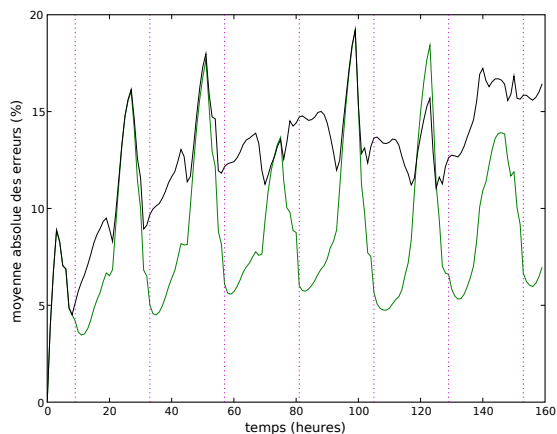
Rouge : La prévision issue du bon modèle ;
Vert : Simulation avec assimilation ;
Noir : La prévision issue du modèle erroné ;
Magenta pointillé : Horaire d'assimilation.

Figure 33 : Prévision des concentrations d'ozone au niveau du sol.

L'assimilation des colonnes simulées d'IASI permet d'obtenir la plupart du temps de meilleurs résultats (En effet, pour $t = 120$ la prévision sans assimilation de données est meilleure). La prévision avec assimilation réalise une erreur moyenne de 8,75 %, alors que sans assimilation, l'erreur moyenne est de 12,78 %.

La figure 35 représente la carte montrant la différence relative des erreurs commises par la simulation avec assimilation par rapport à celle effectuée sans assimilation, à $t = 30$ c'est-à-dire 21 heures après la première assimilation.

Cette expérience expose, sur un exemple de modèle aux données incorrectes, la capacité de correction de l'assimilation des colonnes d'ozone issues des données acquises par IASI.



Noir : Sans assimilation ;
 Vert : Avec assimilation ;
 Magenta pointillé : Horaire d'assimilation.

Figure 34 : Moyenne de la valeur absolue de la différence relative entre la simulation et la référence.

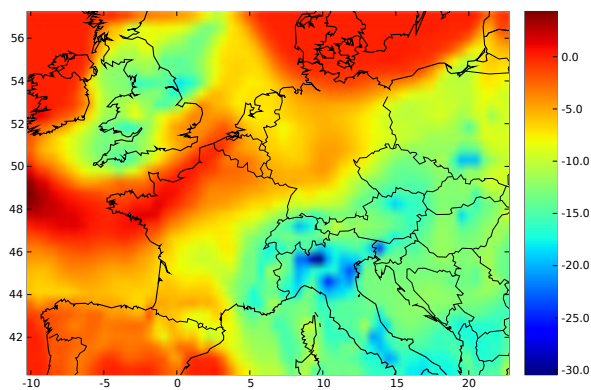


Figure 35 : Erreur relative moyenne entre la simulation et la référence à $t = 30$.

7. CONCLUSION GÉNÉRALE

L'objectif de cette étude était de montrer l'intérêt de l'assimilation de colonnes d'ozone simulées du capteur IASI en vue de préparer l'assimilation des données réelles, lorsque celles-ci seront disponibles. À travers ce document, nous avons exposé les études préliminaires montrant l'intérêt théorique de cette assimilation. Puis, nous avons expliqué la mise en place de l'assimilation des colonnes dans la plate-forme Polyphemus. En suite, le document relate des tests d'assimilation de colonnes calculées à partir de la référence. Il présente la mise en place de la simulation du spectre IASI et le calcul de la colonne d'ozone associée. Enfin, dans ce rapport, nous présentons les tests d'assimilation de colonnes d'ozone issues de la simulation de la mesure du capteur IASI. Les travaux réalisés pour cette étude ont montré l'intérêt de l'assimilation de colonnes d'ozone issues du capteur satellitaire IASI.

8. PERSPECTIVES

Pour poursuivre ce travail, il faut tester d'autres méthodes d'assimilation et notamment des techniques variationnelles comme le 4D-var. L'enjeu est de définir le contrôle : quelles sont les données d'entrée et les paramètres physiques du modèle qu'il est possible de corriger ? Quelle est la sensibilité de l'assimilation des données, dans le modèle, aux observations ?

Il serait également intéressant de travailler sur d'autres capteurs satellites tels que MOPITT, OMI, GOME, GOME2, SCIAMACHY et de travailler sur des espèces chimiques différentes. Un autre objectif est la définition de l'opérateur d'observation qui, pour le cas des aérosols, est non-linéaire. Les objectifs pour la mise en œuvre pratique consisteront en l'optimisation du temps de calcul et, éventuellement, à sa parallélisation. Enfin, un dernier objectif sera de chercher à valider ces résultats grâce à des données terrain.

9. REMERCIEMENTS

Ce travail a été rendu possible grâce à certaines personnes que je tiens à remercier chaleureusement. Tout d'abord, je souhaite remercier madame la directrice de l'unité INRIA Rocquencourt, S. Cluet, pour m'avoir ouvert les portes de son institut, ainsi que son successeur A. Petit. Ensuite, j'aimerais remercier chaleureusement I. Herlin, responsable de l'équipe Clime, autant pour le bon accueil qu'elle m'a réservé au sein de son équipe que pour ses conseils avisés. Je tiens à remercier J.-P. Berroir, mon encadrant de stage, pour le temps qu'il m'a consacré, ses bons conseils et l'aide qu'il m'a apporté lors des difficultés rencontrées. Les conseils avisés de I. Herlin et de J.-P. Berroir m'ont permis d'améliorer la qualité de ce manuscrit, je les en remercie. Je n'oublie pas non plus l'aide de V. Mallet, créateur de Polyphemus, pour ses explications, sa patience et l'aide apportée. Un grand merci à S. Turquety et à C. Clerbaux pour avoir mis à la disposition de notre équipe une partie de leurs travaux sur IASI et d'avoir pris le temps de m'aider à utiliser les logiciels permettant la simulation des colonnes d'IASI. Je remercie enfin tout l'équipe Clime et mon entourage pour leur présence et leur soutien ainsi que pour leurs conseils sur ce rapport.

10. REFERENCES

- [1] F. Bouttier and P. Courtier. Data assimilation concepts and methods. *Meteorological Training Course Lecture Series*, 1999.
- [2] F. Cayla and P. Javelle. IASI instrument overview. *Proc SPIE, Int. Soc. Opt. Eng.*, 2583 :271–281, 1995.
- [3] G. Chalon, F. Cayla, and D. Diebel. IASI : An Advanced Sounder for Operational Meteorology. *Proceedings of the 52nd Congress of IAF*, 2001.
- [4] S. Clough and J. Iacono. Line-by-line calculations of atmospheric fluxes and cooling rates ii : Application to carbon dioxide, ozone, methane, nitrous oxide, and the halocarbons. *J. Geophys. Res.*, 100 :16,519–16,535, 1995.
- [5] R. Daley. Atmospheric Data Analysis. *Cambridge Atmospheric and Space Science Series*, Cambridge University Press, 1991.
- [6] H. Eskes, A. Piters, F. Levelt, M. Allaart, and H. Kelder. Variational assimilation of total-column ozone satellite data in a 2d lat-lon tracer-transport model. *Journal of the Atmospheric Sciences*.
- [7] M. Ghil. Meteorological Data Assimilation for Oceanographers. Part I : Description and Theoretical Framework. *Dyn. of Atmos. Oceans*, 13 :171–218, 1989.
- [8] INRP. Inventaire national des rejets de polluants d’environnement canada. 2003.
- [9] A. Lorenc. Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, 112 :1177–1194, 1986.
- [10] V. Mallet, D. Quélo, and B. Sportisse. Software architecture of an ideal modeling platform in air quality – A first step : Polyphemus. Technical Report 11, CERE, 2005.
- [11] V. Mallet and B. Sportisse. 3-D chemistry-transport model Polair : numerical issues, validation and automatic-differentiation strategy. *Atmos. Chem. Phys. Discuss.*, 4 :1,371–1,392, 2004.
- [12] V. Mallet and B. Sportisse. Data processing and parameterizations in atmospheric chemistry and physics : the AtmoData library. Technical Report 12, CERE, 2005.
- [13] T. Phulpin, F. Cayla, G. Chalon, D. Diebel, and D. Schlüssel. Iasi onboard metop : Project status and scientific preparation. *Proc. 12th International TOVS Study Conference*, 2002.
- [14] L. Rothman, D. Jacquemart, A. Barbe, D. Chris Benner, M. Birk, L. Brown, M. Carleer, C. Chackerian, K. Chance, L. Coudert, V. Dana, V. Devi, J. Flaud, R. Gamache, and Goldman. The hitran 2004 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 96 :139–204, 2005.
- [15] B. Sportisse. Modélisation de la pollution atmosphérique. *Cours de l’Ecole Nationale des Ponts et Chaussées*, 2005.
- [16] S. Turquety, J. Hadji-Lazaro, C. Clerbaux, D. Hauglustaine, S. A. Clough, V. Cassé, and P. Schlüssel. Operational trace gas retrieval algorithm for the infrared atmospheric sounding interferometer. *J. Geophys. Res.*, 109 :D21301, 2004.

11. ANNEXE : DÉMONSTRATION DE L’INTERPOLATION OPTIMALE.

Nous allons démontrer la formule 5 à partir de la formule 4 :

objectif :

$$X_a = \text{Arg min } J \longrightarrow X_a = X_b + K(Y_o - H(X_b))$$

avec $K = B^T(HBH^T + R)^{-1}$

H peut être considéré comme une matrice de transfert. H , B et R sont des matrices définies positives.

$$X_a = \text{Arg min } J$$

Pour trouver le minimum de la fonction J , nous cherchons pour quelle valeur son gradient s’annule.

$$\nabla J(X_a) = 0$$

$$2B^{-1}(X_a - X_b) - 2H^T R^{-1}(Y_o - H(X_a)) = 0$$

$$B^{-1}(X_a - X_b) - H^T R^{-1}(Y_o - H(X_b + X_a - X_b)) = 0$$

H est un opérateur linéaire.

$$B^{-1}(X_a - X_b) - H^T R^{-1}(Y_o - H(X_b) - H(X_a - X_b)) = 0$$

$$B^{-1}(X_a - X_b) - H^T R^{-1}(Y_o - H(X_b)) + H^T R^{-1}H(X_a - X_b) = 0$$

$$(B^{-1} + H^T R^{-1}H)(X_a - X_b) = H^T R^{-1}(Y_o - H(X_b))$$

$$(X_a - X_b) = (B^{-1} + H^T R^{-1}H)^{-1} H^T R^{-1}(Y_o - H(X_b))$$

$$X_a = X_b + \underbrace{(B^{-1} + H^T R^{-1}H)^{-1} H^T R^{-1}}_{\text{matrice de gain } \kappa} \underbrace{(Y_o - H(X_b))}_{\text{innovation}} \quad (9)$$

Modification de l’écriture de K :

objectif :

$$(B^{-1} + H^T R^{-1}H)^{-1} H^T R^{-1} \longrightarrow B^T(HBH^T + R)^{-1}$$

12. ANNEXE : SIGLES UTILISÉS.

$$\begin{aligned} H^T R^{-1}(HBH^T + R) &= H^T R^{-1}HBH^T + H^T R^{-1}R \\ &= H^T R^{-1}HBH^T + H^T \end{aligned} \quad (10)$$

$$\begin{aligned} (B^{-1} + H^T R^{-1}H)BH^T &= B^{-1}BH^T + H^T R^{-1}HBH^T \\ &= H^T R^{-1}HBH^T + H^T \end{aligned} \quad (11)$$

$$(10) \wedge (11) \implies H^T R^{-1}(HBH^T + R) = (B^{-1} + H^T R^{-1}H)BH^T$$

Nous multiplions à droite par $(HBH^T + R)^{-1}$

$$\begin{aligned} H^T R^{-1}(HBH^T + R)(HBH^T + R)^{-1} \\ = (B^{-1} + H^T R^{-1}H)BH^T(HBH^T + R)^{-1} \end{aligned}$$

$$H^T R^{-1} = (B^{-1} + H^T R^{-1}H)BH^T(HBH^T + R)^{-1}$$

Nous multiplions à gauche par $(B^{-1} + H^T R^{-1}H)^{-1}$

$$\begin{aligned} (B^{-1} + H^T R^{-1}H)^{-1}H^T R^{-1} &= \\ (B^{-1} + H^T R^{-1}H)^{-1}(B^{-1} + H^T R^{-1}H) \\ BH^T(HBH^T + R)^{-1} \end{aligned}$$

$$K = (B^{-1} + H^T R^{-1}H)^{-1}H^T R^{-1} = BH^T(HBH^T + R)^{-1} \quad (12)$$

Donc, d'après 9 et 12, nous pouvons conclure que :

$$\begin{aligned} X_a &= X_b + K \underbrace{(Y_o - H(X_b))}_{\text{Innovation}} \\ \text{avec } K &= \underbrace{B^T(HBH^T + R)^{-1}}_{\text{Matrice de gain}} \end{aligned}$$

AER : Atmospheric and Environmental Research inc.

CEREA : Centre d'Enseignement et de Recherche en Environnement Atmosphérique.

Clima : Couplage de la donnée environnementale et des modèles de simulation numérique pour une intégration logicielle.

ECMWF : European Centre for Medium-range Weather Forecasts ; en français : Centre Européen pour la Prévision Météorologique Moyen Terme.

EPS MetOp RAO : European Polar Satellite MetOp Research Announcement of Opportunity.

ESA : European Space Agency.

EUMETSAT : European organisation for the exploitation of METeorological SATellites.

GOME : Global Ozone Monitoring Experiment.

GOME2 : Global Ozone Monitoring Experiment 2.

HITRAN : High-resolution TRANsmission molecular absorption database.

IASI : Infrared Atmospheric Sounding Interferometer ; ou en français : Interféromètre Atmosphérique de Sondage Infrarouge.

INRIA : Institut National de Recherche en Informatique et Automatique.

IPSL : Institut Pierre-Simon Laplace.

LBLRTM : Line-By-Line Radiative Transfer Model.

MetOp : METeorological OPERational.

MOPITT : Measurements Of Pollution In The Troposphere.

OMI : Ozone Monitoring Instrument.

SA : Service d'Aéronomie.

SA-NN : Service d'Aéronomie - Neural Network.

SCIAMACHY : SCanning Imaging Absorption spectroMeter for Atmospheric CartographY.