Les états initiaux pour la prévision numérique en météorologie: incertitude des données et méthode de correction

Claude Fischer

▶ **To cite this version:**

Claude Fischer. Les états initiaux pour la prévision numérique en météorologie: incertitude des données et méthode de correction. Océan, Atmosphère. Université Paul Sabatier - Toulouse III, 2009. tel-00518946

HAL Id: tel-00518946
https://theses.hal.science/tel-00518946

Submitted on 17 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les états initiaux pour la prévision numérique en Météorologie : Incertitude des données et méthodes de correction

Claude Fischer
Météo-France/CNRM - CNRS/GAME,
Habilitation à diriger des recherches, Université P. Sabatier, Toulouse III
soutenue le 26 février 2009

2 mars 2009

Table des matières

1 Introduction **3**
- 1.1 Les principes de la prévision numérique (PN) . . . 3
- 1.2 Les modèles Arpège et Aladin . . . 5
- 1.3 Les erreurs dans les données . . . 6
- 1.4 Le positionnement de mes travaux personnels . . . 8

2 Un exemple algorithmique en guise de fil conducteur **10**
- 2.1 Le filtre de Kalman . . . 10
- 2.2 Un exemple "0-D" . . . 12
- 2.3 Le filtre de Kalman semi-géostrophique : synthèse de l'article "Error growth and Kalman filtering within an idealized barocl
- 2.4 Equivalence avec le 4D-VAR classique et le 4D-VAR à contrainte faible . . . 20

3 La croissance des erreurs de prévision et leur modélisation **24**
- 3.1 La croissance linéaire des erreurs initiales : synthèse de l'article "Linear amplification and error growth in the 2D Eady prob
- 3.2 Les erreurs du modèle . . . 29
- 3.3 La modélisation des covariances d'erreurs de prévision . . . 32

4 Application à un système réel **34**
- 4.1 L'assimilation 3D-VAR Aladin . . . 34
- 4.2 Les erreurs propres au modèle Aladin : synthèse de l'article "Mesoscale background error covariances : recent results obtain
- 4.3 Le réglage des variances d'erreurs de P^b : synthèse de l'article "A posteriori validation applied to the 3D-VAR Arpège and
- 4.4 La prise en considération du système coupleur . . . 42
- 4.5 La prise en compte de la dimension temporelle (vers un 4D-VAR) : synthèse de l'article "Sensitivity of high resolution forec
- 4.6 Les observations innovantes . . . 47

5 Perspectives et travaux prospectifs **49**
- 5.1 Les filtres adaptatifs, le réglage en ligne et le contenu en information . . . 50
- 5.2 L'assimilation 4D-VAR pour la très haute résolution ? . . . 52
- 5.3 La mise en équilibre des conditions initiales, en début de prévision . . . 56

Chapitre 1

Introduction

1.1 Les principes de la prévision numérique (PN)

Passons tout d'abord brièvement en revue les principes généraux de la prévision numérique du temps (PN). Il s'agit de déterminer, à un instant t ultérieur dans le temps, l'état de l'atmosphère telle qu'elle peut se décrire sur une grille aux noeuds régulièrement espacés (les intervalles entre les noeuds formant les mailles, dont la longueur - le pas - sera la résolution, horizontale ou verticale, de la grille considérée). Il s'agit donc d'intégrer dans le temps les lois qui régissent l'évolution temporelle de l'atmosphère. On procède pour cela à une discrétisation des champs dans l'espace et dans le temps (pour le schéma d'avance temporelle choisi). On conçoit aisément de ce qui précède qu'un ingrédient crucial pour la PN sera la condition initiale.

Reprenons les grands thèmes ou défis de la PN :

1. les *lois d'évolution de l'atmosphère* : il s'agit des équations d'un fluide sur une sphère en rotation (pour un modèle de PN global). On se base sur les équations de Navier-Stokes, qui gouvernent à l'échelle du continuum fluide ses champs caractéristiques, et à partir desquelles on peut déduire l'évolution de la quantité de mouvement (loi fondamentale de la dynamique), sa température interne (1er principe de la thermodynamique, équation de l'enthalpie), sa pression (équation d'état du gaz, généralement supposé parfait), sa densité (équation de continuité, conservation de la masse). On ajoute l'évolution de différentes espèces d'hydrométéores (eau liquide nuageuse, glace, etc ...). Différentes simplifications sont alors appliquées, qui vont mener à divers jeux d'équations dans lesquels des phénomènes particuliers auront été filtrés ou simplifiés. Par exemple, les équations semi-géostrophiques que je discute dans le §2.3 sont un jeu d'équations simplifiées et filtrées. Ces jeux d'équations sont adaptés aux phénomènes atmosphériques particuliers que l'on souhaite simuler, en mode recherche ou pour une application opérationnelle ("réaliste"). Une hypothèse particulièrement sensible pour les modèles à haute résolution, généralement à aire limitée, est celle de l'hydrostatisme. Il s'agit alors de négliger le terme d'accélération verticale dans l'équation de la vitesse verticale, ce qui revient à supposer que la force de gravité et le gradient vertical de pression s'équilibrent instantanément : $\partial p/\partial z = -\rho g$. A l'échelle des mouvements de grande échelle (phénomènes de taille supérieure à une dizaine de kilomètres : ondes planétaires, circulations globales méridiennes, dépressions et anticyclones, ...), cette approximation se justifie par une analyse en ordre de grandeur des différents termes dans l'équation du mouvement vertical. Cette hypothèse est communément admise pour tous les modèles de mailles supérieures ou égales à environ 7.5-10 km.

   Aux plus petites échelles (tailles inférieures à 10 km), les mouvements atmosphériques d'intérêt incluent les mouvements de la convection. Or, la convection génère des mouvements verticaux ascendants violents, et elle peut à ce titre être vue comme une déstabilisation de l'équilibre hydrostatique. Dans les mouvements convectifs, des travaux analytiques et numériques ont confirmé que la force de pression n'était pas en équilibre avec la force de gravité ou l'advection verticale. Un premier paramètre-clef est ici le rapport d'aspect géométrique du mouvement : l'écoulement peut devenir fortement non-hydrostatique lorsque les déplacements sur l'horizontale et la verticale deviennent de même ordre de grandeur (Yau,

1979, Dalu *etal.* , 2003). Par ailleurs, un deuxième paramètre-clef semble être le temps caractéristique sur lequel agit le forçage des vitesses verticales (dégagement de chaleur latente par les processus diabatiques de changement de phase de la vapeur d'eau, par exemple). On s'éloignerait plus significativement de l'hydrostatisme lorsque le forçage agit sur des échelles de temps de l'ordre de quelques minutes à 1/2-1 h, et on en resterait assez proche pour des forçages se développant sur quelques heures (Dalu *etal.* , 2003). Ainsi, pour des mailles inférieures à environ 7.5 km, on considère que la simulation explicite de nuages convectifs (de type Cumulonimbus) intenses ou l'existence de fortes pentes (relief très accidenté) peuvent engendrer des forçages aux effets non-hydrostatiques non négligeables.

2. notre représentation de l'atmosphère repose sur *une discrétisation* régulière des champs physiques, dans le temps et l'espace. Les grilles spatiales peuvent avoir des distributions de points de grille régulières, ou variables (exemples d'Arpège, des modèles globaux Canadien ou Allemand). De plus, il peut être très utile de combiner la représentation point de grille avec une représentation spectrale (plus généralement, sur une famille de fonctions aux propriétés algébriques particulières, les rendant attrayantes pour les problèmes de la PN, telles les fonctions harmoniques sphériques[1]). Ainsi, notre connaissance de l'atmosphère est partielle, dans le sens où nous ne disposons pas, dans les modèles, d'une connaissance *continue* des champs physiques. Cependant, la taille du vecteur d'état (c'est à dire l'état discrétisé du modèle) sera déjà poussée (de l'ordre de 10^7 aujourd'hui dans Arpège ou Aladin), ce qui entraîne des contraintes pratiques, notamment en terme de place mémoire et d'espace de stockage pour les algorithmes numériques. Par ailleurs, la discrétisation des données pose la question de la *représentativité* : une mesure physique, réelle, est la valeur en un point (ou groupe de points) d'un champ continu, alors que sa contre-partie modèle est un échantillon discret. Ainsi, des effets de très petite échelle peuvent rester totalement absents du modèle, et on va considérer que les mesures ne sont pas, quant à elles, complètement représentatives des champs discrets du modèle.
3. la paramétrisation des phénomènes de sous-maille : il s'agit, pour les phénomènes physiques non représentés dans la grille du modèle, de tout de même prendre en compte leurs effets sur l'écoulement simulé. Classiquement, les phénomènes de saturation/condensation/évaporation, convection, rayonnement, échanges turbulents et traînées orographiques sont paramétrés dans les modèles de résolution 7.5 km ou plus. Pour des modèles de résolution kilométrique, la convection devient explicite, mais une microphysique compliquée des processus nuageux et précipitants doit être intégrée. On cherchera à paramétrer ces effets *en moyenne dans une maille* de la grille discrète : *les champs modèle sont donc représentatifs (au mieux) de l'état moyen de l'atmosphère sur une maille de la grille*.
4. la PN est très concernée par les questions de stabilité et de précision numériques. Les équations de l'atmosphère, continues, sont discrétisées selon une variété de méthodes spatiales et temporelles. Ces méthodes induisent des propriétés, ou des contraintes, sur la précision (la manière dont la solution converge vers la solution continue, lorsque le pas spatial ou temporel tend vers zéro) et la stabilité (le caractère borné de la solution si le pas de temps augmente ou si la maille spatiale diminue). Notons aussi que le choix du schéma temporel (la discrétisation temporelle), et les éventuels problèmes de stabilité numérique associés, dépendent de la forme des équations retenues (notamment, hydrostatique ou non). De même, dans son principe, chaque méthode d'avance temporelle implique une condition de stabilité, c'est à dire l'existence d'un pas de temps limite au-delà duquel la solution soit explosera au cours des itérations, soit deviendra progressivement absurde (fortes oscillations, erreurs systématiques). Signalons que certains schémas, dits implicites, sont de fait inconditionnellement stables (mais posent d'autres problèmes de résolution numérique par ailleurs).
5. le principe de toute simulation numérique est d'intégrer dans le temps *un état initial du modèle*. Cette idée très générale confère à l'état initial un statut particulier : on conçoit intuitivement que la qualité de la prévision dépendra, au moins pendant un certain temps, de la qualité de cet état de départ (qualité au sens de la comparaison avec l'état "vrai" de l'atmosphère). Cet état initial peut être le produit d'une prévision antérieure (mais dans ce cas, il s'agit plutôt de prolonger une intégration plus ancienne, à partir d'un

[1] les fonctions propres discrètes de l'opérateur Laplacien

état initial plus ancien) ou peut provenir d'un autre système de PN (c'est le cas lorsque l'état initial d'un modèle à aire limitée est obtenu par interpolation de l'état d'un système global, et que la prévision à aire limitée sera une adaptation dynamique à plus haute résolution d'un système global). En toute rigueur, on pourrait aussi imaginer définir l'état initial à partir d'un filtrage spatial d'observations, surtout si on suppose posséder assez d'observations pour couvrir tout le domaine d'intégration (une couverture correcte de mesures de vent, de température, de pression de surface, voire d'humidité). Ce qui nous intéressera dans ce rapport, ce sera l'obtention d'un état initial en combinant un état du modèle avec des observations. Je développerai davantage ce point au §1.3.

Pour conclure, notons que les produits de la PN sont aujourd'hui très variés : champs 3D de l'état de l'atmosphère (vent, température, pression, géopotentiel, humidité relative ...), variables de l'état sensible (température à 2 m, vent à 10 m, flux radiatifs, précipitations, visibilité ...), produits diagnostiques variés (critères de déclenchement de la convection, tourbillon potentiel et autres diagnostics de l'activité cyclonique ...). Les sorties de modèle sont utilisées pour la prévision jusqu'à 5 jours, voire plus en mode ensembliste, pour la prévision immédiate, la comparaison avec ou la reconstitution de données climatologiques, les études climatiques. Dans les dix dernières années, l'aspect déterministe de la PN s'est trouvé augmenté d'un aspect probabiliste, avec l'émergence des prévisions d'ensembles, qui ajoutent leur propre lot de produits (mesures de l'incertitude à partir de la variabilité constatée sur un échantillon, probabilités d'occurence ...). Je reviendrai, de manière beaucoup plus formelle, sur certains de ces aspects dans les prochains chapitres de ce manuscrit.

1.2 Les modèles Arpège et Aladin

Aladin[2] est la version à aire limitée du modèle global Arpège[3] : modèle français de prévision du temps en géométrie étirée (Courtier *etal.* , 1991). Arpège utilise en mode opérationnel un jeu d'équations hydrostatiques, avec une discrétisation mixte : point de grille (sur la sphère) et spectral avec une décomposition en fonctions harmoniques sphériques. Arpège dispose d'un jeu complet de paramétrisations physiques supposées valables jusqu'à des résolutions d'environ 7-8 km : ondes de gravité et traînées orographiques, turbulence verticale, pluies dites stratiformes (résolues à l'échelle de la maille du modèle), pluies dites convectives (sous-maille), microphysique simple des nuages et rayonnement. Le système Arpège a beaucoup bénéficié de la collaboration avec le CEPMMT[4], dans le cadre de la collaboration Arpège/IFS[5].

Aladin, quant à lui, a été développé dans le cadre d'une collaboration internationale entre Météo-France et des Services météorologiques partenaires, dans le but de mieux suivre le comportement des phénomènes de petite échelle tels que : le soulèvement orographique, la convection, la brise de mer ou de forêt etc .... Ce projet a été mené en collaboration avec les pays suivants : *l'Autriche, la Bulgarie, la Slovaquie, la République Tchèque, la Hongrie, la Pologne, la Roumanie et le Maroc*. À ces pays se sont joints *l'Algérie, la Belgique, la Croatie, la Slovénie, le Portugal, la Moldavie (retirée depuis), la Tunisie et la Turquie*. Actuellement, Aladin regroupe 16 pays Européens et d'Afrique du Nord.

Du point de vue technique, la réalisation du projet Aladin a respecté la compatibilité logicielle totale avec le système Arpège, avec un minimum de changement dans le code déjà existant (réutilisation d'environ 80% du code Fortran d'Arpège).

Le modèle Aladin (Bubnovà *etal.* , 1993) est basé sur des équations hydrostatiques utilisant une décomposition horizontale spectrale en série bi-Fourier et une coordonnée hybride sur la verticale (coordonnée de type pression P en altitude et coordonnée $\sigma = \frac{P}{P_{sol}}$ dans les basses couches). En effet, le domaine horizontal d'Aladin, qui n'est autre que la projection du domaine géographique (une portion de sphère) sur un plan cartésien, peut être décomposé en trois régions distinctes (cf figure 1.1), à savoir :

[2] Aire Limitée Adaptation Dynamique et développement INternational

[3] Action de Recherche Petite Echelle Grande Echelle

[4] Centre Européen de Prévisions Météorologiques à Moyen Terme

[5] Integrated Forecast System

– une zone interne *C* (*C*entrale), dite domaine d'intérêt météorologique : elle correspond à la zone où le modèle calcule sa solution uniquement à partir de ses forçages locaux.
– une zone externe *E*, dite d'*E*xtension : dans cette zone, les champs météorologiques qui ne sont pas forcément des champs périodiques horizontalement, sont extrapolés artificiellement en utilisant des fonctions splines cubiques, puis un lissage transversal Laplacien est appliqué. L'objectif est d'assurer la périodisation dans les deux directions zonale et méridienne. Grâce à cette idée dite de *bi-périodisation*, la technique spectrale (Orszag, 1970) en séries bi-Fourier peut être appliquée dans Aladin, bien que son domaine soit limité.
– une zone *I*ntermédiaire *I*, dite zone de relaxation ou zone de couplage : elle se situe aux frontières latérales autour du domaine d'intérêt météorologique *C*. Dans cette zone, le couplage des champs propres à Aladin avec ceux d'Arpège est effectué d'une façon lisse, d'une part pour tenir compte des effets de la circulation à grande échelle, et d'autre part pour éviter la réflexion des ondes dans cette zone de couplage. Cette opération est basée sur le principe de la relaxation donné par la formule de Davies (Davies, 1976) :

$$U_{apres-couplage} = (1-\alpha)U_{avant-couplage} + \alpha U_{frontiere} \tag{1.1}$$

où U est la valeur d'un champ météorologique donné produit par Aladin (prévu ou analysé), et α correspond au coefficient de relaxation. α est nul dans la zone *C*, il vaut 1 dans la zone *E* et varie entre 0 et 1 dans la zone *I* aux frontières. On note que l'opération du couplage est introduite comme dernière étape à chaque pas de temps. Une conséquence de cette approche est qu'une partie des erreurs de prévision Aladin provient des erreurs de prévision Arpège (cf §4.2, et dans une moindre mesure §4.4).

La grille de calcul, quant à elle, est régulière avec une résolution fine. Le nombre de points de grille $NXL\ (resp. NYL)$ est relié à la troncature $M\ (N)$, dans la direction horizontale zonale (méridienne), par la relation : $NXL = 2(M+1)\ (NYL = 2(N+1))$. On rappelle que les champs produits par Aladin sont météorologiquement significatifs uniquement dans la zone (*C+I*), puisque dans la zone d'*E*xtension, les valeurs résultent d'opérations purement mathématiques.

L'état initial dans Aladin peut provenir d'une simple interpolation des champs d'analyse Arpège sur la grille à haute résolution d'Aladin. Cette technique est appelée *Adaptation Dynamique*. Dans ce cas, le modèle va prendre du temps pour développer sa propre dynamique. En mode d'assimilation de données, l'utilisation d'une ébauche permet de toujours repartir d'un état proche d'une prévision antérieure d'Aladin, ce qui modifie considérablement les propriétés au démarrage d'une nouvelle prévision (cf §5.3).

Par ailleurs, une étape de mise en équilibre des champs de l'analyse Aladin est effectuée afin d'éliminer les ondes rapides générées sinon par les ajustements entre les champs initiaux (vent avec pression et température, température et humidité, température sur la verticale, ... voir aussi §5.3). Cette étape se fait par les *Filtres Digitaux* (Lynch et Huang (1992), Lynch *etal.* (1997)). La physique utilisée dans Aladin est la même que celle d'Arpège, avec quelques modifications permettant de prendre en compte la haute résolution.

1.3 Les erreurs dans les données

Il est essentiel de bien comprendre les sources d'erreurs qui viennent entacher la qualité d'une prévision. Dans la section 1.1, j'ai abordé les grandes notions de la prévision numérique. L'importance de la spécification de l'état initial est apparue alors clairement. Comme je l'ai suggéré dans le §1.1, l'état initial d'un système de PN sera généralement obtenu comme une combinaison entre un état du modèle (issu d'une prévision antérieure) et des observations récentes. Quelles sont alors les sources d'erreurs affectant la constitution d'un état initial ?

– les erreurs affectant le modèle : erreurs de discrétisation liées à la précision du modèle, mauvaises relations d'ondes, ondes numériques parasites, imprécisions ou insuffisances des paramétrisations physiques, phénomènes physiques non pris en compte.
– dans un modèle à aire limitée, les erreurs dues à la formulation des conditions limites latérales : mauvaise propagation des ondes entrantes ou sortantes, génération d'ondes parasites.

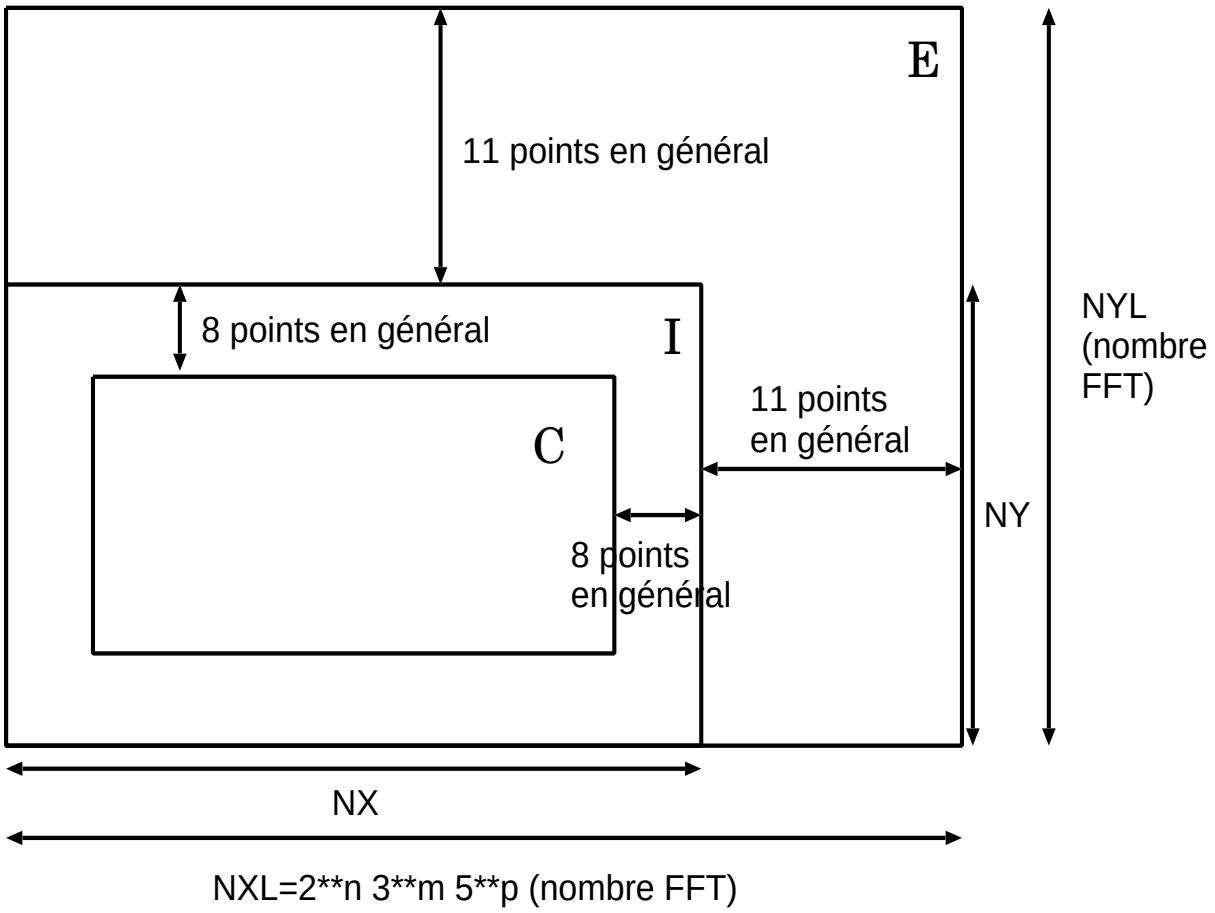


FIG. 1.1 – Géométrie des zones d'intérêt météorologique (C), d'extension (E) et de relaxation aux frontières (I), dans le modèle Aladin.

– dans un système de PN couplé, une partie de l'erreur de la prévision couplée doit provenir du système coupleur. Cependant, dans la mesure où le système coupleur possède un domaine plus vaste que le modèle couplé, il peut "mieux voir" des phénomènes de très grande échelle, ou mieux appréhender des phénomènes (des observations) situés juste en dehors des bords du plus petit domaine, et qui peuvent être utiles pour la prévision ultérieure du système couplé. De plus, il se pose la question de la cohérence des états initiaux et des prévisions qui en découlent, entre les systèmes coupleur et couplé.
– les erreurs de mesure affectant les observations, liées essentiellement aux instruments, voire aux manipulations humaines en aval de la mesure.
– les erreurs de représentativité : les observations ne sont pas représentatives des champs discrétisés du modèle, mais de valeurs ponctuelles des champs continus de l'atmosphère réelle. Les deux ne peuvent donc correspondre, et ce d'autant moins (a priori) que la maille du modèle est grande (la résolution spatiale faible). Formellement, cette différence apparaît sous la forme d'une erreur supplémentaire que l'on ajoute à l'erreur de mesure, et elle dépend donc de la résolution de la grille d'analyse. L'erreur de représentativité contient par ailleurs aussi l'erreur commise lorsqu'une dépendance compliquée existe entre la quantité mesurée et la variable analysée (voir par exemple Lorenc, 1986), et que cette relation ne peut être qu'approchée dans l'opérateur d'observations qui calcule l'équivalent observé à partir des champs modèle (cf §2.1).

Dans la pratique PN, le processus de combinaison entre une prévision antérieure (l'ébauche) et une analyse des données observées, au cours du temps, est appelé "assimilation de données". Le principe consiste à représenter les erreurs de l'état prévu et des observations, et de trouver l'état optimal dans un sens statistique. Il est utile de rappeler que les méthodes d'assimilation classiquement supposent que *les erreurs de biais sont déjà enlevées* (les erreurs modèle et observationnelles sont supposées non biaisées), et la modélisation des erreurs se concentre donc sur les moments d'ordre 2 (les covariances). Le nouvel état initial recherché, l'analyse, correspondra à l'état *de variance d'erreur minimale*, étant données les covariances d'erreurs d'ébauche et d'observations. Ce processus est répété, à l'infini. Il s'agit donc d'un cycle d'assimilation. Dans la section 2.1, je reviens de manière formelle sur les principes de l'assimilation, en mode séquentiel, en introduisant le filtre de Kalman. Je reviens sur certains mécanismes de croissance des erreurs dans les sections 3.1 (erreur initiale) et 3.2 (erreur modèle). Dans le §4.2, je reviens sur la croissance spécifique d'erreurs d'ébauche dans un modèle couplé, par rapport au système coupleur, avec l'exemple d'Aladin par rapport à Arpège. J'adresserai plus particulièrement la question de la prise en compte d'un système coupleur dans l'analyse du système couplé dans la section 4.4.

1.4 Le positionnement de mes travaux personnels

J'ai donc choisi d'organiser la synthèse de mes travaux de recherche autour de la question de la production du "meilleur" état initial possible pour démarrer une prévision numérique de l'évolution de l'atmosphère. En effet, outre l'enjeu scientifique de cette question, que j'ai souligné dans les deux sections précédentes §1.1 et §1.3, il s'agit aussi d'un dénominateur commun possible pour mes activités. Dans le chapitre 2, je reviens ainsi sur certains de mes travaux effectués entre 1995 et 1997 dans l'équipe RECYF du CNRM, sous l'instigation d'Alain Joly et avec la participation de François Lalaurette. Il s'agissait alors d'évaluer, dans un modèle académique, l'impact possible d'observations ciblées, et de réaliser de façon idéalisée des scénarii de ciblage et d'assimilation (voir §2.3). Je fais précéder ce texte d'une présentation plus générale du filtre de Kalman, qui me permet d'introduire les principes et les notations mathématiques que je conserverai par la suite. Le filtre de Kalman est aussi un algorithme intéressant pour qui veut une illustration assez complète des problèmes d'assimilation des données (§2.1). Dans le paragraphe 2.2, je présente un petit exemple très idéalisé, que je reprendrai en guise d'illustration dans d'autres sections (§2.4, 3.2 et 5.1). Enfin, il m'a paru utile de revenir sur les équivalences, et les différences, entre l'assimilation séquentielle (filtrage) et l'analyse variationnelle (modèle adjoint), dans le paragraphe 2.4. En effet, j'aborde dans les chapitres 4 (analyse 3D-VAR Aladin) et 5 (perspectives) les techniques variationnelles et les modèles adjoints.

Dans le chapitre 3, j'aborde la question de la croissance des erreurs dans un système de prévision numé-

rique. Dans le paragraphe 3.1, je passe en revue la croissance linéaire de petites erreurs initiales, en revenant notamment sur l'exemple des vecteurs singuliers dans le problème 2D d'Eady. Bien que n'ayant que peu travaillé sur ce sujet, j'ai souhaité aussi mentionner la question des erreurs modèle (§3.2). Enfin, j'aborde aussi la modélisation des erreurs de prévision dans les analyses (§3.3), ce qui me permet d'introduire et positionner des travaux que j'ai encadrés dans le cadre du développement de l'analyse 3D-VAR Aladin, et que je présente dans le chapitre 4. Dans celui-ci, je passe en revue les travaux de développement et de validation de l'assimilation 3D-VAR dans le modèle Aladin (§4.1), puis des travaux de recherche plus spécifiques sur la croissance des erreurs dans Aladin (§4.2), le réglage des variances d'erreur de prévision (§4.3) et la prise en compte du système coupleur (plus précisément, de son analyse) dans un système couplé (§4.4). J'aborde l'analyse 4D-VAR indirectement, en revenant sur les études menées avec le modèle adjoint d'Aladin, dans des expériences de sensibilité aux conditions initiales (§4.5). Dans le paragraphe 4.6, je termine par un rapide point sur les observations innovantes et leur importance dans un système couplé (importance intrinsèque et au sens d'une plus grande décorrélation statistique entre Arpège et Aladin).

Le chapitre 5 est consacré aux perspectives de recherche et de développement. J'ai choisi de ne pas séparer le contexte général de l'évolution de l'assimilation, notamment dans un Centre National comme le CNRM de Météo-France, des mes "marottes" personnelles. Je fais donc autant un point général sur les challenges de l'assimilation à très haute résolution (échelle convective de quelques kilomètres), en me référant aux activités du projet Arome[6] auquel je participe à CNRM/GMAP[7], et un développement de points plus spécifiques :

- vers une plus grande fusion des méthodes ensemblistes et variationnelles, avec notamment l'idée d'exploiter la notion des "traces ensemblistes" (§5.1). J'illustre mes propos en reprenant l'exemple simplifié du filtre de Kalman du §2.2.
- une analyse 4D-VAR à méso-échelle est-elle réaliste ? Quels sont les problèmes, et comment pourrait-on envisager des applications de recherche, malgré le coût exorbitant de son développement et de son fonctionnement (§5.2) ?
- la question de la mise en équilibre des états initiaux : on peut redouter plusieurs sources d'ondes rapides indésirables dans les prévisions (acoustiques, mais surtout d'inertie-gravité). Des efforts sont faits pour tarir les sources au maximum pendant le processus d'analyse des données lui-même, mais n'aura-t-on pas malgré tout besoin d'une étape de mise en équilibre (par une méthode de filtrage adéquate) supplémentaire, externe à l'analyse et située juste avant la prévision (§5.3) ?

Mes activités à GMAP, que je souhaite pouvoir partager à l'avenir entre de l'encadrement scientifique (motivation de ma HDR) et mes responsabilités dans le cadre de l'animation des coopérations surtout internationales (direction de l'équipe COOPE du GMAP, forte participation au projet Aladin, contacts avec le CEPMMT et Hirlam[8]), me permettront peut-être de défendre ces idées et de les promouvoir dans le cadre de la collaboration.

J'adjoins au rapport de HDR mon Curriculum Vitae, ainsi que les copies de 5 articles : assimilation et ciblage dans un filtre de Kalman semi-géostrophique (cf §2.3, co-signé avec Alain Joly et François Lalaurette), vecteurs singuliers dans le problème d'Eady 2D (§3.1), statistiques d'erreurs NMC dans Aladin/Maroc (§4.2, co-signé avec Wafaa Sadiki et Jean-François Geleyn), réglage a posteriori des variances d'erreur dans les 3D-VAR Arpège et Aladin (§4.3, co-signé avec Wafaa Sadiki), résultats d'expériences de sensibilité aux conditions initiales, à méso-échelle, en employant le modèle adjoint d'Aladin (§4.5, co-signé avec Cornel Soci et Andras Horanyi). Les travaux de Cornel Soci (Bucarest, avril 2004), Wafaa Sadiki (Toulouse, avril 2005) et Vincent Guidard (Toulouse, octobre 2007, cf §4.4) ont donné lieu à des défenses de thèse de Doctorat.

[6] Action de Recherche pour l'Opérationnel à Méso-Echelle

[7] Centre National de Recherches Météorologiques/Groupe de Modélisation et Assimilation pour la Prévision

[8] HIgh Resolution Limited Area Modelling, un consortium de Centres Européens de PN regroupant les pays Scandinaves, l'Irlande, les Pays-Bas, l'Espagne, l'Estonie

Chapitre 2

Un exemple algorithmique en guise de fil conducteur

Dans ce chapitre, je souhaite développer dans les grandes lignes la méthode de filtrage du bruit par filtre de Kalman, méthode qui permet d'appréhender de très nombreux aspects de la problématique de l'assimilation des données pour la prévision numérique du temps. Je développerai un exemple à une variable scalaire, très simple, mais qui permet d'illustrer facilement certaines idées. Je réutiliserai cet exemple "0-D" par la suite (§3.2 et 5.1). Enfin, je poursuis ce chapitre en revenant sur mes travaux effectués au sein de l'équipe RECYF du CNRM/GMME, et dans le cadre d'études théoriques pour la préparation de la campagne de mesure FASTEX. J'ai à cette occasion travaillé directement avec un filtre de Kalman dans un cadre idéalisé, et obtenu une série de résultats qui restent encore d'actualité aujourd'hui. Je clos ce chapitre par quelques rappels sur les ponts qui existent entre le filtre de Kalman et l'analyse variationnelle, sur laquelle je me concentrerai davantage dans les chapitres suivants.

2.1 Le filtre de Kalman

Le filtre de Kalman peut se comprendre comme un processus relativement complet, de contrôle et de filtrage de l'état d'un système dynamique (Kalman, 1960, Kalman et Bucy, 1961). Il s'agit d'un système de contrôle dans la mesure où l'on va chercher à estimer l'état de variance d'erreur (d'analyse) la plus petite, sous l'hypothèse de quelques connaissances a priori sur les erreurs dans les données disponibles (en l'occurence, les moments d'ordre 1 et 2 : les moyennes et les covariances). Plusieurs sources d'information sont disponibles : une ébauche (prévision du système dynamique à partir d'un état antérieur du système) et des observations (des mesures extérieures a priori indépendantes de la prévision dynamique). Le lien fonctionnel entre les champs de type modèle (ébauche, mais aussi analyse en général) et les grandeurs observées se fait par l'opérateur d'observation (il sera noté H_k plus loin). Cet opérateur contient une information sur la physique des observations. S'y ajoute la connaissance des lois physiques de l'écoulement, transcrites dans les équations du modèle. Il s'agit aussi d'un filtrage, spatial à travers la convolution du signal ponctuel des observations par la matrice des covariances d'erreur de prévision, et temporel par l'évolution de cette matrice et la nature séquentielle du filtre.

Dans le détail, le filtre de Kalman peut se décomposer en trois étapes :

1. A partir d'un état Z^a_{k-1} du système dynamique, sur sa grille de représentation (de discrétisation), on effectue une prévision supposée ici *linéaire* :

$$Z^f_k = \mathcal{M}^k_{k-1} Z^a_{k-1} \tag{2.1}$$

Le modèle $\mathcal{M}^k_{k-1}$ est la forme intégrée entre $k-1$ et k du modèle tangent linéaire du système. Cet opérateur linéaire est encore appelé "résolvante" du système. Des exemples dans des cadres idéalisés existent dans la littérature (Joly et Thorpe, 1991). J'ai moi-même calculé la résolvante d'un problème

académique bien connu, le modèle 2D d'instabilité barocline d'Eady (Eady, 1949), sur lequel je reviens brièvement dans la section 3.1 (voir Fischer, 1998).

2. En parallèle à l'avance temporelle de l'état du système, on prévoit également l'évolution des covariances d'erreur de prévision, associées à Z_k^f :

$$P_k^f = \mathcal{M}_{k-1}^k P_{k-1}^a (\mathcal{M}_{k-1}^k)^* + Q_k \tag{2.2}$$

avec $P_{k-1}^a = E[Z_{k-1}^a (Z_{k-1}^a)^t]$ la matrice des covariances d'erreur d'analyse de l'étape $k-1$ du filtre, t étant l'indice de la transposition et $*$ celui de l'adjonction (transposée conjuguée par rapport à un produit scalaire). Q_k représente ici un terme supplémentaire d'erreur du modèle de prévision $\mathcal{M}_{k-1}^k$ (lequel agit entre les étapes $k-1$ et k).

3. L'analyse proprement dite à l'étape k, sur la grille de représentation du signal dynamique :

$$K_k = P_k^f H_k^t (R_k + H_k P_k^f H_k^t)^{-1} \tag{2.3}$$

$$P_k^a = (I - K_k H_k) P_k^f \tag{2.4}$$

$$Z_k^a = Z_k^f + K_k (X_k - H_k Z_k^f) \tag{2.5}$$

avec H_k l'opérateur d'observation, qui permet de passer de l'espace des états du système à celui des observations ($X_k = H_k Z$ pour tout Z). Noter que l'on considère le système *observable*, c'est à dire que l'opérateur Γ constitué de la réunion de l'identité I dans l'espace des Z et de H_k est une injection :

$$Z^1 \neq Z^2 \Rightarrow \Gamma Z^1 = \begin{pmatrix} Z^1 \\ H_k Z^1 \end{pmatrix} \neq \Gamma Z^2 = \begin{pmatrix} Z^2 \\ H_k Z^2 \end{pmatrix} \tag{2.6}$$

De manière plus générale, cette condition exprime le fait que le vecteur complet d'information (Z_k, X_k) contient de l'information sur toutes les composantes de Z, et donc que le problème inverse d'estimation par moindres carrés possède toujours une solution unique. Dans la mesure où nous allons toujours considérer des systèmes d'analyse contenant une ébauche, l'observabilité sera toujours vérifiée. K_k est la matrice de gain de l'analyse, P_k^a est la matrice des covariances d'erreur d'analyse (l'analyse étant Z_k^a). L'une des grandes forces du filtre de Kalman est de donner de façon explicite accès à cette matrice P_k^a, laquelle reste implicite par exemple dans une approche variationnelle. Le terme d'écart entre ébauche et observations est également appelé l'innovation ($X_k - H_k Z_k^f$), car c'est le terme qui contient la seule nouvelle source d'information, propre à modifier nos connaissances statistiques du système à l'étape k.

Plusieurs remarques peuvent être faites ici. Dans le cas où les erreurs sont Gaussiennes, le filtre de Kalman résoud exactement le problème Bayesien : il donne alors exactement la moyenne et la variance de la fonction de probabilité attachée à la variable Z (le minimum de variance d'erreur correspondant alors au maximum de vraisemblance du problème probabiliste) et ces deux moments suffisent alors à décrire la statistique de Z. Nous avons supposé tous les opérateurs comme linéaires. Cela ne sera pas le cas par exemple pour un système de prévision numérique de l'atmosphère. Ainsi, Ghil (1989) élargit-il le filtre de Kalman linéaire à un modèle non linéaire. L'étape 2.1 prend alors simplement une forme non-linéaire ($Z_k^f = M(Z_{k-1}^a)$), et l'on parlera d'un filtre de Kalman étendu (Gauthier *etal.* , 1993). Les opérateurs d'observation peuvent être des algorithmes complexes, non-linéaires, tels que les codes du transfert radiatif pour passer des profils de températures aux radiances (ou aux températures de brillance). Les systèmes d'assimilation météorologiques possèdent donc des composantes non-linéaires. Une conséquence est alors que l'étape 2.2 devient simplement une approximation à l'ordre 1 de la croissance des erreurs. Dans un filtre de Kalman étendu, on suppose que, si les états du système ressentent les effets non-linéaires dans l'intervalle $[k-1, k]$, alors les erreurs restent d'amplitude raisonnable et peuvent être tronquées.

L'avance temporelle de la matrice des covariances d'erreur de prévision, pour relativement simple qu'elle soit à écrire formellement, est une source de complication importante dans des systèmes réels. En effet, la dimension du vecteur d'état étant très grande (de l'ordre aujourd'hui de 10^6 à 10^8), il n'est possible ni de calculer

explicitement, ni de stocker, cette matrice. Une approche très en vogue consiste à faire appel à des méthodes ensemblistes, dans lesquelles un nombre plus ou moins limité de prévisions est utilisé pour échantillonner les covariances d'erreur de prévision (Evensen, 1994, Houtekamer et Mitchell, 1998) :

$$P^f \approx \frac{1}{N-1}\sum_{i=1}^{N}(Z_i^f - \overline{Z^f})(Z_i^f - \overline{Z^f})^t \tag{2.7}$$

où N est la dimension de l'ensemble. Si ces approches sont séduisantes par leur côté "parallélisable" (on peut effectuer les N prévisions les unes à côté des autres sur des grappes de processeurs), ainsi que par l'élimination du modèle adjoint $(\mathcal{M}_{k-1}^k)^*$, elles posent d'autres difficultés. Le faible nombre de prévisions possibles (même 1000 ou 10000 prévisions sont peu par rapport à la taille du problème) induit des problèmes de sous-échantillonage : des corrélations spatiales irréalistes apparaissent à de grandes distances (l'équivalent du repliement de spectre lorsqu'on tronque un signal dans une série de Fourier) et la matrice P^f est déficiente en rang (donc non inversible). Plusieurs solutions ont été proposées pour pallier ces problèmes, telles que l'injection des statistiques échantillonnées dans un modèle mathématique plus sophistiqué imposant notamment le support compact (Gaspari et Cohn, 1999) ou l'addition d'un terme de statistiques modélisé et de rang plein (méthode hybride, voir Etherton et Bishop, 2004). Notons que plusieurs formulations de filtres ensemblistes ont été développées, toutes reposant sur l'équation 2.7 (d'où leur appellation de "filtres racine-carrée", par référence au terme générique $Z_i^f - \overline{Z_i^f}$, voir Tippett *etal.* , 2003).

Le terme d'erreur du modèle peut également soulever quelques remarques. Dans les travaux des années 90, l'hypothèse du modèle parfait a très souvent été utilisée, surtout pour des applications à des systèmes de grande taille. En effet, d'une part, la modélisation des covariances d'erreur est une science en soi, comme nous le verrons dans la section 3.3 pour les covariances d'erreur de prévision P_k^f. D'autre part, identifier, même de façon approximative, l'erreur du modèle n'est pas chose aisée : il faut savoir la séparer de l'erreur déjà contenue dans l'ébauche (et qui est aussi une forme d'erreur du modèle, mais accumulée sur toutes les séquences précédentes du filtre, voir par exemple Trémolet, 2007). Enfin, nous verrons dans le §2.4 que, dans un cadre variationnel, il devient algorithmiquement plus difficile de faire intervenir ce terme. Il n'a donc été réellement pris en compte que de façon sporadique, dans des études académiques et sur des modèles idéalisés assez simples (voir par exemple Dee, 1995, ou Etherton et Bishop, 2004).

Pour finir, rappelons une autre hypothèse fondamentale de ce type de filtre : on va supposer toutes les erreurs *non biaisées*. Cette hypothèse peut sembler assez gratuite dans les équations (il suffit de retirer la moyenne sur X_k et Z) et s'obtient par construction dans des modèles simples, mais elle est presque toujours violée dans les systèmes atmophériques : certaines mesures sont entachées d'erreur de biais, parfois variable dans le temps (données satellite, capteurs au sol), et les erreurs du modèle sont souvent en premier lieu des problèmes de biais. Lorsque ces derniers sont bien connus des équipes de monitoring des systèmes de prévisions, alors ils sont éliminés *en aval* de la production brute du modèle. Cependant, lorsque les erreurs de biais apparaissent avec des nouvelles versions d'un modèle de prévision (meilleur par ailleurs, pour d'autres paramètres), alors ils peuvent constituer une source de problèmes pour l'assimilation. Ainsi, Dee (2003) évoque-t-il l'exemple d'un changement dans la climatologie d'un système de prévision, dû non pas à un authentique changement de type de temps, mais à l'apparition d'une erreur de biais.

2.2 Un exemple "0-D"

J'avais, dans ma note de travail du CNRM (Fischer, 1997) développé un exemple d'une grande simplicité, mais qui permet d'appréhender assez rapidement le fonctionnement du filtre de Kalman. Il suffit de considérer un point (un scalaire), et une observation répétée à pas de temps réguliers. On va s'intéresser surtout à l'évolution des écarts-types d'erreur de prévision et d'analyse. A chaque itération du filtre, on obtient un écart-type d'erreur de prévision noté σ_k. On introduit une dynamique, représentative d'un mode propre linéaire, en supposant que l'erreur croit de façon exponentielle dans le segment de temps (régulier de taille t) $[k-1, k]$:

$\sigma_k = s_{k-1}e^{\lambda t}$. L'observation quant à elle est toujours effectuée avec le même écart-type d'erreur constant ω. Il suffit de partir des équations 2.2 et 2.4 pour obtenir l'écart-type d'erreur d'analyse à chaque itération :

$$s_k = \frac{s_{k-1}\omega e^{\lambda t}}{\sqrt{s_{k-1}^2 e^{2\lambda t} + \omega^2}} \tag{2.8}$$

De même, on a une relation de récurrence pour l'écart-type d'erreur de prévision :

$$\sigma_k = \frac{\sigma_{k-1}\omega e^{\lambda t}}{\sqrt{\sigma_{k-1}^2 + \omega^2}} \tag{2.9}$$

La figure 2.1 illustre le comportement du filtre de Kalman sur cet exemple. Je reviendrai sur ce modèle "0-D" très simple dans les §2.4, 3.2 et 5.1.

2.3 Le filtre de Kalman semi-géostrophique : synthèse de l'article "Error growth and Kalman filtering within an idealized baroclinic flow", par Fischer, Joly et Lalaurette, paru dans Tellus (1998)

Lors de mon passage dans l'équipe RECYF du CNRM/GMME, j'ai eu l'occasion de travailler dans le cadre de la préparation de la campagne de mesure FASTEX. Parmi les objectifs de FASTEX figuraient :

- une meilleure compréhension des mécanismes de développement des cyclogénèses rapides sur l'Atlantique Nord.
- l'évaluation en temps réel des méthodes de ciblage des observations. De fait, en 1995/97, il s'agissait surtout des vecteurs singuliers (Buizza et Palmer, 1995, Molteni *etal.* , 1996), des "bred modes" (Toth et Kalnay, 1997) et des premières ébauches de ce qui allait devenir l'Ensemble Transform Kalman Filter, ETKF (Bishop et Toth, 1999, Bishop *etal.* , 2001).
- la mise en place des systèmes d'observation pour la campagne de mesure, et plus généralement l'évaluation des systèmes d'observation de l'époque pour la prévision à 3, voire 5 jours d'échéance.

Dans ce cadre, nous nous posions la question de l'intérêt relatif d'effectuer des mesures de radiosondages depuis un bateau (plateforme essentiellement fixe sur l'océan, par rapport au passage des systèmes cycloniques), en comparaison à des mesures prises par des dropsondes "perdues" depuis des avions instrumentés (donc des profils le long de segments de mesures, assez mobiles, mais très limités dans le temps). Nous avons ramené cette question à une forme extrêmement dépouillée : comparer dans un cadre idéalisé l'apport d'observations limitées dans l'espace mais régulièrement disponibles tout le long du cycle de vie d'un système barocline, avec des mesures denses dans l'espace, mais très limitées dans le temps par rapport au phénomène que l'on souhaite étudier. Le cadre académique retenu a été celui d'un modèle semi-géostrophique à tourbillon potentiel uniforme, développé auparavant par Alain Joly (voir Hoskins, 1975, pour le développement du système de coordonnées semi-géostrophiques et son utilisation pour comprendre les mécanismes de frontogénèses non-linéaires, et Joly, 1995, pour l'implémentation du modèle lui-même). Nous disposions à la fois du modèle direct, du modèle tangent linéaire et de son adjoint, ainsi que de la modélisation complète de l'évolution de la matrice P_k^f. Ce dernier point est rendu possible par l'hypothèse du tourbillon potentiel uniforme, qui réduit considérablement le nombre de degrés de liberté dans le modèle. En effet, toute la dynamique est connue avec seulement la connaissance du champ de température potentielle à la surface et au toit du modèle (il s'agit en fait d'un modèle à deux niveaux et à une couche unique d'atmosphère). Les champs en question sont pris périodiques dans une direction et tendant vers zéro dans l'autre (c'est donc une géométrie de type canal). Dans ces conditions, nous avons manipulé des vecteurs d'état d'environ 500 éléments, donc des matrices de covariances avec environ 250000 éléments.

J'ai développé l'analyse des données dans ce système (soit les opérations correspondant aux étapes 2.3, 2.4 et 2.5 décrites en section 2.1), en prenant des observations très simples : des températures dans l'espace

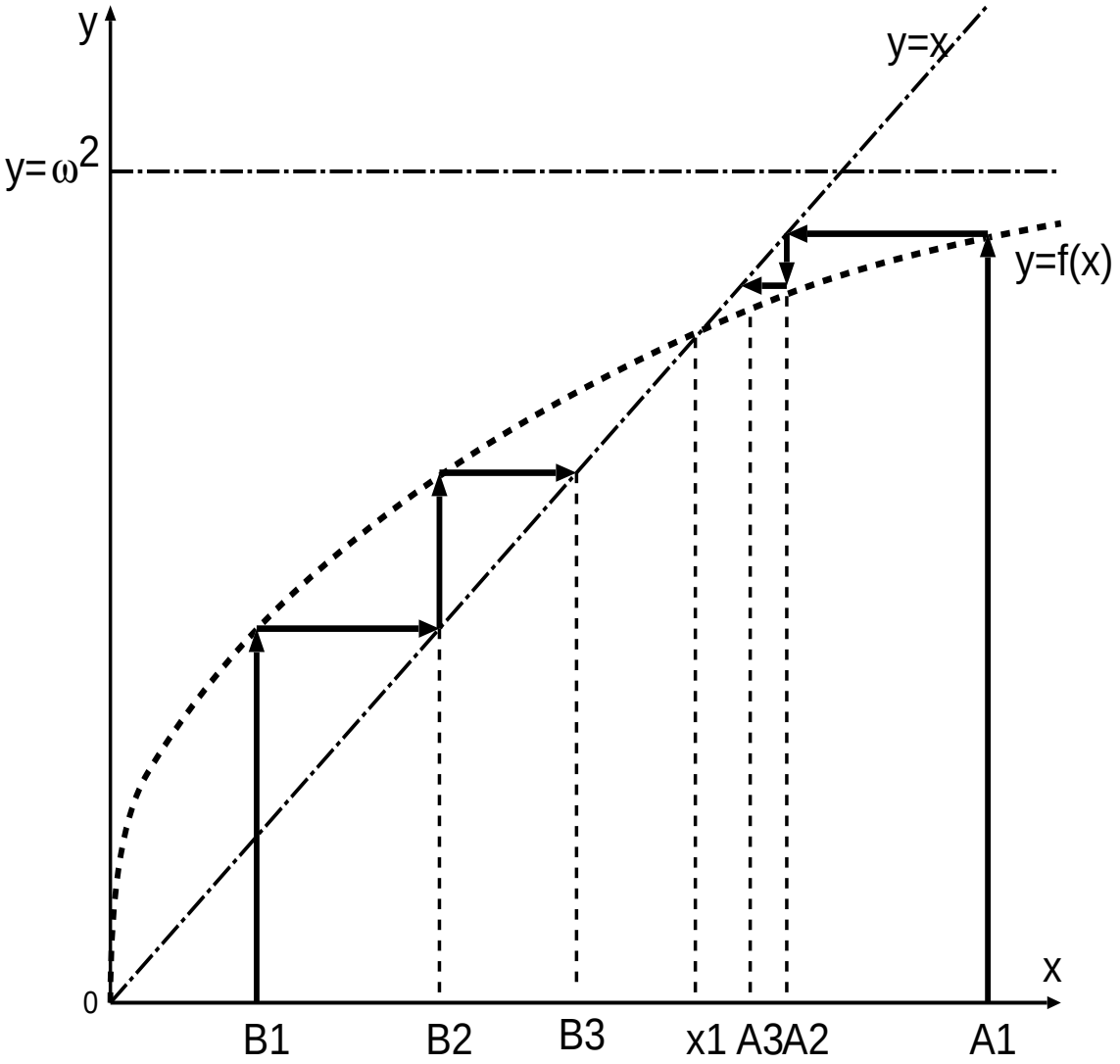


FIG. 2.1 – Evolution itérative des écarts-types d'erreur dans l'exemple à une seule variable. Je montre la convergence de la variance analysée s_k^2 vers sa valeur limite notée x_1. Les deux suites A_n et B_n illustrent la convergence à partir d'une variance initiale élevée et faible, respectivement. La courbe $y = f(x)$ illustre la formule à la limite pour s_k^2, que l'on peut obtenir par la méthode du point fixe : $x_1 = \omega^2 \frac{e^{2\lambda t} - 1}{e^{2\lambda t}}$ (cf Eq. 2.8)

physique du modèle, au sol ou au toit de celui-ci (ou les deux). De plus, dans un souci d'analyser des observations plausibles, j'ai utilisé des valeurs de températures récupérées du modèle lui-même. J'ai donc, dans le jargon, procédé à des expériences "jumelles" ("identical twin experiments", dans lesquelles ce sont en fait des observations simulées par le système lui-même qui sont analysées en guise d'observations, ce qui permet de connaître précisément la vérité à estimer). Les écart-types d'erreur d'observation étaient en général de 1 K (ordre de grandeur des écarts-types contenus dans P_k^f). Je reprends ici les principaux résultats développés dans l'article que je co-signe avec Alain Joly et François Lalaurette (Fischer *etal.* , 1998), paru dans Tellus.

Le contexte expérimental peut se découper en trois phases :

1. on se donne un état de base zonal, supportant le développement d'ondes baroclines (celles-ci viennent s'alimenter dans le réservoir d'énergie potentielle fourni par le gradient transversal de température). L'onde devant produire la cyclogénèse est obtenue comme un vecteur singulier pour la norme de l'enstrophie (le carré du tourbillon) calculé sur 2 jours en se focalisant sur un domaine point-de-grille localisé dans le front thermique zonal. Cette onde avait été précédemment définie par François Lalaurette, pour ses propres travaux sur les ondes baroclines et les vecteurs singuliers. Cette onde ayant néanmoins des caractéristiques très élégantes de perturbations frontales, je l'ai récupérée pour mes propres travaux (!). La figure 2.2 montre le champ de température potentielle obtenu après 8 jours d'évolution de l'onde barocline. A cette échéance, le front est clairement entré dans sa phase non-linéaire (qu'il a quittée environ entre les jours 5 et 6).

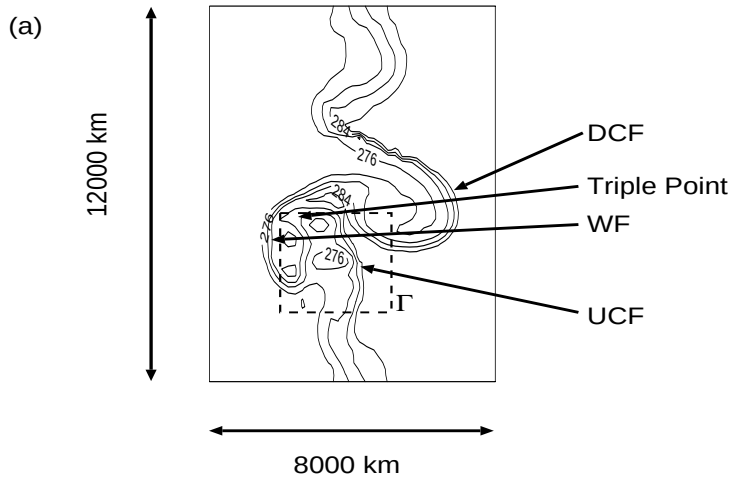
2. en partant de covariances homogènes et isotropes (une hypothèse souvent utilisée pour modéliser les erreurs de l'ébauche, voir section 3.3), je laisse ces covariances évoluer sur deux jours, entre les jours 6 et 8. Le résultat, montré dans la figure 2.3, donne des écarts-types d'erreur de prévision très variables dans l'espace. Les erreurs ont considérablement crû dans les zones frontales (de l'ordre de 12 K, en partant de 1 K initialement, dans le front de surface ; environ 9 K à partir de 2 K au départ, au toit du modèle). Les zones de transition entre les fortes valeurs d'écart-type et celles plus faibles sont assez abruptes. Nous avons donc illustré ici de façon nette l'amplification des erreurs de prévision dans les zones actives d'un système barocline. Même si cet exemple est idéalisé, il illustre le résultat bien connu que les erreurs les plus fortes (et les croissances d'erreurs les plus rapides) sont observées au voisinage des dépressions et des zones frontales actives.

   Les différences étant notables au bout de 2 jours d'évolution dynamique, nous avons décidé que, pour tester l'impact des réseaux d'observations, nous allions démarrer les assimilations avec plusieurs jeux de matrices de covariances d'erreurs : des covariances uniformes et isotropes, des covariances évoluées (entre les jours 5.5 et 6), et des covariances évoluées mais re-normées de sorte à retrouver les variances totales du cas uniforme.

3. enfin, j'ai effectué une série d'assimilations par le filtre de Kalman, en appliquant successivement les équations 2.2, 2.3 et 2.4 (les valeurs observées, identiques au modèle, ne sont en fait pas analysées). Ces assimilations ont été effectuées entre les jours 6 et 8, toutes les 6 heures (soit 8 réseaux). La figure 2.4 montre le champ d'écart-type d'erreur au bout de ces 8 réseaux d'assimilation, dans le cas de deux sources ponctuelles prenant chacune une observation au sol et au toit du modèle. Les points d'observations ont été placés respectivement dans la trace au sol du front, et sous la trace en altitude (d'où aussi leur décalage transversal).

   On constate une réduction très sensible des écarts-types d'erreur, de 20 à 50 % dans toutes les zones frontales. Cette réduction est particulièrement sensible au voisinage des points d'observation, mais elle l'est aussi assez en aval. La réduction est encore un peu plus forte lorsqu'on utilise au départ des covariances déjà dynamiquement evoluées, et cela semble dû à la nature hétérogène, dépendante de l'écoulement, des corrélations (car elle ne semble pas expliquée par la seule re-normalisation des variances).

   J'ai répété ces expériences avec une série d'observations denses dans l'espace, mais très ponctuelles dans le temps (ce qui devait simuler des avions larguant des dropsondes). J'ai pu noter alors que des transects avion qui se contentent de couper le front orthogonalement n'étaient pas très utiles (au sens de la réduction de l'écart-type d'erreur d'analyse au jour 8). De meilleurs résultats sont obtenus en faisant



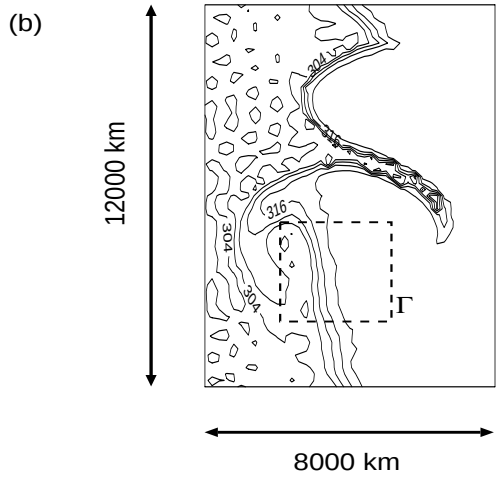


FIG. 2.2 – Température potentielle tous les 4 K, au sol et au toit du modèle semi-géostrophique. La situation correspond au jour 8 de l'évolution de l'onde barocline. (a) surface, (b) toit.

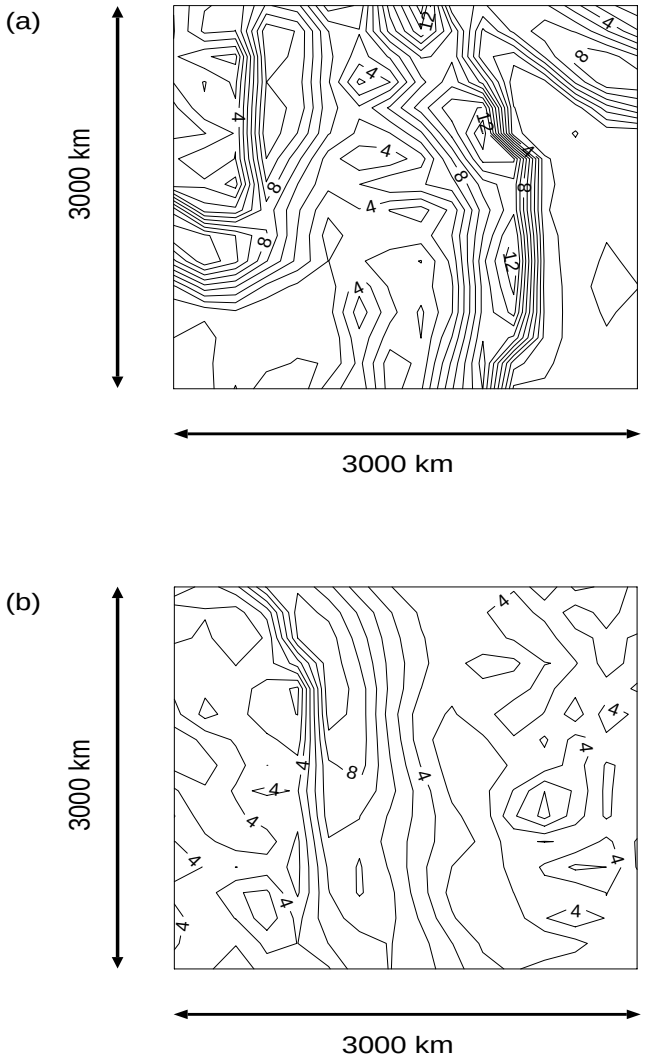


FIG. 2.3 – Ecart-type d'erreur sur la température potentielle tous les 1 K, au sol et au toit du modèle semi-géostrophique. La situation correspond au jour 8 de l'évolution de l'onde barocline, soit après 2 jours d'intégration linéaire des covariances d'erreur. (a) surface, (b) toit.

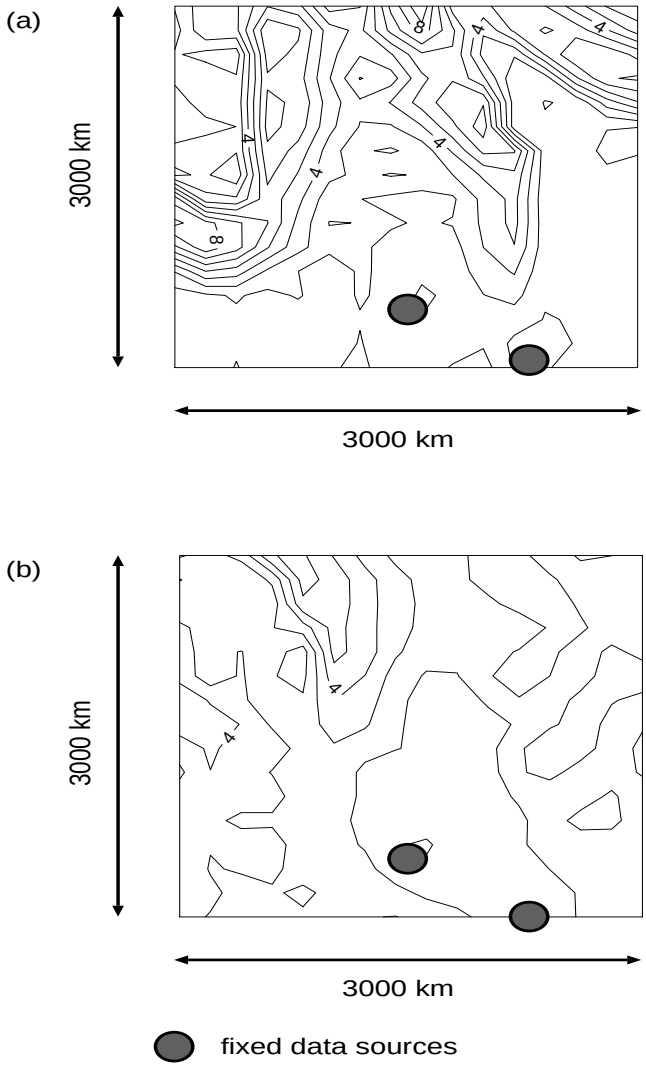


FIG. 2.4 – Ecart-type d'erreur sur la température potentielle tous les 1 K, au sol et au toit du modèle semi-géostrophique. La situation correspond au jour 8 du filtre de Kalman, soit après 2 jours d'assimilation régulière des covariances d'erreur aux points d'observations marqués par des ronds. (a) surface, (b) toit.

passer les points d'observation par des zones sensibles, déterminées objectivement (dans mon cas, j'avais recalculé un vecteur singulier pour la variance de température, et utilisé ses structures pour déterminer le trajet-avion). Cependant, la principale différence que j'ai pu observer venait de l'utilisation ou non des corrélations dynamiquement évoluées (renormées ou non). En effet, encore davantage que pour les deux points de mesures réguliers dans le temps, dans le cas de mesures denses mais ponctuelles dans le temps, il était important d'avoir utilisé, au moment de l'analyse de ces données, des covariances d'erreur riches en information sur l'écoulement.

D'autres expériences de sensibilité ont porté sur la modification des écarts-types d'erreur d'observation et sur le positionnement des sources d'observations. J'ai pu noter que des mesures prises clairement en dehors des zones frontales, ou nettement à l'extérieur des zones de ciblage objectif, n'avaient effectivement que peu d'intérêt.

En conclusion, cette étude, qui illustre tout le potentiel en terme de compréhension, de la méthode du filtre de Kalman, m'a permis d'apporter quelques résultats encourageants pour la préparation et l'exécution de la campagne FASTEX :

- l'intérêt des observations de radiosondages par bateau, effectués régulièrement dans la zone du storm track. Ce résultat a par la suite été retrouvé à partir de données effectivement réanalysées de la campagne FASTEX (Bergot *etal.* , 1999).
- l'utilité d'employer des modèles de covariances d'erreur de prévision qui soient réalistes, et riches en information sur l'écoulement atmosphérique. Or, dans un cadre plus réaliste (comme celui des systèmes opérationnels), où il est hors de question d'accéder directement aux matrices P_k^f ou P_k^a, ce résultat est pour l'instant resté un voeux pieux. Des essais pour employer un filtre de Kalman de rang réduit pour représenter quelques grandes structures dans les corrélations ou les cartes de variances ont échoué (essais notamment au CEPMMT, Mike Fisher). Ces échecs sont peut-être dus au fait que les structures capturées par ces filtres, de très faible dimension, restent trop éloignées de la réalité. Les études les plus récentes se concentrent plutôt sur l'obtention de cartes de variances assez détaillées pour un écoulement donné (cartes de variances "climatologiques", comme dans Belo Pereira et Berre, 2006, ou cartes "du jour" dans le cadre d'une estimation par une méthode de Monte-Carlo comme dans Andersson *etal.* , 2000, dans l'espace des observations : $HP^fH^t \approx \frac{1}{N}\sum_{i=1}^{N} H(P^f)^{1/2}\xi_i(H(P^f)^{1/2}\xi_i)^t$, avec ξ_i un jeu de vecteurs aléatoires, une formule qui rappelle l'équation 2.7).
- l'intérêt pour des approches de ciblage objectif. Sur ce point, les travaux restent encore d'actualité, même à grande échelle. Dans les systèmes opérationnels ou quasi-opérationnels, l'impact des observations ciblées reste assez difficile à mettre en évidence. Je prends pour exemple des bilans récents d'expériences de ciblage, effectuées dans le cadre de la campagne A-TReC (Novembre-Décembre 2003) du programme THORPEX. Dans A-TReC, plusieurs techniques de ciblage (vecteurs singuliers, ETKF) produisent des zones et des mesures de sensibilité de la prévision, en temps réel (Shapiro et Thorpe, 2004). L'objectif est d'obtenir ces diagnostics 48 heures avant l'heure nominale, afin d'avoir le temps de mobiliser les moyens d'observations supplémentaires envisagés (radiosondages supplémentaires, avions instrumentés, capteurs embarqués sur les avions commerciaux, sur les bateaux, dropsondes). Les premiers bilans des expériences d'impact des observations supplémentaires, dans lesquelles on compare les performances des systèmes d'assimilation opérationnels de l'époque (contenant les observations ciblées, car celles-ci étaient présentes sur le GTS[1]) avec des assimilations dans lesquelles ces mesures ont été toutes retirées, donnent une image mitigée sur le bénéfice du ciblage. Environ 30 % des situations donneraient un signal positif, d'après les évaluations faites à Météo-France et au CEPMMT, contre environ 10 à 20 % de cas détériorés et 40 à 50 % de cas neutres (Rabier, 2005). Certes, le fait qu'il y ait des cas négatifs n'est pas surprenant, puisque le principe du ciblage repose toujours sur une amélioration au sens statistique : un nombre plus ou moins marginal de cas détériorés par les observations ciblées est donc normal. Cependant, au vu des efforts algorithmiques, de coût de calcul et de moyens d'observations déployés, les résultats restent médiocres. Plusieurs raisons sont aujourd'hui

[1]Global Transmission System, le réseau mondial d'échange de données

évoquées :

– le poids grandissant des observations satellite, qui réduit l'impact de mesures supplémentaires faites de surcroit sur des zones géographiques réduites,
– la difficulté à échantillonner toute une zone de sensibilité (et la non-cohérence de ces zones de sensibilité entre différents algorithmes de ciblage),
– le manque de maturité des systèmes d'assimilation, notamment dans leur capacité à reproduire dans leurs spécifications les caractéristiques locales (dans le temps et dans l'espace) de l'écoulement.

Ce dernier point vient de la simplicité des modèles de représentation de la matrice P_k^f (j'y reviendrai sous une forme différente dans la section 3.3). Mes résultats académiques confirmeraient plutôt ce soupçon. Je peux aussi signaler que, dans mes expériences, je n'ai fait qu'ajouter des observations là où il n'y en avait aucune par défaut (!), puisque je me suis toujours comparé à la croissance dynamique pure. Mes choix expérimentaux étaient donc, en ce sens, très favorables à l'impact du ciblage.

2.4 Equivalence avec le 4D-VAR classique et le 4D-VAR à contrainte faible

Dans cette section, je souhaite revenir brièvement sur les équivalences entre les méthodes de filtrage séquentiel et les approches variationnelles, dans le cadre de l'analyse optimale des données. En effet, il existe différentes méthodes de résolution du problème aux moindres carrés que représentent l'étape d'analyse et les propagations temporelles des informations (analyses, covariances d'erreur). Le filtre de Kalman, et ses méthodes dérivées ou généralisées, est une approche dans laquelle on va chercher à explicitement inverser l'équation donnant la solution de l'analyse (cf équation 2.3 fournissant le gain). L'étape d'analyse n'est alors rien d'autre qu'une analyse par interpolation optimale (Gandin, 1963), qui, sous une forme localisée dans la grille de calcul, est une approche très ancienne en assimilation de données météorologiques (elle était déjà utilisée dans les analyses 3D des systèmes Emeraude, Péridot et Canari à Météo-France, voir Durand et Bougeault, 1987 ou Ajjaji et Issara, 1994).

Une façon alternative de résoudre le problème est d'adopter une approche variationnelle. Le problème d'inversion consiste alors à plutôt déterminer la solution fournissant le minimum d'une fonction-coût J, laquelle mesure la distance de la solution à l'ébauche et aux observations, normée par les covariances d'erreur respectives :

$$J(Z) = \underbrace{(Z - Z^b)^t (P^b)^{-1} (Z - Z^b)}_{J^b} + \underbrace{(X - HZ)^t R^{-1} (X - HZ)}_{J^o} \tag{2.10}$$

Si on considère un certain nombre N de groupes d'obervations, régulièrement répartis par intervalles de temps t (que l'on indicera k), et toujours sous l'hypothèse d'opérateurs linéaires (modèles $\mathcal{M}_{k-1}^k$ et observations H_k), alors on peut écrire J sous la forme itérative :

$$\left\{ \begin{array}{rcl} J(Z) & = & (Z - Z^b)^t (P^b)^{-1} (Z - Z^b) + \sum_{k=0}^{N} (X_k - H_k Z_k)^t R_k^{-t/2} R_k^{-1/2} (X_k - H_k Z_k) \\ Z_k & = & \mathcal{M}_{k-1}^k Z_{k-1}, \forall k \in [1, N] \\ Z_0 & = & Z^b \text{ au début de la minimisation seulement} \end{array} \right. \tag{2.11}$$

La solution Z_0^{min} est obtenue en cherchant à annuler le gradient de la fonction-coût : $\nabla_{Z_0} J = 0$.

Le système 2.11 présente l'essentiel de l'algorithme dit "4D-VAR", dans lequel on calcule donc la fonction-coût et son gradient par des successions d'applications des opérateurs d'observations et des modèles résolvants, en parcourant toute la fenêtre d'assimilation (Talagrand et Courtier, 1987, Lacarra et Talagrand, 1988, Talagrand, 1989). Si nous identifions les différentes séquences d'observations aux séquences d'analyse du filtre de Kalman, utilisant les mêmes ébauche et statistiques de covariances d'erreur (Z^b, P^b, R_k), les mêmes opérateurs *linéaires* $(H_k, \mathcal{M}_{k-1}^k)$, alors le *résultat final* (Z_N^{min}), en fin de séquence d'assimilation, de l'assimilation 4D-VAR doit correspondre à celui donné par le filtre de Kalman (Rabier *etal.* , 1993).

De manière plus générale, cette équivalence pour *l'analyse finale* entre le filtre de Kalman et le 4D-VAR est vérifiée *tant que les erreurs d'ébauche et d'observations sont décorrélées dans le temps*. En effet, dans le filtre séquentiel, la combinaison linéaire de l'équation 2.5 ne correspond à la solution du problème aux moindres

carrés seulement si les erreurs aux étapes successives k sont effectivement décorrélées. On peut illustrer cette propriété à partir de l'exemple "0-D" du §2.2. En partant d'une ébauche x_1 et d'une observation y_1, la première analyse du filtre de Kalman vaut :

$$\frac{z_1}{s_1^2} = \frac{x_1}{\sigma_1^2} + \frac{y_1}{\omega^2}$$

avec

$$\frac{1}{s_1^2} = \frac{1}{\sigma_1^2} + \frac{1}{\omega^2}$$

pour la variance d'erreur d'analyse. On fait ensuite avancer le filtre à l'instant suivant :

$$x_2 = z_1 e^{\lambda t}$$

et on procède à la deuxième analyse :

$$\frac{z_2}{s_2^2} = \frac{x_2}{\sigma_2^2} + \frac{y_2}{\omega^2}$$

avec $\sigma_2 = s_1 e^{\lambda t}$ et

$$\frac{1}{s_2^2} = \frac{1}{\sigma_2^2} + \frac{1}{\omega^2}$$

Le formalisme du filtre de Kalman ne permet pas de prendre en compte une éventuelle corrélation α/ω^2 des erreurs entre les deux observations y_1 et y_2. Dans le formalisme variationnel, on écrit la fonction :

$$J = (\xi - x_1, \xi - y_1, \xi e^{\lambda t} - y_2) P^{-1} (\xi - x_1, \xi - y_1, \xi e^{\lambda t} - y_2)^t$$

où :

$$P = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \omega^2 & \alpha \\ 0 & \alpha & \omega^2 \end{pmatrix}$$

ξ est l'inconnue du problème variationnel valable pour l'instant d'analyse $k = 1$ ($\xi_1 = \xi$ par définition) ; à l'instant $k = 2$: $\xi_2 = \xi_1 e^{\lambda t}$; α est la covariance d'erreur entre y_1 et y_2. La solution pour ξ vérifie :

$$\nabla_\xi J = \frac{2}{\sigma_1^2}(\xi - x_1) + \frac{2}{\Delta}[\omega^2(\xi(1 + e^{2\lambda t}) - y_1 - y_2 e^{\lambda t}) - 2\alpha(\xi e^{\lambda t} - \frac{y_1 e^{\lambda t} + y_2}{2})] = 0$$

avec $\Delta = (\omega^2 - \alpha)(\omega^2 + \alpha)$. D'une part, on peut comparer, au premier instant d'analyse $k = 1$, et pour $\alpha = 0$, la solution variationnelle ξ_1 avec la solution du filtre séquentiel z_1 :

$$\frac{\xi_1}{s_1^2} - \frac{z_1}{s_1^2} = \frac{\xi_1 e^{2\lambda t} + y_2 e^{\lambda t}}{\omega^2} \neq 0$$

Les deux solutions ne sont pas égales à un instant d'analyse intermédiaire. On vérifie par contre que $\xi_2 = z_2$. En revenant au cas général, ce n'est que lorsque toutes les observations ont été "vues" par chacun des deux systèmes (donc à la fin de la séquence), que les deux méthodes fournissent des états analysés identiques : $Z_N^{min} = \mathcal{M}_0^N Z_0^{min} = Z_N^a$ (attention, les modèles de propagation sont supposés parfaits ici).

D'autre part, on peut vérifier avec l'exemple "0-D" que les deux solutions finales sont différentes si $\alpha \neq 0$. On peut pour cela supposer le modèle dynamique stationnaire ($\lambda = 0$), par simplicité. Après calculs, on obtient :

$$z_2(\frac{1}{\sigma_1^2} + \frac{2}{\omega^2}) = \frac{x_1}{\sigma_1^2} + \frac{2y'}{\omega^2}$$

et

$$\xi_2(\frac{1}{\sigma_1^2} + \frac{2}{\omega^2 + \alpha}) = \frac{x_1}{\sigma_1^2} + \frac{2y'}{\omega^2 + \alpha}$$

avec $y' = (y_1 + y_2)/2$. Les deux analyses z_2 et ξ_2 sont bien différentes pour $\alpha \neq 0$. L'analyse variationnelle permet implicitement, à travers les spécifications de la matrice P, de prendre en compte les corrélations temporelles d'erreurs. Cependant, dans la pratique, par souci de conserver un P "suffisamment diagonal" (par économie de calcul, de place mémoire pour le stockage informatique, par simplicité d'un code etc ...), les corrélations sérielles sont souvent ignorées.

De même, on peut montrer dans des expériences d'assimilation simples (par exemple, avec une seule observation) que le 4D-VAR procède implicitement à une évolution dynamique des structures des covariances d'erreur d'ébauche fournies en entrée d'assimilation (et qui sont généralement assez simples, homogènes et isotropes), comme l'ont illustré Thépaut *etal.* (1996). Il s'agit là, de façon implicite, de l'évolution dynamique des covariances d'erreur du système donnée par l'équation 2.2. De façon plus générale, le 4D-VAR ne renseigne jamais directement sur la forme des covariances d'erreur en cours d'assimilation, au contraire du filtre de Kalman qui donne explicitement accès aux matrices P_k^f et P_k^a.

D'autres différences notables existent entre l'approche par filtrage séquentiel et celle variationnelle. D'un point de vue pratique, la méthode variationnelle nécessite le développement du code adjoint des modèles de propagation (modèles résolvants) et des opérateurs d'observations.

Si, de plus, on prend en compte l'erreur du modèle (le terme Q_k de l'équation 2.2), alors il n'y a plus d'équivalence entre le 4D-VAR "classique" et le filtre de Kalman. Fisher (2004) a ainsi étudié le "4D-VAR à contrainte faible", dans lequel l'erreur du modèle devient un terme de pénalité supplémentaire dans la fonction-coût, et nécessite d'ajouter l'erreur-modèle dans la variable de contrôle Z. En toute rigueur, il faudrait alors ajouter un état-modèle supplémentaire à *chaque* itération, ce qui rend un 4D-VAR à contrainte faible particulièrement coûteux en place mémoire. Je ne m'étendrai pas davantage sur les équivalences entre méthodes de filtrage séquentiel et techniques variationnelles. Il faut retenir qu'il est parfois plus pratique de raisonner en terme de filtrage séquentiel, pour comprendre les propriétés d'une assimilation de données, qu'en terme d'analyse variationnelle.

Pour finir, notons que, dans les systèmes de PN opérationnels, l'assimilation variationnelle est implémentée sous forme *incrémentale* (Courtier *etal.* , 1994). On se ramène alors à une fonction-coût quadratique et à des opérateurs linéaires pour la variable du problème, en procédant formellement à un développement à l'ordre 1 au point d'ébauche :

$$J(Z) = \underbrace{(Z - Z^b)^t (P^b)^{-1} (Z - Z^b)}_{J^b} + \underbrace{(X - \mathcal{H}(Z^b) - H(Z - Z^b))^t R^{-1} (X - \mathcal{H}(Z^b) - H(Z - Z^b))}_{J^o}$$

où H est l'opérateur tangent linéaire de $\mathcal{H}$. On recherche le minimum pour l'incrément $Z - Z^b$. Cette formulation autorise toute une série de "simplifications" entre le calcul de l'innovation $X - \mathcal{H}(Z^b)$ et le traitement de l'incrément $H(Z - Z^b)$ (voir par exemple pour l'application à Arpège, Veersé et Thépaut, 1998, ou Janiskova *etal.* , 1999) : changement de résolution entre le calcul de l'innovation, à haute résolution, et celui de l'incrément, à plus basse résolution ; modèles de propagation non-linéaire et linéaire différents. On peut de plus recalculer le terme $\mathcal{H}(Z)$ après une première minimisation, puis procéder à une deuxième minimisation, et ainsi de suite. Ce procédé itératif aboutit à l'algorithme 4D-VAR multi-incrémental, et permet dans la pratique de tenir compte de non-linéarités dans $\mathcal{H}$.

Par ailleurs, on peut aussi se ramener à une assimilation dite "3D-VAR", dans laquelle la dimension temporelle est "perdue" (poser $\mathcal{M}_{k-1}^k = I$, c'est à dire toutes les propagations d'incréments égales à l'identité, dans le système 2.11). Dans le cas d'une analyse 3D-VAR, on peut aussi noter que la fusion des différents segments d'observations $[k-1, k]$ (aussi appelés les "time-slot") en un seul conduit dans la pratique à des règles de prises de décision concernant les observations légèrement différentes du 4D-VAR. Ainsi, si plusieurs mesures sont faites à l'intérieur d'un même segment, et si l'on décide de ne retenir que la mesure la plus centrale, alors une seule mesure est conservée dans le 3D-VAR, pour l'ensemble de la fenêtre d'assimilation, alors que plusieurs mesures seront gardées en 4D-VAR. En tout, davantage de mesures sont donc susceptibles de rentrer dans une analyse 4D-VAR que dans une analyse 3D-VAR, par le seul effet des règles de prises de décision en début d'analyse (étape dite du "screening" des observations). C'est ce type d'assimilation qui fut

opérationnelle pendant un temps au CEPMMT (Courtier *etal.* , 1998, Rabier *etal.* , 1998) et que j'ai aidé à développer et à porter en mode opérationnel dans le système Aladin/France (Fischer *etal.* , 2005). Je reviendrai sur mes travaux dans le cadre de l'analyse 3D-VAR dans Aladin dans le chapitre 4.

Chapitre 3

La croissance des erreurs de prévision et leur modélisation

J'aborde dans ce chapitre de façon plus spécifique la question de la croissance des erreurs. De fait, je commence par un rapide tour d'horizon sur la croissance des modes linéaires dans des systèmes idéalisés de l'atmosphère, puis j'en arrive à l'estimation des statistiques d'erreur d'ébauche pour l'assimilation (la matrice P^b ou P_0^f du chapitre 2). J'évoque à cette occasion mon travail personnel sur la croissance des vecteurs singuliers dans le problème d'Eady 2D.

3.1 La croissance linéaire des erreurs initiales : synthèse de l'article "Linear amplification and error growth in the 2D Eady problem with uniform potential vorticity", par Fischer, paru dans J.A.S. (1998)

L'étude des perturbations atmosphériques a amené assez tôt les météorologues à étudier les questions de stabilité de l'écoulement atmosphérique. Un exemple parmi d'autre, historique, d'analyse de stabilité est donné par les travaux d'Eady (1949), qui sont une des bases pour la compréhension des instabilités de type barocline. Le cadre général est de se donner un état de base idéalisé, stationnaire (au sens d'un jeu d'équations simplifiées), et de rechercher quelles solutions, sous forme d'un incrément de petite amplitude, peuvent s'y développer de manière exponentielle dans le temps (notamment, donc, en conservant leur structure spatiale). Il s'agit d'une étude de stabilité linéaire, par recherche de possibles modes *normaux*. Eady les fournit pour un écoulement 2D, marqué par un gradient (infini) transversal de température (et donc un cisaillement vertical de vent constant et stationnaire), et sous l'hypothèse d'un tourbillon potentiel uniforme. Eady suppose l'état de base en équilibre géostrophique et hydrostatique. Cet écoulement, contraint entre deux conditions limites rigides (un sol et un toit) admet le développement de modes normaux baroclines (les perturbations se développent en se nourrissant du gradient transversal de température). Les figures 3.1 et 3.2 montrent les taux d'amplification respectivement sur 6 heures et 48 heures des modes normaux d'Eady. On voit notamment qu'il existe un nombre d'onde (les perturbations sont développées en sinus dans le sens longitudinal de l'écoulement) optimal pour lequel le mode d'Eady s'amplifie le plus. Avec des valeurs de grandeur géophysique pour tous les paramètres numériques du problème idéalisé, la longueur d'onde correspondante est de l'ordre de 4000 km. Comme le mode normal conserve sa structure, cette valeur est aussi inchangée si l'on considère l'amplification sur 24 ou 48 heures, au lieu de 6 h. De manière assez similaire, les numériciens ont recherché les modes propres de leurs écoulements dans leurs modèles, en ajoutant à un écoulement de base (toujours en équilibre hydrostatique et géostrophique) une très petite perturbation. Une fois cette perturbation amplifiée, on stoppe la simulation, on retire la perturbation, on la re-normalise (pour rendre à nouveau son amplitude très petite devant celle des champs de l'état de base), puis on la rajoute à nouveau à l'état de base. Cette opération est répétée le nombre de fois qu'il faut pour arriver à dégager du "bruit numérique" un mode quasi-normal, qui va posséder une amplification quasi-exponentielle pendant un temps donné (généralement quelques jours). Ce mode évolue par la suite en un mode

non-linéaire, régime où son amplification devient sub-exponentielle et sa structure, d'amplitude finie, interagit fortement avec l'écoulement de base (voir par exemple les travaux de Joly et Thorpe, 1991, ou Lalaurette *etal.* , 1994). L'analyse par mode normal linéaire ne s'est pas cantonnée aux "grandes instabilités" de l'atmosphère, barocline ou barotrope. Je peux ici rappeler rapidement que cette approche a aussi longtemps été celle des numériciens pour étudier les instabilités dans les zones frontales, à méso-échelle. Ainsi, les modes propres possibles pour l'instabilité symétrique (Bennetts et Hoskins, 1979) ont-ils été étudiés dans des cadres idéalisés (Ducrocq, 1993) et sur des états de base tirés de cas réels (Fischer et Lalaurette, 1995 a et b) avec le modèle Péridot.

Cependant, la contrainte de ne s'intéresser qu'à des modes normaux, de structure constante pendant leur phase de croissance linéaire, s'est avérée être trop forte pour bien comprendre les mécanismes de croissance des erreurs, en régime linéaire. Dans la foulée des travaux de Farrell (1982, 1984), Joly (1995) a posé le problème de la croissance du mode linéaire la plus forte, comme un problème aux conditions initiales. La croissance s'exprime alors par rapport à une norme prédéfinie, et on recherche de fait le vecteur singulier (ou les vecteurs singuliers) du modèle résolvant $\mathcal{M}_{k-1}^{k}$ (voir section 2.1). Ces vecteurs singuliers sont définis comme les vecteurs propres de l'opérateur symétrique $(\mathcal{M}_{k-1}^{k})^{*}\mathcal{M}_{k-1}^{k}$, et ils ont donc des similarités avec les vecteurs propres d'autres opérateurs comme par exemple celui intervenant dans l'équation de croissance des erreurs du filtre de Kalman (équation 2.2). Par exemple, si le terme $P_{k-1}^{a} = I$ (c'est à dire des erreurs d'analyse homogènes et isotropes si on les représente dans l'espace spectral de Fourier), alors les vecteurs propres de P_{k}^{f} (les structures d'erreur qui se seront le plus amplifiées) seront les vecteurs singuliers du modèle résolvant $\mathcal{M}_{k-1}^{k}$ (notés v) auxquels on aura appliquer ce dernier ($\mathcal{M}_{k-1}^{k}v$).

Joly (1995) a appliqué le formalisme des vecteurs singuliers au modèle semi-géostrophique à tourbillon potentiel uniforme, que j'ai déjà introduit en section 2.3. J'ai pour ma part repris ce formalisme pour calculer "à la main" les vecteurs singuliers du problème d'Eady (Fischer, 1998). Je montre sur les figures 3.1 et 3.2 les taux d'amplification des vecteurs singuliers, en comparaison avec celui du mode normal. La définition du vecteur singulier dépend de la norme choisie pour estimer son amplification (ici, je montre l'énergie cinétique). Elle dépend aussi de la durée d'optimisation prescrite, et on obtiendra des vecteurs singuliers différents pour un temps d'amplification de 6 h ou de 48 h. On constate cependant que, pour le problème d'Eady, des taux d'amplification conséquents peuvent être obtenus pour des longueurs d'onde assez courtes, entre 1000 et 4000 km, soit inférieures à la longueur d'onde du mode normal optimal. Cela est surtout vrai pour des temps d'amplification courts (6h à 24h). Pour des temps d'amplification longs (48h ou 72h), le vecteur singulier optimal a une longueur d'onde proche de celle du mode normal, mais on note toutefois qu'il conserve un surplus d'amplification.

Les figures 3.3 et 3.4 montrent les structures du champ de vent transversal (la composante "v") pour le mode normal, le vecteur singulier 6h et le vecteur singulier 48h, respectivement, et pour différentes longueurs d'onde. Le mode normal à 4000 km de longueur d'onde présente une inclinaison vers l'Ouest de son champ de vent, classique des modes linéaires qui vont extraire de l'énergie de l'état de base par des processus de conversion barocline. Le vecteur singulier pour 6h a une structure assez similaire à 4000 km de longueur d'onde, et conserve cette structure inclinée vers l'Ouest pour toutes les autres longueurs d'onde, notamment celles plus courtes (cf pour 2222 et 3000 km). Par contre, le mode normal devient de plus en plus "vertical", au fur et à mesure que la longueur d'onde décroit, ce qui montre qu'effectivement il devient de moins en moins efficace à se développer (de façon barocline, seule conversion autorisée dans ce problème 2D). Enfin, le vecteur singulier à 48h d'échéance présente quasiment un décalage en opposition de phase entre la surface et le toit du problème. Ce déphasage est effectivement celui qui lui permettra, pendant le cours des 48h d'évolution linéaire, de progressivement passer par un décalage en quadrature de phase, lequel est celui qui maximise le taux d'accroissement instantané (Davies et Bishop, 1994).

L'étude du problème d'Eady 2D, et de ses vecteurs singuliers, m'a permis d'appréhender de façon "ludique" la notion de vecteurs singuliers. Il est satisfaisant de constater que mon étude aura inspiré d'autres auteurs, puisque Rivière *etal.* (2001) ont repris ce problème, en y ajoutant un terme d'effet β (variabilité transversale du facteur de Coriolis). Ils ont par ailleurs adopté une analyse mathématique très élégante, en séparant le modèle résolvant en partie symétrique et partie anti-symétrique, ce qui a simplifié leurs développements

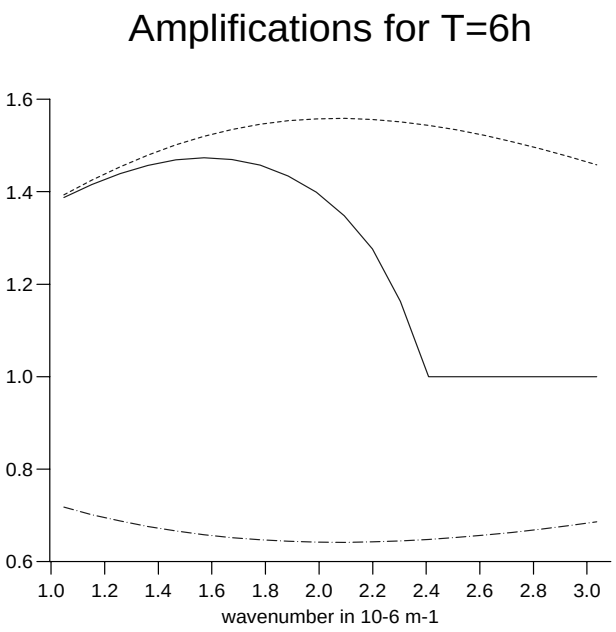


FIG. 3.1 – Taux d'amplification linéaire du mode normal (trait plein), du vecteur singulier croissant (tireté) et du vecteur singulier décroissant (pointillé), sur 6 h, en fonction du nombre d'onde horizontal.

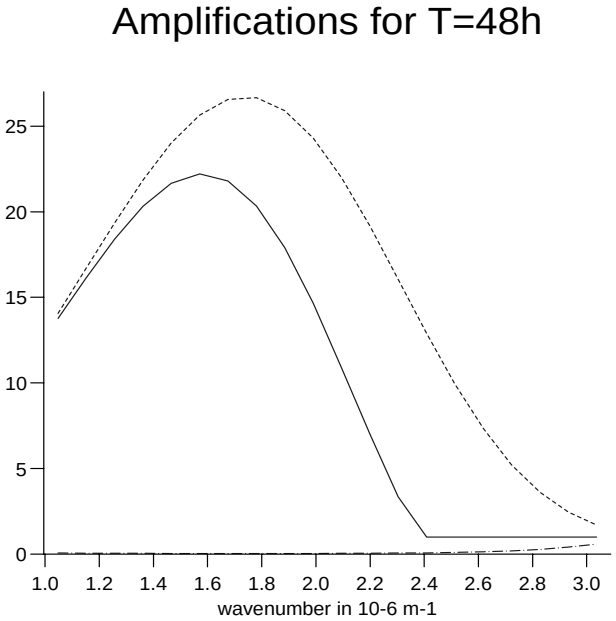


FIG. 3.2 – Taux d'amplification linéaire du mode normal (trait plein), du vecteur singulier croissant (tireté) et du vecteur singulier décroissant (pointillé), sur 48 h, en fonction du nombre d'onde horizontal.

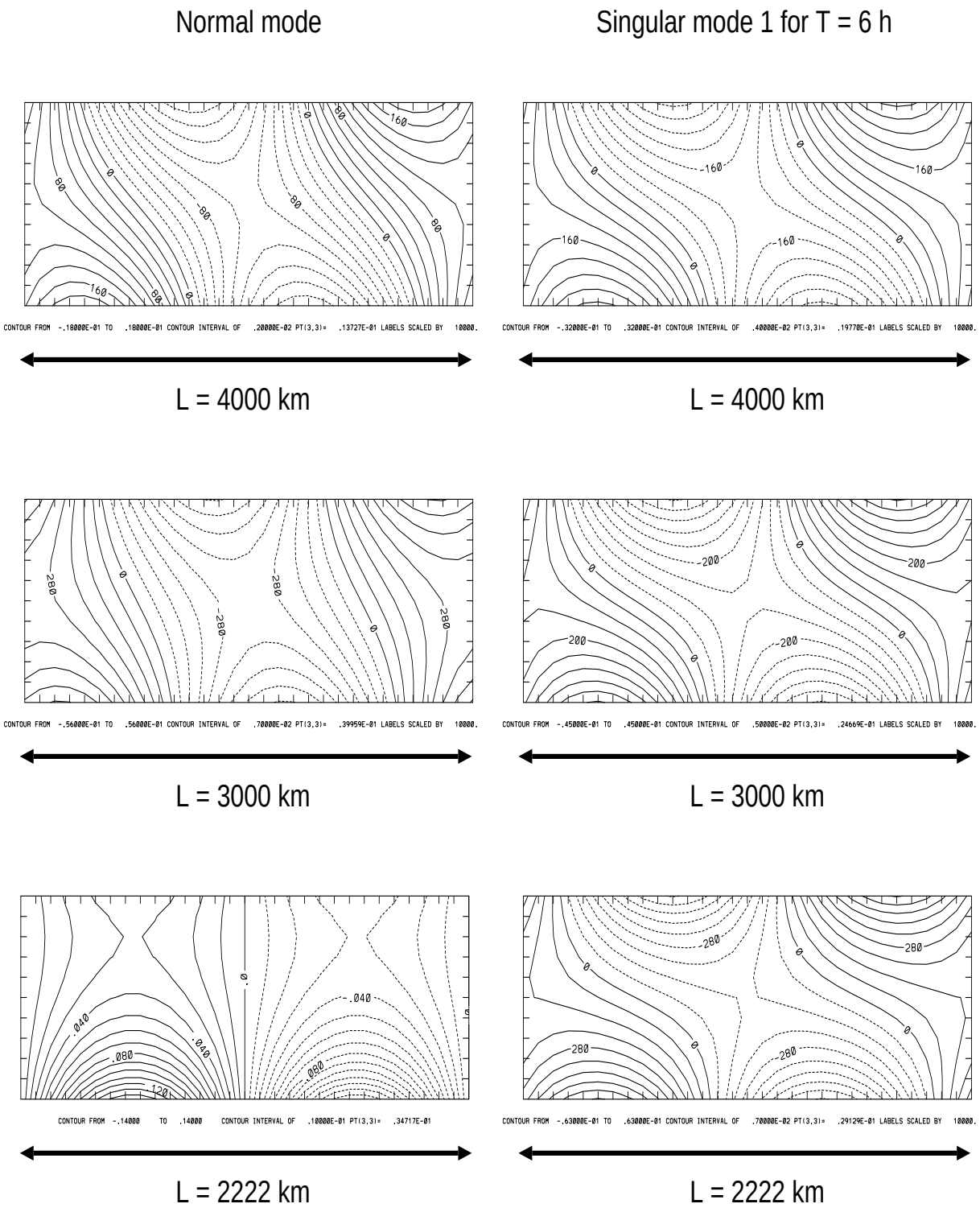


FIG. 3.3 – Structure du vent méridien v_g à l'état initial, pour le mode normal (à gauche) et pour le vecteur singulier (à droite), et pour différentes longueurs d'onde. Le temps d'optimisation est de 6h.

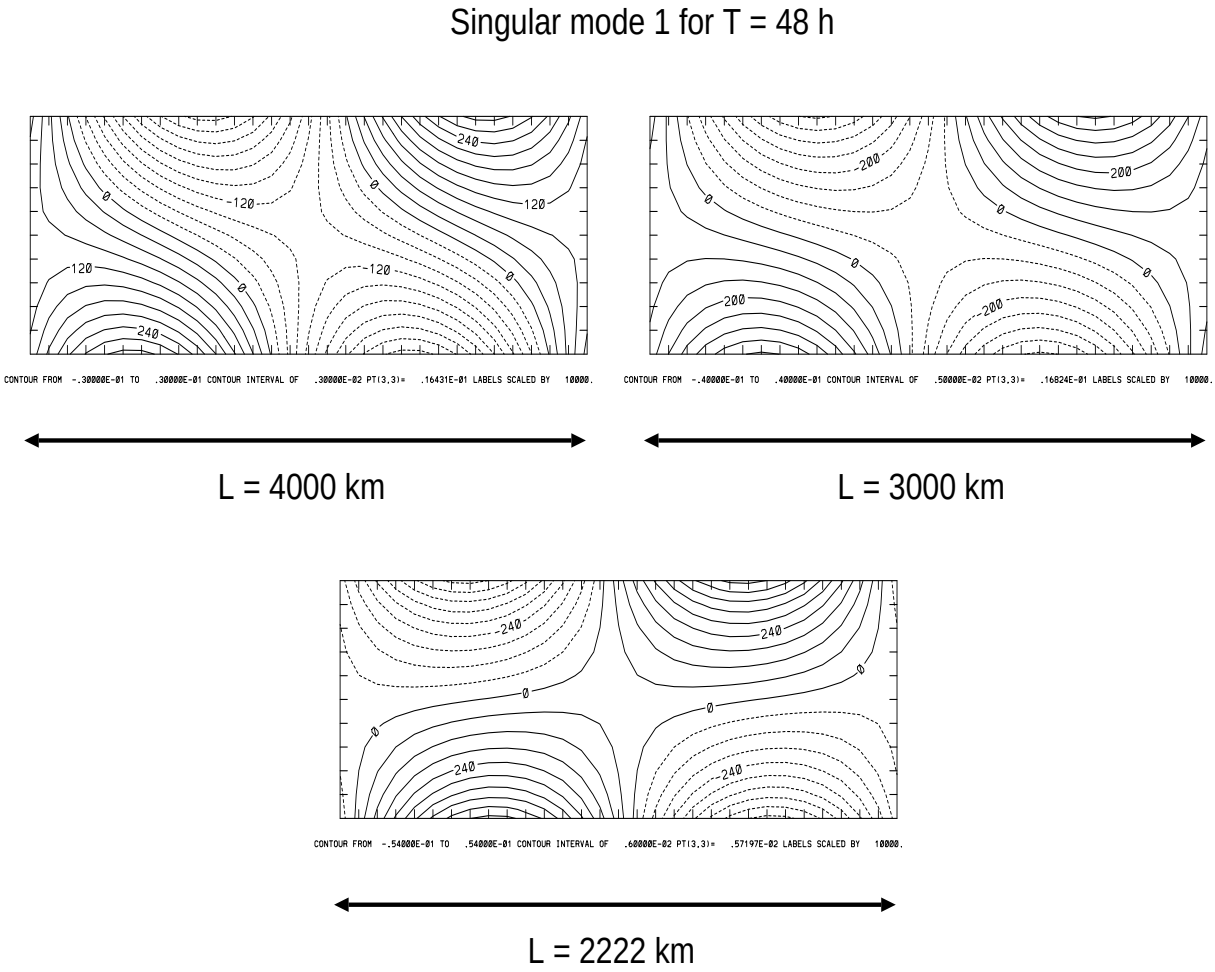


FIG. 3.4 – Structure du vent méridien v_g à l'état initial, pour le vecteur singulier et pour différentes longueurs d'onde. Le temps d'optimisation est de 48h.

analytiques.

Je signale aussi l'approche par "bred modes" (modes "ensemencés"), très prisée notamment outre-Atlantique (Toth et Kalnay, 1997). Il s'agit alors de laisser croître des perturbations en cours d'assimilation, en les renormant à chaque nouveau réseau d'analyse par rapport à une solution de contrôle. De cette manière, les bred modes s'obtiennent comme des modes de Lyapunov : ils sont en effet les modes *non-linéaires* qui peuvent s'amplifier à travers les réseaux d'analyse, contraints par ces derniers à travers la remise à jour systématique. La renormalisation est assez empirique, et son effet principal est de ramener les perturbations à des intensités petites. De plus, on procède à des recombinaisons (addition/soustraction des perturbations entre elles) afin de se donner les moyens d'explorer d'éventuelles directions nouvelles de croissance dans l'espace des phases du modèle. Il y a donc une part d'empirisme dans cette approche. Elle a le mérite de ne pas demander de très gros développement algorithmique (pas de modèle adjoint, pas d'algorithme de diagonalisation comme Lanczos pour les vecteurs singuliers). Un prolongement plus rigoureux semble être l'Ensemble Transform (Bishop et Toth, 1999), dans la mesure où la renormalisation se fait alors en appliquant une approximation de la matrice des covariances d'erreur d'analyse, qui doit effectivement être la norme adéquate pour la remise à jour des perturbations (par exemple : diminuer leur amplitude dans les zones de faible erreur d'analyse estimée, et les laisser plus intenses dans les zones pauvres en observations). Des études d'intercomparaison entre bred modes et vecteurs singuliers ont été réalisées (Majumdar *etal.* , 2002, en utilisant la technique de l'Ensemble Transform pour les bred modes), mais sans pouvoir mener à des conclusions fermes.

Dans le monde "réel" des systèmes de prévision, les vecteurs singuliers sont devenus un outil de référence pour l'initialisation des prévisions d'ensemble. Le CEPMMT notamment a beaucoup investi dans cette méthode (Buizza, 1992, Molteni *etal.* , 1996, Palmer *etal.* , 1998, Gelaro *etal.* , 1998) et le calcul des vecteurs singuliers est une composante essentielle pour l'initialisation de son système de prévisions d'ensemble "EPS"[1]. Cette technique a aussi été transposée à Météo-France dans le cadre du système "PEARP"[2]. On notera que cette approche des vecteurs singuliers répond surtout à la question de comment représenter ou estimer la structure des erreurs initiales du système, celles contenues dans le terme P_k^a à un moment donné de l'assimilation. Pour aller plus loin, et notamment pour représenter de façon approximative les erreurs du modèle, il faut ajouter des ingrédients supplémentaires. Bien que je n'aie que peu travaillé sur ces sujets, je vais toutefois rapidement évoquer des diagnostics possibles de l'erreur de modèle ainsi que l'approche stochastique dans les paramétrisations physiques dans la section 3.2.

3.2 Les erreurs du modèle

L'erreur modèle a été pendant un temps négligée dans les problèmes d'assimilation de données. Ainsi, les assimilations variationnelles ont-elles été développées et étudiées sous l'hypothèse du modèle parfait, pendant la durée de la fenêtre d'assimilation. Qu'implique cette hypothèse pour les propriétés de l'assimilation ?

L'hypothèse du modèle parfait appliquée au modèle résolvant $\mathcal{M}_{k-1}^k$ signifie que la solution recherchée, comme le "meilleur" compromis entre ébauche et observations (au sens de l'équation 2.5), et le long de la fenêtre d'assimilation (cad pour $k = 0, N$) doit exactement vérifier les équations du modèle. Cette propriété, associée à d'autres hypothèses, notamment sur la répartition (homogène) des observations de même (bonne) qualité, peut éventuellement être exploitée.

En effet, on peut considérer que, dans le modèle $\mathcal{M}_{k-1}^k$, sont contenus à la fois des modes stables (au taux de croissance linéaire $\lambda < 0$) et instables ($\lambda > 0$). L'analyse 4D-VAR d'un modèle parfait sera mieux "fittée" à la vérité (au sens d'une variance d'erreur au minimum $E[(Z_k^a - Z_k^v)^t(Z_k^a - Z_k^v)]$ petite, où $(Z_k^v)_{k=0,N}$ est la vérité) à droite de la fenêtre d'assimilation (k proche de N) pour les modes stables, et à gauche de la fenêtre d'assimilation (k proche de 0) pour les modes instables. Dans un modèle atmosphérique ou géophysique simplifié, on aura à la fois des modes instables et des modes stables, et donc la variance d'erreur d'analyse sera minimale quelque part à l'intérieur de la fenêtre d'assimilation. Cette propriété est par exemple illustrée dans

[1] Ensemble Prediction System

[2] Prévision d'Ensemble ARPège

les résultats expérimentaux obtenus avec un modèle quasi-géostrophique par Swanson, Vautard et Pirès (1998). La courbe de la variance d'erreur d'analyse au cours du temps ($k = 0, N$), quant à elle, aura une forme convexe (courbure positive) tant que la fonction-coût (2.10) reste à peu près quadratique (elle l'est exactement quand tous les opérateurs sont linéaires). Par ailleurs, au minimum, on a la relation d'orthogonalité suivante (Talagrand, 1998), *lorsque les erreurs d'observations sont décorrélées entre les instants* k :

$$\forall k, E[(X_k - H_k Z_k^v)^t (X_k - H_k Z_k^v)] = E[(Z_k^a - Z_k^v)^t H_k^t H_k (Z_k^a - Z_k^v)] + E[(X_k - H_k Z_k^a)^t (X_k - H_k Z_k^a)] \quad (3.1)$$

Sous l'hypothèse que les erreurs d'observations sont homogènes (constantes) et décorrélées dans la fenêtre d'assimilation, cette égalité implique une courbure négative (concave) pour la courbe de la variance des écarts observations moins analyse ($E[(X_k - H_k Z_k^a)^t (X_k - H_k Z_k^a)]$), le long de la fenêtre d'assimilation. On retrouve sous forme simplifiée cette propriété dans notre modèle "0-D" des §2.2 et 2.4, avec la relation très simple suivante pour un cas sans ébauche (pas de x_1 dans les formules du §2.4) et pour $\alpha = 0$:

$$(\xi_2 - y_2) = -e^{-\lambda t}(\xi_1 - y_1)$$

avec, rappelons-le, y_1 et y_2 les deux observations aux instants $k = 1, 2$, et ξ_1 et ξ_2 les deux valeurs au minimum de la fonction-coût aux deux instants d'observation. On voit que $E[(\xi_2 - y_2)^2] > E[(\xi_1 - y_1)^2]$ lorsque $\lambda < 0$ (mode stable), donc $E[(\xi_2 - \xi_2^v)^2] < E[(\xi_1 - \xi_1^v)^2]$ pour $\lambda < 0$, et vice versa lorsque $\lambda > 0$ (mode instable).

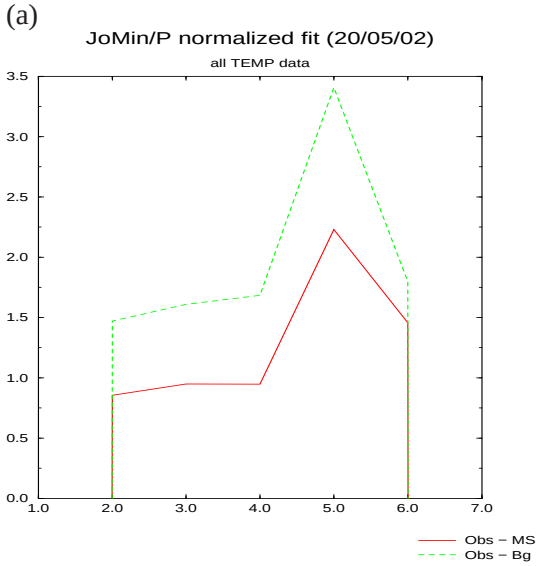
Dans le cas d'un modèle imparfait et pour un système d'assimilation "réel" (non idéalisé), l'impact de la non-prise en compte des erreurs-modèles pourrait alors se voir a posteriori, en traçant la moyenne sur un grand nombre de réseaux des écarts au carré observations moins analyse, divisés par les variances d'erreur d'observations et le nombre total de ces observations (le "normalized fit", chez Ménard et Daley, 1996). Une courbe de type parabolique (convexe), de courbure positive, indiquera l'existence d'erreur-modèle non prise en compte. J'ai tracé ce type de courbes à partir des fichiers d'observations du 4D-VAR/Arpège (voir figures 3.5 a, b et c). La comparaison des courbes pour des données de radiosondages (figure 3.5a), des données avion (figure 3.5b) et des radiances satellitales (figure 3.5c) ne montre aucun comportement systématique. J'en conclus qu'il n'est pas possible, pour un 4D-VAR évoluant sur des fenêtres de 6 h, de mettre en évidence de manière systématique (statistique) l'erreur-modèle (ces travaux avaient été réalisés dans le cadre d'un projet PATOM). Un problème supplémentaire ici est que les observations ne sont pas équitablement distribuées dans les différents time-slot (les instants k) de l'assimilation. Le CEPMMT, qui utilise à présent des fenêtres d'assimilation sur 12 h, et pourrait souhaiter augmenter encore cet intervalle (24 h), étudie ce problème à présent avec beaucoup d'intérêt. Les travaux de Mike Fisher et de Yannick Trémolet visent ainsi à représenter l'erreur-modèle comme une contrainte faible supplémentaire (fonction-coût additionnelle), avec une composante supplémentaire dans le vecteur de contrôle (Fisher, 2004, Trémolet, 2007).

Par une approche un peu différente, basée sur l'étude des écarts entre les analyses vérifiantes et les prévisions pour différentes échéances, Orrell avait également conclu à l'existence bien sensible de l'erreur-modèle (Orrell et Smith, 2000). Ses travaux visent surtout l'erreur dans le modèle non-linéaire. Notons que l'erreur-modèle est en réalité déjà prise en compte dans certaines applications, de manière plus ou moins heuristique, et pour les modèles non-linéaires. C'est bien ainsi qu'il faut comprendre l'application d'une physique "stochastique" dans le modèle IFS du CEPMMT dans le cadre de sa prévision d'ensemble EPS (Buizza *etal.* , 1999, Palmer, 2001).

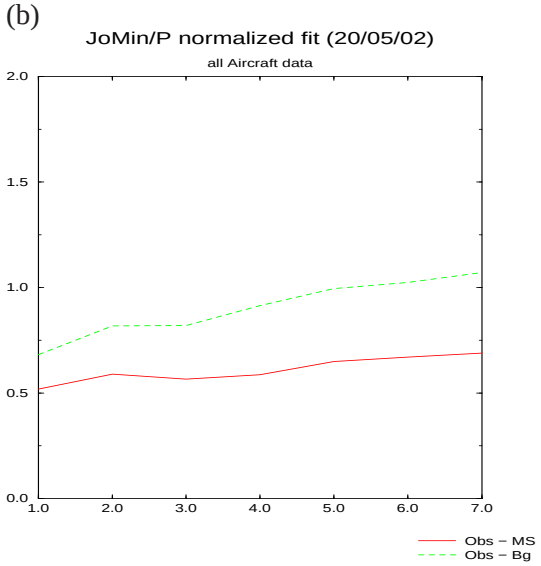
En effet, dans les méthodes ensemblistes, l'erreur-modèle est implicitement prise en compte lorsque l'on échantillonne sur l'ensemble des membres (prévus). Ainsi, l'équation 2.7 ne fait-elle pas la différence entre la croissance d'une erreur initiale (mesurée comme la dispersion entre les états initiaux des membres) et celle apportée en plus par l'erreur-modèle (que celle-ci amplifie ou amortisse la dispersion totale), sur la durée de la prévision (généralement, 6 h). Ce constat pourrait plaider dans un avenir plus ou moins proche pour un intérêt accru en faveur des méthodes d'ensemble.

Enfin, les services opérationnels n'ont pas attendu que la communauté des assimilateurs de données s'intéresse à ce type de problème. Les erreurs-modèles, surtout leurs composantes systématiques (les biais), sont

(a)



(b)



(c)

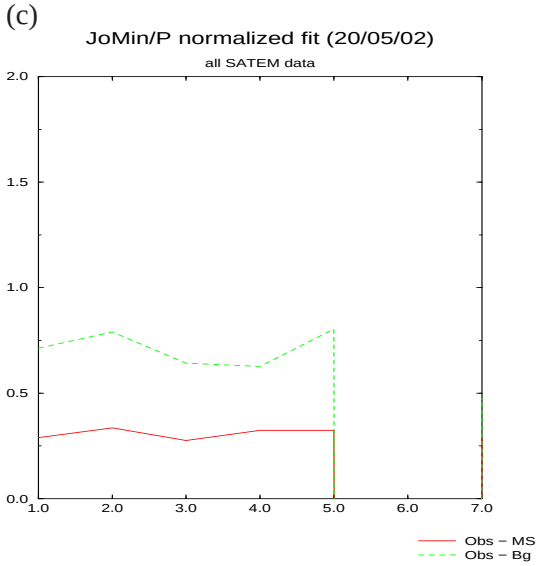


FIG. 3.5 – Evolution temporelle, dans la fenêtre d'assimilation du 4D-VAR Arpège, des écarts au carré observations moins ébauche (trait tireté) et observations moins analyse (trait continu) normés par les variances d'erreur d'observations et le nombre total d'observations par timeslot (tous champs mesurés confondus). (a) radiosondage ($P_{total} \approx 70000$) ; (b) avion ($P_{total} \approx 36000$) ; (c) radiances satellite ($P_{total} \approx 23000$).

bien connus et parfois corrigés en aval de la production brute du modèle (adaptation statistique des sorties de modèle par des régressions linéaires reliant certains champs du modèle, surtout diagnostiques comme la température à 2 m, à des prédicteurs observés). Les prévisionnistes aussi "flairent" les comportements systématiques : le modèle Arpège a été longtemps taxé de prévoir trop de cyclogénèses rapides (modèle trop "cyclonique"), et aujourd'hui Aladin est parfois critiqué de produire des "Aladinades" (en fait, des très fortes précipitations, très localisées, et irréalistes en général, de sorte que les prévisionnistes n'hésitent pas à les identifier et à les filtrer dans leur travail quotidien). Bien sûr, ce type d'erreur-modèle peut conduire à des non-détections dans le cas où l'événement "aberrant" se produit réellement.

3.3 La modélisation des covariances d'erreurs de prévision

Pour clore ce chapitre sur la forme et la croissance des erreurs de prévision, je souhaite faire un tour d'horizon des trois principales méthodes de calcul des statistiques de covariances d'erreur (la matrice P^b ou P_k^f du chapitre 2). Cette section ne fait pas spécialement référence à des travaux personnels, bien que j'ai eu l'occasion de travailler sur ce sujet et de calculer à plusieurs reprises des statistiques pour le système Aladin (voir Sadiki *etal.* , 2000 ou Široka *etal.* , 2003). Il me permet plutôt de situer cette problématique dans le cadre plus large des travaux dans la communauté internationale, et dans la communauté IFS/Arpège/Aladin en particulier. Je reviendrai sur certains aspects de manière plus personnelle dans tout le chapitre 4. De fait, il faut séparer trois aspects dans cette problématique :

- l'échantillonnage des erreurs : ce dernier se fait toujours en calculant des moyennes sur des différences, donc en fait sur des dispersions entre des jeux de prévisions, que l'on va supposer représentatives des erreurs.
- la modulation, ou le réglage a posteriori, des variances d'erreur : il s'agit de déterminer au moins la variance totale d'erreur de prévision, afin de mieux équilibrer le poids respectif de l'ébauche et des observations dans l'analyse.
- la projection des statistiques estimées dans une base de fonctions, qui seront le support des fonctions de structures (des corrélations) de la matrice P^b. Ces fonctions de structure dicteront très directement comment les innovations seront propagées sur la grille de l'analyse. Elles contrôlent donc le filtrage spatial de l'information apportée par les innovations.

La plus ancienne méthode de calcul pour P^b a consisté à effectuer des statistiques sur des séries temporelles de vecteurs d'innovation $X - \mathcal{H}(Z^b)$ (Hollingsworth et Lönnberg, 1986, Lönnberg et Hollingsworth, 1986). Une fois obtenues les statistiques dans l'espace des observations, rangées par groupes de distances horizontale et verticale, celles-ci sont ajustées avec des fonctions analytiques simples (de type Gaussiennes). On obtient ainsi à la fois un échantillonage de P^b et une modélisation de ses fonctions de structure. Cependant, le fait de travailler avec l'innovation implique qu'il faut savoir séparer de manière heuristique l'erreur d'ébauche de l'erreur d'observation. Dans la pratique, la séparation se fait en supposant que seules les erreurs d'ébauche sont corrélées spatialement, et on sépare les deux variances σ_b^2 et σ_o^2 en prolongeant par continuité la courbe des covariances vers l'axe du zéro : l'intersection fournit le σ_b^2, et le reste donne le σ_o^2. On note ici cette propriété très fondamentale des systèmes d'assimilation, et plus particulièrement des informations fournies par les innovations : pour passer d'une statistique sur l'innovation aux statistiques séparées de l'ébauche et des observations, une hypothèse supplémentaire sur ces deux statistiques est nécessaire. Il s'agit en général de reporter toutes les auto-corrélations spatiales sur P^b.

Dans le but de ne considérer d'emblée que des échantillons issus du modèle, Parrish et Derber (1992) ont échantillonné des différences entre deux échéances de prévision différentes, mais valables pour la même heure de vérification, par exemple des différences entre des prévisions 48 h moins 24 h, valables pour le réseau 0 UTC. Cette méthode, d'abord mise en oeuvre au NCEP[3] Américain (alors le NMC[4]), a été baptisée la méthode du NMC. Elle permet effectivement de se concentrer seulement sur des données modèle. L'impact

[3]National Centers for Environmental Prediction
[4]National Meteorological Center, ancêtre du NCEP

des assimilations est ressenti indirectement par le fait que les prévisions courtes (les P24) doivent être plus proches de la réalité que les prévisions longues, puisque 24 h d'assimilation de données supplémentaires ont été effectuées. Cependant, cette méthode a été reconnue comme produisant des portées d'auto-corrélation trop longues (Van Tuyl et Daley, 1999) dans les zones sans observations. En effet, la prise en compte des effets des analyses est très partielle, et peut être dominée par les défauts du modèle, ce d'autant plus que ce sont des prévisions assez longues, P24 et P48, qui sont échantillonnées. Cela favorise l'extension spatiale des effets dispersifs et tend à une surestimation des portées horizontales des erreurs (Berre *etal.* , 2006). De plus, même l'estimation des variances totales reste assez incertaine et, dans la pratique, tous les Centres opérationnels ont ré-ajusté a posteriori les variances totales (Arpège 3D-VAR et 4D-VAR ont ainsi longtemps possédé un facteur de réduction des variances totales d'erreur d'ébauche de 0.81). Notons que les fonctions de structure sur lesquelles sont projetées les statistiques d'erreurs sont généralement des fonctions dans l'espace spectral, avec des hypothèses d'homogénéité et d'isotropie. Ainsi, la matrice P^b d'Arpège est-elle développée dans l'espace des harmoniques sphériques (ce qui fut aussi jusqu'à récemment le cas de l'IFS, voir Courtier *etal.* , 1998). Cette approche a aussi été transposée au modèle Aladin, dans l'espace bi-Fourier, par Berre (2000). L'homogénéité (invariance des fonctions de structure par translation) et l'isotropie (invariance par rotation) ont pour effet de produire des incréments d'analyse circulaire, dans le cas d'une seule observation (voir par exemple Široka *etal.* , 2003, pour Aladin/Centre Europe, ou Gustafsson *etal.* , 2001, pour le système Hirlam). Ce n'est qu'assez récemment que des modélisations alternatives, dans des familles de fonctions différentes, ont été proposées. Je ne m'étendrai pas beaucoup plus sur ce sujet, sur lequel je n'ai pas eu d'activité personnelle, mais on peut citer les travaux de Mike Fisher au CEPMMT avec des familles d'ondelettes non-orthogonales (Fisher, 2003) et les travaux d'Alex Deckmyn pour Aladin (Deckmyn et Berre, 2005).

La troisième technique d'échantillonnage est tirée des méthodes ensemblistes. Il s'agit d'exécuter un ensemble d'analyses et de prévisions, même de taille modeste (une dizaine de membres), pendant une assez longue période de temps (sur 30 à 60 jours). Les analyses diffèrent alors par les perturbations apportées sur les jeux d'observations (normées par leurs erreurs R_k), puis, au bout de quelques réseaux, aussi par leurs ébauches qui auront divergé (Houtekamer *etal.* , 1996, Fisher, 1999, Belo Pereira et Berre, 2006). Fisher (1999) et Berre *etal.* (2006) ont montré que cette approche par "system simulation" produit des dispersions de prévisions dans lesquelles l'effet des analyses (donc des observations qui viennent contrôler le système) est assez immédiat : les corrélations des dispersions sont exactement celles recherchées pour les erreurs, et les variances sont doubles. Cependant, dans la pratique, l'estimation des variances totales n'est toujours pas parfaite, et actuellement on enfle par exemple les variances totales d'ébauche d'Arpège d'un facteur 2.25. De manière indirecte, Stefanescu *etal.* (2006) ont calculé des statistiques d'ensemble pour le 3D-VAR d'Aladin, en effectuant simplement des prévisions par adaptation dynamique à partir des ensembles d'analyses et de prévisions d'Arpège. Lorsque nous avons testé ces statistiques en mode pré-opérationnel, nous avons également constaté la nécessité de réajuster a posteriori les variances totales (de fait, Aladin/France opérationnel actuel utilise le même facteur d'amplification des variances d'ébauche qu'Arpège : 2.25). Signalons également que la nécessité d'une inflation (plus généralement, d'une multiplication a posteriori) des variances d'erreurs d'ébauche a été reconnue dans tous les travaux d'estimation de P^b dans des méthodes de filtrage entièrement ensemblistes (voir par exemple Anderson, 2007, Etherton et Bishop, 2004). Je reviendrai dans la section 4.3 sur le réglage a posteriori des variances d'erreurs d'ébauche dans le cadre de l'analyse Aladin.

Chapitre 4

Application à un système réel

Je présente ici une vue d'ensemble des travaux que j'ai encadrés, ou co-encadrés, dans le cadre du développement de l'analyse variationnelle pour le système à aire limitée Aladin. Les éléments, personnels ou en vue d'ensemble, que j'ai développés dans les chapitres 2 et 3 doivent permettre au lecteur de situer ces travaux dans un contexte plus large, à la fois des innovations qui se font dans la communauté R&D et dans le cadre de la collaboration toujours forte entre l'IFS (CEPMMT), Arpège (CNRM/GMAP) et Aladin (partenaires du projet Aladin).

4.1 L'assimilation 3D-VAR Aladin

Le système Aladin (Radnoti *etal.* , 1995, Horanyi *etal.* , 1996), tout comme par exemple le système Hirlam, repose sur une combinaison de calculs effectués en point de grille (les termes non-linéaires de la prévision, l'interpolation au point d'origine des trajectoires du schéma Semi-Lagrangien, l'interpolation aux points d'observation des valeurs du modèle etc...) et dans l'espace spectral défini par une décomposition bi-Fourier (diffusion horizontale, partie verticale du Semi-Implicite, convolution par les fonctions de structure de P^b). Bien que la version hydrostatique d'Aladin dispose de ses versions tangente linéaire et adjointe (mais, jusqu'à peu, seulement pour un schéma d'avance temporelle Eulérien à trois pas de temps), nos efforts se sont concentrés sur le développement de l'analyse 3D-VAR, et non pas directement sur un 4D-VAR. Ce choix, qui peut sembler naturel dans la mesure où c'est également ainsi que sont apparues les analyses variationnelles globales dans l'IFS et dans Arpège, était aussi dicté par l'impossibilité de dégager des moyens humains pour les questions très spécifiques au 4D-VAR (tangent linéaire et adjoint du schéma Semi-Lagrangien à deux pas de temps, physique simplifiée pour Aladin, étude d'une stratégie multi-incrémentale). Le problème du coût numérique exorbitant d'un 4D-VAR sur un domaine d'environ 3000 par 3000 km, de résolution 10km, s'est aussi posé. Comme je vais le montrer dans la suite de ce chapitre, le développement d'une analyse 3D-VAR permet déjà d'aborder toute une série de questions scientifiques, communes aux systèmes globaux ou spécifiques aux systèmes à aire limitée.

La stratégie de développement retenue, que nous évoquons dans Fischer *etal.* (2005), a consisté à réutiliser beaucoup des développements de code et des innovations scientifiques du projet IFS/Arpège. Cette approche est certes dictée par des considérations de forces de travail, mais elle se justifie également par des considérations plus techniques :

- la géométrie des opérateurs d'observations ne varie pas significativement entre des résolutions de l'ordre de 20-30 km, et 10 km. On continue à considérer des opérateurs verticaux pour les radiosondages ou les profils de radiances. Tout au plus peut-on prendre en considération la dérive des radiosondages, par exemple, mais ce point ne justifie pas de changer toute une algorithmique.
- la physique des opérateurs d'observations, tant qu'elle ne fait pas intervenir (trop) les paramétrisations physiques du modèle lui-même, sera toujours la même. Ainsi, nous profitons dans Aladin de toute l'infrastructure d'installation des codes de transfert radiatif développée surtout au CEPMMT

pour IFS/Arpège (y compris les remises à jour). Une divergence de cette règle peut apparaître pour des opérateurs ayant besoin de champs fournis par la physique du modèle, et si cette physique diffère entre les modèles globaux et aire limitée. Cela n'est pas le cas entre Arpège et Aladin/France, mais le sera par exemple entre Arpège et Arome (Ducrocq *etal.* , 2005), puisque ce dernier aura une description assez complète de la microphysique. Cependant, là encore, une divergence ne se fera sentir que sur des opérateurs qui réellement utiliseront ces champs différents (on peut penser à des champs prognostiques d'eau nuageuse, d'eau précipitante, pour la prise en compte des nuages dans le transfert radiatif ou l'assimilation des réflectivités radar).

– nous conservons la formulation incrémentale développée dans IFS/Arpège. Nous ne l'utilisons pas dans le 3D-VAR pour réduire la résolution de l'incrément (nous ne faisons pas de l'incrémental spatial), mais nous adoptons tout le code et le flot de données qui lui est associé, notamment en terme de linéarisation des opérateurs d'observations et d'organisation de la boucle interne de la minimisation.
– en retour, des développements faits dans Aladin 3D-VAR peuvent être récupérés assez facilement dans Arpège (cela pourrait s'appliquer à l'assimilation des données du capteur SEVIRI de Météosat 8, ou aux données de vent radar plus tard).

Travailler avec un système à domaine limité pose deux questions très spécifiques : comment traiter l'existence des bords latéraux dans l'analyse, et comment coupler cette analyse à aire limitée avec le système global (le coupleur). Dans la mesure où les fonctions de structure pour Aladin ont été inspirées de ce qui se fait dans les systèmes globaux (voir section 3.3), nous avons donc des corrélations homogènes, identiques sur tout le domaine. Cependant, la bipériodisation des champs, qui permet ensuite la transformation des champs point de grille en champs spectraux bi-Fourier (Machenhauer, 1979) implique de travailler, pour toutes les opérations spectrales, dans une géométrie en forme de tore. Notamment, les convolutions par les fonctions de structure de P^b se font dans cette géométrie, et la conséquence en est que les incréments d'analyse, propagés sur la grille, sont bipériodiques. Ils traversent donc les bords physiques du domaine, se propagent à travers la zone de bipériodisation des champs, et peuvent ré-apparaître par le bord opposé (voir dans Široka *etal.* , 2003, figure 6, par exemple). Si l'on tente de forcer ces incréments à zéro près des bords, alors on crée des gradients d'incréments aberrants, à travers les relations multi-variées qui existent entre les prédicteurs (Derber et Bouttier, 1999, Berre, 2000), ce qui est illustré dans Gustafsson *etal.* (2001). Dans le système Hirlam, cette difficulté liée à la bipériodisation des champs avant le passage en spectral est partiellement contournée en prenant des zones de bipériodisation grandes par rappport aux portées des corrélations horizontales. Dans Aladin, nous avons essayé de prescrire des statistiques de P^b ayant implicitement des portées assez courtes, et nettement plus courtes que celles données par la méthode du NMC notamment. Ainsi, Široka *etal.* (2003) proposent une version modifiée de la méthode du NMC, la version "lagged-NMC". Dans cette dernière, les prévisions courtes (à 12 h d'échéance dans Aladin) sont calculées en utilisant les *mêmes* conditions à la limite latérale que les prévisions longues (à 36 h ici). Les conditions limites étant identiques dans les deux jeux de prévisions, les dispersions échantillonnées sont surtout sensibles aux échelles moyennes et petites représentées dans le domaine de calcul, car les plus grandes échelles sont contraintes dans les deux jeux de prévision par les mêmes données aux bords (ce dernier résultat avait été suggéré par les résultats de comparaison entre des statistiques de données de couplage et des statistiques de prévisions, dans Sadiki *etal.* , 2000). Certes, avec des statistiques "lagged-NMC", nous avons obtenu des portées plus courtes, mais bien sûr le problème intrinsèque des incréments bipériodiques reste posé. Par ailleurs, il faut signaler que la méthode "lagged-NMC", si elle semble bien corriger les portées horizontales en les réduisant par rapport à la méthode NMC standard, ne permet plus d'obtenir les variances totales de l'erreur d'ébauche. En effet, les variances totales estimées par cette méthode sont bien trop petites. Ce constat avait notamment motivé le travail de thèse mené par Wafaa Sadiki sur le réglage a posteriori des variances d'erreur, que je développerai davantage dans la section 4.3. Les statistiques "lagged-NMC" sont aujourd'hui quasiment abandonnées dans la communauté Aladin, au profit surtout de statistiques tirées d'ensemble d'analyses Arpège (voir section 3.3).

La question du couplage de l'analyse variationnelle à aire limitée avec le système global est abordée dans le cadre du travail de FCPLR (Formation Complémentaire Par La Recherche) de Vincent Guidard (co-encadré par moi et par François Bouttier). L'approche, que je détaillerai dans la section 4.4, consiste à utiliser l'analyse

du système coupleur, projetée dans la grille du domaine limité et tronquée aux échelles représentatives, comme une nouvelle source d'information, au même titre que l'ébauche Aladin et que les observations. Cette technique peut être comprise, dans un sens plus large, comme l'une des techniques possibles de "blending" (c'est à dire littéralement de "mélange") des champs de l'analyse globale avec ceux de l'ébauche du modèle à aire limitée.

Où en sommes-nous des résultats expérimentaux de l'assimilation 3D-VAR ? Ce système est devenu opérationnel le 25 juillet 2005, sur la base de plusieurs semaines de chaîne en double, pour comparer ses performances avec celles de l'adaptation dynamique à partir de l'analyse Arpège. Je montre dans les figures 4.1 et 4.2 les scores obtenus pendant ces tests, pour l'humidité relative à 2 m (réseaux de surface Synop[1] et Radome[2]) et pour la température (radiosondages). On constate en général les points suivants :

- l'assimilation améliore considérablement les scores à 6 h et à 12 h pour l'humidité relative et la température à 2 m mais aussi les précipitations (non montré). Cette amélioration reste sensible aux échéances ultérieures, 18 h voire 24 h. Ce bénéfice est dû à plusieurs ingrédients : l'assimilation des humidités relatives et des températures à 2 m observées (elles ne sont pas assimilées dans Arpège), le meilleur spin-up (ou équilibre initial, voir §5.3) du modèle en assimilation (qui se traduit notamment par des niveaux inférieurs du modèle moins (sur-)chargés en humidité).
- effectivement, les précipitations prévues à 12 h et à 18 h d'échéance sont plus réalistes sur les études de cas. Par contre, à 6 h d'échéance, le spin-up du modèle reste sensible et les pluies prévues à très courte échéance, lors du dernier réseau d'assimilation pour une échéance ultérieure, restent critiquées par les prévisionnistes.
- les scores aux radiosondages montrent que l'assimilation se comporte un peu mieux que l'adaptation dynamique pour le vent (sachant que l'adaptation dynamique du vent est un point fort d'Aladin en général). Sur la température, on constate que les prévisions tirées de l'assimilation soit s'éloignent, soit se rapprochent des mesures des sondages. Cela est dû à plusieurs facteurs : nous assimilons une grande quantité de températures de brillance via les radiances du capteur SEVIRI (un éloignement des radiosondages n'est donc pas forcément un signe de plus mauvaise prévision) ; nous savons qu'il existe un biais chaud dans la stratosphère du modèle (que le 3D-VAR ne corrige pas très bien) ; dans la couche-limite enfin, le 3D-VAR produit un léger biais positif de température qui reste assez inexpliqué à l'heure actuelle.

4.2 Les erreurs propres au modèle Aladin : synthèse de l'article "Mesoscale background error covariances : recent results obtained with the limited area model ALADIN over Morocco", par Sadiki, Fischer et Geleyn, paru dans M.W.R. (2000)

Une question assez spécifique, qui s'est posée dans le cadre de l'assimilation à aire limitée, était de comprendre quelles étaient les erreurs que l'on devait chercher à corriger. En effet, dans un système comme Aladin, couplé avec Arpège, si les erreurs étaient essentiellement dictées par le modèle coupleur (erreurs initiales projetées d'Arpège, erreurs ne rentrant par la suite que par les conditions limites latérales), alors il suffirait d'améliorer l'assimilation globale, donc ici le 4D-VAR Arpège, ainsi que la *formulation* du couplage latéral, sans se préoccuper du tout de l'utilisation directe des données dans une analyse Aladin. Toute amélioration dans Arpège apporterait alors de façon mécanique une amélioration de la prévision Aladin. Cependant, de manière un peu intuitive, on imagine bien que l'augmentation de résolution (d'un facteur 2 en général entre Arpège et Aladin/France) et la diffusion horizontale (d'origine numérique) réglée différemment, vont amener le modèle Aladin à décrire des échelles (des structures) plus fines qu'Arpège, avec des forçages de petite échelle spécifiques (une orographie plus précise, une description des sols et des côtes plus fine). Pour vérifier cette idée, Wafaa Sadiki, dans la première partie d'un travail de recherche autour du 3D-VAR Aladin, a comparé

[1] Observations de surface pour le réseau mondial d'observation, transmises sur le GTS

[2] Réseau Français dense de capteurs de surface

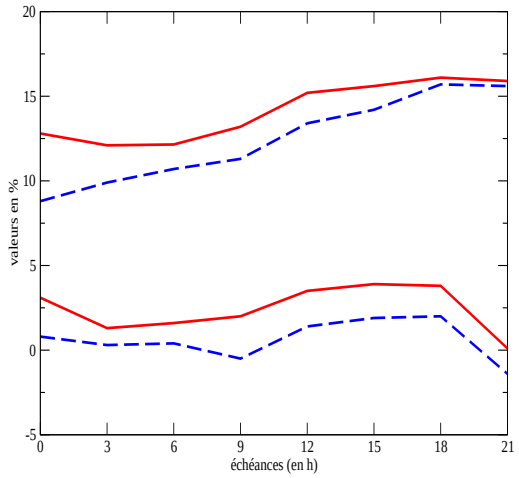


FIG. 4.1 – Scores des prévisions tirées à partir de l'assimilation 3D-VAR (bleu tireté) et de l'adaptation dynamique (rouge continu), en fonction des échéances, par rapport aux mesures d'humidité relative à 2 m. Courbes du bas : biais, courbes du haut : écart quadratique moyen.

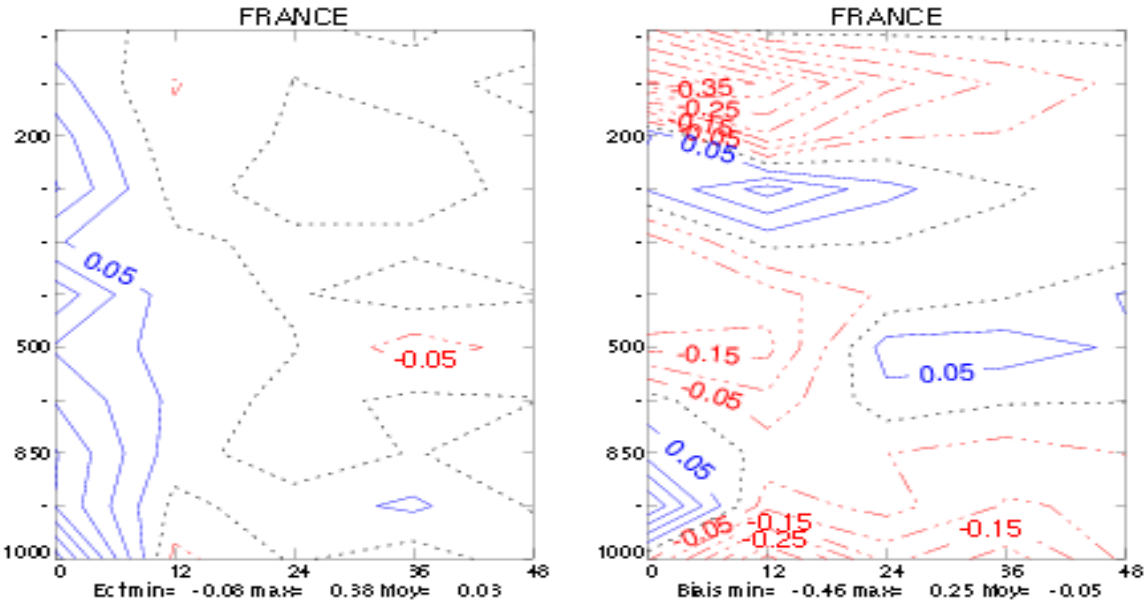


FIG. 4.2 – Scores des prévisions tirées à partir de l'assimilation 3D-VAR et de l'adaptation dynamique, en fonction des échéances, et en terme de différences de différences : bleu (continu) si le 3D-VAR est plus proche des observations, rouge (pointillé) si c'est l'adaptation dynamique. A gauche : pour l'écart-type aux observations ; à droite : pour le biais aux observations.

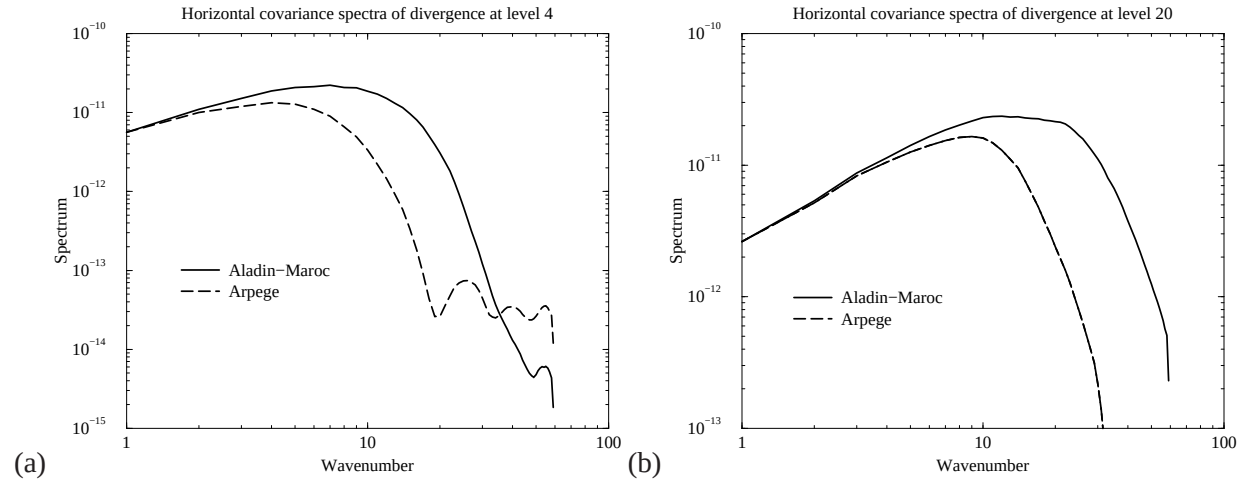


FIG. 4.3 – Spectres des covariances horizontales pour la divergence. Trait plein : données du modèles Aladin/Maroc ; trait tireté : données de couplage. (a) niveau modèle 4, (b) niveau modèle 20. L'unité est le s^{-2}.

des statistiques d'erreur de type NMC en utilisant deux échantillons de données : des champs coupleurs (donc des champs Arpège interpolés sur une grille Aladin), et des prévisions propres au modèle Aladin. En guise d'exemple, je montre ici la figure 4.3 qui illustre l'un des principaux résultats publiés dans Sadiki *etal.* (2000). On constate que le spectre horizontal des variances d'erreur est "rempli" (complet sur tous les nombres d'onde) pour des données de prévision, mais il est tronqué pour des données de couplage. Ainsi, la dispersion dans la méthode NMC nous renseigne sur toute une gamme d'échelle sur laquelle le modèle Aladin génère sa propre dynamique, ainsi que ses propres erreurs. De manière assez générale, nous avons constaté que les spectres de variances d'Arpège et d'Aladin se séparaient l'un de l'autre environ autour du nombre d'onde 10, soit autour de 300 km de longueur d'onde. Wafaa Sadiki a obtenu ses résultats à partir de données d'Aladin/Maroc, mais ce résultat a été confirmé sur Aladin/France. D'autres approches, en étudiant un modèle académique 2D shallow water (Vincent Guidard) ou en cyclant simplement des petits ensembles de prévisions Aladin (Robert Miller et François Bouttier), ont confirmé également ce résultat. Au sens des erreurs de prévision, dans un système couplé du type Arpège/Aladin, il existe des sources d'erreur de prévision propres à la dynamique interne du modèle couplé, qu'il y a lieu de corriger. Ces erreurs propres se situent dans une gamme d'échelle intermédiaire, entre les échelles fortement contraintes par le couplage aux bords (des grandes échelles) et celles où la diffusion horizontale d'origine numérique (typiquement en laplacien puissance 2 ou 3) se fait sentir. Pour des domaines de l'ordre de 3000 km de côtés, des résolutions entre 10 et 16 km, et des rapports de couplage entre les résolutions Arpège et Aladin entre 2 et 3, nous avons constaté que l'intervalle des longueurs d'onde concerné était typiquement $[300 - 400km, 80 - 100km]$. Notons que la limite inférieure doit dépendre de l'intensité de la diffusion numérique dans le modèle (qui contrôle les ondes courtes "énergétiques" en queue de spectre), ainsi que de la méthode d'échantillonnage (la queue du spectre est moins énergétique si on échantillonne des prévisions 6h, comme dans la méthode ensembliste, plutôt que des prévisions longues de 12 et 36 h, comme dans la méthode NMC). On obtient cependant une idée grossière des échelles de dynamique d'erreur propre au modèle couplé.

Notons au passage que le travail de Wafaa Sadiki a également permis de comparer les statistiques d'erreur (de type NMC) entre les données d'Aladin-Maroc et celles déjà publiées par Loïk Berre avec Aladin/France. Wafaa Sadiki a montré que les statistiques sur le Maroc étaient somme toute assez semblables à celles sur la France, bien que son domaine d'application soit nettement plus au Sud (donc latitudes plus basses). Son travail a montré qu'il n'était pas nécessaire de revenir dans le détail sur la formulation de la matrice P^b pour le Maroc, et c'est effectivement avec une formulation très proche d'Aladin/France que nos collègues Marocains testent aujourd'hui de manière quotidienne l'assimilation 3D-VAR, sur un domaine au demeurant plus vaste que celui étudié par Wafaa (il recouvre toute la moitié Nord du continent Africain).

Comme je l'aborderai assez rapidement dans la section 4.6, une deuxième justification pour une analyse

de données à échelle fine, dans un modèle comme Aladin, est la présence d'observations denses, peu ou pas assimilées dans Arpège. En effet, sur une grille d'analyse plus fine, on peut espérer assimiler des réseaux d'observations plus denses, et représenter des structures de l'écoulement plus fines que dans le système global. Par ailleurs, à aire limitée, on peut espérer assimiler de plus grandes quantités d'observations locales, sans avoir à payer le prix d'une assimilation globale (dans laquelle par exemple il faut aussi traiter les Tropiques, les Antipodes etc.).

4.3 Le réglage des variances d'erreurs de P^b : synthèse de l'article "A posteriori validation applied to the 3D-VAR Arpège and Aladin data assimilation systems", par Sadiki et Fischer, paru dans Tellus (2005)

J'ai déjà discuté différentes méthodes d'échantillonnage des covariances d'erreurs de prévision dans la section 3.3, ainsi que deux approches de la méthode NMC au §4.1 ("standard" et "lagged"). Dans la section précédente, §4.2, j'ai illustré la croissance propre d'erreurs dans le modèle couplé. Une question importante, primordiale avant de tester un cycle d'assimilation en grandeur réelle, est d'obtenir, au moins au premier ordre, une évaluation des variances totales d'erreur de prévision et d'observation. Ce sont en effet ces variances totales qui seront les paramètres ajustables les plus élémentaires pour fixer le poids respectif de l'ébauche et des observations. Dans la deuxième partie de son travail de thèse, Wafaa Sadiki s'est intéressée à ce problème, en utilisant plusieurs approches (Sadiki et Fischer, 2005) :

- comparaison avec des prévisions forcées seulement par des analyses Arpège (supposées proches de la "vérité")
- application de la méthode d'échantillonnage des vecteurs d'innovation (méthode de Lönnberg-Hollingsworth)
- diagnostic dit "de la fonction-coût au minimum" (Talagrand, 1998)
- diagnostics dit "du S_o et du S_b", qui sont une première généralisation du diagnostic précédent (Desroziers et Ivanov, 2001)

Notons que la plupart de ces approches reposent sur l'utilisation du vecteur d'innovation $X - \mathcal{H}(Z^b)$, seule source indépendante d'information a posteriori. A priori, on s'est donné des hypothèses sur la forme des statistiques d'erreur (fonctions de structure de P^b, R diagonale), voire des matrices "toutes faites" (obtenues par les méthodes citées au §3.3). Des diagnostics listés ci-dessus, le dernier cité a été particulièrement étudié, dans la mesure où il fournit une séparation objective entre les variances d'erreur de l'ébauche et des observations (moyennant cependant toujours une hypothèse de séparation entre P^b et R ; typiquement, on va supposer que les erreurs d'observation sont toutes décorrélées entre elles : R diagonale). Passons brièvement en revue les diagnostics mis en jeu. L'espérance mathématique de la fonction-coût totale au minimum prend une valeur très simple :

$$E[J(Z^a)] = E[(X - HZ^b)^t (HP^bH^t + R)^{-1}(X - HZ^b)] = p$$

où p est le nombre total d'observations. Une formule plus générale, qui permet d'exprimer l'espérance mathématique de la fonction-coût au minimum pour toute sous-partie du vecteur d'information (l'union ébauche + observations, $(Z^b|X)$), est développée et discutée dans Talagrand (1998), Chapnik *etal.* (2004) ou encore Sadiki et Fischer (2005). Lorsqu'on sépare le vecteur d'information en vecteur d'observations et en vecteur d'ébauche, on obtient les deux expressions suivantes :

$$E[J^o(Z^a)] = p - Tr(HK)$$

$$E[J^b(Z^a)] = Tr(KH) = Tr(HK)$$

Il est alors intéressant de considérer, pour un système d'assimilation variationnelle, les deux rapports suivants :

$$S_o = \frac{E[J^o(Z^a)]}{Tr(I - HK)} \tag{4.1}$$

$$S_b = \frac{E[J^b(Z^a)]}{Tr(KH)} \tag{4.2}$$

(par rapport à la bibliographie, un facteur 2 est omis car un facteur 1/2 a été omis dans la définition de la fonction-coût 2.10 et 2.11 ; l'indice k des itérations temporelles est également omis car ces formules sont valables autant en 3D-VAR qu'en 4D-VAR). Pour des systèmes aux statistiques d'erreur correctement spécifiées a priori, les deux fractions 4.1 et 4.2 doivent être égales à 1. Chapnik *etal.* (2004) ont par ailleurs montré qu'appliqués de façon itérative, en ré-injectant les nouvelles variances d'erreur dans le système d'assimilation pour un réseau donné ($\sigma_b^2 \rightarrow S_b\sigma_b^2$, $\sigma_o^2 \rightarrow S_o\sigma_o^2$), ces diagnostics permettent d'obtenir les estimations les plus vraisemblables des variances totales, étant données les informations dont on dispose (les statistiques a priori et les vecteurs d'innovation du réseau en cours). Une difficulté particulière pour cette méthode est cependant l'estimation des Traces (les dénominateurs). Desroziers et Ivanov ont proposé une méthode de Monte-Carlo, dérivée des travaux de Girard (1987), afin d'estimer ces termes dans l'espace des observations (on estime $Tr(H_kK_k)$). Cette estimation nécessite donc a priori un grand nombre d'observations, sur un réseau donné. Une approche alternative a été proposée par Bernard Chapnik, dans le cadre de son travail de thèse, qui consiste à perturber les observations et l'ébauche par leur statistique a priori, ce qui permet de se placer de facto dans les conditions d'optimalité des relations 4.1 et 4.2. L'analyse résultante permet ensuite de calculer les termes de $Tr(H_kK_k)$ recherchés. Wafaa Sadiki a montré que l'on pouvait, moyennant des hypothèses d'ergodicité, utiliser des séries temporelles de données d'analyses successives. Cela permet d'augmenter le vecteur d'observation (et plus généralement, le vecteur d'information), et de mettre "bout à bout" les éléments perturbés de façon aléatoire, pour estimer la Trace. Cette technique est d'autant plus utile dans un système à aire limitée, lorsqu'on veut estimer un terme tel que S_o pour un type d'observation précis, mais dont le volume de données sera forcément limité, sur un domaine limité.

Les figures 4.4 et 4.5 montrent les séries temporelles des termes S_o et S_b sur un mois d'analyses successives (non cyclées), pour les statistiques NMC standard et lagged, respectivement. Plusieurs résultats ont été tirés de ces travaux :

- la validation pratique de cette approche dans Aladin. Elle avait déjà été appliquée au 4D-VAR Arpège (à peu près en même temps) par Chapnik *etal.* (2004). Notamment, nous avons retrouvé le résultat déjà largement suggéré par notre expérience concrète, selon lequel les variances d'erreur d'ébauche étaient sur-estimées par la méthode NMC standard, et très sous-estimées par la méthode NMC lagged.
- la comparaison avec les mêmes séries temporelles à partir du 3D-VAR Arpège a montré que les termes d'observation (le S_o) semblaient surtout contrôlés par les propriétés intrinsèques des erreurs d'observation (leurs "vraies" variances, éventuellement des corrélations omises) et moins par les systèmes desquels ils étaient extraits (Arpège ou Aladin). Ainsi, le S_o varie entre 0.6 et 0.9 dans Arpège et Aladin, avec des observations assez identiques dans les deux systèmes (pas de radiances dans Aladin, cependant).
- le terme d'ébauche dépend par contre en premier lieu des statistiques choisies, et donc aussi du système d'assimilation. Ainsi, S_b vaut environ 0.5 dans Aladin avec NMC standard, mais 1.7 avec NMC lagged. Dans Arpège, ce terme vaut environ 1.1.
- une fois les nouvelles valeurs injectées dans les analyses, les diagnostics a posteriori se rapprochent de leur valeur théorique (notamment, la valeur au minimum de la fonction-coût).
- on peut stratifier assez facilement S_o par type d'observation, type de champs observé, voire par couche atmosphérique. Dans ce cas, et en utilisant l'ergodicité temporelle (donc en étudiant des séries temporelles d'analyses), on peut obtenir des profils verticaux de corrections (par exemple, pour les radiosondages), ou des résultats plus surprenants comme l'indication que les variances d'erreur des vents zonaux et méridiens des données avion ne devraient pas être rigoureusement égales.
- dans Fischer *etal.* (2005), nous avons utilisé les résultats de Wafaa Sadiki afin de calibrer les variances d'erreur d'ébauche des statistiques ensemblistes. Sans refaire tous les calculs, il est possible assez aisément de re-caler les variances d'erreur afin de conserver des variances d'erreur totales assez proches entre les trois jeux de statistiques (le réglage suggère en effet de rapprocher les variances totales entre NMC standard et NMC lagged, et on fait de même pour les statistiques d'ensemble).

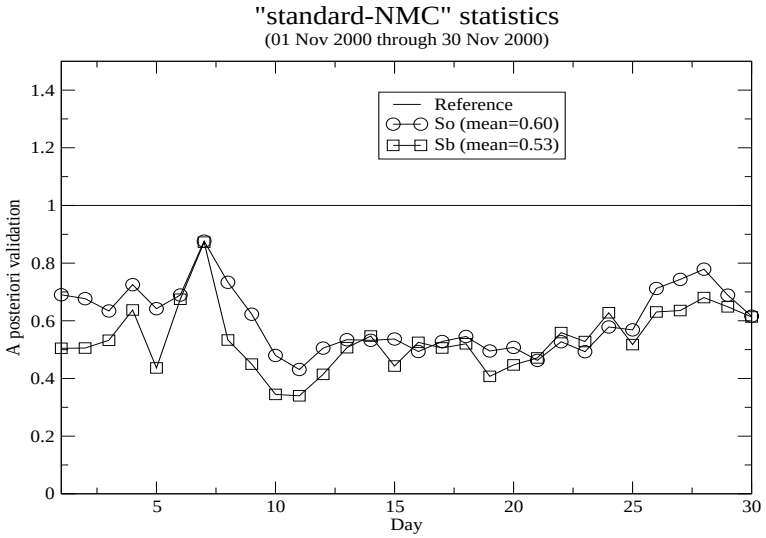


FIG. 4.4 – Séries temporelles du terme de réglage des variances d'erreur d'observations S_o (labels cercles) et du terme de réglage des variances d'erreur d'ébauche S_b (labels carrés). Ces valeurs sont obtenues avec l'analyse 3D-VAR Aladin avec statistiques NMC standard pour P^b.

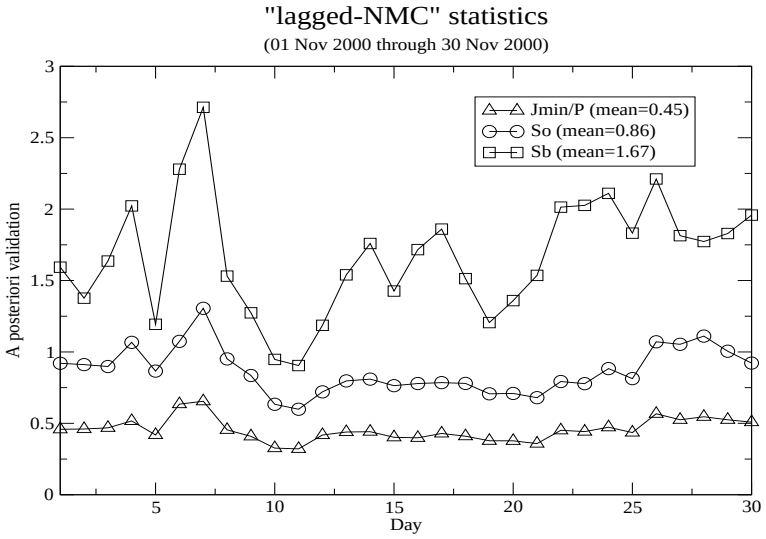


FIG. 4.5 – Séries temporelles de la valeur dans l'analyse finale de la fonction-coût $J(Z^a)$, divisée par le nombre total d'observations (labels triangles), du terme de réglage des variances d'erreur d'observations S_o (labels cercles) et du terme de réglage des variances d'erreur d'ébauche S_b (labels carrés). Ces valeurs sont obtenues avec l'analyse 3D-VAR Aladin avec statistiques NMC lagged pour P^b.

Il est à noter que, depuis les travaux de Wafaa, une approche sensiblement différente a été élaborée par Gérald Desroziers et ses collaborateurs (Desroziers *etal.* , 2005). Dans cette nouvelle approche, il suffit de calculer les covariances entre les différences analyse-moins-ébauche, observation-moins-ébauche et observation-moins-analyse, dans l'espace des observations, afin d'obtenir une estimation a posteriori des covariances d'erreur (projetées dans l'espace des observations). Ces termes sont très facilement accessibles après analyses, en relisant les fichiers d'observation mis à jour. Les diagnostics donnent accès à des estimations à la fois des covariances d'erreurs d'ébauche, d'observation et d'analyse. Desroziers *etal.* (2005) ont aussi montré que l'on pouvait ré-injecter les résultats, afin de corriger les matrices P^b et R a priori, de façon itérative, et lorsque les portées sont très différenciées (idéalement, toute la corrélation spatiale des erreurs d'innovation est portée par P^b, et R diagonale). Il n'est alors pas nécessaire de faire appel à une méthode de Monte-Carlo. Cependant, si on souhaite avoir accès aux termes de Traces, qui sont aussi une mesure de la quantité d'information apportée par les différents éléments dans l'analyse (l'ébauche, les observations), alors la méthode de Girard (ou celle de Chapnik) reste d'actualité. Je reviendrai sur ce point dans le paragraphe 5.1.

4.4 La prise en considération du système coupleur

Dans le modèle de prévision à aire limitée, les données du système coupleur interviennent éventuellement pour fournir les conditions initiales (si le modèle couplé ne dispose pas de sa propre analyse), et surement à travers la spécification des conditions aux limites latérales du domaine. Plusieurs types de formulation du couplage existent. Dans les modèles de type opérationnel, la méthode de relaxation des champs (Davies, 1976) reste la plus utilisée. Dans les modèles de recherche, de type "cloud resolving model", des formulations plus complexes existent (conditions radiatives, grid-nesting). Je ne m'étendrai pas davantage sur ces méthodes, bien qu'il soit évident que les erreurs du modèle à aire limitée, notamment près des bords et à grande échelle, sont, d'une manière généralement mal comprise, dépendantes de ces formulations.

Par contre, la question se pose de savoir comment utiliser l'information du système coupleur, notamment sa dernière analyse de données (du réseau en cours), dans le système couplé. En effet, prenons l'exemple du 3D-VAR Aladin. Tant que ce modèle tourne simplement avec sa propre assimilation de données, en cyclant sur sa propre ébauche, il n'est contraint par le système coupleur qu'au travers des conditions limites latérales en cours de prévision. Or, l'analyse globale peut utiliser d'autres sources d'observations que l'analyse à aire limitée, notamment toutes les observations situées à l'extérieur du domaine limité. De plus, l'analyse globale est souvent plus développée et a priori performante pour les grandes structures de l'écoulement atmosphérique (4D-VAR Arpège). Enfin, intuitivement, on peut penser qu'une prise en compte de l'analyse globale dans l'état initial à aire limitée doit rendre cet état initial plus cohérent près des bords (et pour les grandes échelles) avec les conditions limites latérales (qui, elles, proviennent toujours du réseau le plus récent du système global).

Une première approche pour "mélanger" une ébauche à aire limitée (une prévision) avec une analyse globale projetée sur le domaine limité, a été développée par nos collègues d'Europe Centrale, dans le cadre du projet Aladin. Il s'agit de la technique dite "du blending par filtres digitaux" (Brožkovà *etal.* , 2001). Dans cette technique, on combine l'ébauche Aladin avec l'analyse Arpège, sur une grille de résolution plus faible qu'Aladin et après filtrage digital des ondes courtes. L'incrément ainsi obtenu est ramené à la résolution nominale d'Aladin, et ajouté à l'ébauche. Il s'agit donc d'une correction incrémentale, mais sans la notion de contrôle optimal : les parts respectives de l'ébauche et de l'analyse dans l'incrément sont fixées de manière très implicite à travers l'effet des filtres digitaux. On pourrait parler d'une méthode de "blending" empirique. Cette technique est aujourd'hui opérationnelle dans le modèle Aladin-CE ("Aladin Central Europe"). Nous l'avons testée dans Aladin/France, avec un certain succès sur un cas de la campagne MAP (la POI 14, voir Guidard *etal.* , 2006). A cette occasion, nous avons pu apprécier de manière qualitative (ce n'était qu'une étude de cas unique) l'apport du cyclage d'une analyse 3D-VAR (meilleures estimations des quantités de précipitations à 6 h et 12 h d'échéance, structures d'altitude légèrement modifiées). Nous avons aussi pu estimer l'apport de la prise en compte de l'analyse du système coupleur Arpège, par la méthode du blending par filtres digitaux. Le blending augmente encore un peu la qualité des structures d'altitude (notamment, l'orientation de l'axe et des gradients d'un talweg d'altitude), ainsi que le réalisme des précipitations au sol (on se compare aux réseaux de

pluviomètres, aux échos radar et aux analyses Arpège et CEPMMT vérifiantes).

Cependant, dans le cadre d'Aladin/France, nous avons souhaité aller plus loin, et reformuler l'idée originale du blending dans le cadre variationnel. Ce travail a fait partie du travail de FCPLR (Formation Complémentaire par la Recherche) de Vincent Guidard (Guidard et Fischer, 2008). L'idée développée par Vincent consiste à modifier sensiblement le problème de recherche de l'état le plus probable. En effet, la probabilité que l'on cherche à maximiser est conditionnelle : elle suppose que l'on a observé les valeurs X (affectées d'une certaine incertitude). Dans le cas d'une analyse d'un système couplé, nous pouvons re-formuler cette probabilité : nous allons rechercher l'état le plus probable, ayant observé X et sous l'hypothèse que le système coupleur nous fournit l'analyse $\mathcal{P}(Z^A)$, où Z^A est l'analyse dans la grille globale, et $\mathcal{P}$ est la projection (non-linéaire) de cet état sur une grille à aire limitée tronquée. On peut, de manière classique, se ramener à une fonction-coût faisant intervenir les covariances d'erreur des différentes sources d'information : l'ébauche aire limitée, les observations et l'analyse globale. Dans le cas du système Arpège/Aladin, Vincent Guidard a vérifié quelle était l'erreur que l'on commet lorsque, de surcroit, on néglige les covariances entre l'ébauche Aladin et l'analyse Arpège (les deux sont corrélées, au sens strict, à travers les conditions limites latérales et donc, les ébauches respectives du réseau précédent et toute "l'histoire" des deux cycles d'assimilation). La figure 4.6 montre les spectres de variances de l'ébauche Aladin, de l'analyse Arpège, et du terme croisé (covariance $E[Z^b(Z^A)^t]$). On voit que les covariances croisées sont pratiquement d'un ordre de grandeur inférieures aux variances d'ébauche et d'analyse, pour les grandes échelles du domaine limité. Par ailleurs, on retrouve des courbes similaires à celles obtenues par Sadiki *etal.* (2000) avec les données d'Aladin/Maroc : les variances d'erreur d'ébauche et d'analyse (donc, des données de couplage Arpège) ont des ordres de grandeur identiques dans les grandes échelles (comparer les figures 4.6 et 4.3). L'implémentation pratique de la nouvelle contrainte vers l'analyse Arpège a donc été réalisée sous la forme d'une nouvelle contrainte faible, un terme supplémentaire dans le problème variationnel :

$$J(Z) = \underbrace{(Z - Z^b)^t (P^b)^{-1} (Z - Z^b)}_{J^b} + \underbrace{(X - \mathcal{H}(Z))^t R^{-1} (X - \mathcal{H}(Z))}_{J^o} + \underbrace{(\mathcal{P}'(Z) - \mathcal{P}(Z^A))^t V^{-1} (\mathcal{P}'(Z) - \mathcal{P}(Z^A))}_{J^k} \tag{4.3}$$

avec $\mathcal{P}'$ la projection (non-linéaire) de la grille nominale d'Aladin vers la grille tronquée. Le terme des covariances d'erreur d'analyses Arpège, V, est obtenu dans les travaux de Vincent de manière cohérente avec les méthodes d'échantillonnage des erreurs d'ébauche. En effet, la méthode ensembliste fournit, dans le cadre d'Arpège, à la fois des échantillons d'ensembles d'analyses et d'ébauches globales (voir Belo Pereira et Berre, 2006). Il suffit donc de projeter les analyses d'ensemble dans la grille Aladin tronquée, puis de calculer les covariances des dispersions fournies par cet échantillon, pour accéder à une estimation des covariances d'erreur d'analyse Arpège, dans la grille d'Aladin. Cette estimation sera a priori aussi précise que peut l'être celle des covariances d'erreur d'ébauche à partir de la même méthode. Je souligne donc ici une force réelle de l'approche ensembliste : elle autorise de manière assez élégante de traiter dans un même cadre conceptuel et technique (la méthode ensembliste et les dispersions qui en résultent) un emboîtement de systèmes d'assimilation.

La figure 4.7 illustre une cartographie des variances d'erreur d'analyse Arpège, dans la grille d'Aladin/France. Pour son implémentation pratique, le terme J^k a été représenté sous forme spectrale, en faisant l'hypothèse de corrélations homogènes et isotropes (nous nous sommes donc laissés guidés par les spécifications de J^b). La figure 4.7 montre, qu'au moins pour les variances, on pourrait ajouter une variabilité spatiale (ce qui est tout à fait autorisé par la formulation spectrale retenue pour V). Une remarque identique est valable pour le terme J^b, au demeurant. De manière plus générale, nous espérons qu'à l'avenir, le terme de couplage J^k pourra suivre les évolutions du terme J^b lorsque cela sera possible : cartographies des variances, fonctions de structure modifiées (ondelettes, ...), nouveaux échantillons ensemblistes.

Les résultats expérimentaux montrent que le terme de rappel vers l'analyse globale Arpège permet effectivement, de manière assez mécanique, d'améliorer les scores de l'analyse Aladin là où celle-ci est significativement moins bonne que le 4D-VAR global :

– faible réduction des erreurs de biais et d'écart quadratique moyen sur la pression réduite au niveau de

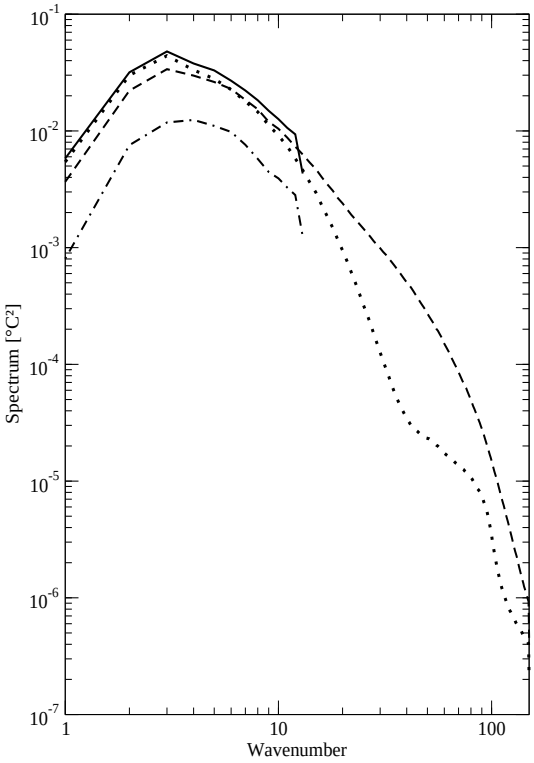


FIG. 4.6 – Spectres des variances horizontales pour la température, au niveau 22 du modèle : ébauche Aladin (trait tireté), analyse Arpège (trait pointillé), analyse Arpège tronquée sur la grille Aladin à plus basse résolution ($\mathcal{P}(Z^A)$) (trait continu), covariance croisée entre ébauche Aladin et analyse Arpège, tronquées (trait mixte).

la mer

– réduction de l'erreur de biais sur la température dans la stratosphère du modèle

L'assimilation avec J^k produit des ébauches légèrement plus proches des observations, ce qui entraîne une petite augmentation du nombre d'observations retenu dans l'étape du screening lors des tests d'amplitude sur les innovations. Nous avons aussi observé des impacts sur les précipitations à toutes les échéances. Cet effet reste cependant assez difficile à comprendre, et pourrait provenir de plusieurs raisons (des grandes échelles, notamment des talwegs d'altitude et des profils de stabilité verticale, plus réalistes ; une plus grande cohérence entre l'état initial analysé et les conditions limites latérales, d'où peut-être moins d'ondes parasites potentiellement générées près des bords). Cependant, un effet négatif, pervers, du J_k est que si les systèmes coupleur et couplé possèdent une erreur de biais similaire, alors le J_k va amplifier ce biais dans le système couplé. Nous avons observé cet effet pour un biais chaud de température dans la couche-limite d'Aladin (Chapnik et Fischer).

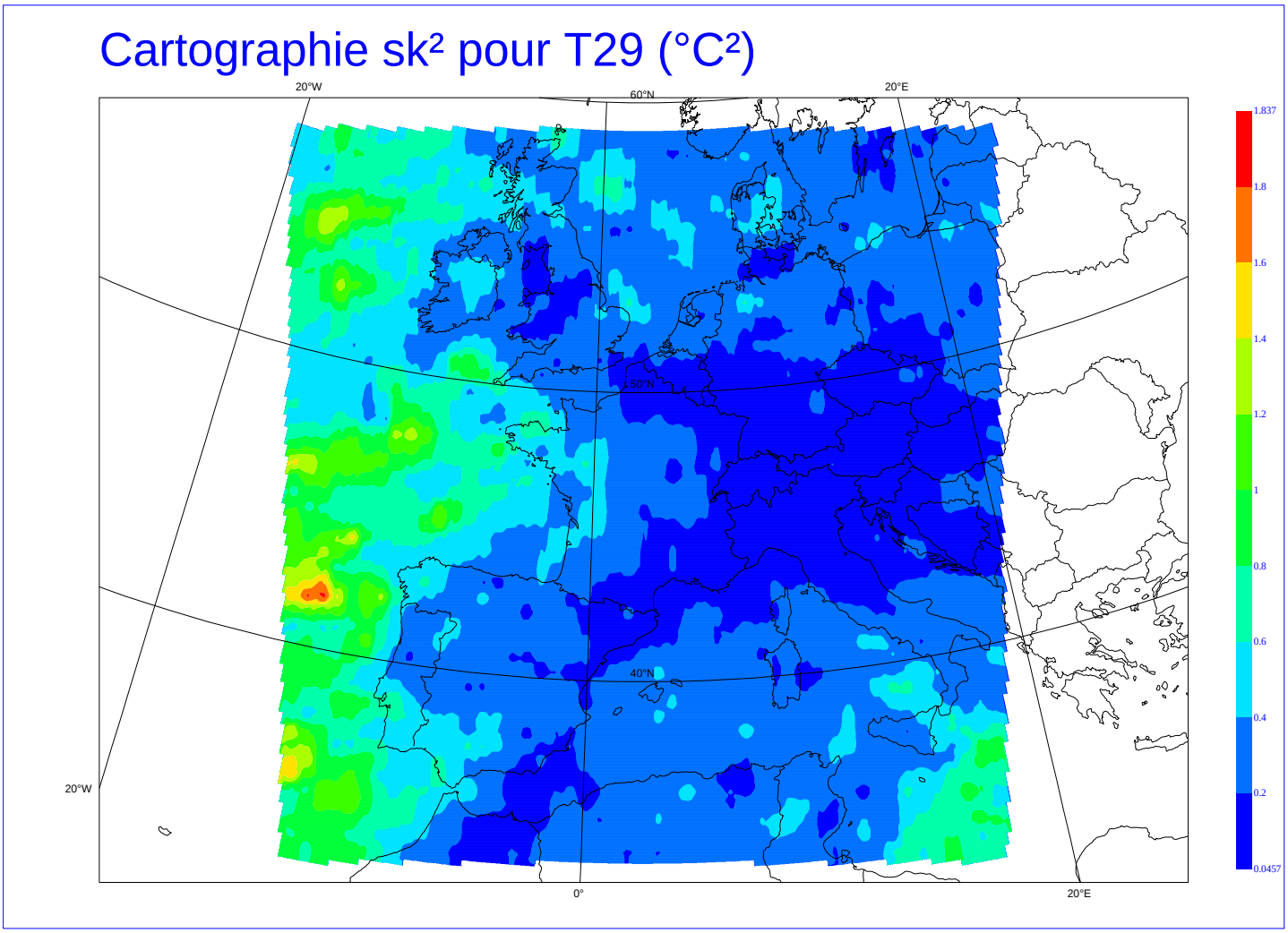


FIG. 4.7 – Cartographie des variances d'erreur de l'analyse Arpège, tronquée, estimée à partir d'un échantillon ensembliste. Le champ est la température au niveau 29 du modèle.

4.5 La prise en compte de la dimension temporelle (vers un 4D-VAR) : synthèse de l'article "Sensitivity of high resolution forecasts using the adjoint technique at the 10 km scale", par Soci, Fischer et Horányi, paru dans M.W.R. (2006)

Un objectif somme toute assez ancien du projet Aladin consistait à se doter des outils numériques préparant ou permettant des expériences d'analyses 4D-VAR en aire limitée. Dans ce but, les versions tangente linéaire et adjointe d'Aladin ont été codées, en partant du code Arpège. Ce travail permet notamment de mener des études de recherche sur les propriétés de la solution adjointe à haute résolution.

Ainsi, grâce à un financement dans le cadre du projet Européen "Alatnet", j'ai encadré avec Andras Horanyi le travail de thèse (Roumanie) de Cornel Soci. Nous avons analysé la sensibilité du modèle Aladin, à des résolutions de 10-12 km, par rapport à des corrections de ses conditions initiales, obtenues à partir de champs de gradients délivrés par le modèle adjoint. Cette approche est donc tout à fait comparable à celle de Rabier *etal.* (1996) dans le cadre de l'IFS. Nous disposions en fait de la version Eulérienne du modèle adjoint, et des versions Eulérienne et Semi-Lagrangienne du modèle direct (le changement du schéma d'advection, à pas de temps égal, induisant en toute rigueur des différences par exemple sur les pluies prévues). Le principe des expériences de Cornel, publiées dans Soci *etal.* (2006), était de faire une prévision de contrôle, d'effectuer le calcul adjoint le long de la trajectoire fournie par cette dernière, puis d'appliquer le gradient ainsi obtenu comme une perturbation aux conditions initiales. Le gradient est calculé en partant de la différence entre la prévision de contrôle et une analyse vérifiante Arpège, avec pour fonction-coût l'énergie totale sèche ou humide.

Nous avons testé une série de cas (d'autres sont illustrés dans Soci *etal.* , 2003). Beaucoup se sont avérés assez neutres face aux perturbations imposées (dans la mesure où l'impact sur les prévisions ultérieures était soit négligeable, soit un mélange d'impressions positives et négatives). Un cas cependant s'est révélé assez marquant. Les figures 4.8 et 4.9 montrent les pluies cumulées sur 6 heures (entre les échéances 6 et 12 h) de ce cas (le 03/05/2001). La figure 4.8 montre les précipitations du contrôle, avec un noyau principal de pluies sur le Centre-Ouest, d'échelle méso-α (cf le rectangle "A" de la figure 4.8). Ce noyau semble provoqué par une forte activité frontale, dans laquelle se développe de la convection, mais les pluies simulées dans le contrôle sont nettement sur-estimées : 27 mm/6h prévues contre environ 7 mm/6h observées. Dans le run perturbé (figure 4.9), le champ de pluie est modifié avec l'apparition d'un deuxième noyau de précipitations, de plus petite échelle, et situé au Nord-Est de Paris. Ce noyau peut correspondre à des pluies observées par le réseau des pluviomètres de surface au Nord de Paris (cf le rectangle "C" de la figure 4.8), et cette solution modifiée se rapproche effectivement plus de la réalité observée. On note aussi que le système précipitant "A" est beaucoup moins intense dans le run perturbé. Il n'est certes pas possible, à partir de cet unique exemple, d'envisager que les perturbations issues du modèle adjoint d'Aladin, à plus forte raison un 4D-VAR, aient de manière plus systématique un tel impact positif. Malgré tout, j'en retire quelques enseignements qui me paraissent utiles et justifier au moins un travail "de veille" sur l'algorithme 4D-VAR :

- la perturbation obtenue sur les conditions initiales n'a que peu dépendu dans cet exemple de la forme initiale du gradient dans le modèle adjoint, 6 heures plus tard. Il semble donc que, sur ce cas et dans un 4D-VAR, la présence d'observations, même de qualité assez médiocre, aurait permis d'initier le développement du gradient dans la zone de sensiblité.
- la qualité de la correction a dépendu de la résolution de la trajectoire : lorsque le modèle adjoint était intégré le long d'une trajectoire filtrée à basse résolution, l'impact de la perturbation qui en a résulté était nettement moins positif.
- dans cette situation précise, il semble que *seules deux solutions* étaient possibles dans le modèle direct : le système précipitant unique, assez vaste (figure 4.8) et le système double, avec un noyau plus concentré vers le Nord-Est (figure 4.9). Ce résultat peut être encourageant, si on considère qu'avec quelques bonnes observations on espère "basculer" de la solution initiale vers la deuxième solution. Il peut aussi nous inciter à la prudence, dans une optique multi-incrémentale par exemple, car il y a un risque que les boucles externes alternent d'une solution à l'autre, sans converger. Enfin, on peut de façon un peu

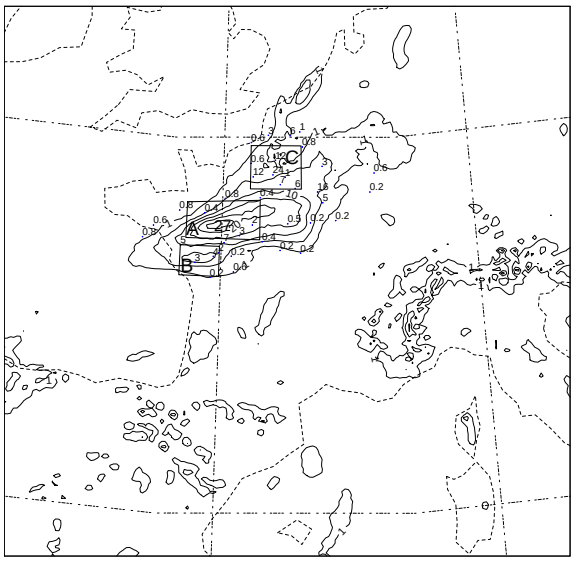


FIG. 4.8 – Précipitations cumulées sur 6 heures, entre 6 et 12 h d'échéance, pour le 03/05/2001. Prévision de contrôle sans perturbation initiale par un gradient. Les petits chiffres indiquent les cumuls de pluie observés. La valeur "27" dans le carré "A" indique le maximum des cumuls modélisés par Aladin.

provocante considérer qu'un petit ensemble de prévisions pouvait avoir une chance raisonnable de capter ces deux solutions, la zone de sensibilité (calculée sur 6 h) qu'il aurait fallu perturber se situant dans la Région Lyonnaise, et la forme de cette perturbation n'ayant apparemment pas un rôle crucial.

Je reviens dans la section 5.2 sur les nombreuses questions que soulèvent les projets d'analyse 4D-VAR pour la haute résolution, notamment lorsqu'on se projette à encore plus fine résolution (de 10 km vers 2.5 km, par exemple).

4.6 Les observations innovantes

Pour clore ce chapitre 4, je souhaite souligner simplement l'importance et l'utilité qu'il y a à travailler sur des sources d'observations innnovantes, pour une analyse à aire limitée, comparée à une analyse globale. D'un point de vue formel, il est plus aisé de supposer l'analyse à aire limitée décorrélée de celle du système coupleur, comme par exemple pour les travaux sur le terme J^k (voir §4.4). Pour des diagnostics du type S_b ou S_o, il est probablement aussi plus aisé de considérer les sources d'information comme décorrélées (si on voulait un terme "S_k" ?). Les deux systèmes seront d'autant plus innovants l'un par rapport à l'autre que les sources d'observations seront nombreuses et différenciées entre les deux systèmes.

Sur un plan pratique, l'analyse à aire limitée doit se concentrer sur l'assimilation des données denses, à haute fréquence. Elle doit surtout permettre d'améliorer la prévision à l'échelle régionale et aux courtes échéances, par opposition à l'échelle planétaire et aux prévisions sur plusieurs jours. Ce système doit donc valoriser les réseaux d'observations d'un pays, ou d'un groupe de pays.

Dans un avenir très proche, le modèle à aire limitée contiendra une microphysique évoluée, et sera capable de produire des champs d'ébauche des variables microphysiques. Dans ce cadre, des observations inconnues des systèmes globaux actuels devront être assimilées (réflectivités radar, bogus nuageux). Cette remarque pointe bien sûr directement vers le projet Arome à Météo-France, mais aussi vers les efforts au Royaume-Uni (combinaison d'une analyse 3D-VAR des champs atmosphériques classiques avec des analyses nuageuses et de précipitations du système MOPS), aux Etats-Unis (projet WRF) et au Japon (travaux sur un 4D-VAR non-hydrostatique).

Dans le cadre de l'assimilation 3D-VAR d'Aladin/France, nous avons ainsi pu insérer deux types d'ob-

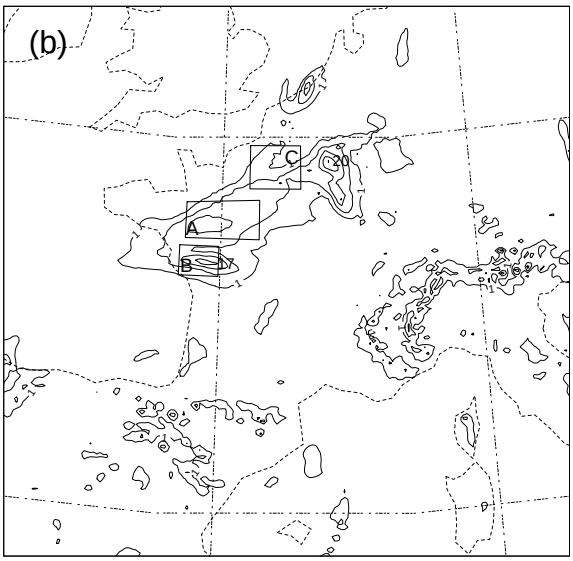


FIG. 4.9 – Précipitations cumulées sur 6 heures, entre 6 et 12 h d'échéance, pour le 03/05/2001. Prévision de sensibilité avec perturbation initiale par un gradient.

servations innovantes, non assimilées par Arpège :

- les radiances de 6 canaux infrarouge et vapeur d'eau du capteur SEVIRI, à bord de Météosat 8 (cf Montmerle *etal.* , 2007)
- les températures et les humidités relatives mesurées à 2 m, des réseaux Synop et Radome (travaux menés par Ludovic Auger)

Ces observations sont désormais partie intégrante de l'assimilation Aladin opérationnelle à Météo-France.

Chapitre 5

Perspectives et travaux prospectifs

J'aborde dans ce chapitre final différentes perspectives de travail, que je me proposerais de poursuivre, sous la forme de travaux encadrés ou dans le cadre de collaborations (internationales). Il est d'usage dans les manuscrits d'habilitation à diriger des recherches de s'en tenir à des éléments de réflexion critiques, de portée générale, et d'en déduire quelques thématiques sans entrer dans les détails (les détails étant également du ressort des doctorants, dont le premier travail consiste effectivement à participer, voire à proposer, le plan de recherche pour leur sujet).

Pour ma part, je situe la poursuite de mon travail de recherche, et surtout d'encadrement de recherches, dans le cadre de l'assimilation de données, en gardant en ligne de mire les grands projets structurants pour mon ... employeur, Météo-France. Ainsi, il est naturel de me placer dans la logique d'un passage progressif, mais volontariste, des considérations à grande échelle (échéances 2-5 jours, structures baroclines et barotropes de grande échelle, fenêtre d'assimilation de 6 à 24 h) vers des considérations de petite échelle (échéances 6-24 h, phénomènes convectifs et frontogénétiques transitoires, fenêtre d'assimilation entre 1/4 h et 3 h). En terme de projets, cela revient à une transition d'Arpège/Aladin vers Arome (Ducrocq *etal.* , 2005).

Plusieurs thématiques nouvelles vont prendre de l'importance dans le cadre d'un système à très haute résolution comme Arome :

- les modèles décriront de manière prognostique des champs microphysiques (eau liquide et solide nuageuse, des variétés d'hydrométéores, eau liquide ou solide précipitante). La question de bien comprendre quels champs peuvent être intégrés dans la variable de contrôle (le "Z" de mes notations dans ce rapport de HDR) est posée. On peut penser qu'il faut faire correspondre cette variable le plus possible aux champs accessibles par les observations. Dans ce cadre, l'eau totale précipitante est un candidat potentiel, car accessible indirectement par les réflectivités radar. Ducrocq *etal.* (2002) ont notamment montré qu'un bénéfice supplémentaire pouvait être obtenu pour des simulations avec le modèle Méso-NH, lorsque l'eau précipitante était ajustée aux observations. Cependant, je dois souligner qu'au premier ordre, les travaux dans le groupe CNRM/GMME ont mis en évidence *l'importance primordiale* de corriger au mieux *le champ d'humidité relative 3D* dans les conditions initiales. D'autres observations, très indirectes il est vrai, peuvent aussi apporter une information correctrice utile : ainsi, Guidard *etal.* (2006) ont-ils assimilé des bogus de nuages, sous la forme de profils d'humidité relative obtenus empiriquement à partir d'imagerie satellite Météosat combinée à une classification nuageuse. En filigrane, la question de comprendre comment les autres variables microphysiques s'ajustent à des corrections sur des champs tels que l'humidité relative 3D ou l'eau totale précipitante se pose : existera-t-il des problèmes de spin-up spécifiques ? des problèmes de cohérence thermodynamique dans les changements de phase ?
- comment représenter des liens statistiques entre les erreurs sur les champs "microphysiques" de l'analyse (en fait : humidité spécifique ou relative 3D, eau totale précipitante, éventuellement eau totale nuageuse) et les champs traditionnels de la dynamique du modèle (vent, température, pression) ? de tels liens sont-ils du reste nécessaires, ou est-ce que le modèle procède "très vite" (significativement plus vite que la fréquence de l'assimilation) aux ajustements ? Depuis les travaux de thèse de Loïk

Berre (Berre, 2000), l'analyse 3D-VAR Aladin possède une corrélation multi-variée entre l'humidité 3D et les champs dynamiques du modèle, dans la variable de contrôle. Cependant, cette relation n'assure pas les équilibres thermodynamiques, et nous procédons par exemple dans Aladin/France à la fois à un ré-étalonnage de l'humidité spécifique en début de prévision, avec la contrainte de borner l'humidité relative entre 0 et 100 % et à un ajustement par des filtres digitaux (voir §5.3). Autrement, des cas de fortes précipitations localisées, dans les tous premiers instants du modèle, peuvent apparaître. Si celles-ci disparaissent rapidement, leur trace dans le champ de divergence par exemple demeure ensuite dans les premières heures de prévision.

- des efforts importants ont commencé pour explorer des nouvelles sources d'observation. J'ai déjà cité les réflectivités radar (travaux engagés dans le cadre d'Arome par Olivier Caumont et Véronique Ducrocq à CNRM/GMME, Eric Wattrelot et François Bouttier à CNRM/GMAP, et collaboration à laquelle je participe avec nos partenaires Aladin, Marian Jurašek, et Hirlam, Günther Haase) ; les bogus de nuages (travaux de Mathieu Nuret à CNRM/GMME) ; les vents radar (collaboration souhaitée avec le groupe Hirlam) ; le GPS[1] terrestre (Brenot *etal.* , 2005, Healy et Thépaut, 2006). L'assimilation Aladin nous a permis déjà de nous différencier un peu d'Arpège dans ce domaine (cf §4.6).
- des questions assez pratiques, sur la fréquence des analyses à très haute résolution, la largeur des fenêtres d'assimilation des observations, se posent. Elles seront probablement en partie résolues en prenant en compte le coût réel des analyses de données et des prévisions, en fonction aussi du calculateur. L'exemple de référence dans ce domaine serait probablement l'analyse horaire "Rapid Update Cycle" (RUC) du NCEP (Devenyi et Benjamin, 2003), mais on souhaite dans le cadre d'Arome aboutir à un cycle d'assimilation continu.

Je ne m'étendrai pas plus sur ces questions. Elles sont effectivement dans certains cas déjà adressées par des actions en cours, et certaines constituent des éléments du plan scientifique et technique du projet Arome, qu'il serait déplacé de revisiter dans ce rapport.

Sur le plan des collaborations internationales, les travaux envisageables (qu'ils soient étiquetés "Arome" en général, ou que ce soient des thématiques plus personnelles) doivent aussi être re-situés dans le contexte international. L'assimilation de données à aire limitée profite de l'apport technique initial du projet Aladin. Elle profite aussi des efforts et des connaissances, ainsi que des futures applications locales, que nos partenaires Aladin maitrisent à présent. Il faut souhaiter que l'un ou l'autre axe de recherche puisse être significativement poussé par ce partenariat. Le rapprochement avec le consortium Hirlam peut aussi fournir des forces de travail, et des idées nouvelles. Cette communauté est notamment très intéressée par la problématique d'un 4D-VAR aire limitée, à l'échelle régionale aujourd'hui, et à l'échelle convective pour des travaux d'ici 2 à 5 ans. Enfin, je rappelle la collaboration toujours d'actualité avec le CEPMMT, qui peut encore nous apporter des solutions algorithmiques ou techniques (en terme de codes) innovantes, lorsqu'elles peuvent s'appliquer à une grande gamme d'échelles spatiales ou temporelles.

Pour poursuivre dans mes perspectives, je souhaite à présent développer plus spécifiquement 3 thèmes de recherche.

5.1 Les filtres adaptatifs, le réglage en ligne et le contenu en information

Dans l'article Sadiki et Fischer (2005), nous avons proposé, comme perspective à l'estimation par une méthode de Monte-Carlo des Traces du type $Tr(H_k K_k)$, de reporter cette technique dans un cadre ensembliste. En effet, comme je l'ai souligné notamment dans la section 3.3, il devient à présent possible de faire tourner des petits ensembles, hors assimilation de données pour l'instant, afin d'estimer les covariances d'erreur d'ébauche. L'idée serait à présent d'utiliser directement les éléments de ces ensembles (donc, en gros, 10 éléments par réseau d'analyse) pour produire des estimations des Traces, puis des termes S_b et S_o. Dans le cas où, pour un réseau donné, on considère que l'ensemble d'analyses est obtenu à partir d'un ensemble d'ébauches, et des observations constantes, Sadiki et Fischer ont proposé la formule suivante, par identification :

[1] Global Positioning System : il s'agit de délais dans les temps de réception des signaux électromagnétiques

$$S_b = \frac{E[J^b(Z^a)]}{-(\Delta Z^b)^t (P^b)^{-1} (\Delta Z^a - \Delta Z^b)} \tag{5.1}$$

A partir de cette idée de départ, plusieurs pistes me semblent s'ouvrir :

- vérifier la formule 5.1 à partir de données réelles ou idéalisées, ainsi que re-valider l'approche formelle proposée.
- on peut se demander si la possibilité d'accéder à $Tr(H_k K_k)$ par un petit ensemble ouvre la porte à d'autres applications. En effet, cette trace représente le contenu informatif du groupe d'observations que l'on est en train d'analyser (Fisher, 2003, Cardinali *etal.* , 2004), ce que l'on peut comprendre implicitement en se rappelant que : $Tr(P^a(P^b)^{-1}) = n - Tr(HK)$, pour tout réseau k, avec n la taille du vecteur de contrôle. Peut-on alors utiliser des estimations ensemblistes des Traces pour réajuster *en ligne* (c'est à dire en cours d'assimilation) les variances totales d'erreur ? Peut-on utiliser ces estimations pour déterminer de façon assez directe des critères de sélection des observations, voire des stratégies de ciblage ?
- si je croise ces idées avec les travaux en cours sur des nouvelles familles de fonctions de structure, du type ondelettes, dans lesquelles les spectres de variances (les diagonales de P^b dans l'espace fonctionnel) renseignent non plus sur les variances totales, mais aussi sur les portées et leur dépendance géographique (Deckmyn et Berre, 2005), alors on peut imaginer dans quelques années des méthodes d'ajustement assez "souples" des structures de P^b à partir d'information ensembliste somme toute assez réduite (je n'envisage en effet ici pas des gros ensembles, comme ceux qui semblent nécessaires dans les filtres ensemblistes "purs", voir §2.1).
- en filigrane, les idées listées ci-dessus suggèrent de "marier" l'approche ensembliste avec l'analyse variationnelle, mais en conservant l'approche variationnelle incrémentale comme le coeur du système d'assimilation. Les données ensemblistes n'interviennent qu'en terme de diagnostics a posteriori, pour l'ajustement des paramètres réglables de l'analyse.

J'illustre simplement la notion de réglage en ligne, en repartant de l'exemple du filtre de Kalman pour un champ scalaire "0-D", que j'ai présenté dans la section 2.2. Je suppose que je commets une erreur dans la spécification de mon filtre : soit une erreur sur une variance d'erreur en début de filtre (erreur sur $\sigma_b = \sigma$ à l'étape $k = 0$), soit une erreur sur le taux de croissance dynamique de l'erreur de prévision (je spécifie un λ qui n'est pas celui ressenti par l'analyse de l'innovation, λ_v pour "vrai taux"). Les formules de récurrence du filtre 0-D doivent être un peu revues, surtout pour pouvoir introduire les deux taux de croissance λ (celui affecté typiquement à l'équation 2.2) et λ_v (celui ressenti dans la croissance des innovations successives, donc dans les équations d'analyse 2.5, 2.4 et 2.3).

A chaque itération k du filtre, le modèle (faux) donne une croissance $\sigma_k = s_{k-1} e^{\lambda t}$. La vraie croissance d'erreur d'ébauche est, d'après mes hypothèses précédentes : $\widetilde{\sigma_k} = s_{k-1} e^{\lambda_v t}$. On peut de plus imaginer que l'écart-type d'erreur d'observation est mal spécifié : ω faux, $\widetilde{\omega}$ effectivement ressenti dans les innovations, et je les prends constants pour toutes les itérations k. Le diagnostic a posteriori présenté au §4.3 devient :

$$S_b = \frac{\widetilde{\sigma_k}^2 + \widetilde{\omega}^2}{\sigma_k^2 + \omega^2} = S_o$$

Je réinjecte ce coefficient dans les variances d'erreurs et je refais l'analyse. De la sorte, j'obtiens finalement 3 filtres que je peux comparer : le filtre exact (aucune erreur a priori) se ramène à l'exemple présenté au §2.2 ; le filtre corrigé a posteriori possède ses propres relations de récurrence que je ne précise pas ici ; le filtre faux (erreurs de spécification a priori, non corrigées par le terme $S_b = S_o$). La figure 5.1 montre l'évolution des écarts-type d'erreur d'ébauche σ_k et d'observation ω_k (ce dernier dépend de l'itération du filtre dans le cas du filtre corrigé, car j'applique aussi le terme S_o à ω), en fonction des itérations k, et pour une erreur de spécification initiale de σ seulement ($\lambda = \lambda_v$ et $\omega = \widetilde{\omega} = 1$). On constate que le filtre corrigé se situe entre le filtre exact et le filtre faux, que les trois filtres convergent effectivement avec le temps (le système perd sa mémoire des conditions initiales), et que la correction faite sur σ se paye ici sur ω, qui était en fait bien spécifié, et qui est donc modifié de façon indue (mais nous n'en savons rien a priori). En effet, le principe de la méthode

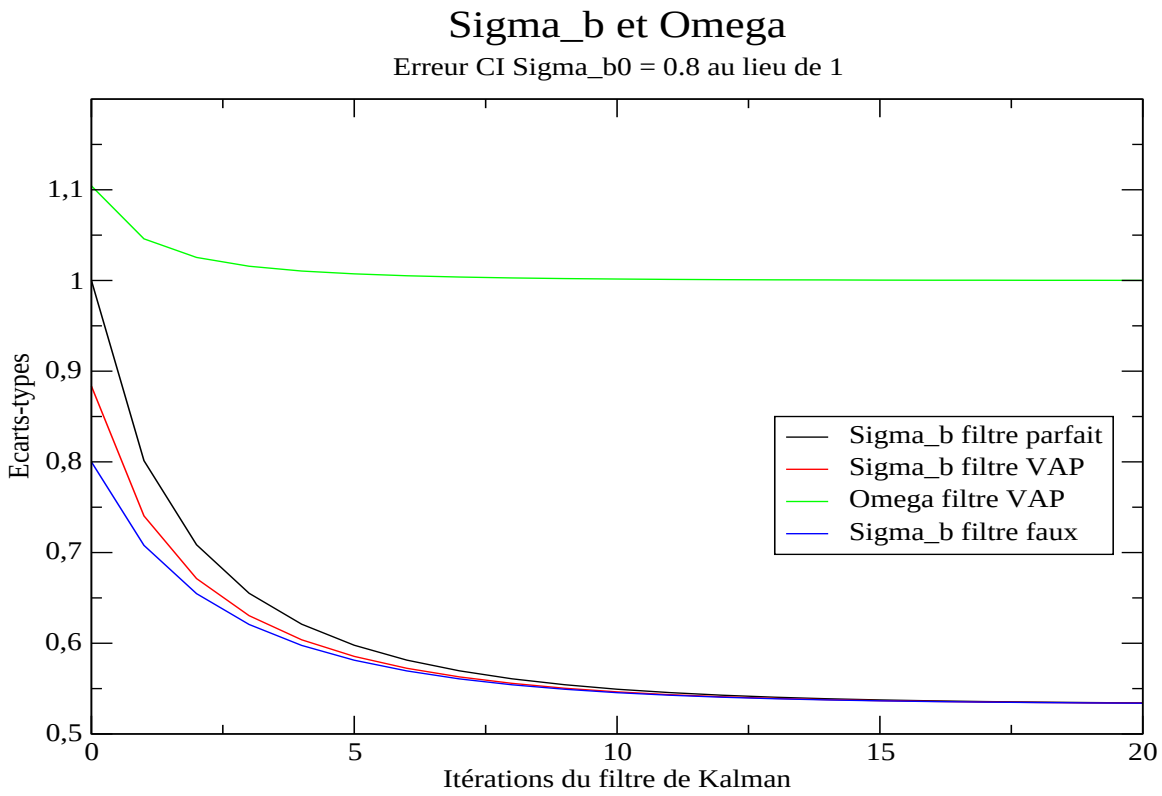


FIG. 5.1 – Evolution au cours des itérations du filtre 0-D des écarts-types d'erreurs d'ébauche (filtres exact, corrigé et faux) et d'observation (filtre corrigé). On suppose ici une erreur de spécification a priori de la variance d'erreur d'ébauche initiale.

reposant sur une mesure du vecteur d'innovation $X - HZ^b$, il manque des degrés de liberté pour pouvoir séparer exactement le réglage côté ébauche de celui côté observations. Dans le cas où la croissance de l'erreur par le modèle résolvant est mal spécifiée ($\lambda \neq \lambda_v$), les trois filtres divergent au cours du temps (figure 5.2). Le filtre corrigé se situe toujours entre l'exact et le faux, sans toutefois se rapprocher de la solution exacte car le terme d'observation, également "corrigé", est légèrement modifié, à tort toujours.

Voilà donc un exemple d'une grande simplicité, mais qui illustre l'idée de l'ajustement en ligne de manière concrète. Notons que le défaut de spécification de la croissance dynamique de l'erreur pourrait dans la pratique provenir d'une erreur du modèle résolvant (filtre de Kalman classique, tel que présenté dans la section 2.1) ou d'une erreur d'échantillonnage des données (dans un filtre ensembliste, ou pour un P^b ensembliste).

5.2 L'assimilation 4D-VAR pour la très haute résolution ?

Quelles peuvent être les forces d'une analyse 4D-VAR à l'échelle convective ? La nature très rapide (violente) et intermittente de la convection fait qu'il semble illusoire d'envisager de très longues fenêtres d'assimilation : on se contenterait surement de fenêtres allant de 15 mns à 3 heures. Cependant, des fenêtres d'assimilation courtes demanderaient, dans l'hypothèse d'une assimilation continue et contigüe dans le temps des observations, d'effectuer l'analyse très souvent (tous les quarts d'heure, toutes les heures ...). On se heurte donc alors au coût élevé qu'un 4D-VAR à haute résolution ne manquerait pas de représenter. Dans ce cadre, on pourrait imaginer de tester l'approche (multi-)incrémentale géographique (plusieurs analyses, avec des résolutions horizontales de plus en plus élevées) comme cela se pratique avec succès dans les modèles globaux

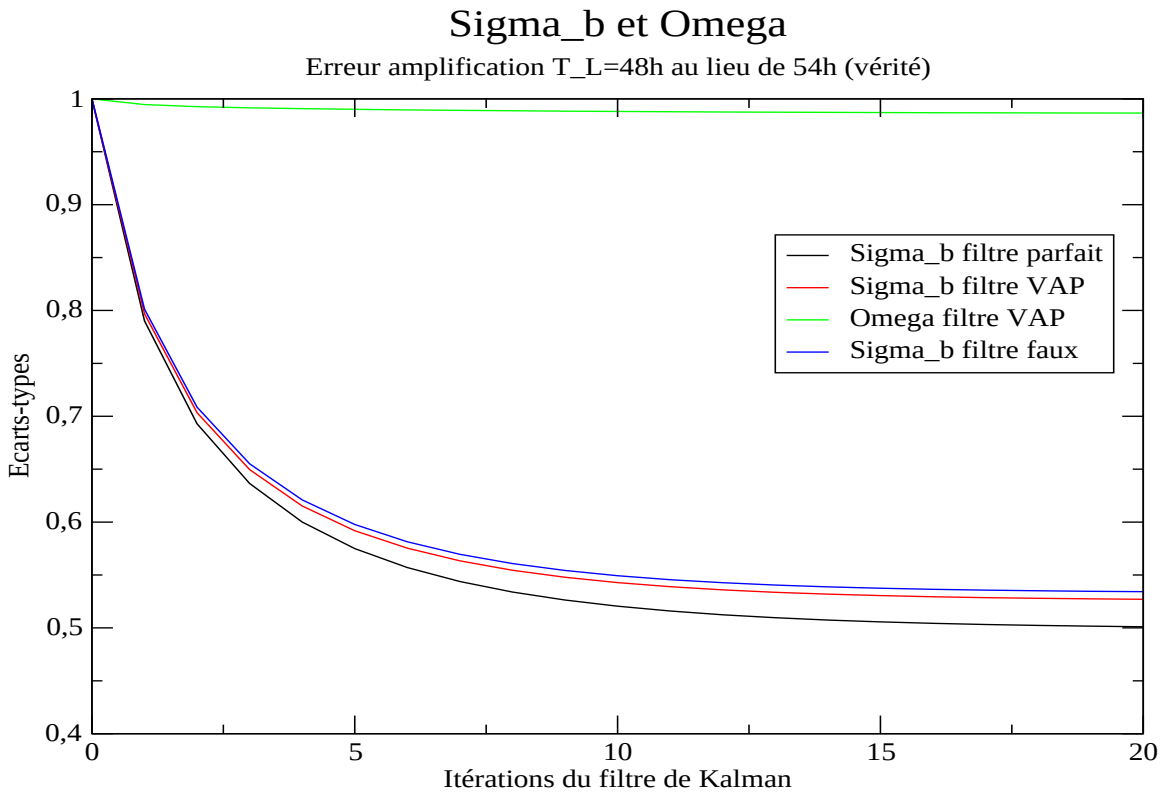


FIG. 5.2 – Evolution au cours des itérations du filtre 0-D des écarts-types d'erreurs d'ébauche (filtres exact, corrigé et faux) et d'observation (filtre corrigé). On suppose ici une erreur de spécification a priori de la croissance dynamique (linéaire) des erreurs de prévision.

(Arpège, IFS ... voir §2.4) ou de grande échelle à aire limitée (Hirlam). Cependant, il reste à prouver qu'une analyse à la résolution dégradée puisse apporter des corrections utiles sur les structures de la divergence du vent dans les basses couches, sur l'humidité 3D, sur certains champs microphysiques. Peut-on espérer en effet que la couverture nuageuse et le champ de précipitation s'affineront de façon cohérente pendant les boucles externes d'une telle assimilation incrémentale ? Notre expérience des modèles numériques doit nous inciter à la plus grande prudence. Les précipitations sont généralement très sensibles à la résolution du modèle, à paramétrisations physiques identiques (quand elles ne le sont pas tout simplement au pas de temps). Par ailleurs, différentes résolutions horizontales (voire verticales) nous amènent presque automatiquement à devoir considérer des paramétrisations physiques adaptées à chaque boucle externe : convection paramétrée à 10 km et 5 km, convection explicite à 2 ou 1 km. La microphysique ne sera pas la même non plus. Il est difficile de penser que de telles différences n'affecteront pas de manière significative les solutions des analyses, et donc le calcul des innovations successives. Ainsi, nous constatons qu'un 4D-VAR (multi-)incrémental à très haute résolution n'est pas un algorithme à l'application triviale, en terme de recherches autant qu'en terme d'ingéniérie.

Par ailleurs, d'autres développements pour un 4D-VAR à très haute résolution ne sont pas accessibles à court ou moyen terme : les paramétrisations physiques posent ainsi un réel problème (il faudrait travailler sur l'adjoint d'une microphysique). Dans Aladin, notre expérience en matière de solution adjointe et paramétrisations physiques repose entièrement sur le travail réalisé par Cornel Soci (voir §4.5). Ce travail bénéficiait largement de l'existant dans Arpège et ne touchait pas encore concrètement l'échelle convective.

Cependant, en mode recherche, il semble inversement frustrant de ne pas poursuivre de développements et de travaux sur une analyse 4D-VAR. En effet, des jeux de données observées à haute fréquence temporelle sont disponibles : observations de surface horaires, voire plus, mesures de vents radar à haute fréquence. Ces mesures fourniraient donc potentiellement les contraintes observationnelles que l'on pourrait espérer assimiler de façon continue dans le temps.

Ainsi, d'autres centres développent déjà ce type de fonctionnalité : Japon (Honda *etal.* , 2005), le système ARPS (Martin et Xue, 2005), WRF[2]. Plusieurs travaux récents faisant appel à un 4D-VAR à méso-échelle ont été publiés, en utilisant le modèle MM5[3] et son adjoint (qui existe aussi en version non-hydrostatique) typiquement à des résolutions de 30-15 km. Park et Županski (2003) ont synthétisé ces travaux, et effectué un tour d'horizon des problèmes que pouvait poser un 4D-VAR à haute résolution. Je souhaite ici résumer leur vision, avec quelques commentaires personnels :

- inclure *des observations à haute densité*, afin de décrire des structures de haute résolution dans la contrainte observationnelle. En supplément, la question de la corrélation spatiale et temporelle des erreurs se (re-)pose alors.
- l'impact des *non-linéarités* doit être surveillé : fonction-coût à minima multiples, à plateau, à zone interdite. L'aspect irrégulier de la fonction-coût peut entraîner des difficultés sérieuses de convergence des minimiseurs ou des techniques multi-incrémentales (boucles internes et externes, respectivement). Quelles sont les origines de ces non-linéarités ? :
  - l'importance accrue des phénomènes physiques à seuil
  - la plus grande complexité des opérateurs d'observation (radiances satellite à haute densité spatiale, réflectivité radar, ...)
  - une dynamique non-linéaire plus exacerbée : rétro-action rapide d'une instabilité sur son environnement, effects advectifs, turbulence.

  Cependant, la présence significative de non-linéarités n'implique pas forcément des limitations drastiques pour un algorithme 4D-VAR, au point de le rendre inutile. Ainsi, Pirès *etal.* (1996) ont montré sur des systèmes dynamiques chaotiques idéalisés (Hénon, 1976 et Lorenz, 1963) que l'on pouvait considérablement améliorer la qualité d'une analyse 4D-VAR et de sa prévision ultérieure, en procédant par assimilations successives des observations avec des fenêtres d'assimilation progressivement allongées (algorithme baptisé "QSVA" pour Quasi Static Variational Algorithm). Soulignons néanmoins deux hypothèses importantes dans leur travail : la disponibilité d'observations denses, régulières dans le

[2]Weather Research Forecasting system : modèle communautaire de PN aux USA

[3]Mesoscale Model version 5 : modèle de recherche aux USA, ancêtre de WRF

temps et de bonne qualité, ainsi que l'hypothèse d'un modèle parfait. Appliqué à un modèle atmosphérique de PN, un tel algorithme QSVA devrait à tout le moins, dans un premier temps, être testé dans un cadre idéalisé (observations simulées et expériences "jumelles", comme décrites au §2.3, absence par construction d'erreur modèle), car les hypothèses précédentes semblent très contraignantes ici. De manière plus élémentaire encore, il serait intéressant de bien comprendre le comportement d'un 4D-VAR multi-incrémental (à résolution constante cependant) constitué uniquement par une succession de minimisations et de re-calcul de trajectoires. Ce type d'approche permettrait d'évaluer l'importance et les conséquences pratiques des différentes sources de non-linéarité décrites ci-dessus. On pourrait aussi tester dans ce cadre l'ajout progressif de nouvelles observations (au fur et à mesure de leur disponibilité, par exemple).

- des travaux plus importants pour obtenir *l'adjoint d'un processus physique*. Les deux approches les plus fréquemment employées pour contourner le problème d'un processus physique discontinu sont de différencier séparément de part et d'autre de la discontinuité, ou de lisser et rendre continue la description paramétrée du processus (exemple de la physique simplifiée d'Arpège, voir Janiskova *etal.* , 1999).
- la question de la largeur de la fenêtre d'assimilation, et du *temps de mise à disposition du résultat de l'analyse*. Un 4D-VAR à haute résolution doit améliorer la prévision à courte échéance (quelques heures), et donc être d'autant plus disponible très rapidement. Cependant, il faut ajouter au temps d'exécution de l'algorithme le temps d'acquisition des observations. Dans la pratique, il est probable qu'en fait, les observations soient centrées légèrement en avant du milieu de la fenêtre d'assimilation (ce qui est par exemple le cas dans le RUC, cf Devenyi et Benjamin, 2003).
- *la description adéquate de l'erreur modèle*. Park et Županski privilégient les erreurs de biais, ce qui semble raisonnable. On ne peut effectivement pas, dans un système de PN, faire abstraction des erreurs systématiques. La description des moments d'ordre 2, dans les systèmes de grande échelle, commence tout juste (travaux de Yannick Trémolet et Mike Fisher au CEPMMT : Fisher *etal.* , 2005, Trémolet, 2007). Cette problématique pourrait renforcer l'intérêt pour une combinaison d'une analyse variationnelle avec un petit ensemble de prévisions (cf mon §5.1) si on sait séparer dans l'échantillon l'erreur d'analyse de l'erreur-modèle. Par ailleurs, il pourrait être intéressant de revenir sur les diagnostics a posteriori d'erreur-modèle, dont j'ai esquissé le principe et l'utilisation au §3.2. En tout état de cause, il semble primordial de d'abord travailler sur les méthodes permettant d'isoler et de décrire (avant de la quantifier) l'erreur-modèle : corrélations spatiales, temporelles, répartition géographique, champs les plus concernés. Ce n'est qu'ensuite que l'on pourra s'attacher à la prendre en compte dans l'assimilation.
- la prise en compte des *erreurs aux conditions limites latérales*. Les propositions classiques consistent à introduire toute la longueur des points de conditions limites dans la variable de contrôle, et, en toute rigueur, pour toutes les itérations. Cela est à nouveau assez coûteux en mémoire, et pose le problème d'avoir à mélanger une correction exprimée dans un espace spectral pour les conditions initiales internes avec une correction exprimée en points de grille sur les bords. Une approche à tester est de simplement rajouter un deuxième terme J^b valable à la fin de la fenêtre d'assimilation, afin de contraindre spatialement et temporellement la trajectoire optimale entre deux états du système coupleur (Nils Gustafsson, communication personnelle). Cette idée, déjà testée au Japon, se rapproche de la notion de "J^k" présentée au §4.4. De manière plus générale, on pourrait envisager de généraliser le J^k en ajoutant la contrainte (faible) de l'analyse (ou de la trajectoire) Arpège en plusieurs points de la fenêtre d'assimilation du système à haute résolution.
- *l'assimilation des précipitations*. Park et Županski insistent, à juste titre je pense, sur les questions de déclenchement d'ondes rapides (ondes d'inertie-gravité très courtes) liées aux déséquilibres dans les incréments (forcer des incréments de température et d'humidité, sans contre-partie sur la divergence ou le vent vertical). Je reviens sur ce point dans ma section 5.3.

Des travaux de recherche d'impacts d'observations, utilisant le 4D-VAR MM5 avec des résolutions de 30-15 km, sont régulièrement présentés par la communauté Américaine. Ainsi, Ha (2005), compare l'assimilation

directe de délais GPS obliques avec l'assimilation de profils inversés de contenu en eau précipitable, en utilisant des données de la campagne IHOP 2002. Une comparaison d'un tel algorithme 4D-VAR avec une analyse séquentielle à haute fréquence, de type "Rapid Update Cycle", sera probablement incontournable. On utilisera plutôt une analyse 3D-VAR ou 3D-VAR FGAT[4] pour ce dernier.

Dans le cadre d'Aladin ou d'Arome, les travaux algorithmiques et de codage restent importants : adjoint du schéma Semi-Lagrangien (en s'inspirant de ce qui a été fait pour l'IFS), code non-hydrostatique (environ 1 homme.année ?). En mode recherche uniquement, il semble envisageable, sous l'hypothèse que les développements algorithmiques dans la dynamique du modèle sont réalisés, d'utiliser un "petit" 4D-VAR à formulation incrémentale (mais pas au sens de la résolution) disposant d'une turbulence dans son modèle adjoint (pour représenter à peu près correctement la couche-limite et les échanges sol/atmosphère) dans le but d'assimiler des données de vent radar à haute fréquence ou des données de vent ou de température des réseaux de surface (vent à 10 m, température à 2 m).

5.3 La mise en équilibre des conditions initiales, en début de prévision

Dans cette section, je souhaite faire le point et esquisser des idées de travaux sur le problème de la mise en équilibre initial des états analysés. En effet, dans tout modèle numérique, le démarrage de la prévision à partir d'un état discrétisé qui ne soit pas rigoureusement une copie conforme d'un état prévu du modèle (en l'occurence, un état instantané, dans la grille de discrétisation du modèle, avec éventuellement les champs adéquats de l'instant précédent afin de pouvoir exactement reproduire les tendances du modèle à l'instant courant) va générer des ondes "parasites". Ces ondes sont dues à la remise en équilibre de l'état du modèle, par rapport à ses équations discrétisées. La remise en équilibre fait appel, au sens du modèle, aux ajustements géophysiques bien connus (hydrostatique, géostrophique). Les ondes générées sont de type acoustiques (dans un modèle compressible), de type inertie-gravité (avec généralement, dans des cas de déséquilibre marqué, l'excitation d'un mode de gravité externe, affectant toute l'épaisseur de l'atmosphère du modèle), de type mixte (incluant aussi des effets barotropes ou baroclines). Ces ondes se différencient les unes des autres par leur signature en terme de perturbations par rapport à l'état de base (pris comme une solution stationnaire ou aux caractéristiques pré-définies, par exemple en terme de tourbillon potentiel), en terme de fréquence temporelle, de structure d'onde (longitudinale, transversale), de vitesse de phase. Elles sont nuisibles dans un système d'assimilation, car elles correspondront à des structures non météorologiques, générées par exemple par les incréments d'analyse et capables de se maintenir dans le système si elles ne sont pas éliminées, et au contraire re-générées à chaque nouvelle analyse. Ainsi, si un même type d'innovations force systématiquement la même réponse du modèle, avec un temps d'amortissement qui ne soit pas significativement plus faible que la fréquence d'assimilation, alors les ondes parasites seront entretenues dans le processus d'assimilation.

L'expérience du numéricien, plus particulièrement dans le cadre d'un modèle à aire limitée comme Aladin, tend à séparer plusieurs types de déséquilibres :

- les déséquilibres contenus dans les incréments d'analyse : des incréments de température ou de pression, non balancés par des incréments de vent ; des incréments de vent, de vorticité ou de divergence, non balancés par des modifications sur la température ou la pression.
- l'impact des conditions limites latérales : des incohérences entre la solution intérieure, contenant les incréments d'analyse, et les conditions limites latérales, interpolations d'une analyse d'un système coupleur. Ce type d'incohérence peut produire des ondes de gravité externes ou quasi-externes (de mode vertical 1 ou 2) qui vont traverser très rapidement le domaine intérieur, à partir de la zone de relaxation de Davies. J'avais identifié ce type de phénomène dans Aladin, et ces problèmes ont aussi été étudiés dans le cadre de la formulation du couplage par Aidan McDonald (voir par exemple McDonald, 2005, ou ses rapports internes dans le cadre du projet Hirlam).
- des effets d'ajustement très rapides, sur quelques pas de temps seulement, dans les zones orographiques (éventuellement, là où les pentes sont abruptes), peut-être dus à des problèmes de spécification de la

[4]First Guess at Appropriate Time, un algorithme intermédiaire entre le 3D-VAR et le 4D-VAR

condition à la limite inférieure du modèle, pour la discrétisation au niveau le plus bas.

- un autre ajustement, sous forme d'un spin-down rapide suivi d'un spin-up de une à plusieurs heures, peut se produire si un déséquilibre important existe entre des champs microphysiques (eau nuageuse, eau précipitante) et des champs de la dynamique (notamment divergence ou vitesses verticales). Ce phénomène est notamment discuté dans Chen et Huang (2006) pour la mise en équilibre initial du modèle MM5 non-hydrostatique. Ce déséquilibre ne poserait aucun problème tant que toutes les variables en jeu ne sont pas analysées, car elles resteraient alors en équilibre local prescrit par l'ébauche.
- des descriptions différentes de la surface peuvent générer des mises en équilibre plus ou moins longues, en fonction de l'inertie des paramètres de surface concernés.
- hors assimilation, une source notoire de déséquilibre des champs initiaux d'un modèle reste tout processus d'interpolation de grille, typiquement d'une grille globale sur la sphère vers une grille à aire limitée en projection plane. Notamment, la difficulté à prescrire la pression de surface, et donc la pression des niveaux verticaux z ou hybrides, engendre des incohérences dans les champs dynamiques. Ainsi, une prévision du modèle Aladin en mode d'adaptation dynamique (démarré à partir d'une analyse Arpège) est toujours précédée d'un filtrage digital des champs (voir ci-dessous).

Les deux premiers problèmes sont certainement à surveiller dans le cadre d'une assimilation de données à haute résolution, et à haute fréquence temporelle. Dans les modèles de prévision numérique actuels, les ondes parasites sont filtrées soit par modes normaux non-linéaires (le principe est alors d'annuler les tendances des ondes rapides, obtenues comme les modes propres rapides d'un système linéarisé simplifié du modèle complet, cf Machenhauer, 1977, Marquet et Giard, 1992), soit par filtres digitaux (Lynch et Huang, 1992, Lynch *etal.* , 1997). Les filtres digitaux sont utilisés dans Aladin (et disponibles dans Arome). Par contre, les modes normaux non-linéaires n'ont jamais été développés dans Aladin, car ils posent de gros problèmes algorithmiques. En effet, si l'on peut espérer transposer les systèmes linéaires simplifiés (sous la forme de systèmes barotropes équivalents) au formalisme spectral bi-Fourier d'Aladin, il resterait à résoudre des difficultés nouvelles : facteur d'échelle variable dans le plan de projection (d'où des opérateurs de dérivation contenant le facteur d'échelle plus compliqués), conditions aux limites latérales imposées et ajoutant des contraintes et des termes nouveaux. De plus, seuls les 5 premiers modes normaux sont en général assujettis au filtrage, qui est itératif et donc a un coût. Signalons que les modes normaux ont été abandonnés dans des modèles tels qu'Arpège, l'IFS ou Hirlam.

La formulation du filtre digital consiste en un produit de convolution entre une fonction de filtre prédéfinie et une série d'états du modèle, nécessitant une intégration directe et éventuellement à rebours du modèle. Dans ces intégrations, les conditions limites latérales interviennent naturellement et les filtres digitaux "perçoivent" donc à la fois des ondes générées dans le domaine et sur ses bords. De fait, dans Aladin, les filtres digitaux semblent efficacement amortir autant les ondes de type "intérieur" que les ondes "de bord" (de couplage). A l'opposé, on reproche parfois aux filtres digitaux d'être un simple filtre temporel numérique, et de ne pas faire la différence entre des ondes réellement parasites et celles météorologiques, notamment à méso-échelle où les ondes d'ajustement numérique ont les mêmes temps caractéristiques que les phénomènes météorologiques tels les systèmes convectifs (de 45 mns à 1h30 typiquement). Il s'agit du reste ici d'un problème de fond important, qui rend difficile l'application d'un filtrage basé sur une séparation en fréquence temporelle : *à haute résolution, les mouvements de haute fréquence incluent autant des phénomènes météorologiques (convection, advection, ...) que des ondes parasites*. Enfin, je signale que l'expérience des numériciens semblent aussi montrer qu'un modèle hydrostatique (tel qu'Aladin) possède un temps de mise en équilibre plus long qu'un modèle non-hydrostatique aux équations compressibles (Arome) ou anélastiques (Méso-NH) : on évoque en effet des temps d'ajustement de l'ordre de 3 à 6 heures pour les premiers, et seulement de 1 heure pour les seconds. Bien sûr, je rappelle que les paramétrisations physiques participent aussi largement à ces temps de mise en équilibre.

Quels types de travaux peut-on alors envisager dans le cadre de la mise en équilibre initial des modèles de haute résolution, dotés d'une assimilation de données ? Certes, il est plus adéquat, pour des travaux de type thèse, d'envisager d'abord des modèles simplifiés. Cependant, dans cette problématique, il me semble qu'il faut travailler d'emblée avec le système numérique qu'on vise à balancer, puisque les propriétés de mise en équilibre dépendent beaucoup du type d'équations, des paramétrisations physiques, de l'assimilation, du

couplage. Néanmoins, les travaux idéalisés, que l'on peut trouver dans la littérature, peuvent servir de guides :

- effectuer un tour d'horizon du "bestiaire" des ondes et des types d'ajustement que le modèle en question peut produire. Dans un modèle analytique, Chagnon et Bannon (2005) ont étudié la réponse d'un système compressible à différentes inclusions de perturbations (réchauffement, masse, quantité de mouvement, divergence, vent rotationnel). Chaque type de perturbations produit une famille d'ondes données, davantage de type acoustique, de type gravité (flottabilité), de type mixte etc... De même, la durée pendant laquelle on impose un certain forçage joue un rôle dans la réponse. Ce type d'étude peut aussi être envisagé sur un modèle numérique, par exemple d'abord dans une version simplifiée bi-dimensionnelle (les modèles Aladin-NH et Arome disposent ainsi d'une version plan vertical). La compréhension du type d'ondes générées, en fonction des durées de forçage, peut nous renseigner sur les types de balances qui sont les plus cruciales : une réponse forte en terme d'ondes de gravité à une perturbation de température peut indiquer un problème potentiel pour l'assimilation à haute fréquence de radiances satellite ou de réflectivités radar.
- évaluer les signatures des problèmes de déséquilibre dans les modèles 3D. Des diagnostics assez simplistes en apparence peuvent déjà fournir des renseignements précieux : séries temporelles de pression de surface ou de divergence en certains points du modèle, écarts quadratiques moyens de ces champs sur tout le domaine, cumul sur tout le domaine des flux de précipitations totales. Des tracés de champ, entre des prévisions démarrant d'analyses et d'autres démarrant d'ébauches, peuvent fournir une assez bonne vision de la structure spatiale des ondes "parasites". A ma connaissance, ces diagnostics sont parfois appliqués de façon occasionnelle, mais rarement de manière systématique (Chen et Huang fournissent néanmoins un exemple récent de mise en application de ces diagnostics avec MM5, voir Chen et Huang, 2006).
- tirer les enseignements de ces études, en terme de modélisation des relations de balance dans la matrice P^b ou en terme d'application des filtres digitaux (choix des fréquences de coupure, des fonctions de filtre, des champs filtrés). Une plus grande sophistication dans la modélisation de P^b est une piste déjà instruite par nos collègues Canadiens du SMC (Service Météorologique Canadien) et de l'UQAM (Université du Québec à Montréal), sous l'initiative de Luc Fillion et Peter Zwack. Ainsi, Pagé, Fillion et Zwack (2007) proposent de résoudre l'équation de balance non-linéaire en ω ("NLB"), qui relie une vitesse verticale diagnostique en coordonnée pression aux tendances physiques de la température, à la tendance du tourbillon agéostrophique et du géopotentiel, et à l'advection de température. Le champs ω_{NLB} ainsi obtenu est effectivement très proche de celui du modèle 3D, notamment pour des simulations de convection profonde à 2.5 km de résolution. Le calcul de cette vitesse verticale diagnostique nécessite cependant une méthode assez complexe : passage en coordonnée pression (puis retour), filtrage point de grille des champs, méthode itérative. On pourrait instruire la possibilité de dériver à partir de l'équation de balance non-linéaire une relation de régression (linéarisée) approximative, dans la grille du modèle et en ne gardant que les termes de forçage dominants.

En filigrane, la question qui se pose ici est de comprendre si oui ou non, à partir éventuellement d'une plus grande sophistication dans la modélisation de P^b, on a toujours besoin d'un filtrage externe, en début de prévision (notamment, dans le cas d'Aladin ou d'Arome, par filtres digitaux). L'enjeu serait donc d'augmenter la complexité de P^b afin de s'affranchir d'une procédure de filtrage basé sur une séparation en fréquence temporelle non justifiée à haute résolution. L'élimination d'une procédure de mise en équilibre initial externe de type filtre digital évite également la problématique de la fréquence de l'assimilation par rapport à la fenêtre temporelle du filtre numérique.

Remerciements

Je remercie en premier lieu les membres du jury de HDR pour leur disponibilité à évaluer ma candidature. Olivier Talagrand a eu la gentillesse d'accepter d'oeuvrer en tant que directeur de HDR. Il a, à ce titre, évaluer avec patience et précision mon mémoire et contribué indéniablement à le parfaire. Je suis reconnaissant à Yvon Lemaître, Andrew Lorenc et François-Xavier Le Dimet pour leur rôle d'experts, ainsi qu'à Frank Roux et Jean Pailleux pour avoir participé au jury.

La préparation, et plus généralement la possibilité scientifique, à défendre une habilitation à diriger des recherches est le fruit d'une carrière de plusieurs années. Dans mon cas, celle-ci s'est jusqu'à présent entièrement déroulée au CNRM. Bien que n'ayant jamais occupé un poste fixe de chercheur, à temps plein, j'ai tout de même bénéficié de possibilités matérielles pour mener à bien des travaux expérimentaux, des publications et des encadrements. Cela a été rendu possible grâce à la confiance et au soutien de plusieurs personnalités :

- Jean-Claude André, ancien Directeur du CNRM, et Philippe Bougeault, ancien responsable de l'unité GMME, qui m'ont soutenu à leur niveau dans mes premières années de stage IT et FCPLR
- Alain Joly, chef de l'équipe RECYF et l'un des principaux coordinateurs de la campagne de mesure FASTEX, et Christophe Billard, responsable de la formation IM à l'Ecole Nationale de la Météorologie
- Jean-François Geleyn et François Bouttier, responsables ancien et actuel de l'unité GMAP
- Dominique Giard, ancienne responsable de l'équipe EXT/COOPE du GMAP

Je considère que la qualité scientifique que le jury de HDR voudra bien accorder à mon travail et aux idées que je défends ici, doit beaucoup à plusieurs rencontres et collaborations scientifiques. Je tiens en premier lieu à citer François Lalaurette, qui avait dirigé mes travaux de thèse, et qui m'avait guidé avec conviction et compétence dans mon "apprentissage de la recherche". J'associerai volontiers à François les noms de quelques scientifiques dont le contact fut bénéfique et sympathique : P. Ola Persson, Chris Snyder, Patrick Mascart, Jean-Noël Thépaut, Jean Pailleux, Nils Gustafsson.

Je profite aussi de ces quelques lignes pour remercier les quelques stagiaires, ENM ou Aladin pour l'essentiel, qui m'ont fait confiance et ont travaillé sous ma supervision. Je tiens en particulier à souligner tout le plaisir que j'ai eu à encadrer les travaux de thèse de Wafaa Sadiki (DMN/Maroc), Cornel Soci (ANM/Roumanie) et Vincent Guidard (Météo-France).

Enfin, je remercie les membres des équipes MC2, RECYF, COME (aujourd'hui MOANA et Méso-NH) du GMME et mes collègues du GMAP pour les excellentes relations que j'ai pu avoir et que j'ai encore avec eux tous.

Bibliographie

[1] Radi Ajjaji and Samir Issara, *Introduction de l'analyse canari du modèle global arpège dans le modèle à domaine limité aladin*, Master's thesis, Ecole Nationale de la Météorologie-Toulouse, 1994.

[2] J.L. Anderson, *An adaptive covariance inflation error correction algorithm for ensemble filters*, Tellus **59A** (2007), 210–224.

[3] E. Andersson, M. Fisher, R. Munro, and A. McNally, *Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence*, Quart. J. Roy. Meteor. Soc. **126** (2000), 1455–1472.

[4] M. Belo Pereira and L. Berre, *The use of an ensemble approach to study the background error covariances in a global NWP model*, Mon. Wea. Rev. **134** (2006), 2466–2489.

[5] D.A. Bennetts and B.J. Hoskins, *Conditional symmetric instability : a possible explanation for frontal rainbands*, Quart. J. Roy. Meteor. Soc. **105** (1979), 945–962.

[6] Thierry Bergot, Gwenaëlle Hello, Alain Joly, and Sylvie Malardel, *Adaptive observations : a feasibility study*, Mon. Wea. Rev. **127** (1999), 743–765.

[7] L. Berre, *Estimation of synoptic and mesoscale forecast error covariances in a limited area model*, Mon. Wea. Rev. **128** (2000), 644–667.

[8] L. Berre, S.E. Ştefănescu, and M. Belo Pereira, *The representation of the analysis effect in three error simulation techniques*, Tellus **58A** (2006), 196–209.

[9] C.H. Bishop, B.J. Etherton, and S.J. Majumdar, *Adaptive sampling with the ensemble transform Kalman filter. Part I : Theoretical aspects*, Mon. Wea. Rev. **129** (2001), 420–436.

[10] C.H. Bishop and Z. Toth, *Ensemble transformation and adaptive observations*, J. Atmos. Sci. **56** (1999), 1748–1765.

[11] H. Brenot, V. Ducrocq, A. Walpersdorf, C. Champollion, and O. Caumont, *GPS zenith delay sensitivity evaluated from high-resolution NWP simulations of the 9th September 2002 flash flood over southeastern France*, J. Geoph. Res. (2005), –, accepted.

[12] R. Brožkovà, D. Klarić, S. Ivatek-Šahdan, J.-F. Geleyn, V. Cassé, M. Širokà, G. Radnóti, M. Janoušek, K. Stadlbacher, and H. Seidl, *DFI blending : An alternative tool for preparation of the initial conditions for LAM*, Research activities in atmospheric and oceanic modelling (H. Ritchie, ed.), WMO, April 2001, CAS/JSC WGNE Report n. 31, pp. 1.7–1.8.

[13] R. Bubnová, A. Horányi, and S. Malardel, *International project ARPEGE-ALADIN*, EWGLAM newsletter n. 22 (Institut Royal Météorologique de Belgique, ed.), 1993, pp. 117–130.

[14] R. Buizza, *Unstable perturbations computed using the adjoint technique*, ECMWF internal report, ECMWF, 1992, ECMWF Research Dept. Technical Memorandum no. 189.

[15] R. Buizza, M. Miller, and T.-N. Palmer, *Stochastic simulation of model uncertainties in the ECMWF ensemble prediction system*, Quart. J. Roy. Meteor. Soc. **125** (1999), 2887–2908.

[16] R. Buizza and T. N. Palmer, *The singular-vector structure of the atmospheric global circulation*, J. Atmos. Sci. **52** (1995), 1434–1456.

[17] C. Cardinali, S. Pezzulli, and E. Andersson, *Influence matrix diagnostic of a data assimilation system*, Quart. J. Roy. Meteor. Soc. **130** (2004), 2767–2786.

[18] J.M. Chagnon and P.R. Bannon, *Adjustment to injections of mass, momentum and heat in a compressible atmosphere*, J. Atmos. Sci. **62** (2005), 2749–2769.

[19] B. Chapnik, G. Desroziers, F. Rabier, and O. Talagrand, *Properties and first application of an error statistics tuning method in a variational assimilation*, Quart. J. Roy. Meteor. Soc. **130** (2004), 2253–2275.

[20] M. Chen and X.-Y. Huang, *Digital filter initialization for MM5*, Mon. Wea. Rev. **134** (2006), 1222–1236.

[21] P. Courtier, E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher, *The ECMWF implementation of three-dimensional variational assimilation (3D-VAR). I : Formulation*, Quart. J. Roy. Meteor. Soc. **124** (1998), 1783–1807.

[22] P. Courtier, C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas, *The ARPEGE project at Météo-France*, Proceedings of the 1991 ECMWF seminar, ECMWF, 9-13 September 1991, available from the European Centre for Medium-Range Weather Forecasts, pp. 193–231.

[23] P. Courtier, J.-N. Thépaut, and A. Hollingsworth, *A strategy for operational implementation of 4D-VAR using an incremental approach*, Quart. J. Roy. Meteor. Soc. **120** (1994), 1367–1387.

[24] G.A. Dalu, M. Baldi, R.A. Pielke Sr., and G. Leoncini, *Mesoscale nonhydrostatic and hydrostatic pressure gradient forces - Theory*, J. Atmos. Sci. **60** (2003), 2249–2266.

[25] H.C. Davies, *A lateral boundary formulation for multi-level prediction models*, Quart. J. Roy. Meteor. Soc. **102** (1976), 405–418.

[26] H.C. Davies and C.H. Bishop, *Eady edge waves and rapid development*, J. Atmos. Sci. **51** (1994), 1930–1946.

[27] A. Deckmyn and L. Berre, *A wavelet approach to representing background error covariances in a Limited Area Model*, Mon. Wea. Rev. **133** (2005), 1279–1294.

[28] D. Dee, *Detection and correction of model bias during data assimilation*, Seminar on recent developments in data assimilation for atmosphere and ocean, ECMWF, 8-12 September 2003, available from the European Centre for Medium-Range Weather Forecasts, pp. 65–74.

[29] D.P. Dee, *On-line estimation of error covariance parameters for atmospheric data assimilation*, Mon. Wea. Rev. **123** (1995), 1128–1145.

[30] J. Derber and F. Bouttier, *A reformulation of the background error covariance in the ECMWF global data assimilation system*, Tellus **51A** (1999), 195–221.

[31] G. Desroziers, L. Berre, B. Chapnik, and P. Poli, *Diagnosis of observation, background and analysis-error statistics in observation space*, Quart. J. Roy. Meteor. Soc. **131** (2005), 3385–3396.

[32] G. Desroziers and S. Ivanov, *Diagnosis and adaptive tuning of information-error parameters in a variational assimilation*, Quart. J. Roy. Meteor. Soc. **127** (2001), 1433–1452.

[33] D. Devenyi and S.G. Benjamin, *A variational assimilation technique in a hybrid isentropic-sigma coordinate*, Meteor. Atmos. Phys. **82** (2003), 245–257.

[34] V. Ducrocq, *Adiabatic and viscous simulations of symmetric instability : structure, evolution and energetics*, J. Atmos. Sci. **50** (1993), 23–42.

[35] V. Ducrocq, F. Bouttier, S. Malardel, T. Montmerle, and Y. Seity, *Le projet AROME*, La Houille Blanche **2** (2005), 39–43.

[36] V. Ducrocq, D. Ricard, J.-P. Lafore, and F. Orain, *Storm-scale numerical rainfall prediction for five precipitating events over France : On the importance of the initial humidity field*, Wea. Forec. **17** (2002), 1236–1256.

[37] Y. Durand and P. Bougeault, *L'analyse objective PERIDOT*, Tech. report, Direction de la Météorologie Nationale/EERM, 1987, Note de travail n. 193.

[38] E.T. Eady, *Long waves and cyclone waves*, Tellus **1** (1949), 33–52.

[39] B.J. Etherton and C.H. Bishop, *Resilience of hybrid ensemble/3DVAR analysis schemes to model error and ensemble covariance error*, Mon. Wea. Rev. **132** (2004), 1065–1080.

[40] G. Evensen, *Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics*, J. Geoph. Res. **99** (1994), no. C5, 10143–10162.

[41] Brian F. Farrell, *The initial growth of disturbances in a baroclinic flow*, J. Atmos. Sci. **39** (1982), 1663–1686.

[42] ———, *Modal and non-modal baroclinic waves*, J. Atmos. Sci. **41** (1984), 668–673.

[43] C. Fischer, A. Joly, and F. Lalaurette, *Error growth and Kalman filtering within an idealized baroclinic flow*, Tellus **50A** (1998), 596–615.

[44] C. Fischer and F. Lalaurette, *Meso-beta scale circulations in realistic fronts. I : steady basic state*, Quart. J. Roy. Meteor. Soc. **121** (1995), 1255–1283.

[45] ———, *Meso-beta scale circulations in realistic fronts. II : frontogenetically forced basic states*, Quart. J. Roy. Meteor. Soc. **121** (1995), 1285–1321.

[46] C. Fischer, T. Montmerle, L. Berre, L. Auger, and S.E. Ştefănescu, *An overview of the variational assimilation in the ALADIN/France numerical weather-prediction system*, Quart. J. Roy. Meteor. Soc. **131** (2005), 3477–3492.

[47] Claude Fischer, *Estimation d'erreurs de prévision et d'analyse dans un modèle semi-géostrophique à l'aide d'un filtre de Kalman*, Tech. report, Météo-France/CNRM, 1997, Note de centre GMME n. 48.

[48] ———, *Linear amplification and error growth in the 2D Eady problem with uniform potential vorticity*, J. Atmos. Sci. **55** (1998), 3363–3380.

[49] M. Fisher, *Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems*, ECMWF internal report, ECMWF, 2003, ECMWF Research Dept. Technical Memorandum no. 397.

[50] ———, *On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation*, ECMWF internal report, ECMWF, 2004, ECMWF Research Dept. Technical Memorandum no. 447.

[51] ———, *Background error covariance modelling*, Seminar on recent developments in data assimilation for atmosphere and ocean, ECMWF, 8-12 september 2003, available from the European Centre for Medium-Range Weather Forecasts, pp. 45–64.

[52] M. Fisher, M. Leutbecher, and G.A. Kelly, *On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation*, Quart. J. Roy. Meteor. Soc. **131** (2005), 3235–3246.

[53] Mike Fisher, *Background error statistics derived from an ensemble of analyses*, ECMWF internal report, ECMWF, 1999, ECMWF Research Dept. Memorandum File R43.2/MF/79.

[54] L.S. Gandin, *Objective analysis of meteorological fields*, translated from Russian 1965 ed., Israel program for scientific translations, 1963.

[55] G. Gaspari and S.E. Cohn, *Construction of correlation functions in two and three dimensions*, Quart. J. Roy. Meteor. Soc. **125** (1999), 723–757.

[56] P. Gauthier, P. Courtier, and P. Moll, *Assimilation of simulated wind lidar data with a Kalman filter*, Mon. Wea. Rev. **121** (1993), 1803–1820.

[57] R. Gelaro, R. Buizza, T.N. Palmer, and E. Klinker, *Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors*, J. Atmos. Sci. **55** (1998), 1012–1037.

[58] M. Ghil, *Meteorological data assimilation for oceanographers. Part I : description and theoretical framework*, Dyn. Atmos. Oc. **13** (1989), 171–218.

[59] D. Girard, *A fast Monte-Carlo cross-validation procedure for large least square problems with noisy data*, Technical report RR 687-M, IMAG, IMAG, Université de Grenoble, France, 1987, 22 pp.

[60] V. Guidard and C. Fischer, *Introducing the coupling information in a limited area variational assimilation*, Quart. J. Roy. Meteor. Soc. **134** (2008), 723–735.

[61] V. Guidard, C. Fischer, M. Nuret, and A. Džiedžic, *Evaluation of the ALADIN 3D-VAR with observations of the MAP campaign*, Meteor. Atmos. Phys. **92** (2006), 161–173.

[62] N. Gustafsson, L. Berre, S. Hörnquist, X.-Y. Huang, M. Lindskog, B. Navascués, K.S. Mogensen, and S. Thorsteinsson, *Three-dimensional variational data assimilation for a limited area model. Part I : General formulation and the background error constraint*, Tellus **53A** (2001), 425–446.

[63] S.Y. Ha, *Four-dimensional variational data assimilation of ground-based GPS water vapor in the June 12th IHOP case*, 11th Conference on Mesoscale Processes and 32nd Conference on Radar Meteorology, American Meteorological Society, 24-29 October 2005.

[64] S.B. Healy and J.-N. Thépaut, *Assimilation experiments with CHAMP GPS radio occultation measurements*, Quart. J. Roy. Meteor. Soc. **132** (2006), 605–623.

[65] M. Hénon, *A two-dimensional mapping with a strange attractor*, Comm. Math. Phys. **50** (1976), 69–77.

[66] A. Hollingsworth and P. Lönnberg, *The statistical structure of short-range forecast errors as determined from radiosonde data. Part I : the wind field*, Tellus **38A** (1986), 111–136.

[67] Y. Honda, M. Nishijima, K. Koizumi, Y. Ohta, K. Tamiya, T. Kawabata, and T. Tsuyuki, *A pre-operational variational data assimilation system for a non-hydrostatic model at the Japanese Meteorological Agency : formulation and preliminary results*, Quart. J. Roy. Meteor. Soc. **131** (2005), 3465–3475.

[68] A. Horányi, I. Ihász, and G. Radnóti, *ARPEGE-ALADIN : A numerical weather prediction model for Central-Europe with the participation of the Hungarian Meteorological Service*, Idöjárás **100** (1996), 277–301.

[69] B.J. Hoskins, *The geostrophic momentum approximation and the semi-geostrophic equations*, J. Atmos. Sci. **32** (1975), 233–242.

[70] P.L. Houtekamer, L. Lefaivre, J. Derome, H. Ritchie, and H.L. Mitchell, *A system simulation approach to ensemble prediction*, Mon. Wea. Rev. **124** (1996), 1225–1242.

[71] P.L. Houtekamer and H.L. Mitchell, *Data assimilation using an ensemble Kalman filter technique*, Mon. Wea. Rev. **126** (1998), 796–811.

[72] M. Janisková, J.N. Thépaut, and J.F. Geleyn, *Simplified and regular physical parametrizations for incremental four-dimensional variational assimilation*, Mon. Wea. Rev. **127** (1999), 26–45.

[73] Alain Joly, *The stability of steady fronts and the adjoint method : nonmodal frontal waves*, J. Atmos. Sci. **52** (1995), 3081–3108.

[74] Alain Joly and Alan J. Thorpe, *The stability of time-dependent flows : an application to fronts in developing baroclinic waves*, J. Atmos. Sci. **48** (1991), 163–182.

[75] R.E. Kalman, *A new approach to linear filtering and prediction problems*, Trans. ASME **82D** (1960), 35–45.

[76] R.E. Kalman and R.S. Bucy, *New results in linear filtering and prediction theory*, Trans. ASME **83D** (1961), 95–108.

[77] J.-F. Lacarra and O. Talagrand, *Short-range evolution of small perturbations in a barotropic model*, Tellus **40A** (1988), 81–95.

[78] F. Lalaurette, C. Fischer, and J.P. Cammas, *Location and interaction of upper and lower troposphere adiabatic frontogenesis*, Mon. Wea. Rev. **122** (1994), 2004–2021.

[79] P. Lönnberg and A. Hollingsworth, *The statistical structure of short-range forecast errors as determined from radiosonde data. Part II : the covariance of height and wind errors*, Tellus **38A** (1986), 137–161.

[80] A.C. Lorenc, *Analysis methods for numerical weather prediction*, Quart. J. Roy. Meteor. Soc. **112** (1986), 1177–1194.

[81] E.N. Lorenz, *Deterministic nonperiodic flow*, J. Atmos. Sci. **20** (1963), 130–141.

[82] P. Lynch, D. Giard, and V. Ivanovici, *Improving the efficiency of a digital filtering scheme*, Mon. Wea. Rev. **125** (1997), 1976–1982.

[83] P. Lynch and X.-Y. Huang, *Initialization of the HIRLAM model using a digital filter*, Mon. Wea. Rev. **120** (1992), 1019–1034.

[84] B. Machenhauer, *On the dynamics of gravity oscillations in a shallow water model with applications on normal mode initialization*, Beitr. Phys. Atmos. **50** (1977), 253–271.

[85] ———, *The spectral method. Numerical methods used in atmospheric models*, GARP publication series n. 17, 1979, pp. 121–275.

[86] S.J. Majumdar, C.H. Bishop, R. Buizza, and R. Gelaro, *A comparison of ensemble transform Kalman filter targeting guidance with ECMWF and NRL total energy singular vector guidance*, Quart. J. Roy. Meteor. Soc. **128** (2002), 2527–2549.

[87] P. Marquet and D. Giard, *L'initialisation dans arpège*, Tech. report, Météo-France/CNRM, 1992, Note de travail Arpège n. 23.

[88] W.J. Martin and M. Xue, *A comparison of sensitivity analyses from three methods : An adjoint, a very large ensemble, and a new method of random perturbations*, 11th Conference on Mesoscale Processes and 32nd Conference on Radar Meteorology, American Meteorological Society, 24-29 October 2005, pp. –.

[89] A. McDonald, *Transparent lateral boundary conditions for baroclinic waves : a study of two elementary systems of equations*, Tellus **57A** (2005), 171–182.

[90] Richard Ménard and Roger Daley, *The application of Kalman smoother theory to the estimation of 4DVAR error statistics*, Tellus **48A** (1996), 221–237.

[91] F. Molteni, R. Buizza, T.N. Palmer, and T. Petroliagis, *The ECMWF ensemble prediction system : methodology and validation*, Quart. J. Roy. Meteor. Soc. **122** (1996), 73–119.

[92] T. Montmerle, F. Rabier, and C. Fischer, *Relative impact of polar-orbiting and geostationary satellite radiances in the Aladin-France numerical weather prediction system*, Quart. J. Roy. Meteor. Soc. **133** (2007), 655–671.

[93] D. Orrell and L. Smith, *Model error in weather forecasting : does chaos matter ?*, Proceedings of the AGU (San Francisco), AGU, 2000, poster, p. 5 pp.

[94] Steven A. Orszag, *Transform method for the calculation of vector-coupled sums : application to the spectral form of the vorticity equation*, J. Atmos. Sci. **27** (1970), 890–895.

[95] C. Pagé, L. Fillion, and P. Zwack, *Diagnosing summertime mesoscale vertical motion : implications for atmospheric data assimilation*, Mon. Wea. Rev. **135** (2007), 2076–2094.

[96] T.N. Palmer, *A nonlinear dynamical perspective on model error : a proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models*, Quart. J. Roy. Meteor. Soc. **127** (2001), 279–304.

[97] T.N. Palmer, R. Gelaro, J. Barkmeijer, and R. Buizza, *Singular vectors, metrics and adaptive observations*, J. Atmos. Sci. **55** (1998), 633–653.

[98] S.K. Park and D. Županski, *Four-dimensional variational data assimilation for mesoscale and storm-scale applications*, Meteor. Atmos. Phys. **82** (2003), 173–208.

[99] D. Parrish and J. Derber, *The National Meteorological Center's spectral statistical interpolation analysis system*, Mon. Wea. Rev. **120** (1992), 1747–1763.

[100] C. Pires, R. Vautard, and O. Talagrand, *On extending the limits of variational assimilation in nonlinear chaotoc systems*, Tellus **48A** (1996), 96–121.

[101] F. Rabier, *Review of results from A-TREC*, THORPEX/ICSC-5, WMO, 2005, Doc. 2.2.2.

[102] F. Rabier, P. Courtier, J. Pailleux, O. Talagrand, and D. Vasiljevic, *A comparison between four-dimensional variational assimilation and simplified sequential assimilation, relying on three-dimensional variational analysis*, Quart. J. Roy. Meteor. Soc. **119** (1993), 845–880.

[103] F. Rabier, E. Klinker, P. Courtier, and A. Hollingsworth, *Sensitivity of forecast errors to initial conditions*, Quart. J. Roy. Meteor. Soc. **122** (1996), 121–150.

[104] F. Rabier, A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth, and F. Bouttier, *The ECMWF implementation of three-dimensional variational assimilation (3D-VAR). II : Structure functions*, Quart. J. Roy. Meteor. Soc. **124** (1998), 1809–1829.

[105] G. Radnóti, R. Ajjaji, R. Bubnová, M. Caian, E. Cordoneanu, K. von der Emde, J.-D. Gril, J. Hoffman, A. Horányi, S. Issara, V. Ivanovici, M. Janoušek, A. Joly, P. LeMoigne, and S. Malardel, *The spectral limited area model ARPEGE-ALADIN*, PWPR report series n. 7 - WMO TD n. 699, 1995, pp. 111–118.

[106] G. Rivière, B.L. Hua, and P. Klein, *Influence of the β-effect on non-modal baroclinic instability*, Quart. J. Roy. Meteor. Soc. **127** (2001), 1375–1388.

[107] W. Sadiki and C. Fischer, *A posteriori validation applied to the 3D-VAR Arpège and Aladin data assimilation systems*, Tellus **57A** (2005), 21–34.

[108] W. Sadiki, C. Fischer, and J.-F. Geleyn, *Mesoscale background error covariances : recent results obtained with the limited area model ALADIN over Morocco*, Mon. Wea. Rev. **128** (2000), 3927–3935, Notes and correspondences.

[109] M. Shapiro and A. Thorpe, *THORPEX international science plan*, THORPEX/ICSC, WMO, 2004, Version 3.

[110] M. Širokà, C. Fischer, V. Cassé, R. Brožkovà, and J.-F. Geleyn, *The definition of mesoscale selective forecast error covariances for a limited area variational analysis*, Meteor. Atmos. Phys. **82** (2003), 227–244.

[111] C. Soci, C. Fischer, and A. Horányi, *Sensitivity of high resolution forecasts using the adjoint technique at the 10 km scale*, Mon. Wea. Rev. **134** (2006), 772–790.

[112] C. Soci, A. Horányi, and C. Fischer, *Preliminary results of high resolution sensitivity studies using the adjoint of the ALADIN mesoscale numerical weather prediction model*, Idöjárás **107** (2003), 49–65.

[113] S.E. Ştefănescu, L. Berre, and M. Belo Pereira, *The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited-area analysis*, Mon. Wea. Rev. **134** (2006), 3456–3478.

[114] K. Swanson, R. Vautard, and C. Pires, *Four-dimensional variational assimilation and predictability in a quasi-geostrophic model*, Tellus **50A** (1998), 369–390.

[115] O. Talagrand, *Four-dimensional variational assimilation*, ECMWF internal report, ECMWF, 1989, ECMWF Lecture Series.

[116] ———, *A posteriori evaluation and verification of analysis and assimilation algorithms*, Workshop on diagnosis of data assimilation systems, ECMWF, 2-4 november 1998, available from the European Centre for Medium-Range Weather Forecasts, pp. 17–28.

[117] Olivier Talagrand and Philippe Courtier, *Variational assimilation of meteorological observations with the adjoint vorticity equation. I Theory*, Quart. J. Roy. Meteor. Soc. **113** (1987), 1311–1328.

[118] J.N. Thépaut, P. Courtier, G. Belaud, and G. Lemaître, *Dynamical structure functions in a four-dimensional variational assimilation : a case study*, Quart. J. Roy. Meteor. Soc. **122** (1996), 535–561.

[119] M.K. Tippett, J.L. Anderson, C.H. Bishop, T.M. Hamill, and J.S. Whitaker, *Ensemble square root filters*, Mon. Wea. Rev. **131** (2003), 1485–1490, Notes and correspondence.

[120] Z. Toth and E. Kalnay, *Ensemble forecasting at NCEP and the breeding method*, Mon. Wea. Rev. **125** (1997), 3297–3319.

[121] Y. Trémolet, *Model error estimation in 4D-Var*, ECMWF internal report, ECMWF, 2007, ECMWF Research Dept. Technical Memorandum no. 520.

[122] Andrew H. Van Tuyl and Roger Daley, *Estimation of forecast error covariances over an oceanic region using NORPEX data*, 13th Conference on Numerical Weather Prediction, Amer. Meteorol. Soc., 1999, pp. 15–19.

[123] F. Veersé and J.-N. Thépaut, *Multiple-truncation incremental approach for four-dimensional variational data assimilation*, Quart. J. Roy. Meteor. Soc. **124** (1998), 1889–1908.

[124] M.K. Yau, *Perturbation pressure and cumulus convection*, J. Atmos. Sci. **36** (1979), 690–694.

Liste des tableaux

Table des figures

1.1 Géométrie des zones d'intérêt météorologique (C), d'extension (E) et de relaxation aux frontières (I), dans le modèle Aladin

2.1 Evolution itérative des écarts-types d'erreur dans l'exemple à une seule variable. Je montre la convergence de la variance ana
2.2 Température potentielle tous les 4 K, au sol et au toit du modèle semi-géostrophique. La situation correspond au jour 8 de l'
2.3 Ecart-type d'erreur sur la température potentielle tous les 1 K, au sol et au toit du modèle semi-géostrophique. La situation c
2.4 Ecart-type d'erreur sur la température potentielle tous les 1 K, au sol et au toit du modèle semi-géostrophique. La situation c

3.1 Taux d'amplification linéaire du mode normal (trait plein), du vecteur singulier croissant (tireté) et du vecteur singulier décr
3.2 Taux d'amplification linéaire du mode normal (trait plein), du vecteur singulier croissant (tireté) et du vecteur singulier décr
3.3 Structure du vent méridien v_g à l'état initial, pour le mode normal (à gauche) et pour le vecteur singulier (à droite), et pour d
3.4 Structure du vent méridien v_g à l'état initial, pour le vecteur singulier et pour différentes longueurs d'onde. Le temps d'optim
3.5 Evolution temporelle, dans la fenêtre d'assimilation du 4D-VAR Arpège, des écarts au carré observations moins ébauche (tra

4.1 Scores des prévisions tirées à partir de l'assimilation 3D-VAR (bleu tireté) et de l'adaptation dynamique (rouge continu), en
4.2 Scores des prévisions tirées à partir de l'assimilation 3D-VAR et de l'adaptation dynamique, en fonction des échéances, et en
4.3 Spectres des covariances horizontales pour la divergence. Trait plein : données du modèles Aladin/Maroc ; trait tireté : donne
4.4 Séries temporelles du terme de réglage des variances d'erreur d'observations S_o (labels cercles) et du terme de réglage des v
4.5 Séries temporelles de la valeur dans l'analyse finale de la fonction-coût $J(Z^a)$, divisée par le nombre total d'observations (la
4.6 Spectres des variances horizontales pour la température, au niveau 22 du modèle : ébauche Aladin (trait tireté), analyse Arpè
4.7 Cartographie des variances d'erreur de l'analyse Arpège, tronquée, estimée à partir d'un échantillon ensembliste. Le champ
4.8 Précipitations cumulées sur 6 heures, entre 6 et 12 h d'échéance, pour le 03/05/2001. Prévision de contrôle sans perturbation
4.9 Précipitations cumulées sur 6 heures, entre 6 et 12 h d'échéance, pour le 03/05/2001. Prévision de sensibilité avec perturbati

5.1 Evolution au cours des itérations du filtre 0-D des écarts-types d'erreurs d'ébauche (filtres exact, corrigé et faux) et d'observ
5.2 Evolution au cours des itérations du filtre 0-D des écarts-types d'erreurs d'ébauche (filtres exact, corrigé et faux) et d'observ