



ECOLE NATIONALE
DE LA MÉTÉOROLOGIE

ENM



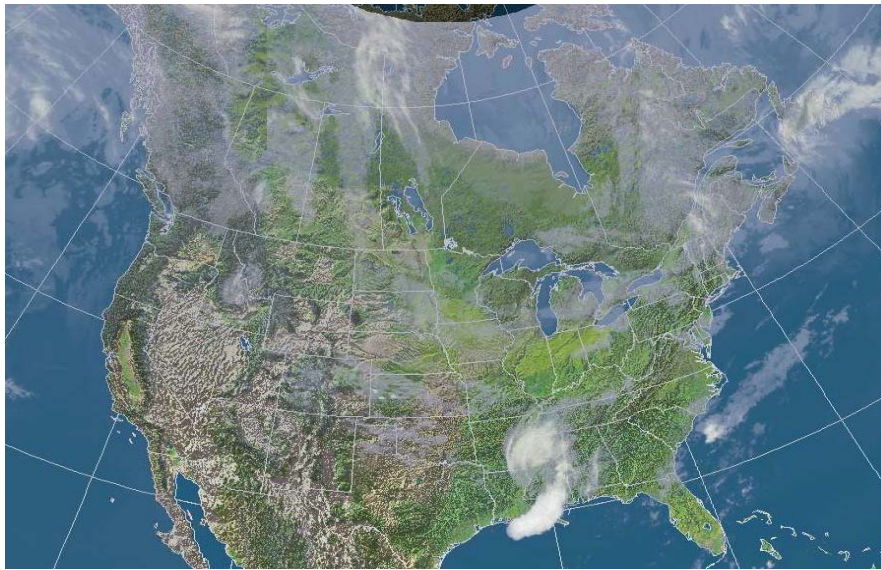
Meteorological Service of Canada
Environment Canada

Service Météorologique du Canada
Environnement Canada



Rapport de stage de fin d'études n° 931

*Mise en oeuvre d'une analyse variationnelle atmosphérique
gérée par le coupleur PALM : une étude de faisabilité*



Amal EL AKKRAOUI

Stage d'approfondissement effectué au Service Météorologique du Canada (SMC)
Sous la direction de

Dr. Pierre GAUTHIER (SMC)
Samuel BUIS (CERFACS)

Promotion d'élèves Ingénieurs Météorologistes de classe I 2001/2004
Janvier – Juin 2004

Remerciements

Parce que ce travail a pu être mené à bien dans de bonnes conditions grâce à la contribution de plusieurs personnes, je tiens à remercier tous ceux qui m'ont aidée, suivie et encouragée tout au long de mon stage. Mes remerciements s'adressent tout d'abord au Dr. Pierre GAUTHIER, mon directeur de stage. C'est un privilège de partager son expérience en matière d'assimilation de données. Aussi, grâce aux communications trans-atlantique, j'ai pu bénéficier de l'aide précieuse de mon encadrant Samuel BUIS du CERFACS, Toulouse. Sans oublier bien sûr, Simon PELLERIN, dont la disponibilité, les suggestions et les critiques ont guidé la progression de mon travail. Je leur suis reconnaissante.

Mes remerciements s'adressent aussi à Thierry MOREL et Michel VALIN pour leur soutien et leur aide, ainsi qu'à Jean PAILLEUX, mon parrain de stage pour les discussions intéressantes qu'il m'a permis d'échanger avec lui. C'est un honneur pour moi.

Pour joindre l'utile à l'agréable, j'ai été accueillie au sein d'une équipe chaleureuse et sympathique. Je garderai un bon souvenir de tous les gens du SMC, leur sympathie et leur bonne humeur. Je les remercie donc tous, sans oublier mes amis qui m'ont aidée à supporter le froid canadien, Aude LEMONSU, Ilenea SADETA, Amaury LABURTHER, Nicolas CHAPELON et GianPaolo BALSAMO.

Enfin, je tiens particulièrement à dédier ce travail à mon père, qui m'a toujours soutenue, guidée et encouragée. Je lui suis reconnaissante ainsi qu'à tous les membres de ma famille.

A tous, MERCI.

Résumé

Le logiciel PALM (Projet d'Assimilation par Logiciel Multi-méthodes) a été développé au CERFACS (Centre Européen de Recherche et Formation Avancée en Calcul Scientifique) pour permettre la gestion d'applications complexes de manière modulaire et parallèle par assemblage de composants élémentaires indépendants. L'implémentation dans PALM du schéma opérationnel d'assimilation de données 3D-Var incrémental du Service Météorologique du Canada (SMC) a été réalisée puis validée dans le cadre de ce travail.

Cette implémentation permet de faire évoluer le code initial vers d'autres méthodes d'assimilation, notamment le 3D-PSAS qu'on étudie ici. Dans le but de résoudre le problème de l'assimilation dans l'espace des observations au lieu de celui du modèle, profitant ainsi de la réduction de la taille du problème ainsi que du nombre de ses degrés de liberté, le 3D-PSAS est formulé à partir du 3D-var et implémenté dans PALM.

La validation des résultats de ce schéma a permis de mettre en exergue - dans un cadre opérationnel - l'équivalence démontrée théoriquement entre le 3D-Var et le 3D-PSAS. En effet, la comparaison des incréments d'analyse et de leurs spectres d'énergie montre que l'on retrouve le même état analysé après la convergence des deux méthodes.

D'autre part, la convergence du 3D-PSAS est comparable à celle du 3D-Var et affiche un résultat plus intéressant après le conditionnement de sa hessienne en convergeant en 15 itérations contre 115 pour un 3D-Var lancé dans les mêmes conditions. Par contre, la fonction objective du PSAS n'étant pas interprétable physiquement, la courbe des valeurs de la fonction coût du 3D-Var à chaque point de la minimisation PSAS présente un comportement singulier en passant par des maxima lors des premières itérations. Le conditionnement de la hessienne permet de diminuer l'amplitude de ces maxima sans les enlever.

Pour cette raison, une nouvelle approche est proposée, dans laquelle la minimisation du problème 3D-Var est effectuée dans l'espace des observations à travers un simple changement d'espace : le 3D-Var-dual. On vise ici à réaliser un compromis entre le 3D-Var (optimisation liée directement à la variance d'erreurs d'analyse) et le 3D-PSAS (réduction de la taille et du nombre de degrés de liberté du problème). On étudie la convergence de cette nouvelle fonction ainsi que le conditionnement de sa hessienne.

Table des matières

Résumé	2
Préambule	5
Introduction	7
1 Etude théorique	8
1.1 L'assimilation de données	8
1.2 Le 3D-Var	9
1.3 Le 3D-Var incrémental	10
1.3.1 Formulation du 3D-Var incrémental	10
1.3.2 Le préconditionnement	11
1.3.3 La résolution des incréments	12
1.4 L'approche duale : le 3D-PSAS	12
2 Implémentation dans PALM	15
2.1 PALM	15
2.2 Le schéma d'assimilation de données au SMC	17
2.3 Implémentation du 3D-Var opérationnel du SMC dans PALM	18
2.3.1 Méthodologie	19
2.3.2 Modularisation	20
2.3.3 Les unités PALM	22
2.3.4 Les étapes de l'algorithme	22
2.4 Implémentation du 3D-PSAS dans PALM	25
2.5 Difficultés et contraintes	26
2.5.1 OPEN-MP/MPI	26
2.5.2 Partage de mémoire	28
2.5.3 Allocations dynamiques	28
3 Résultats et discussions	29
3.1 Validation du 3D-Var PALMé	29
3.2 Résultats du 3D-PSAS	32
3.2.1 Les incréments d'analyse	32
3.2.2 Dualité	37

3.2.3	Pré-conditionnement de la hessienne du 3D-PSAS	39
4	Optimisation du 3D-Var dans un espace réduit	41
4.1	Formulation théorique	41
4.2	Implémentation dans PALM	42
4.3	Résultats du 3D-Var-Dual	42
5	Discussions et perspectives	46
5.1	Les schémas PALMés	46
5.2	Le 3D-PSAS	48
5.3	Le 3D-Var-dual	49
	Conclusion	50
	Annexe	52
	Bibliographie	54

Préambule

Le Service Météorologique du Canada (SMC)

Créé en 1871, le Service Météorologique du Canada a pour mission de fournir des prévisions et des alertes sur les phénomènes météorologiques extrêmes, surveiller la qualité de l'air et prévoir l'état des glaces, des vagues et des eaux intérieures navigables.

Servant le deuxième grand pays au monde en superficie, le SMC couvre, par un réseau étendu d'observations, tout le territoire canadien, même en Arctique. D'autre part, cette étendue géographique est à l'origine d'une régionalisation du service météorologique : 14 centres régionaux et plusieurs bureaux spécialisés se partagent la responsabilité de la couverture territoriale en matière de veille, de prévision et de recherche en météorologie. La coordination entre ces centres est du ressort de l'administration nationale du SMC qui est située en grande partie à Downsview (Ontario).

La Direction de la Recherche Météorologique (MRB) a pour mandat de fournir au Service Météorologique du Canada l'information scientifique nécessaire pour améliorer les prévisions météorologiques et environnementales. Cette information englobe la compréhension des processus atmosphériques et l'établissement de systèmes numériques (modèles) perfectionnés et de technologies de télédétection pour faire les prévisions du temps.

La Direction de la Recherche en Météorologie comprend trois divisions :

1. Assimilation des données et météorologie satellitaire (ARMA) ;
2. Division de la recherche en prévision numérique (RPN) ;
3. Division de la recherche sur la physique des nuages.

Lors de mon stage, j'ai été accueillie au sein de l'équipe ARMA (Assimilation de données) de la Division de l'assimilation de données et de météorologie satellitaire (Dorval, CANADA) dont la mission consiste entre autres à mettre au point des techniques d'assimilation de données météorologiques.

Outre l'expertise en assimilation de données des chercheurs de cette division, j'ai pu aussi bénéficier du soutien et du support technologique assuré par l'équipe Modélisation du Climat et de son Changement Global au CERFACS pour la composante PALM.

L'équipe Modélisation du Climat et de son Changement Global

L'équipe « Modélisation du Climat et de son Changement Global » a été créée en 1990. Depuis 1998, elle fait partie du Centre National de la Recherche Scientifique (CNRS) à travers l'Unité de Recherche Associée URA1875. L'équipe mène des activités de recherches théoriques et appliquées dans le domaine de la modélisation du climat. Ces activités sont réalisées à travers la participation à des programmes Français ou internationaux abordant les différents thèmes dont l'étude de la variabilité climatique globale et régionale, le développement du logiciel OASIS pour le couplage des modèles représentant les différentes composantes du système climatique ainsi que le développement du logiciel PALM pour les problèmes d'assimilation de données.

Introduction

Le logiciel PALM, développé au CERFACS pour les besoins du projet MERCATOR, permet de gérer des applications de calcul scientifique de façon modulaire et parallèle. Dans la mesure où plusieurs méthodes d'assimilation de données utilisent en général les mêmes opérateurs (l'opérateur d'observation, les matrices des covariances des erreurs d'ébauche ou d'observation, le modèle direct, le modèle adjoint, ...), on peut construire ces opérateurs de manière à ce qu'ils soient indépendants de l'application dans laquelle ils interviennent, garantissant ainsi une modularité qu'on peut exploiter en vue de l'implémentation avec PALM.

L'agencement et la construction de la séquence d'exécution de ces opérateurs définit ensuite l'opération ou l'algorithme utilisé.

Dans ce contexte, le présent travail (proposé dans le cadre d'une coopération entre le SMC et le CERFACS) vise, dans un premier temps, à évaluer la faisabilité technique de l'implémentation dans PALM du schéma opérationnel d'assimilation de données 3D-Var incrémental du SMC. Puis, à exploiter, dans un deuxième lieu, la modularité réalisée afin de construire un autre algorithme : Le 3D-PSAS.

Ainsi, la première partie de ce document s'attache à présenter les aspects théoriques de quelques méthodes d'assimilation de données, notamment le 3D-Var, sa variante incrémentale 3D-Var incrémental et le 3D-PSAS. On insiste sur la formulation de leurs fonctions objectives ainsi que leurs propriétés respectives.

On s'intéresse ensuite aux implémentations de ces méthodes dans PALM (chapitre 2) à travers la description de la méthodologie suivie en vue de la modularisation du code initial. Aussi est-il question de discuter des difficultés et contraintes rencontrées au cours de ce travail.

La validation du schéma 3D-Var PALM^é, ainsi que les résultats du 3D-PSAS, font l'objet du troisième chapitre. On y traite de la convergence de ce schéma ainsi que de la dualité 3D-Var/3D-PSAS.

Enfin, une nouvelle approche est présentée dans le chapitre 4 visant à trouver un compromis entre le 3D-Var et le 3D-PSAS. On étudie sa convergence ainsi que le conditionnement de sa hessienne.

Chapitre 1

Etude théorique

1.1 L'assimilation de données

Les systèmes numériques de prévision du temps reposent, en grande partie, sur une bonne définition de leur état initial traduisant l'état réel de l'atmosphère. Le but des méthodes d'assimilation de données est de fournir cet état initial dit **état analysé** en tenant compte et des observations in situ ou satellitaires de grandeurs physiques atmosphériques ou océanographiques et des connaissances théoriques des lois physiques régissant l'atmosphère.

Bien que les données observées soient le meilleur moyen d'approcher la réalité atmosphérique, elles sont mal réparties spatialement et souvent ne sont pas liées directement aux variables du modèle.

Dans le cas où l'état du modèle est surdéterminé par les observations, le problème de l'analyse se réduit à une simple interpolation vers les points de grille du modèle. Autrement, (problème sous déterminé -au moins localement-), il est nécessaire de faire intervenir en plus une autre source d'information plus complète et mieux répartie spatialement tel un jeu de données climatologiques ou le résultat d'un état antérieur du modèle appelé **champ d'essai** ou **ébauche**.

Une assimilation idéale devrait donc permettre d'intégrer toute l'information disponible en tenant compte des incertitudes qui lui sont associées. Il existe plusieurs méthodes d'assimilation de données qui proposent différentes manières d'aboutir à ce compromis, tels que l'interpolation optimale (IO), les méthodes variationnelles (3D-Var et 4D-Var) ou duales (3D-PSAS et 4D-PSAS), la méthode point de selle...

Notations

L'état d'un modèle donné et l'ensemble des observations peuvent être représentés respectivement par un vecteur x de dimension n et un vecteur y de dimension p . Dans toute la suite de ce document, on adoptera les notations suivantes :

x la variable de contrôle contenant toutes les variables analysées (dimension n);
 x^t l'état vrai de l'atmosphère (dimension n);
 x^b le champ d'essai ou ébauche (dimension n);
 x^a l'état analysé du modèle (dimension n);
 y le vecteur des observations (dimension p);
 y' le vecteur d'innovation ou misfit (dimension p);
 \mathcal{H} l'opérateur d'observation non linéaire (de dimension pxn);
 H l'opérateur d'observation linéaire tangent (de dimension $n \times p$);
 B la matrice des covariances d'erreur du champ d'essai ($x^b - x^t$) (dimension $n \times n$);
 R la matrice des covariances d'erreur d'observation ($y - H(x^t)$) (dimension $p \times p$).

Dans ce qui suit, on présente quelques méthodes d'assimilation sur lesquelles porte l'étude, notamment le 3D-Var, le 3D-Var incrémental et le 3D-PSAS.

1.2 Le 3D-Var

L'assimilation variationnelle peut se traduire en terme de minimisation d'une fonction coût qui tient compte des écarts, d'une part, entre l'état du modèle et les observations, et d'autre part, entre l'état du modèle et le champ d'essai.

Dans l'analyse variationnelle tridimensionnelle, dite **3D-Var**, on cherche ainsi, l'état le plus vraisemblable du modèle à partir de la connaissance des lois de probabilité respectives des erreurs d'observation et de l'ébauche de l'état dont on dispose.

Si on suppose, d'une part, que les erreurs d'observation et de prévision sont non corrélées, et d'autre part, que les densités de probabilité de ces erreurs sont des fonctions gaussiennes, alors le meilleur estimé linéaire non biaisé est obtenu en déterminant l'état x du modèle qui minimise la variance d'erreur d'analyse. Ceci correspond au minimum de la fonction coût suivante :

$$J(x) = \underbrace{\frac{1}{2}(x - x^b)^T B^{-1}(x - x^b)}_{J_b} + \underbrace{\frac{1}{2}(\mathcal{H}(x) - y)^T R^{-1}(\mathcal{H}(x) - y)}_{J_o}$$

Dans cette formulation, on distingue deux termes pondérés par les matrices des covariances d'erreurs d'ébauche et d'observation. En effet, celles-ci définissent une norme dans l'espace du modèle et dans l'espace des observations respectivement, permettant de calculer la distance de l'état du modèle à l'ébauche (J_b) et aux observations (J_o).

Remarques :

1. La qualité de l'analyse est tributaire d'une bonne connaissance des covariances d'erreurs d'observation et du champ d'essai :

- les corrélations des erreurs d'ébauche distribuent l'information provenant des observations dans un rayon d'influence autour de ces observations. Ceci est dominant dans des zones où les observations sont éparses.
- les covariances d'erreurs d'ébauche spécifient la manière avec laquelle les valeurs observées vont être lissées eu égard à la différence entre la densité des observations et la résolution du modèle de prévision.
- un effet de balance se traduit en terme de contraintes physiques liant les différentes variables du modèle¹.

2. L'opérateur d'observation sert à relier l'état du modèle aux observations. Il s'agit au minimum d'un opérateur d'interpolation spatiale, mais il peut également être un opérateur d'interpolation physique lorsque les observations sont de natures différentes des champs pronostiques du modèle (ex. les radiances satellitaires).

La minimisation :

Au minimum de la fonction coût, son gradient s'annule. On cherche alors la valeur x pour laquelle le gradient de cette fonction s'annule ($\nabla J = 0$).

Au lieu de tenter de résoudre directement l'équation $\nabla J = 0$ (principe de l'interpolation optimale), l'analyse variationnelle utilise une approche itérative se basant sur la valeur du gradient (la pente) pour converger plus rapidement, et ce en utilisant des algorithmes dits de **descente** tels le gradient conjugué ou quasi-Newton [12].

Ces algorithmes reposent sur le principe selon lequel le gradient fournit de l'information permettant de trouver un nouveau point $x^{(k)} = x^{(k-1)} - \lambda \nabla J(x^{(k-1)})$. En répétant le processus, on réussit à converger graduellement vers le point où J atteint son minimum et où le gradient s'annule avec une précision suffisante.

1.3 Le 3D-Var incrémental

1.3.1 Formulation du 3D-Var incrémental

La formulation incrémentale est une variante des méthodes variationnelles. On reprend la même expression de la fonction coût à l'exception que tous les opérateurs d'observation sont remplacés par leur linéaire tangent ($H \equiv \frac{\delta \mathcal{H}}{\delta x}(x^b)$).

On pose $x = x^b + \delta x$. En linéarisant autour de x on a : $\mathcal{H}(x) = \mathcal{H}(x^b) + H \delta x$, et la fonction coût peut s'écrire comme suit :

$$J(\delta x) = \frac{1}{2} \delta x^T B^{-1} \delta x + \frac{1}{2} [H \delta x - y']^T R^{-1} [H \delta x - y']$$

¹Une observation de température à un point donné peut être lissée pour produire une correction du champ de vent géostrophique.

où $y' = y - \mathcal{H}x^b$ est le vecteur des innovations (ou le misfit). Il représente la différence entre les observations et l'état du modèle ramené aux points d'observation.

Et $\delta x = x - x^b$ est l'incrément. Sa valeur optimale δx^a (dite **incrément d'analyse**) sera ajoutée au terme de l'ébauche x^b afin de fournir l'état analysé x^a .

$$x^a = x^b + \delta x^a$$

L'incrément est supposé être suffisamment petit, de sorte que la solution de cette estimation linéaire constitue une bonne approximation du problème non linéaire. D'autre part, comme discuté dans Laroche et Gauthier (1998) [13], des itérations externes sont nécessaires pour redéfinir les opérateurs d'observation linéarisés et corriger l'écart aux observations pour un δx d'une amplitude finie.

1.3.2 Le préconditionnement

Pour une fonction coût parfaitement isotrope, une seule itération de l'algorithme de descente permet de trouver le minimum, et ce à partir d'un point de départ arbitraire. Dans ce cas, on dit que le problème est parfaitement conditionné. Dans la pratique, il est nécessaire d'introduire un changement de variables qui permet de reformuler la fonction coût pour la rendre plus isotrope. On peut alors, normaliser l'espace des solutions possibles par une norme adéquate, idéalement, la hessienne de la fonction coût :

$$\nabla \nabla_{\delta x} J = B^{-1} + H^T R^{-1} H$$

Cette matrice, de taille égale à celle de la matrice B , correspond également à l'inverse de la matrice des covariances des erreurs d'analyse. En pratique, cet opérateur n'est pas connu, mais peut toutefois être approximé tel que proposé par Andersson et al. (2000) [15]. C'est pourquoi on utilise souvent un préconditionnement par la matrice des covariances des erreurs d'ébauche uniquement, et on définit la nouvelle variable de contrôle du 3D-Var comme :

$$v = B^{-\frac{1}{2}} \delta x.$$

Dans l'espace de cette variable de contrôle normalisée, la fonction coût J_b et son gradient deviennent triviaux :

$$J_b = \frac{1}{2} v^T v \quad \nabla_v J_b = v$$

Ainsi, on peut finalement écrire la fonction coût du 3D-var incrémental sous la forme :

$$J(v) = \frac{1}{2} v^T v + \frac{1}{2} [HB^{\frac{1}{2}}v - y']^T R^{-1} [HB^{\frac{1}{2}}v - y']$$

Son gradient est :

$$\nabla_v J = (I + B^{\frac{1}{2}} H^T R^{-1} H B^{\frac{1}{2}})v - B^{\frac{1}{2}} H^T R^{-1} y'$$

Cette formulation tout en étant semblable à la forme initiale, présente l'intérêt de prendre comme variable de contrôle l'incrément d'analyse δx plutôt que l'état x du modèle. Le vecteur des observations est aussi remplacé par y' , l'écart entre l'observation et le champ d'essai.

1.3.3 La résolution des incréments

La résolution des incréments d'analyse dépend dans une grande mesure des longueurs de corrélation utilisées dans les covariances des erreurs de prévision. Ces longueurs, de l'ordre de quelques centaines de kilomètres, déterminent le rayon d'influence d'une observation, ce qui fait que les incréments nécessitent une résolution inférieure à celle du champ d'essai qui inclut des échelles beaucoup plus fines (Laroche et al.(1999) et Gauthier(2003)).

D'autre part, le premier terme de l'expression de $J(\delta x)$ adapte la structure de l'incrément aux covariances d'erreur d'ébauche qui sont spatialement lissées. Par conséquent, l'incrément d'analyse est lui aussi lissé spatialement. Ce qui justifie l'utilisation d'une faible résolution pour représenter δx , B et H afin de diminuer le coût de la minimisation.

Par contre, pour faire un calcul des innovations ($y - H(X_b)$), on peut se référer au champ d'essai à sa pleine résolution.

Ainsi, l'approche incrémentale permet d'utiliser l'état du modèle à sa pleine résolution pour le comparer aux observations tout en n'ayant à calculer les incréments qu'à une résolution horizontale moindre. Une fois l'incrément obtenu, il est interpolé horizontalement sur la grille à haute résolution et ajouté à l'ébauche.

1.4 L'approche duale : le 3D-PSAS

Formulation du 3D-PSAS

Da Silva et al. (1995) et Cohn et al. (1998) [7] ont proposé une approche équivalente au 3D-Var basée sur la minimisation de la fonction coût dans l'espace des observations plutôt que dans l'espace du modèle. C'est ce qu'on appelle une approche duale ou PSAS (Physical-space Statistical Analysis System), (Courtier 1997 [2]).

Vu que le nombre des observations reste en général faible par rapport à la taille du modèle (surtout dans le domaine océanographique), ces méthodes duales offrent l'avantage de réduire la dimension de la variable de contrôle et par conséquent, le nombre de degrés de liberté du problème.

En partant de l'expression de la fonction coût du 3D-Var :

$$J(\delta x) = \frac{1}{2}\delta x^T B^{-1}\delta x + \frac{1}{2}[H\delta x - y']^T R^{-1}[H\delta x - y']$$

et son gradient :

$$\nabla_{\delta x} J = (B^{-1} + H^T R^{-1} H)\delta x - H^T R^{-1} y'$$

Deux espaces se côtoient ici : le vecteur des incréments d'analyse appartient à l'espace du modèle, tandis que le vecteur des innovations est exprimé dans l'espace des observations.

Le lien entre ces deux espaces est assuré par l'opérateur d'observation H . En effet, $HB^{\frac{1}{2}}v = H\delta x^a$ est le vecteur des incréments d'analyse exprimés dans l'espace des observations.

Au minimum, le gradient s'annule, et on obtient :

$$\delta x^a = BH^T(HBH^T + R)^{-1}y'$$

Pour formuler le 3D-PSAS, on définit un vecteur de contrôle w dans l'espace des observations solution du système linéaire : $(HBH^T + R)w = y'$.

L'incrément δx est lié à w par la relation $\delta x = BH^T w$.

$HBH^T w$ est le vecteur des incréments d'analyse ramené aux points d'observation.

Le système linéaire à résoudre sera donc :

$$F(w) = \frac{1}{2}w^T(R + HBH^T)w - w^T y'$$

et le gradient :

$$\nabla_w F = (R + HBH^T)w - d$$

On introduit le préconditionnement $u = R^{\frac{1}{2}}w$. L'expression finale de la fonction coût du 3D-PSAS et son gradient sont :

$$F(u) = \frac{1}{2}u^T(I + R^{-\frac{1}{2}}HBH^T R^{-\frac{1}{2}})u - u^T R^{-\frac{1}{2}}y'$$

$$\nabla_u F = (I + R^{-\frac{1}{2}}HBH^T R^{-\frac{1}{2}})u - R^{-\frac{1}{2}}y'$$

Courtier (1997) [2] a démontré l'équivalence des deux approches 3D-Var et PSAS à la convergence. Les hessiennes des fonctions objectives ont en effet le même spectre complété éventuellement par des 1. On se trouve alors à résoudre le même problème mais dans un espace différent.

Remarque :

La fonction $J(x)$ minimisée par le 3D-Var est directement liée à la distribution de probabilité "a posteriori" et son minimum correspond au mode de cette distribution (état le plus probable). Dans le cadre du 3D-Var, le terme **fonction coût**, originaire des théories économiques, est donc pleinement justifié car on cherche effectivement à minimiser les écarts aux observations et au modèle comme lorsque l'on cherche à minimiser des dépenses.

Par contre, l'optimisation du problème d'assimilation dans le PSAS relève d'une autre nature puisque la fonction à minimiser $F(u)$ n'a pas d'interprétation immédiate. On peut désormais parler de **fonction objective**.

Chapitre 2

Implémentation dans PALM

2.1 PALM

La philosophie de PALM

PALM (Projet d'Assimilation par Logiciel Multi-méthodes) est un logiciel développé au sein de l'équipe "Global Change and Climate Modelling" au CERFACS pour gérer le système d'assimilation de données du projet océanographique opérationnel **MERCATOR**. Après le succès de cette implémentation, il s'est avéré que le concept de PALM pouvait être exploité dans plusieurs autres domaines, notamment l'assimilation des données atmosphériques.

La philosophie de PALM consiste à décomposer toute méthode d'assimilation en une suite d'opérations élémentaires en séparant celles qui relèvent du problème physique de celles concernant l'algèbre, ce qui permet aux scientifiques de se concentrer sur la physique du problème à traiter. Ces opérations sont traduites en terme de composants indépendants appelés **UNITES PALM**.

Les unités PALM sont des sous-routines ou fonctions (Fortran, C ou C++) capables de traiter les données qu'elles reçoivent et/ou d'en produire d'autres de façon indépendante de l'application dans laquelle elles interviennent. Ces données, appelées **objets**, sont échangées par appel à des primitives PALM (`Palm_Put` et `Palm_Get`) et sont parfaitement identifiées dans la carte d'identité de l'unité dans laquelle l'utilisateur spécifie les types et les dimensions des objets manipulés.

Pour construire l'application, l'utilisateur assemble les unités, dans l'interface graphique PrePALM, au sein de séquences d'exécution qui peuvent inclure tous les types de structures de contrôles fortran (`do`, `while`, `if`, `select case`, ...) : les branches. Les unités contenues dans les branches peuvent être des codes parallèles, et plusieurs branches peuvent être exécutées en parallèle. Ce double niveau permet d'exploiter au maximum le parallélisme intrinsèque d'une application.

Les échanges de données entre unités sont spécifiés par la création dans PrePALM de

communications, représentées par des traits liant les objets des unités concernées (voir figure 2.1). Le logiciel PALM, via un processus dédié à son driver, se charge de gérer les séquences d'exécution des unités et les échanges de données.

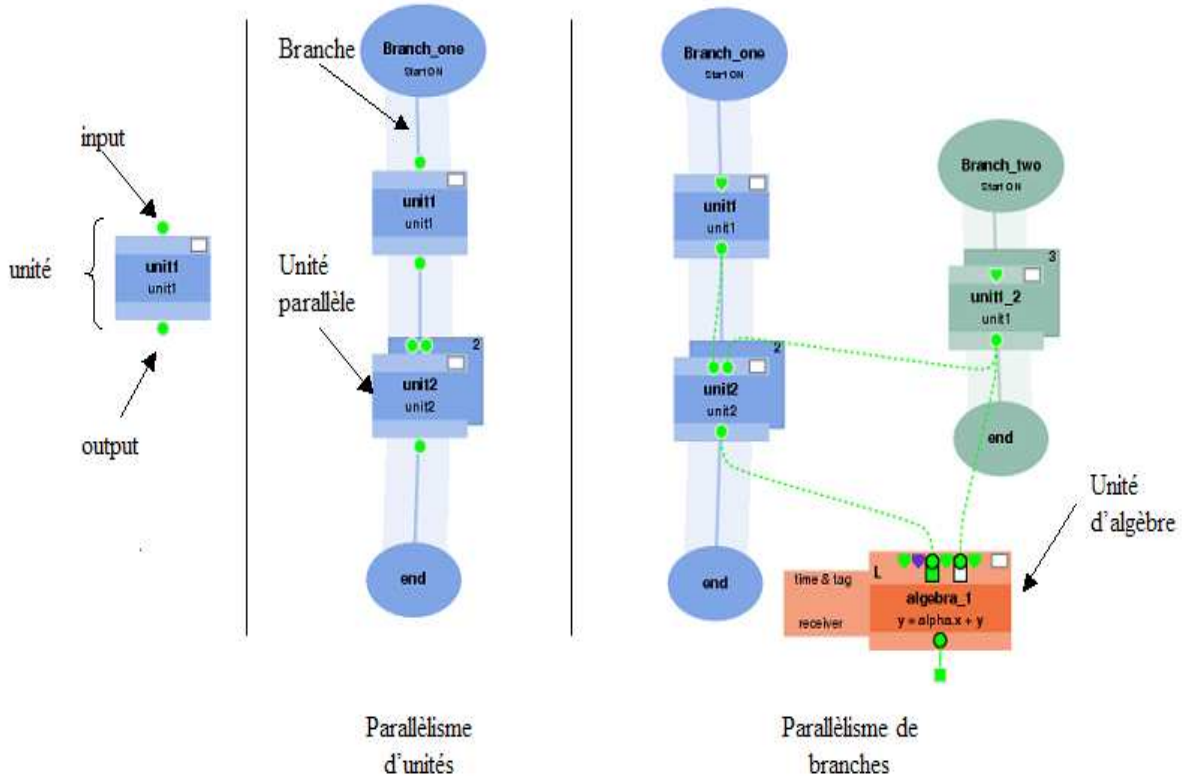


FIG. 2.1 – Exemples de schémas PALM : une unité indépendante (à gauche), une branche munie de deux unités dont une est parallèle (au milieu) et un schéma avec deux branches pouvant tourner en parallèle avec une unité d’algèbre (à droite).

En exploitant le fait que plusieurs méthodes d’assimilation de données utilisent en général les mêmes opérateurs (l’opérateur d’observation, les matrices des covariances d’erreurs, le modèle direct, le modèle adjoint...), on peut construire ces opérateurs de manière à ce qu’ils soient indépendants de l’application dans laquelle ils interviennent.

L’agencement et la construction de la séquence de lancement de ces opérateurs-unités définit ensuite l’opération ou l’algorithme utilisé. Cette approche donne alors un caractère **évolutif** à toute application de ce genre, de même que le passage d’une méthode à une autre s’effectuera sans changement conséquent dans le code de base.

2.2 Le schéma d'assimilation de données au SMC

Le schéma d'assimilation de données au Service Météorologique du Canada (SMC) est un 3D-Var incrémental opérationnel depuis juin 1997. Un modèle global multi-échelles **GEM** (Global Environmental Multiscale) permet de produire des prévisions¹ à moyen et long terme sur une grille uniforme de 0.9° de résolution² (401x200) et 28 niveaux verticaux définis en terme de la coordonnée verticale $\eta = \frac{P-P_T}{P_S-P_T}$, où P_S est la pression de surface et $P_T = 10hPa$.

Les incréments d'analyse sont produits dans une grille gaussienne à une résolution de T108 (240x120).

Comme dans la plupart des centres météorologiques opérationnels, une prévision à courte échéance transportant vers le moment présent l'information fournie par les observations antérieures est utilisée comme ébauche. Il s'agit d'une intégration de 6 heures à partir de la dernière analyse initialisée. Les analyses sont effectuées aux heures synoptiques (T) avec des observations recueillies sur une fenêtre temporelle $T \pm 3h$. Le cycle global 3D-var incrémental du SMC est présenté dans la figure 2.2.

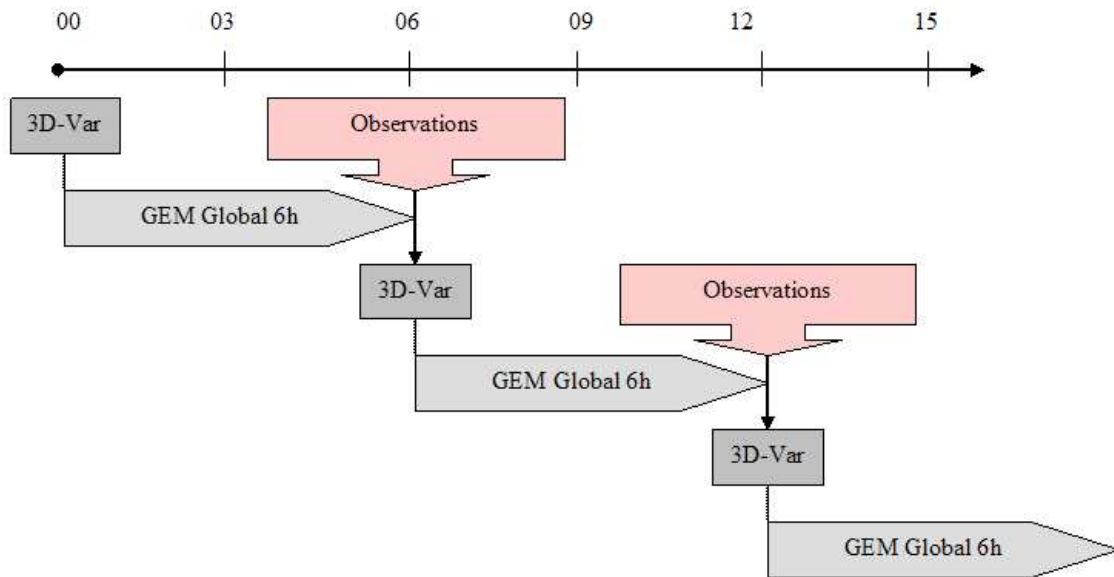


FIG. 2.2 – Cycle d'analyse au Service Météorologique du Canada (SMC).

Le processus de l'assimilation 3D-Var au SMC est le suivant :

¹Il s'agit des prévisions 72h (12 UTC) et 240h (00 UTC).

²C'est la résolution du champ d'essai.

- **Première étape :**
 - Calcul du misfit à haute résolution : $y' = \mathcal{H}(x^b)^{HR} - y$. (HR pour Haute Résolution) ;
 - Définition des linéaires tangents des opérateurs d'observation.
- **Deuxième étape :** Minimisation et calcul des incréments à basse résolution, et corrections des grandes échelles.
- **Troisième étape :** Interpolation et calcul de l'état analysé.

$$(x^a)^{HR} = (x^b)^{HR} + \delta x^{HR}$$

δx^{HR} est obtenu par interpolation de l'incrément basse résolution sur la grille haute résolution.

2.3 Implémentation du 3D-Var opérationnel du SMC dans PALM

La faisabilité scientifique et technique de l'implémentation dans PALM des chaînes d'assimilation de données a été testée, validée puis appliquée opérationnellement dans le cadre du projet MERCATOR ainsi que dans d'autres projets de recherche.

Par ailleurs, l'utilisation du logiciel PALM ne représente pas en soi une quantité de travail conséquente lorsque l'on implémente une chaîne d'assimilation ou que l'on souhaite coupler des codes indépendants. Dans notre cas, il s'agit d'implémenter dans PALM une application déjà existante. C'est-à-dire de la découper en unités élémentaires puis de coupler ces unités dans PALM. Ce travail peut être plus délicat selon le degré de modularité de la chaîne existante et notamment l'utilisation des passages de données entre sous-routines par blocs communs (COMMON³) fortran.

Le but de ce travail était donc, dans un premier temps, d'évaluer la faisabilité technique de l'implémentation dans PALM sans réécriture conséquente de cette chaîne et de choisir la stratégie adéquate pour son découpage. D'autre part, la taille du problème traité dans le cas d'une chaîne opérationnelle d'assimilation de données météorologiques est très importante et la contrainte concernant son coût d'exécution en temps de calcul très forte. Il sera donc essentiel de vérifier que le coût d'utilisation en temps de calcul du logiciel PALM ne dégrade pas les performances de la chaîne.

Dans ce qui suit, on présente la démarche entreprise pour l'implémentation dans PALM ainsi que les contraintes et problèmes rencontrés. Les résultats de cette étude préliminaire sur la faisabilité d'un tel exercice feront l'objet du chapitre suivant.

³Common : c'est un bloc de stockage en mémoire pour que plusieurs unités de programmes puissent partager les mêmes données sans utiliser d'arguments.

2.3.1 Méthodologie

La méthodologie utilisée au cours du projet a été dictée par plusieurs facteurs dont, principalement, la structure existante du 3D-Var, le niveau de modularité exigé par la ré-implémentation d'un PSAS à partir des mêmes unités de calcul que le 3Dvar, certains impératifs de PALM et, bien sûr, les échéances du projet.

La programmatisation du 3D-Var définit un ensemble d'opérations via des sous-routines dont la modularité est en général assez bien circonscrite. Un programme principal (cnt0) organise l'exécution du 3D-Var du SMC en 3 phases principales dont une, la minimisation, a fait l'objet de réaménagement plus important pour sa migration vers PALM.

1. Le pré-traitement :

La première étape de l'exécution initialise un ensemble de matrices et vecteurs qui seront à la base des opérations lors de la seconde phase de la minimisation. Parmi ces paramètres "statiques" on retrouve notamment le vecteur d'innovation, les matrices de covariances d'erreurs d'ébauche et d'erreur d'observations, le masquage de certaines données observées dû à leur positionnement réel par rapport à la topographie de l'ébauche, etc. Cette phase est commune aux différents algorithmes développés dans ce projet. Elle fut donc ré-implémentée sous l'unité **SETUP3D** et représentée sur le canevas PALM de la figure 2.3.

2. La minimisation :

Dans l'implémentation du 3D-Var du SMC, la minimisation est subdivisée en 3 étapes :

- L'évaluation de la fonction coût et de son gradient, avant la minimisation, via un appel au simulateur. Ce simulateur fera l'objet de 2 versions (3D-Var et 3D-PSAS), chacune associée aux 2 algorithmes développés.
- L'appel au minimiseur. Le minimiseur utilisé dans la version 3Dvar est M1QN3. La première étape du développement fut de sortir le minimiseur de la branche principale en faisant l'usage de la boîte d'algèbre de PALM.
- La dernière étape de la minimisation consiste à récupérer les incréments d'analyse et de les "communiquer" à l'étape de post-traitement.

3. Le post-traitement :

Tout comme l'étape du pré-traitement, la phase de post-traitement demeure la même d'un algorithme à l'autre. Cette étape se charge de faire l'interpolation des incréments d'analyse et leur addition aux champs d'essais. Certains ajustements sont faits, en particulier aux champs d'humidité.

Dans tous les algorithmes, cette phase a été implémentée sous l'unité **POSTMIN3D** (figure 2.3).

En vue de l'implémentation dans PALM, ces étapes sont retraduites d'abord en terme d'algorithme détaillé puis en plusieurs entités accomplissant une tâche particulière du processus d'assimilation.

Redéfinition du simulateur :

La suite du développement a concerné principalement la définition du simulateur et de ses différentes versions algorithmiques (3D-Var, 3D-PSAS, ...) et le choix du découpage adéquat en opérateurs élémentaires.

Partons de l'expression de la fonction coût du 3D-Var et de son gradient :

$$J(v) = \frac{1}{2}v^T v + \frac{1}{2}[HB^{\frac{1}{2}}v - y']^T R^{-1}[HB^{\frac{1}{2}}v - y']$$

$$\nabla_v J = (I + B^{\frac{1}{2}}H^T R^{-1}HB^{\frac{1}{2}})v - B^{\frac{1}{2}}H^T R^{-1}y'$$

$$= v + B^{\frac{1}{2}}H^T R^{-1}(HB^{\frac{1}{2}}v - y')$$

On remarque qu'on peut a priori identifier les premières unités PALM, à savoir H , B et R , et qu'on peut d'ores et déjà les isoler, puis les replacer de manière à retrouver la structure de l'algorithme de base.

2.3.2 Modularisation

Afin de réutiliser les mêmes opérateurs, les efforts de développement ont été axés sur la modularisation des opérateurs d'observations.

L'opérateur d'observation H :

Dans la version initiale du 3D-Var, les opérateurs d'observation sont appliqués pour chaque type d'observations en même temps que le calcul de la fonction coût J_o . Pour chaque famille d'observations, on calcule $\frac{HB^{\frac{1}{2}}v - y'}{\sigma}$, σ étant le vecteur des écart-type des erreurs d'observation. On calcule alors la contribution de chaque famille d'observation au terme J_o de la fonction coût.

$$J_o = J_o(raob) + J_o(satem) + J_o(humsat) + J_o(airep) + J_o(satwind) + J_o(surf) + J_o(tov) + J_o(goes)...$$

Afin d'implémenter les opérateurs intervenant dans le calcul de la fonction coût comme des entités informatiques indépendantes, on souhaite construire un seul opérateur d'observation GLOBAL et indépendant du calcul de la fonction coût.

Ainsi, pour chaque famille d'observations, les sous-routines sont modifiées à cet égard.

D'autre part, pour un modèle météorologique, cet opérateur est une combinaison de trois opérations élémentaires :

- La transformée spectrale inverse (S^{-1}) : permet d'obtenir l'état du modèle dans la grille gaussienne à tous les niveaux, après pré-conditionnement de cet état par $B^{\frac{1}{2}}$;
- L'interpolation horizontale aux lieux d'observation (h) : permet d'obtenir l'état du modèle sous forme de profils verticaux aux points d'observation.
- La transformation (V) de l'état profilé du modèle dans l'espace physique des observations (radiances TOVS ou GOES, etc.).

On peut donc écrire : $H = VhS^{-1}$.

Aussi, l'opérateur H est-il implémenté sous forme de trois unités (S_1, H_newbilin et H_eval) assurant chacune une étape des opérations décrites ci-dessus.

Le contrôle de qualité :

Par ailleurs, les opérateurs tel que définis dans la programmathèque du 3Dvar actuel incluent, pour certaines familles d'observations, un contrôle de qualité variationnel (QCvar) lors de l'étape de l'évaluation de la fonction coût. En sortant l'évaluation de la fonction coût des opérateurs d'observations, il s'en suit que le QCvar doit aussi faire l'objet d'une nouvelle implémentation imbriquée dans l'opérateur de calcul de la fonction coût.

Étant donné les questions scientifiques soulevées par l'utilisation d'un QCvar dans un algorithme PSAS et l'échéance trop courte du développement d'un PSAS PALM^é, la décision fut prise, assez tôt dans le projet, d'ignorer le QCvar dans les versions PALM^{ées} du 3Dvar et du PSAS.

Le masque des données assimilées :

Les communications du vecteur d'observation dans le 3D-Var PALM^é sont sur-dimensionnées du fait que la version du logiciel PALM utilisée dans le cadre de ce travail ne dispose pas de la fonctionnalité de définition dynamique de la taille des objets. Elles contiennent donc toutes les observations, même celles ne contribuant pas à l'évaluation de la fonction coût et de son gradient. Le réaménagement des opérations dans le version PALM^{ée} du 3D-Var a nécessité le développement d'un masque global servant à filtrer les observations non assimilées⁴.

⁴rejets dus à des inconsistances entre la surface de l'état du modèle et la hauteur réelle d'une observation par exemple

2.3.3 Les unités PALM

Pour construire le schéma PALM^é du 3D-Var opérationnel du SMC (voir figure 2.3), on a créé les unités suivantes :

- **Setup3d** : initialise les quantités statiques "au sens itératif du terme" comme le vecteur d'innovation, les matrices des covariances d'erreur d'ébauche et d'observation et le masque des données non assimilées...
- **Minimize** : Afin d'isoler le minimiseur de la branche principale du simulateur, cette étape a du être fragmentée en trois sections statiques branchées via des GOTO : une initialisation des quantités nécessaires au minimiseur, une section assurant les communications avec celui-ci et une troisième section récupérant le résultat à la convergence.
- **S_1, H_newbilin et H_eval** : trois unités qui constituent l'opérateur d'observation représentant respectivement le changement d'espace et le préconditionnement par $B^{\frac{1}{2}}$, l'interpolation horizontale puis la transformation des états profilés vers l'espace physique des observations. Les modifications apportées à cet opérateur sont entreprises dans l'unité H_eval.
- **Yprim** : extrait le vecteur des observations.
- **SQRT_R_1** : applique $1/\sigma$ à un vecteur pris comme entrée.
- **S_1_T, H_newbilin_T et H_eval_T** : représentent l'adjoint de l'opérateur d'observation. Celui-ci est construit de la même manière que l'opérateur direct.
- **postmin3d** : récupère l'incrément d'analyse, effectue le calcul de l'état analysé et termine le processus de l'assimilation.

2.3.4 Les étapes de l'algorithme

Le schéma PALM du 3D-Var est présenté dans la figure 2.3. Les boîtes oranges représentent les unités d'algèbre fournies par la bibliothèque d'algèbre de PALM, dont le minimiseur M1QN3 (algebra_1). La branche PALM (trait vert) structure la séquence d'appel aux unités (rectangles verts) à l'aide des structures de contrôle Fortran (DO,IF...).

La boucle de minimisation est repérée par un cercle blanc dans la branche et un cercle noir pour la fin. Les unités avec un demi-cercle à côté sont appelées dans la boucle.

Afin de mieux comprendre ce schéma, on peut rappeler la fonction coût du 3D-Var et son gradient.

$$J(v) = \frac{1}{2}v^T v + \frac{1}{2}[HB^{\frac{1}{2}}v - y']^T R^{-1}[HB^{\frac{1}{2}}v - y']$$

$$\begin{aligned}\nabla_v J &= (I + B^{\frac{1}{2}}H^T R^{-1}HB^{\frac{1}{2}})v - B^{\frac{1}{2}}H^T R^{-1}y' \\ &= v + B^{\frac{1}{2}}H^T R^{-1}(HB^{\frac{1}{2}}v - y')\end{aligned}$$

Les unités présentées ci-dessus, sont aménagées dans le canevas PALM (PrePALM) afin de reconstruire les étapes de l'algorithme du 3D-var PALM qu'on développe ici.

1. Initialisation : Initialisation des paramètres, calcul du vecteur d'innovation et du masque des données assimilées (unité setup3d) ;
2. Minimisation :
 - Le vecteur de contrôle v (de dimension n) est initialisé à zéro au début de la minimisation dans l'unité setup3d et envoyé par l'unité minimize_1 ;
 - La fonction coût :
 - Calcul de l'incrément $d = HB^{\frac{1}{2}}v$ (unités S_1, H_newbilin et H_eval) ;
 - Calcul de $(d - y')$ (unités Yprim et algebra_6) ;
 - Calcul de $d^n = R^{-\frac{1}{2}}(d - y')$ (unité SQRT_R_1) ;
 - Calcul de $J_o = d^{nT}d^n$ (algebra_8) ;
 - Calcul de $J_b = \frac{1}{2}v^T v$ (algebra_2) ;
 - Calcul de $J = J_o + J_b$ (algebra_3).
 - Le gradient :
 - Calcul de $R^{-1}(d - y')$ à partir de $R^{-\frac{1}{2}}(d - y')$ (unité SQRT_R_1) ;
 - Calcul de $\nabla_v J_o = B^{\frac{1}{2}}H^T R^{-1}(d - y')$ (unités H_eval_T, H_newbilin_T et S_1_T) ;
 - Calcul de $\nabla_v J = v + \nabla_v J_o$ (algebra_4).
 - Le minimiseur M1QN3 (algebra_1) évalue la nouvelle valeur du v à laquelle la fonction et/ou son gradient doivent être réévalués. Il tourne en parallèle avec la branche jusqu'à la fin de la minimisation (à la convergence ou après atteinte des critères d'arrêt) ;
3. Calcul de l'incrément d'analyse (d) issu de la dernière itération ;
4. Fin de l'assimilation : interpolations et calcul de l'état analysé x^a (unité postmin3d).

On notera qu'aucune corrélation d'erreurs d'observation n'est considérée dans le système 3D-Var du SMC.

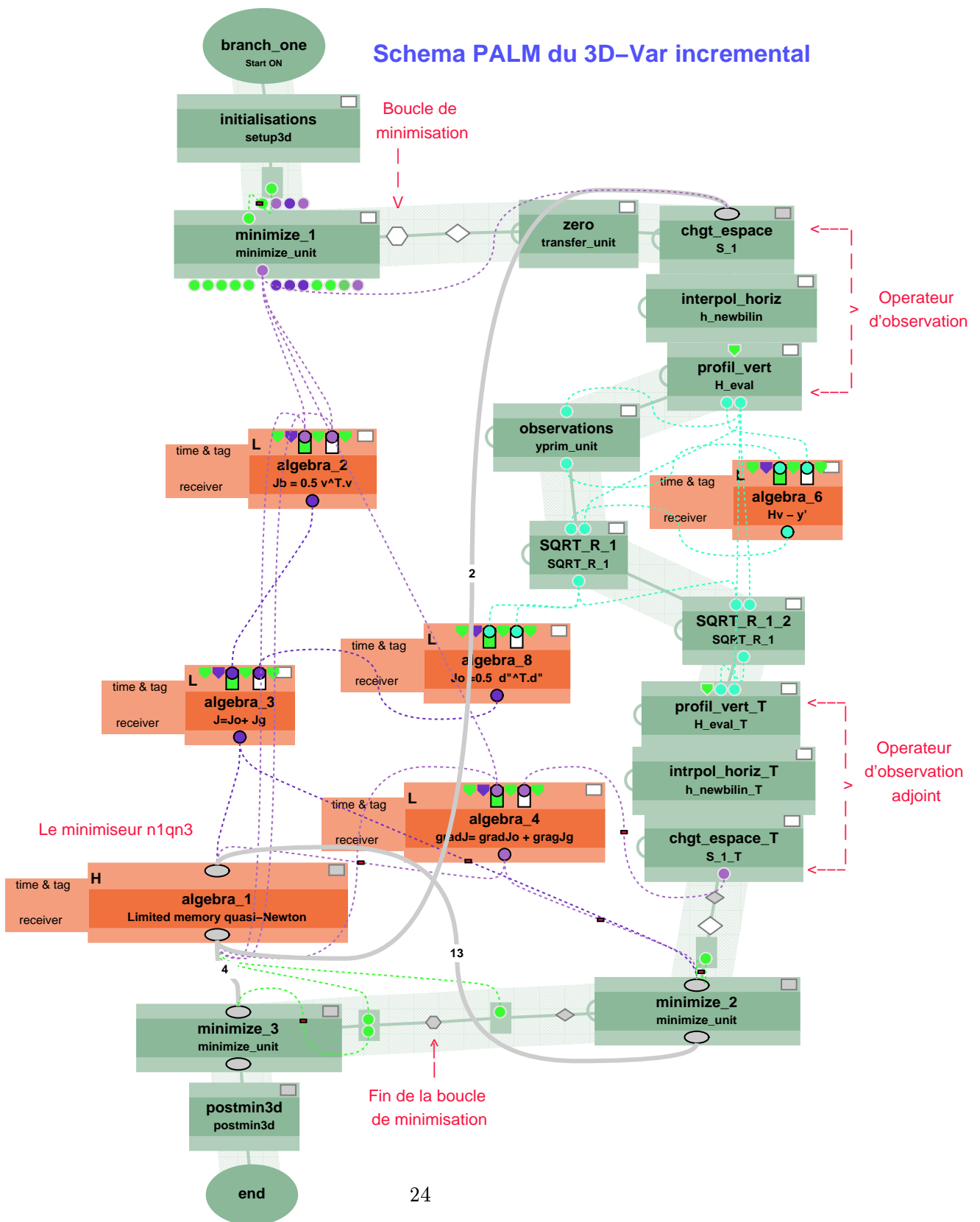


FIG. 2.3 – Schéma PALM du 3D-Var incrémental

2.4 Implémentation du 3D-PSAS dans PALM

Etant donné que l'on retrouve des opérateurs communs au 3D-Var et au PSAS, l'idée consiste à exploiter les mêmes unités déjà construites, en rajouter quelques-unes éventuellement, et les coupler différemment pour construire le nouvel algorithme PSAS.

Le vecteur de contrôle est maintenant défini dans l'espace des observations, soit de taille égale au nombre total des observations **assimilées** (p). La taille des vecteurs à communiquer au minimiseur variera donc d'une assimilation à l'autre. Pour permettre des communications dynamiques, il est nécessaire de communiquer des vecteurs sur-dimensionnés avec leurs masques de données assimilées (fournis par l'unité setup3d). Un vecteur de "travail" est défini ayant comme taille le nombre total de données observées disponibles (p').

Aussi a-t-on besoin de construire des unités de conversion permettant de concaténer ou d'extraire le vecteur de travail à partir du vecteur de contrôle et vice-versa (opérations d'injection et de projection effectuées par les unités nx2nz et nz2nx).

Rappelons les formules de la fonction objective du 3D-PSAS et de son gradient :

$$F(u) = \frac{1}{2}u^T(I + R^{-\frac{1}{2}}HBH^T R^{-\frac{1}{2}})u - u^T R^{-\frac{1}{2}}y'$$

$$\nabla_u F = (I + R^{-\frac{1}{2}}HBH^T R^{-\frac{1}{2}})u - R^{-\frac{1}{2}}y'$$

Ainsi, pour implémenter le 3D-PSAS dans PALM, on suit les étapes de l'algorithme suivant.

1. Initialisation : Initialisation des paramètres, calcul du vecteur d'innovation y' , du masque des données assimilées ainsi que du vecteur $d' = R^{-\frac{1}{2}}y'$;
2. Minimisation :
 - Le vecteur de travail u' : de dimension p' ;
 - Calcul de $d'' = u' + R^{-\frac{1}{2}}HBH^T R^{-\frac{1}{2}}u'$, $u'^T d'$;
 - La fonction coût $F = \frac{1}{2}u'^T d'' - u'^T d'$ et son gradient $\nabla_u F = d'' - d'$;
 - Construction des vecteurs qui ne tiennent compte que des données assimilées. (concaténation du gradient et du vecteur de travail);
 - Le minimiseur évalue la nouvelle valeur du vecteur de contrôle u (de taille p) à laquelle la fonction et/ou son gradient doivent être réévalués;
 - Reconstruction du vecteur u' à partir de u donné par le minimiseur;
3. Calcul de l'incrément d'analyse $d = B^{\frac{1}{2}}H^T R^{-\frac{1}{2}}u'$ issu de la dernière itération;
4. Fin de l'assimilation : interpolations et calcul de l'état analysé x^a .

Pour l'implémentation dans PALM de ce schéma (voir figure 2.4), les seules unités à construire sont celles des conversions de vecteurs⁵.

Remarque

Le critère de sélection des observations assimilées (le contrôle de qualité) est le même pour le 3D-Var et le PSAS. Ce qui nous place dans un cas optimal pour la comparaison des deux méthodes.

2.5 Difficultés et contraintes

On a recensé dans cette partie les difficultés et contraintes techniques rencontrées lors de l'implémentation dans PALM de la chaîne d'assimilation de données 3D-Var incrémental du Service Météorologique du Canada (SMC).

2.5.1 OPEN-MP/MPI

Le SMC dispose d'un supercalculateur 28 IBM pSeries 690 (120 nœuds de 8PEs⁶ par nœud). Il est utilisé dans un environnement MPI pour les applications opérationnelles ainsi qu'à des fins de recherche sur l'atmosphère.

Le 3D-Var du SMC fait appel à OpenMP pour la gestion du parallélisme dans certaines portions des calculs (transformées spectrales, transfert radiatif). Comme Open-MP requiert un partage de mémoire entre les «threads»⁷ des portions de calcul parallèle, cela implique que l'exécution du 3D-var doit être confinée à un nœud (8PEs) de l'IBM (voir annexe).

L'utilisation de PALM pour la chaîne 3D-VAR implique, non seulement, un parallélisme du type «multi-threading», mais aussi de tâches multiples MPI. Dans la version PALMée du 3D-Var et du 3D-PSAS, trois tâches sont requises : une pour la branche algorithmique (3D-Var ou 3D-PSAS), une pour le minimiseur (opération algébrique de haute priorité) et une troisième partagée entre le driver et les opérations algébriques à basse priorité.

Par ailleurs, pour tirer parti des calculs multi-threads dans la version PALMée du 3D-Var (ou du 3D-PSAS), il est essentiel d'associer plus d'un thread à la branche algorithmique.

La séquence et la distribution des ressources des travaux sur l'IBM du SMC sont assurées par le logiciel LoadLeveler (LL). Ce dernier associe symétriquement un nombre de threads par tâche aux travaux soumis. Afin d'associer un seul thread aux deux tâches d'algèbre, des ré-assignations explicites ont dû être faites au moment du lancement de l'exécutable et dans le code de branche de manière à déjouer la symétrie imposée par LoadLeveler.

⁵Il existe des unités d'algèbre qui peuvent faire cette opération

⁶PE : Processeur Élémentaire

⁷thread : processus léger, correspondant à l'exécution d'un petit programme, ou d'une routine d'un programme plus gros, indépendamment de celui-ci (on parle alors de multithread).

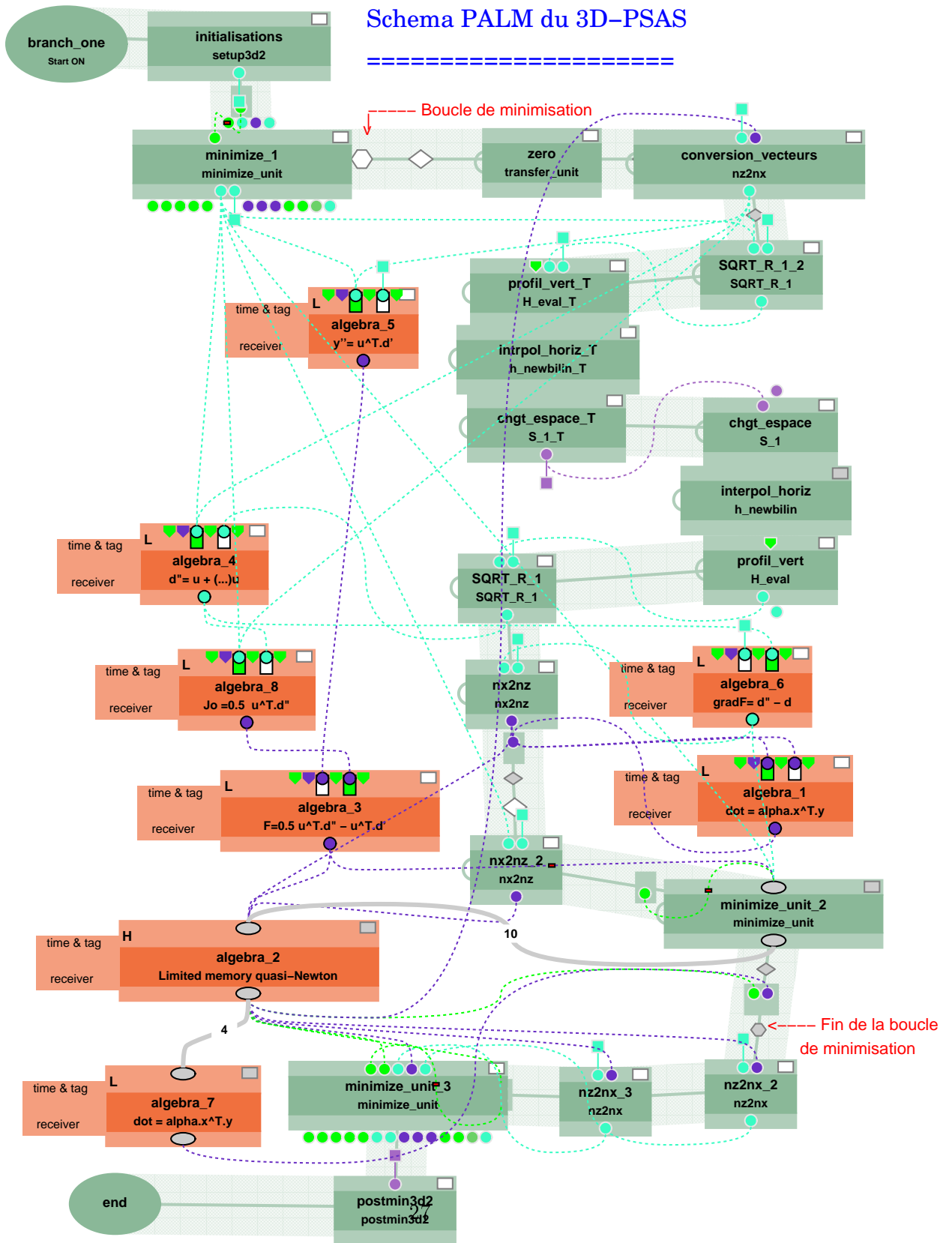


FIG. 2.4 – Schéma PALM du 3D-PSAS

2.5.2 Partage de mémoire

Tel que mentionné plus haut, le 3D-Var du SMC ne comporte pas de décomposition de données traitées en multi-tâches. Cette structure des données ne permet pas d'implémenter une application sur deux branches. Chose qui peut s'avérer utile pour exploiter le parallélisme de branches offert par PALM. Il s'en suit que l'implémentation dans PALM du 3D-Var et du 3D-PSAS ne fait usage que d'une seule branche principale de calcul.

D'autre part, étant donné que l'on utilise la boîte d'algèbre faisant appel au minimiseur M1QN3, ce dernier a besoin d'un certain nombre de paramètres qui, dans la configuration initiale, sont passés dans des commons. On utilise alors un procédé prévu par M1QN3 qui consiste à construire trois vecteurs entier, logique et double précision servant de "cheval de Troie" pour transporter toutes ces variables concaténées.

2.5.3 Allocations dynamiques

Dans la plupart des applications, les dimensions du problème qu'on traite sont connues à l'avance. PALM-Research ne permet pas de traiter le cas où l'on a besoin de définir un objet dont la taille sera déterminée à l'intérieur de l'application elle-même. Cette contrainte a été soulevée lors de l'implémentation du 3D-PSAS vu que l'on a besoin du nombre exact des données assimilées pour définir le vecteur de contrôle.

En général, on peut contourner ce problème par un sur-dimensionnement des objets échangés (c'est le cas du 3D-Var). Par contre, pour le PSAS, la solution n'est pas aussi triviale. Dans la mesure où seul le minimiseur a besoin de la taille exacte du vecteur de contrôle, on peut définir une unité qui remplace la boîte d'algèbre M1QN3 et qui a comme entrée supplémentaire un vecteur surdimensionné contenant le masque des données assimilées. A l'intérieur de cette unité, on peut appliquer le masque puis appeler le minimiseur avec le vrai vecteur de contrôle.

Dans la présente implémentation, la solution du sur-dimensionnement des vecteurs a été retenue pour le 3D-PSAS vu que l'on connaît la taille de notre problème.

Notons par ailleurs que cette fonctionnalité (définir les dimensions des objets dynamiquement) sera intégrée dans PALM-MP (voir page 47) d'ici avril 2005 pour les objets non distribués et avril 2006 pour les objets distribués.

Chapitre 3

Résultats et discussions

Tous les résultats présentés ici sont obtenus avec une expérience numérique complète au sein de laquelle tous les types d'observation sont représentés et contenant toutes les données retenues pour la passe opérationnelle du 15 janvier 2004, soit **280 217** observations dont **125 159** sont assimilées. Le vecteur de contrôle du modèle est de taille **21 670 950**.

3.1 Validation du 3D-Var PALM_é

La faisabilité technique de l'implémentation dans PALM du schéma 3D-Var du SMC étant mise en évidence dans le chapitre précédent, il s'agit ici de valider cette implémentation tant au niveau de l'optimisation du temps de calcul que des ressources exploitées.

Dans la mesure où les modifications apportées au code 3D-Var initial en vue de l'implémenter dans PALM ne sont pas de nature à changer l'algorithme de base¹, on ne s'attend pas à avoir une amélioration du temps de calcul nécessaire pour faire une analyse 3D-Var. On cherche alors à mesurer le surcoût en temps et en moyens de calcul engendré par l'utilisation de PALM, notamment au niveau des communications.

Dans la figure 3.1, sont présentés les résultats de l'analyseur de performances de PALM, qui mesure le temps de calcul et de communications d'une application PALM.

Les trois tâches (driver, branche et minimiseur) sont analysées séparément (proc0, proc1 et proc2). La légende des couleurs pour les unités utilisées dans l'application est présentée tout en haut du schéma. En dessous, les barres horizontales montrent la séquence d'exécution des unités au sein de chaque tâche MPI. Enfin, tout en bas, les camemberts montrent la répartition du temps de calcul par unité au sein de chaque tâche.

¹sauf pour l'opérateur d'observation.

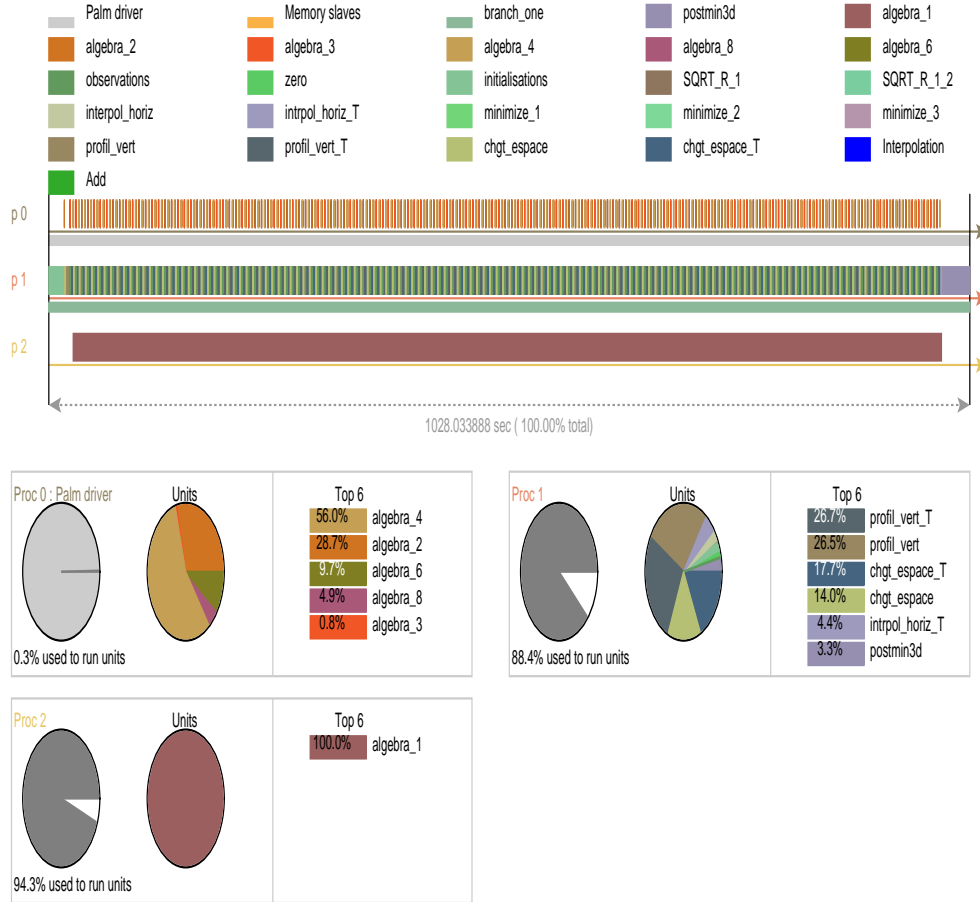


FIG. 3.1 – Résultats de l’analyseur de performances de PALM pour le schéma 3D-Var.

La branche (proc1), organise la séquence de lancement et d’exécution des unités. Celles-ci se partagent le temps de calcul de manière disproportionnée en fonction de la charge de calcul nécessaire. Ainsi, les opérateurs adjoint et direct d’observation par exemple, consomment presque 75% du temps de la branche (26.7% pour l’unité profil_vect_T, 26.5% pour l’unité profil_vect ...).

Le minimiseur M1QN3 (proc2) utilise 94,3% du temps total de l’exécution. Toutefois, étant en «communication directe²», il consomme ce temps en **attente bloquante** sans utiliser les ressources qui lui sont allouées. D’autre part, le driver (proc0) de PALM s’occupe de gérer la séquence temporelle de lancement des unités, les communications, le stockage éventuel - temporaire ou permanent (BUFFER) - des objets, et d’exécuter les unités d’algèbre basse priorité. Celles-ci n’étant pas nombreuses dans notre cas, le driver consacre seulement 0.3% du temps à cette tâche.

²après avoir trouvé le point de la minimisation pour une itération donnée, le minimiseur se met en **attente** de la fonction coût et/ou de son gradient correspondant à ce point et ne rend la main que lorsqu’il converge.

Ainsi, les deux tâches consacrées, dans l'application PALMée, au driver et au minimiseur sont dans un cas sous utilisées et consommées en attente bloquante dans l'autre.

On présente dans le tableau 3.1 une estimation des temps de calcul utilisés par chaque méthode sachant que nos expériences sont, autant que possible, lancées dans des conditions analogues. Le temps mur représente la différence en secondes entre le début et la fin de l'exécution de l'application. Le temps CP est le temps consommé par la tâche considérée. Les tâches représentées dans le tableau 3.1 sont associées à la branche des simulateurs du 3D-Var, du 3D-Var PALMé et du 3D-PSAS respectivement. Comme ces tâches nécessitent 4 threads chacune, l'utilisation d'OPEN-MP réduit d'environ un facteur 4 le temps mur requis (Temps par CPU).

Méthode	3D-Var	3D-Var palmé	3D-PSAS
Temps mur (s)	1003	1006	949
Temps CP secs (s)	4002	4007	3783
Temps par CPU (s)	1000.5	1001.75	945.7
Nombre d'itérations	126	126	100

TAB. 3.1 – Comparaison des temps de calcul nécessaires pour faire une analyse pour le 3D-Var, le 3D-Var PALMé et le 3D-PSAS

Ainsi, la somme des ressources utilisées est 6x1000 s pour le 3D-Var PALMé (6 CPU utilisés) contre 4x1000 s pour la version non PALMée. Rappelons que PALM nécessite 2 threads de plus pour le driver et la partie algébrique et pour le minimiseur M1QN3. Notons ici que si le minimiseur était en "communication renversée", ce surcoût disparaîtrait puisque le driver s'occupera aussi de l'exécution de M1QN3.

Toutefois, à quelques secondes près, les différences en temps de calcul sont minimes, ce qui était prévisible étant donné le nombre réduit de communications dans le schéma PALM. En effet, on s'est contenté dans ce travail de recenser les échanges de variables "dynamiques" qui sont passées en "common" puis les échanger via des communications explicites entre les unités, réduisant ainsi le nombre de communications établies.

D'autre part, dans le but de préserver le code initial aussi intact que possible, le calcul se trouve parfois dédoublé pour PALM (dans le cas de l'application de l'opérateur d'observation par exemple), ce qui engendre un léger coût supplémentaire.

Somme toute, la gestion du schéma d'assimilation de données 3D-Var incrémental du SMC par PALM engendre un surcoût qui reste négligeable au niveau du temps mur, mais nécessite le déploiement de deux CPU additionnels. D'autre part, le découpage entrepris dans le code initial du 3D-Var permet une migration plus facile vers d'autres méthodes d'assimilation. Des améliorations restent aussi à apporter au schéma PALMé en exploitant d'autres possibilités offertes par PALM au niveau du parallélisme et de la gestion de mémoire.

3.2 Résultats du 3D-PSAS

Tel qu'indiqué dans le tableau 3.1 et la figure 3.2, le 3D-PSAS converge en 100 itérations contre 126 pour le 3D-Var.

Dans Courtier (1997) [2], on montre que le nombre de conditionnement est le même pour le 3D-Var et le 3D-PSAS. On s'attend donc à ce que ces algorithmes convergent dans un nombre d'itérations comparable. Nos résultats confirment donc cette analyse.

Par contre, le nombre d'itérations étant variable d'une situation à une autre, une seule expérience n'est pas suffisante pour conclure à un taux de convergence en faveur du PSAS.

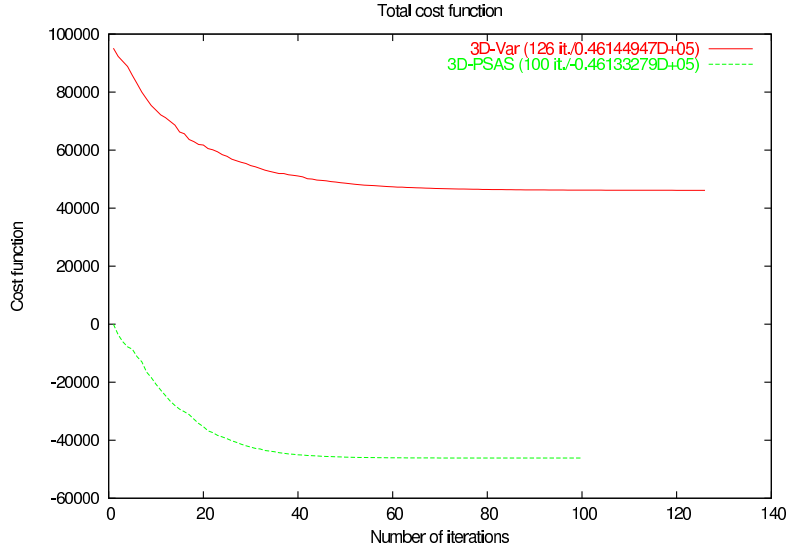


FIG. 3.2 – Graphe des fonctions objectives du 3D-Var et du 3D-PSAS.

3.2.1 Les incréments d'analyse

L'équivalence démontrée théoriquement entre le 3D-Var et le 3D-PSAS est ici mise en exergue par la comparaison des incréments d'analyse après la convergence. La figure 3.3 est une représentation des incréments d'analyse obtenus après la convergence des deux méthodes pour le champ de température au niveau 500 hPa.

Les différences entre les deux champs sont négligeables, d'autant plus qu'on a pu établir une corrélation de 0.999. Il en est ainsi pour les autres champs à différents niveaux. Ce qui nous mène à une première conclusion : les incréments 3D-Var et PSAS sont équivalents.

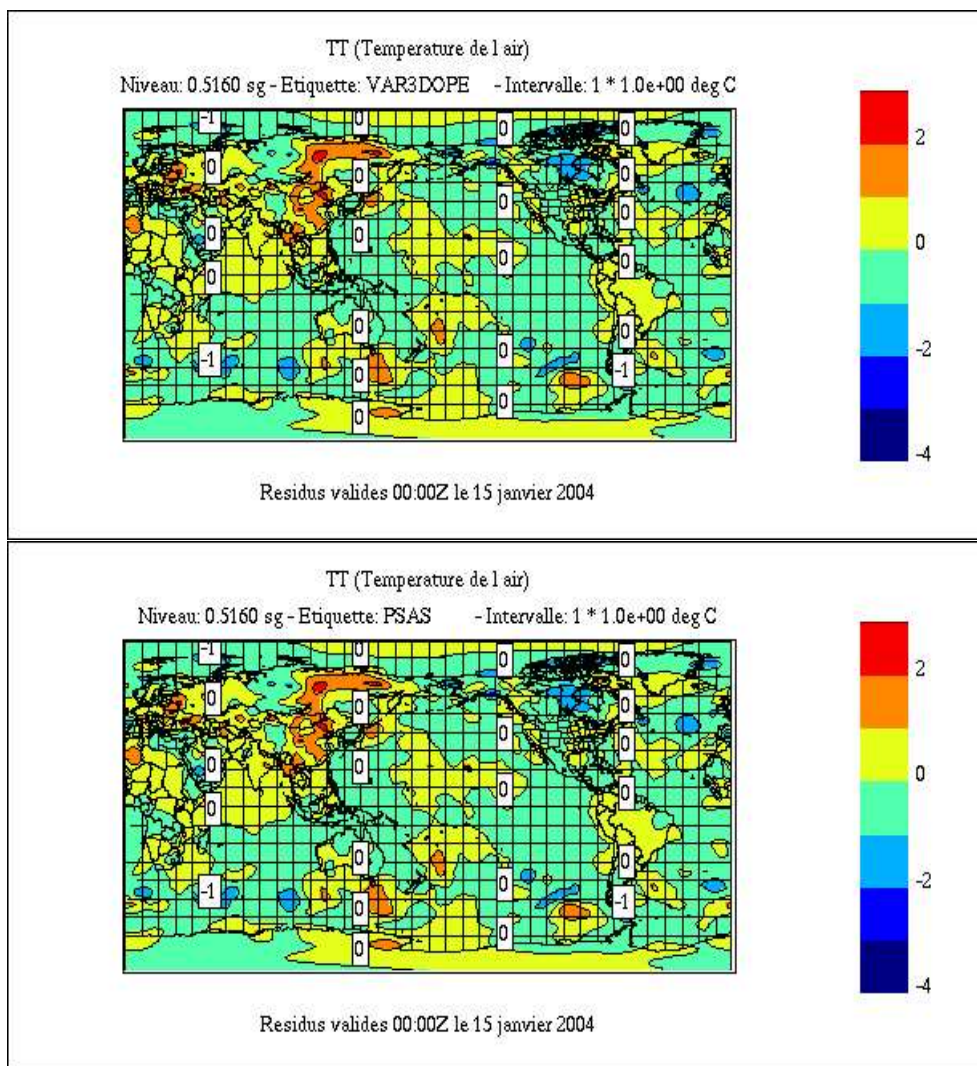


FIG. 3.3 – Les incréments d’analyse du 3D-Var (en haut) et du 3D-PSAS (en bas) pour le champ de température au niveau 500 hPa.

Les spectres d’énergie

Pour confirmer le résultat obtenu ci-dessus, on étudie les spectres d’énergie des incréments d’analyse 3D-Var et 3D-PSAS. On peut ainsi déceler les petites différences de comportement selon les nombres d’onde. Dans la figure 3.4 on a superposé les spectres d’énergie des incréments 3D-Var (en rouge) et 3D-PSAS (en vert). Les courbes sont identiques.

Remarque

Notons que les corrections sont apportées autour du nombre d'onde 10, ce qui représente une signature conforme à celle observée pour le spectre d'énergie des erreurs du modèle.

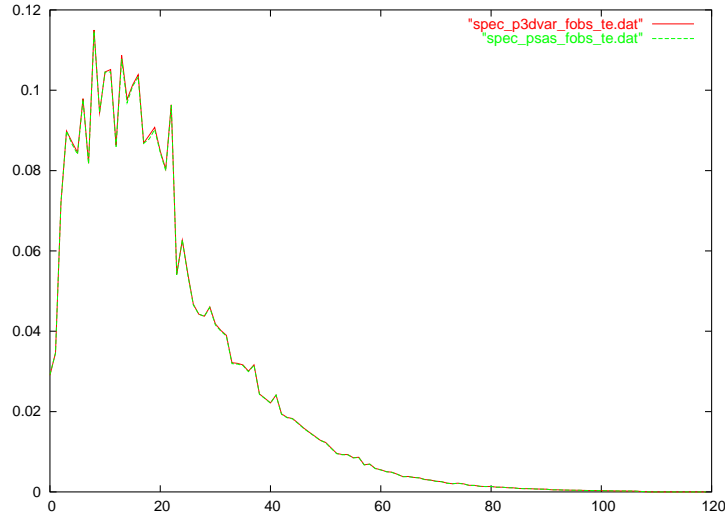


FIG. 3.4 – Spectres d'énergie des incréments d'analyse 3D-Var (en rouge) et 3D-PSAS (en vert). En abscisse, les nombres d'onde et en ordonnée l'énergie totale.

Les prévisions

A partir de deux analyses 3D-PSAS et 3D-Var, on a lancé une prévision 120 h dans le but de valider les résultats contre les observations de radiosondage³ recueillies à l'échéance 120 h. Les écarts-type et biais des champs "Observation - Prévision" sont tracés dans les figures 3.5 et 3.6 pour les hémisphères Nord et Sud respectivement. On distingue les champs de vent (UU, VV), de géopotential (GZ), de température (TT) et d'humidité spécifique.

A gauche sont tracés les niveaux verticaux de pression (16 niveaux) et à droite le nombre d'observations de radiosondage par niveau.

Les courbes du 3D-Var et 3D-PSAS pour l'hémisphère Nord se distinguent difficilement, avec un léger avantage pour le 3D-Var, surtout dans les niveaux moyens de l'atmosphère. Dans l'hémisphère Sud (figure 3.6) où le nombre de données observées est faible, le PSAS se démarque dans les basses couches.

Toutefois, l'impact sur la prévision n'est pas certain du fait que l'on valide contre un nombre peu élevé d'observations et un seul cas météorologique.

³Réseau OMM des sites classés de radiosondages

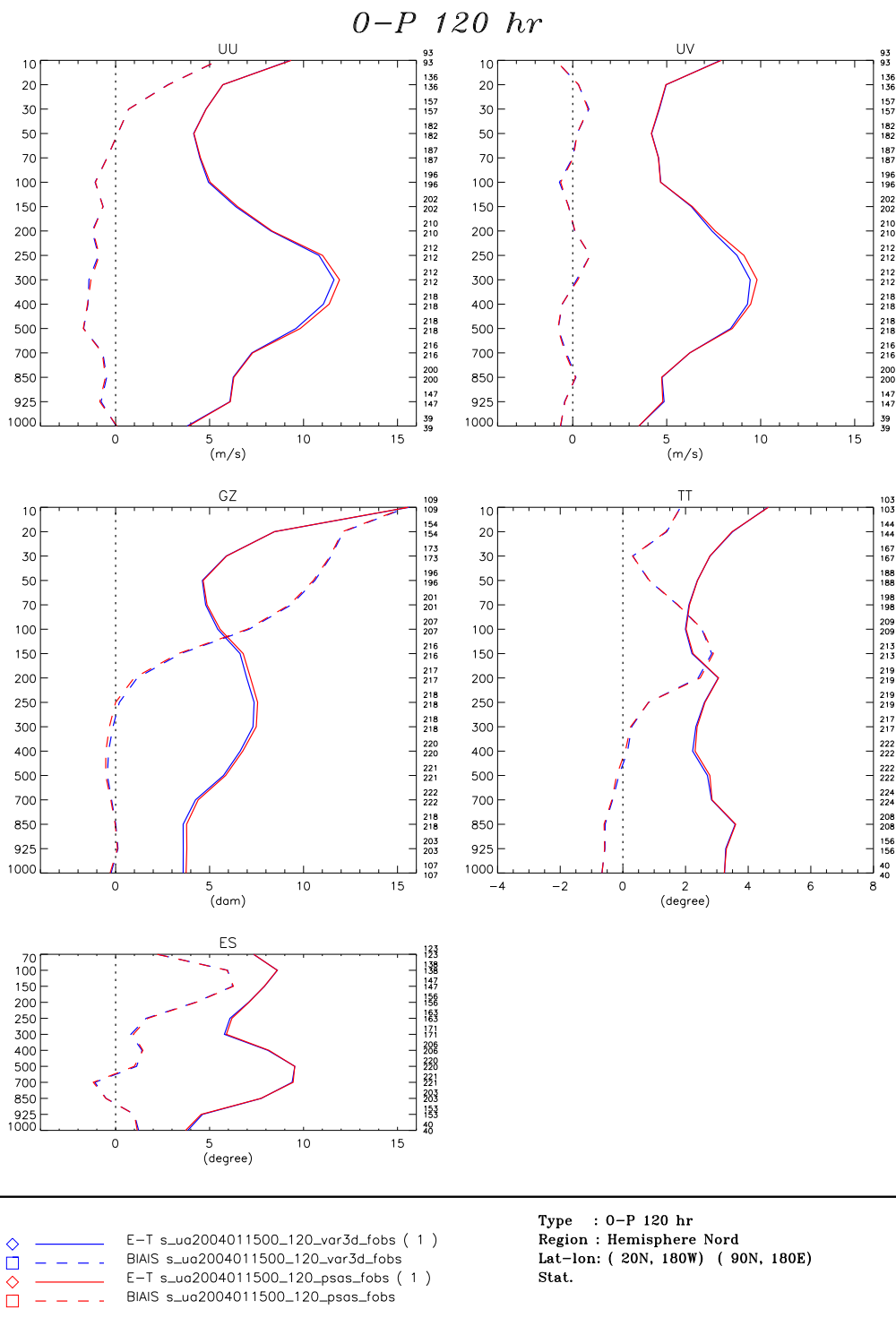


FIG. 3.5 – Prévision 120 h pour l'hémisphère Nord à partir d'une assimilation 3D-Var et 4D-PSAS

0-P 120 hr

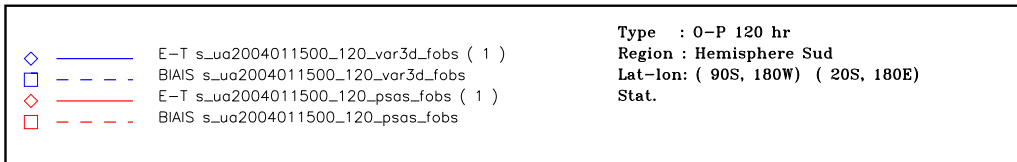
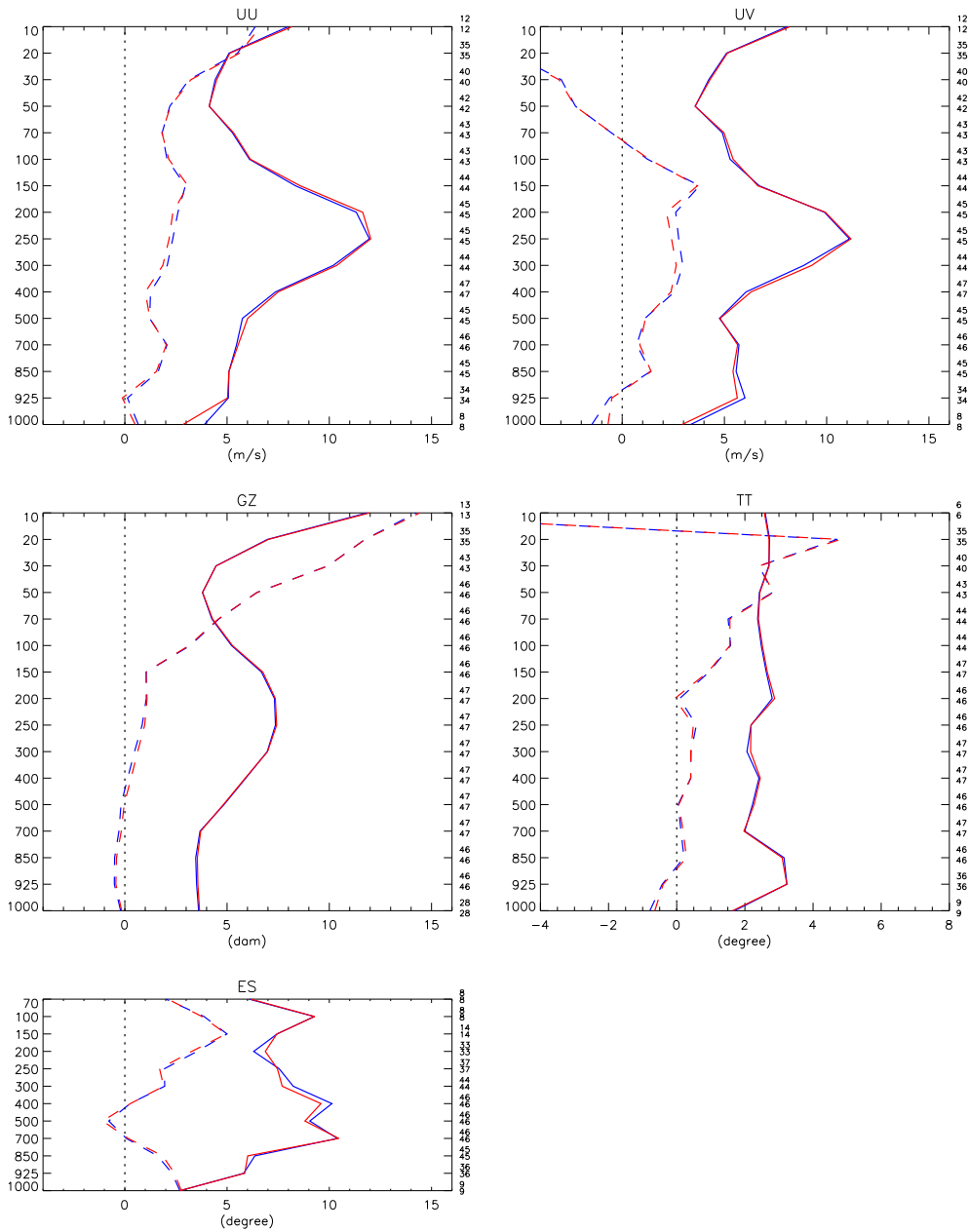


FIG. 3.6 – Prévision 120 h pour l'hémisphère Sud à partir d'une assimilation 3D-Var et 4D-PSAS

En effet, une autre situation (date différente) pourrait montrer des différences de même grandeur en faveur du PSAS. Pour tirer des conclusions fermes de ces vérifications, il aurait fallu lancer des cycles d'analyse sur quelques semaines.

D'autre part, on ne s'attend pas à une amélioration de la prévision vu que les deux méthodes 3D-Var et PSAS sont théoriquement équivalentes.

Ainsi, on arrive à traduire l'équivalence théorique entre le 3D-Var et le 3D-PSAS dans un cadre opérationnel avec un jeu complet de données observées. Les incréments d'analyse sont équivalents, il en est de même pour les prévisions faites à partir d'une analyse 3D-Var ou 3D-PSAS. On vérifie résolument les objectifs théoriques du 3D-PSAS à savoir : l'équivalence avec le 3D-Var ainsi que la réduction des dimensions et des degrés de liberté du problème de l'assimilation.

3.2.2 Dualité

En minimisant la fonction coût dans l'espace des observations (3D-PSAS), on perd la signification physique de la fonction objective qu'on cherche à optimiser. On ne peut désormais plus parler de réduction des écarts entre l'état du modèle et les observations.

On peut chercher un équivalent 3D-Var à chaque point de la minimisation PSAS. Autrement dit, à chaque itéré $w^{(k)}$, on calcule $\delta x^{(k)} = BH^T w^{(k)}$, et on évalue la fonction coût du 3D-Var pour chacun de ces points. En traçant la courbe des valeurs de la fonction coût du 3D-Var au cours de l'optimisation de la fonction objective du 3D-PSAS (figure 3.7), on remarque un comportement singulier se traduisant par plusieurs maxima.

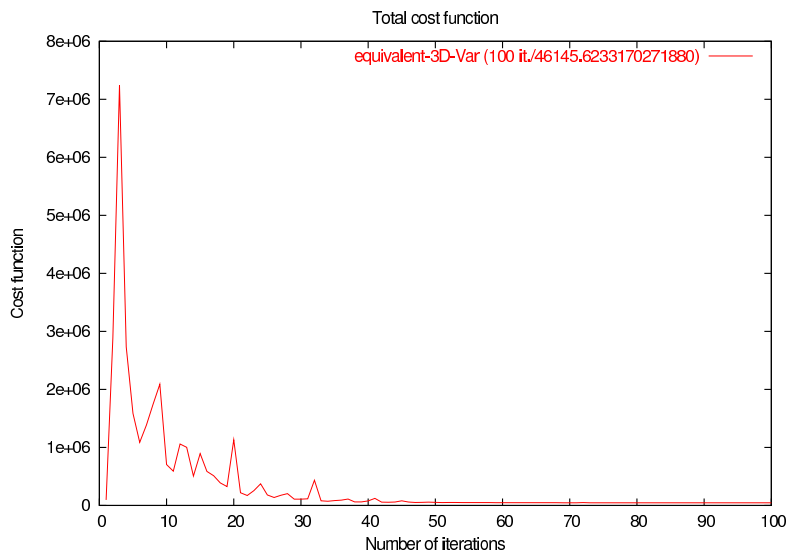


FIG. 3.7 – Courbe de l'équivalent 3D-var au cours de la minimisation PSAS .

A première vue, on peut y voir une ressemblance avec les courbes des simulations 3D-Var, où le minimiseur explore plusieurs chemins à la recherche de sa direction de descente. Ce qui se traduit par des maxima de la fonction coût, surtout lors de la première simulation. Toutefois, dans la figure 3.7, ce sont les itérés qui sont représentées, les simulations étant filtrées préalablement (le pic maximum des simulations atteint 10^8).

En raison de l'amplitude du premier pic, la figure 3.8 représente les fonctions 3D-Var, 3D-PSAS et l'équivalent 3D-Var avec une échelle en ordonnées dont le maximum est réduit à 3.10^6 . Les pics persistent tout au long des soixante premières itérations en diminuant d'amplitude.

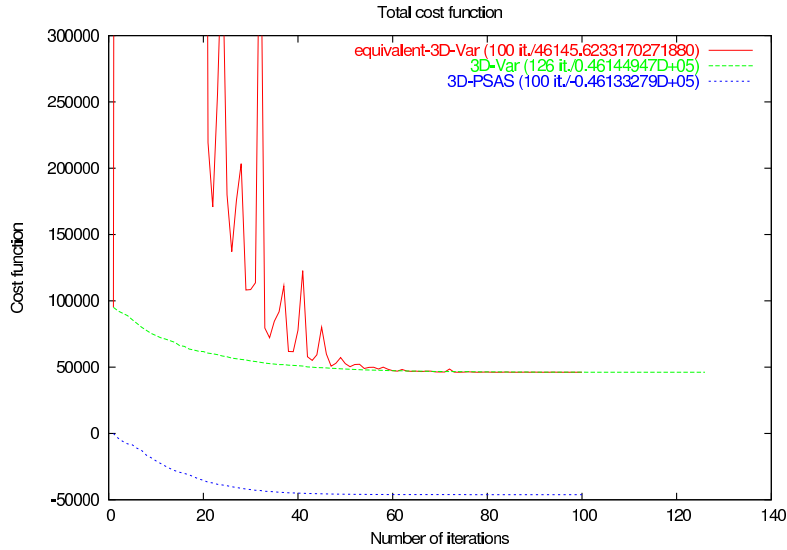


FIG. 3.8 – Courbe de l'équivalent 3D-var au cours de la minimisation PSAS, échelle tronquée à 3.10^6 .

Le passage de cette fonction par des maxima porte à croire qu'au cours de la minimisation PSAS, on ne minimise pas forcément la variance des erreurs d'analyse. Un comportement similaire a été observé par T. Lagarde (2000) [1] dans son étude du 4D-PSAS où il a observé un seul maximum. S.Louvel (1999) [8] a pour sa part observé un palier, et lors d'une précédente étude [16] portant sur le 3D-PSAS et 4D-PSAS dans un modèle shallow water [16], on a observé un maximum lors des premières itérations de la minimisation 3D-PSAS. Toutefois, le problème ici est de taille différente puisque ces précédentes études utilisaient un modèle simplifié. La complexité de la fonction objective pourrait être à l'origine de l'apparition de plusieurs pics au lieu d'un seul.

On peut toujours penser que dans le cas du PSAS, la recherche de la direction de descente est la cause de ce comportement singulier. C'est une signature du PSAS dont l'explication est encore sujet de recherche.

3.2.3 Pré-conditionnement de la hessienne du 3D-PSAS

Tel que discuté dans Gilbert et LeMaréchal (1989) [17], le minimiseur construit une approximation de la matrice hessienne à partir des écarts entre les différents itérés et l'évaluation du gradient en ces différents points. Cette approximation de la hessienne permet de mieux pré-conditionner la minimisation. Le démarrage à chaud consiste à reprendre la minimisation en considérant l'information sur la hessienne obtenue d'une minimisation précédente.

Des expériences conduites au SMC ont montré que ce procédé pouvait réduire de manière significative le nombre d'itérations. Après six analyses, le nombre d'itérations nécessaires est réduit de moitié par rapport à une analyse 3D-Var qui est démarrée à froid (sans pré-conditionnement). Ceci illustre clairement que la minimisation peut bénéficier d'un meilleur pré-conditionnement que celui défini par rapport au seul terme d'ébauche. Cette question a récemment été examinée par Andersson et al. (2000) [15] pour le 3D-Var.

Pour les besoins de la présente étude, on a effectué le pré-conditionnement de la hessienne du 3D-PSAS pour la même expérience en reprenant **la même analyse**, mais en ré-utilisant l'information sur la hessienne. La comparaison avec le même exercice effectué avec le 3D-Var est très intéressante (figure 3.9).

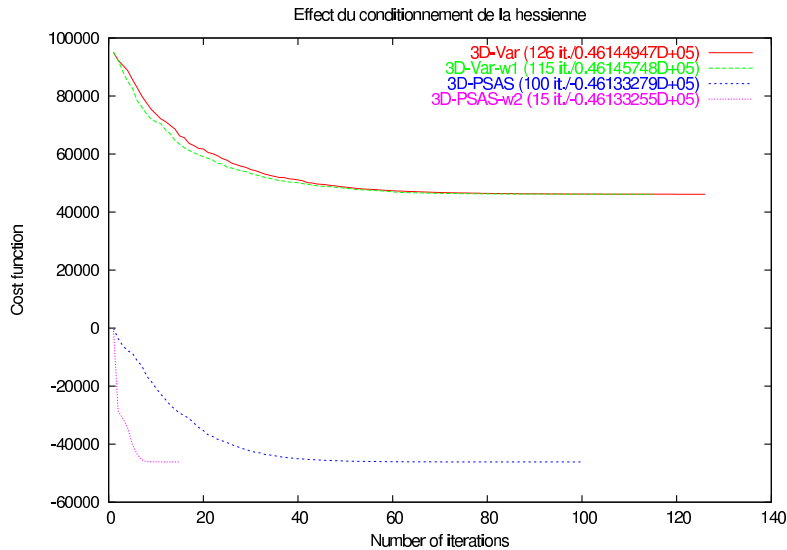


FIG. 3.9 – Effet du pré-conditionnement de la hessienne sur les fonctions objectives 3D-Var et 3D-PSAS.

Après pré-conditionnement de sa hessienne, le PSAS converge vers le même état analysé au bout de 15 itérations (contre 100 itérations sans pré-conditionnement). Ce résultat

est d'autant plus intéressant que pour le 3D-Var le pré-conditionnement de la hessienne prend encore 115 itérations pour converger (contre 126 initialement).

D'autre part, l'impact du pré-conditionnement de la hessienne est aussi visible sur la courbe de l'équivalent 3D-Var au cours de la minimisation PSAS. En effet, l'amplitude des pics diminue significativement, mais on garde toujours une signature marquée lors des premières itérations (figure 3.10).

Toutefois, il faut souligner que dans le cas du PSAS, la possibilité de réexploiter la hessienne n'est pas envisageable pour lancer des cycles d'analyse consécutifs car le vecteur de contrôle est dans l'espace des observations, ce qui fait que sa taille - aussi bien que les dimensions de la hessienne - sont vouées au changement d'une analyse à l'autre.

Par ailleurs, ce résultat ne perd pas tout son intérêt puisqu'il permet de relancer la recherche dans la voie d'un meilleur pré-conditionnement pour le PSAS.

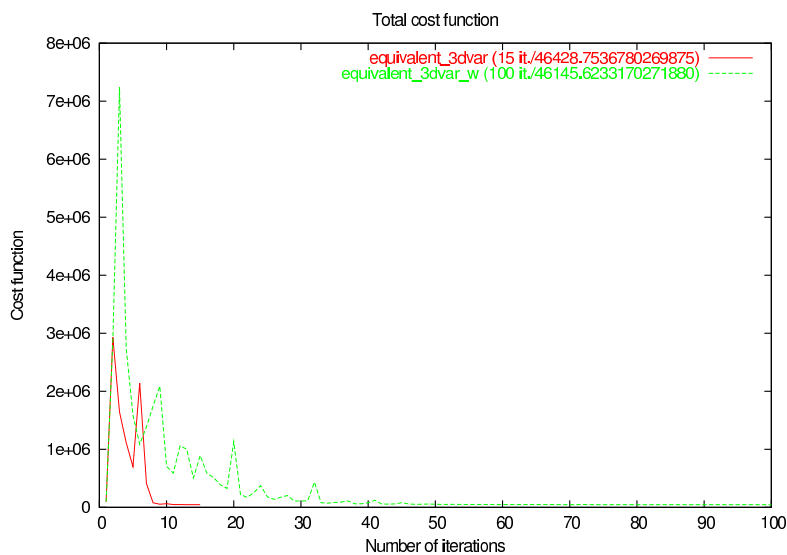


FIG. 3.10 – Effet du conditionnement de la hessienne sur l'équivalent 3D-Var au cours de la minimisation 3D-PSAS.

Dans cette optique, on peut par exemple chercher à optimiser le problème du PSAS en tirant profit de la hessienne du 3D-Var. Il s'agit donc de repérer un lien entre les deux hessiennes.

Courtier (1997) [2] a expliqué comment ces deux hessiennes se retrouvent avec le même spectre complété éventuellement par des 1. Le problème est donc équivalent.

Dans quelle mesure peut-on utiliser l'information contenue dans la hessienne du 3D-Var pour améliorer une minimisation PSAS, ou du moins une minimisation du 3D-Var dans l'espace des observations ?

Chapitre 4

Optimisation du 3D-Var dans un espace réduit

4.1 Formulation théorique

On propose ici une nouvelle approche pour résoudre le problème de l'assimilation variationnelle 3D-VAR dans l'espace des observations. Il s'agit en effet d'exploiter l'équivalence des schémas 3D-Var et PSAS à la convergence pour établir un changement de variable ramenant le vecteur de contrôle du modèle vers un espace réduit et faisant en sorte que la minimisation se fasse dans ce nouvel espace. On parlera désormais d'un **3D-Var-dual**.

Ceci nous permettra alors d'étudier le comportement de la fonction coût du 3D-Var dans l'espace des observations sans avoir à reformuler tout le problème de l'assimilation dans cet espace (c'est le cas du PSAS).

On a établi en section 1.4 la relation $\delta x^a = BH^T w$, δx^a étant l'incrément d'analyse du 3D-Var et w le vecteur de contrôle du PSAS à la convergence.

En posant le changement de variable $\delta x = BH^T w$, on peut démontrer que la fonction coût du 3D-Var exprimée en fonction de w (au lieu de δx) s'écrit :

$$J(w) = \frac{1}{2} w^T H B H^T w + \frac{1}{2} [H B H^T w - y']^T R^{-1} [H B H^T w - y']$$

et que son gradient peut s'écrire comme suit :

$$\begin{aligned} \nabla_w J &= H B H^T w + H B H^T R^{-1} [H B H^T w - y'] \\ &= H B^{\frac{1}{2}} [B^{\frac{1}{2}} H^T w + B^{\frac{1}{2}} H^T R^{-1} (H B H^T w - y')] \end{aligned}$$

Cette formulation est intéressante dans la mesure où l'on peut toujours parler d'écarts aux observations et à l'ébauche tout en se plaçant dans l'espace des observations. Chose qui n'était pas possible dans le cas du PSAS.

On choisit un pré-conditionnement semblable à celui du 3D-PSAS, soit $u = R^{\frac{1}{2}}w$. La nouvelle fonction coût et son gradient deviennent :

$$\begin{aligned}
J(u) &= \frac{1}{2}u^T R^{-\frac{1}{2}} H B H^T R^{-\frac{1}{2}} u + \frac{1}{2} [H B H^T R^{-\frac{1}{2}} u - y']^T R^{-1} [H B H^T R^{-\frac{1}{2}} u - y'] \\
\nabla_u J &= R^{-\frac{1}{2}} H B H^T R^{-\frac{1}{2}} u + R^{-\frac{1}{2}} H B H^T R^{-1} [H B H^T R^{-\frac{1}{2}} u - y'] \\
&= R^{-\frac{1}{2}} H B^{\frac{1}{2}} [B^{\frac{1}{2}} H^T R^{-\frac{1}{2}} u + B^{\frac{1}{2}} H^T R^{-1} (H B H^T R^{-\frac{1}{2}} u - y')] \\
&= R^{-\frac{1}{2}} H B H^T R^{-\frac{1}{2}} \underbrace{[(I + R^{-\frac{1}{2}} H B H^T R^{-\frac{1}{2}})u - R^{-\frac{1}{2}} y']}_{\text{gradient de la fonction du 3D-PSAS}}
\end{aligned}$$

4.2 Implémentation dans PALM

Le schéma PALM du 3D-Var-Dual est présenté dans la figure 4.2. Pour cette implémentation, les unités utilisées sont celles qui ont servi à l'implémentation du 3D-VAR et du 3D-PSAS. Pour construire cette nouvelle application, seul le schéma de couplage de ces unités change, pas leur code source. Cet exemple illustre l'intérêt de la gestion évolutive d'une application avec le logiciel PALM.

4.3 Résultats du 3D-Var-Dual

Convergence

Le 3D-Var-dual est formulé pour tirer parti et des avantages du 3D-Var (optimisation liée directement à la variance d'erreur d'analyse) et du 3D-PSAS (réduction des dimensions et des degrés de liberté du problème).

Aussi, sachant que l'on minimise la même fonction (3D-Var) mais dans deux espaces différents, s'attend-on à retrouver le même état analysé après la convergence. Toutefois, le 3D-Var-dual atteint le nombre maximal d'itérations (fixé à 400) sans aboutir au point de la convergence du 3D-var (figure 4.1, courbes bleue et rouge).

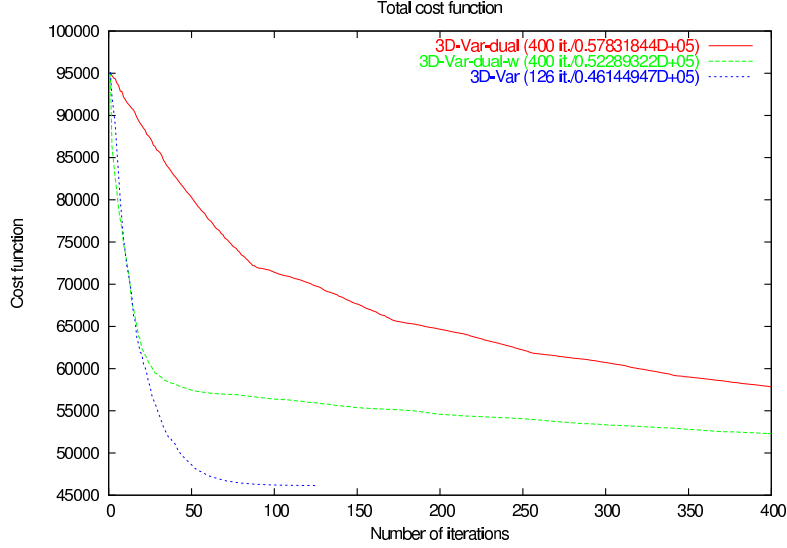


FIG. 4.1 – Courbe des fonctions coût 3D-Var-dual, 3D-Var-dual après pré-conditionnement de sa hessienne (3D-Var-dual-w) et 3D-Var

En réalisant la même expérience avec un nombre réduit d’observations (une observation par famille), on arrive au même état analysé du 3D-Var après deux fois plus d’itérations, soit 24 au lieu de 12 pour le 3D-Var. La dimension du problème serait-elle en cause ?

Par ailleurs, le pré-conditionnement de la hessienne du 3D-Var-dual tend à ramener la fonction coût vers le point de convergence du 3D-var (figure 4.1, courbe verte), ce qui laisse supposer qu’il s’agirait d’un problème de conditionnement de la hessienne.

Le pré-conditionnement de la hessienne

Pour pré-conditionner le 3D-var-dual, on se sert de la matrice $R^{\frac{1}{2}}$. Idéalement, on pré-conditionne par rapport à la hessienne de la fonction coût (voir section 1.3.2). Dans notre cas, cette matrice est de la forme :

$$J_u'' = HBH^T[I + R^{-1}HBH^T] = HBH^T R^{-1}[R + HBH^T]$$

Au lieu de $R^{\frac{1}{2}}$, on peut conditionner par rapport au deuxième terme de la hessienne, soit $B^{\frac{1}{2}}H^T$. Mais cela revient à reconsidérer le problème dans l’espace du modèle. On se ramène alors au 3D-Var classique.

Par ailleurs, la hessienne de 3D-var-dual peut aussi s’écrire sous la forme suivante :

$$J_u'' = R^{-\frac{1}{2}}HB^{\frac{1}{2}}\underbrace{[I + B^{\frac{1}{2}}H^T R^{-1}HB^{\frac{1}{2}}]}_{J_{3dvar}''}B^{\frac{1}{2}}H^T R^{-\frac{1}{2}}$$

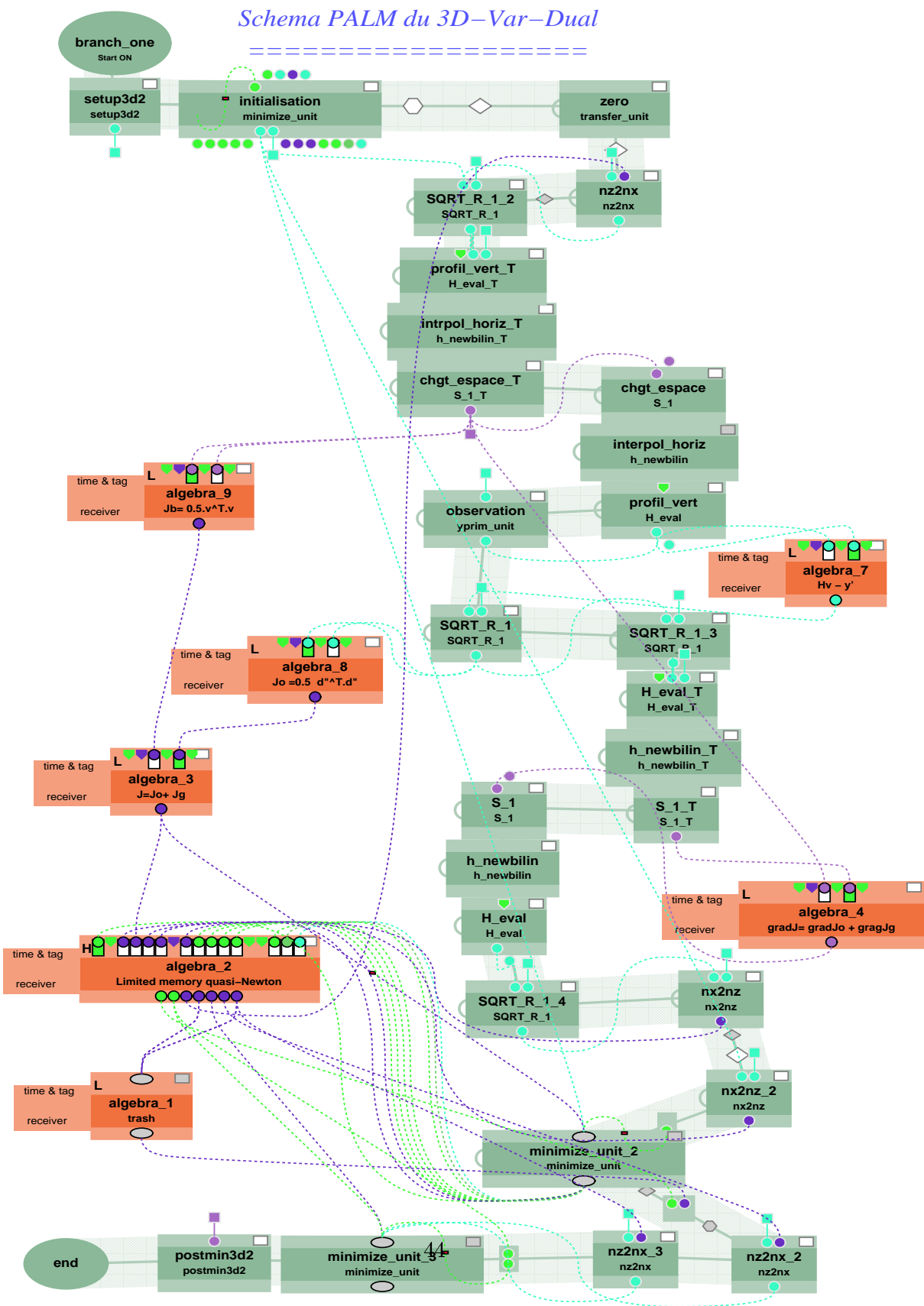


FIG. 4.2 – Schéma PALM du 3D-Var-dual

On peut alors exploiter la hessienne du 3D-Var après un changement de variable la ramenant à l'espace des observations pour conditionner le problème du 3D-Var-dual. Cette solution est à envisager dans une étude future.

Chapitre 5

Discussions et perspectives

Au terme de cette étude, un certain nombre de points n'a pas pu être développé. Dans ce chapitre, on présente quelques axes sur lesquels une étude ultérieure pourrait se baser afin de valider davantage les résultats obtenus ou d'en expliquer d'autres. On y discutera aussi les perspectives futures du projet PALM ainsi que de l'étude des méthodes duales.

5.1 Les schémas PALMés

Modularité

Cette première étude de faisabilité a permis de mettre l'accent sur certaines contraintes liées à la modularité de la structure informatique du code 3D-var du SMC. Le découpage en unités élémentaires en vue de l'implémentation dans PALM n'a concerné que les unités "prioritaires" en limitant le nombre de variables à échanger. En effet, seules les variables dites **dynamiques** sont recensées et établies en tant qu'objets communiqués explicitement entre les unités, tandis que les variables **statiques** sont encore échangées dans des blocs communs (common). Cette démarche est d'autant plus importante qu'elle garantit une modularité pouvant être exploitée dans l'évolution du 3D-Var indépendamment de son implémentation dans PALM.

Aussi, à travers les modifications apportées au code 3D-Var, a-t-on voulu conserver la structure de l'algorithme initial, les sous-routines devant être fonctionnelles dans les deux versions. Ceci entraîne l'utilisation d'un double jeu de variables dynamiques nécessaires pour le 3D-Var initial (qui restent déclarées dans des blocs communs) et pour la version PALMée (communiquées explicitement). L'indépendance totale des unités est donc mise en cause par cette solution qui induit aussi un coût supplémentaire pour la version PALMée.

Ainsi, dans l'optique de continuer dans la voie de la gestion du code 3D-Var par PALM, ces facteurs peuvent être pris en compte pour garantir à la fois la modularité du code 3D-Var permettant une migration facile vers d'autres méthodes d'assimilation de données, et une implémentation mieux adaptée au contexte de PALM.

PALM-MP

Dans le cadre de ce travail, on a utilisé PALM-Research. Une nouvelle version de PALM a été récemment développée au sein de l'équipe " Global Change and Climate Modeling " au CERFACS : PALM-MP.

PALM-MP est une version MPMD (Multiple Program Multiple Data) du coupleur PALM qui permet de réaliser des applications parallèles dont les composantes logicielles sont des programmes exécutables indépendants. Il se distingue en cela de la version PALM-Research qui ne fait que simuler un MPMD en regroupant les différentes composantes en un seul exécutable. Autrement dit, au lieu de considérer chaque unité PALM comme un exécutable par exemple, on peut construire un **bloc** d'unités dans le même programme exécutable afin de partager la mémoire ou optimiser l'application (figure 5.1).

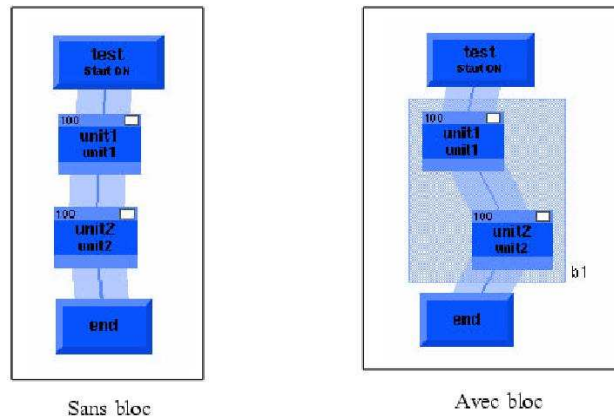


FIG. 5.1 – Les blocs dans PALM-MP.

Si plusieurs unités sont des sous-programmes devant partager des variables globales (common ou module), il est utile de déclarer un bloc englobant ces unités.

D'autre part, contrairement à PALM-Research, les branches PALM-MP ne monopolisent aucun processeur de l'application : elles sont concurrentes. Il devient donc possible de faire tourner plusieurs branches sur le même processeur.

Il serait alors intéressant de porter les implémentations réalisées dans le cadre de ce travail (3D-Var, 3D-PSAS et 3D-Var-dual) dans PALM-MP, sachant que le portage peut se faire sans modifications importantes des versions PALM-Research.

Notons par ailleurs que toutes les fonctionnalités présentes dans PALM-Research sont aujourd'hui implémentées dans PALM-MP, sauf la boîte à outil d'algèbre, dont l'implémentation sera terminée d'ici fin mars 2005. D'autre part, de nouvelles fonctionnalités ont été développées dans PALM-MP, tel que le monitoring temps réel des applications dans PrePALM, ou la gestion de sous-objets (gestion d'objets définis comme des ensembles de sous parties d'objets existants).

5.2 Le 3D-PSAS

3D-PSAS Vs 3D-Var

A priori, le 3D-Var et le 3D-PSAS ne semblent pas se distinguer l'un par rapport à l'autre quand aux résultats de la minimisation, leurs incréments sont équivalents, et les prévisions qui en découlent le sont aussi.

L'un des principaux avantages du 3D-Var est que sa fonction objective est directement liée à la loi de probabilité du contrôle (ie. à la variance d'analyse), ce qui fait que même avec une optimisation incomplète on peut obtenir des résultats satisfaisants et surtout interprétables physiquement. Par contre, la dimension de l'espace de contrôle peut croître à l'introduction des erreurs du modèle et des erreurs de forçage.

Par ailleurs, les performances de l'algorithme 3D-PSAS peuvent être tributaires de la taille de l'espace des observations. Mais, du fait que l'on cherche à optimiser une fonction objective qui n'a pas de sens physique, on s'assure uniquement de la vraisemblance du résultat final de l'optimisation. Ce qui fait que si cette optimisation n'est que partielle, on risque d'obtenir de mauvais résultats.

Le contrôle de qualité

En assimilation de données, on suppose en général que les données observées ont des erreurs non biaisées et ne sont pas affectées par des problèmes majeurs nuisant au fonctionnement normal de l'appareil de mesure. Pour pouvoir satisfaire cette hypothèse, un contrôle de qualité est effectué, d'abord, avant l'analyse pour éliminer les données aberrantes contenant des erreurs grossières (on parle de **background check**) puis au cours de l'analyse afin de détecter les observations ayant une forte probabilité d'être erronées (**QC-VAR**).

Pour le QC-VAR, on se base sur les techniques de probabilité conditionnelle : quelle est la probabilité pour qu'une donnée soit fautive sachant que son équivalent modèle (issu du champ d'essai) est vrai, et vice-versa. Au-delà d'un seuil défini, la donnée sera rejetée.

Dans la configuration actuelle au SMC, le QC-VAR est activé au cours de la minimisation (après quelques itérations¹ avec un jeu complet d'observations) afin d'attribuer un poids à chaque donnée observée selon qu'elle concorde ou non avec l'état de l'analyse préliminaire correspondant à l'itération courante de l'analyse variationnelle.

Or, dans le cas du PSAS, on n'est pas sûr qu'au cours de l'optimisation, on minimise bien la variance d'erreur d'analyse (Les pics observés dans la courbe de l'équivalent 3D-Var de la fonction PSAS en sont le parfait exemple). Pour cette raison, on ne tient pas compte, dans les implémentations réalisées dans le cadre de ce travail, du contrôle de

¹si la minimisation se fait en 100 itérations, on active le QC-VAR après 30 itérations par exemple.

qualité. Notons que pour le 3D-Var, la réactivation de ce contrôle est facile, tandis que le 3D-PSAS nécessite la définition d'un nouveau procédé pour détecter les données à rejeter.

5.3 Le 3D-Var-dual

Bien que l'étude du 3D-Var-dual n'ait pas été concluante quant aux attentes exprimées d'après les formulations théoriques, le concept de chercher un compromis entre le 3D-Var et le 3D-PSAS reste intéressant en soit.

Le travail à ce sujet n'est certainement qu'à ses débuts. En effet, voir l'effet du pré-conditionnement du 3D-Var-dual par une hessienne formulée à partir de celle du 3D-Var est une piste d'étude à envisager ultérieurement. D'autre part, les valeurs des paramètres de la minimisation (df1 : la diminution attendue de la fonction coût lors de la première itération, dxmin : la distance minimale entre deux points de la minimisation...) sont ceux du 3D-Var. On ne s'étonnera pas de voir un impact sur les résultats en adaptant ces paramètres au cas du 3D-Var-dual.

Conclusion

Au terme de ce travail, on a pu établir la faisabilité technique de l'implémentation du schéma opérationnel 3D-Var incrémental du Service Météorologique du Canada avec PALM. A travers la stratégie choisie de découpage et de modularisation de la programmation initiale, cette implémentation est réalisée sans changements conséquents dans la version de base, bien que des modifications aient été apportées aux procédures d'application des opérateurs d'observation et de traitement des données TOVS.

Le coût de la version PALMée reste minimale en temps mur, ce qui n'est pas surprenant vu le nombre limité de communications établies qui concernent essentiellement les variables dynamiques. Par contre, au niveau des ressources informatiques exploitées, PALM nécessite deux processeurs élémentaires additionnels pour le driver et le minimiseur. D'autres améliorations sont à apporter au schéma PALMée en vue d'exploiter toutes les possibilités offertes par PALM, ainsi que pour garantir une modularité permettant de migrer plus facilement d'une méthode d'assimilation à une autre.

D'autre part, l'étude du 3D-PSAS permet de vérifier, dans un cadre pratique, l'équivalence démontrée théoriquement entre le 3D-Var et le 3D-PSAS. Les incréments d'analyse étant presque identiques, les prévisions lancées à partir des analyses des deux méthodes sont, comme on s'y attend, neutres et ne permettent pas de conclure en faveur de l'une ou de l'autre des méthodes. Aussi, le conditionnement de la hessienne du 3D-PSAS permet-il de retrouver le même état analysé en 15 itérations contre 100 itérations sans pré-conditionnement. Ce résultat est d'autant plus intéressant que le 3D-Var nécessite encore 115 itérations pour converger vers cet état (contre 124 initialement).

Le fait de travailler dans l'espace des observations oblige les dimensions des hessiennes et des vecteurs de contrôle à changer d'une analyse à l'autre, ce qui limite l'intérêt du raffinement itératif de la hessienne du PSAS. Toutefois, ce résultat reste encourageant car, avec un bon pré-conditionnement de la hessienne du 3D-PSAS, on peut réduire considérablement le nombre d'itérations requis, et donc le temps de calcul. Ceci, peut être intéressant dans le cas du 4D-PSAS.

La fonction objective du 3D-PSAS n'ayant pas de sens physique, on ne s'assure pas, lors de son optimisation, de minimiser en même temps la variance d'erreur d'analyse. L'équivalence entre le PSAS et le 3D-Var n'étant valable qu'à la convergence, une minimisation non complète du 3D-PSAS peut donner des résultats aberrants.

D'autre part, en utilisant cette équivalence "conditionnelle" comme changement de

variable (3D-Var-dual), ramenant le problème du 3D-Var à l'espace des observations sans en reformuler la fonction coût, permet "théoriquement" de réaliser un compromis entre une optimisation liée directement à la variance d'erreur d'analyse (3D-Var) et une diminution des dimensions du problème (3D-PSAS).

L'étude de cette méthode n'est certainement qu'à ses débuts. Sa convergence lente trahit un problème dans le conditionnement de sa hessienne, d'autant plus que tous les paramètres de la minimisation sont adaptés au cas du 3D-Var.

Ce travail ouvre donc des perspectives intéressantes au niveau de l'étude de l'apport de la nouvelle version de PALM : PALM-MP, ainsi que de l'étude du pré-conditionnement du 3D-PSAS et du 3D-Var-dual.

Annexes

MPI/Open-MP

Open-MP

Un programme Open-MP est exécuté par un processus unique qui active des processeurs légers (threads) à l'entrée d'une région parallèle. Chaque thread exécute une tâche composée d'un ensemble d'instructions.

Un programme Open-MP est donc une alternance de régions séquentielles et de régions parallèles qui peuvent être exécutées par plusieurs tâches à la fois. Les tâches peuvent se partager le travail contenu dans une région parallèle. Ceci consiste essentiellement à :

- exécuter une boucle par répartition des itérations entre les tâches ;
- exécuter plusieurs sections de code mais une seule par tâche ;
- exécuter plusieurs occurrences d'une même procédure par différentes tâches.

Open-MP Vs MPI

Ce sont deux modèles complémentaires de parallélisation. MPI est un modèle multiprocessus dont le mode de communication entre les processus est explicite (la gestion des communications est à la charge de l'utilisateur), tandis que Open-MP est un modèle multi-tâches dont le mode de communication entre les tâches est implicite (la gestion des communications est à la charge du compilateur).

MPI est utilisé en général sur des machines multiprocesseurs à mémoire distribuée, alors que Open-MP est mieux adapté aux machines multiprocesseurs à mémoire partagée.

Sur une grappe de machines indépendantes (nœuds) multiprocesseurs à mémoire partagée, la mise en œuvre d'une parallélisation à deux niveaux (MPI et Open-MP) dans un même programme peut être un atout majeur pour les performances parallèles du code.

Table des figures

2.1	Exemples de schémas PALM : une unité indépendante (à gauche), une branche munie de deux unités dont une est parallèle (au milieu) et un schéma avec deux branches pouvant tourner en parallèle avec une unité d'algèbre (à droite).	16
2.2	Cycle d'analyse au Service Météorologique du Canada (SMC).	17
2.3	Schéma PALM du 3D-Var incrémental	24
2.4	Schéma PALM du 3D-PSAS	27
3.1	Résultats de l'analyseur de performances de PALM pour le schéma 3D-Var.	30
3.2	Graphe des fonctions objectives du 3D-Var et du 3D-PSAS.	32
3.3	Les incréments d'analyse du 3D-Var (en haut) et du 3D-PSAS (en bas) pour le champ de température au niveau 500 hPa.	33
3.4	Spectres d'énergie des incréments d'analyse 3D-Var (en rouge) et 3D-PSAS (en vert). En abscisse, les nombres d'onde et en ordonnée l'énergie totale.	34
3.5	Prévision 120 h pour l'hémisphère Nord à partir d'une assimilation 3D-Var et 4D-PSAS	35
3.6	Prévision 120 h pour l'hémisphère Sud à partir d'une assimilation 3D-Var et 4D-PSAS	36
3.7	Courbe de l'équivalent 3D-var au cours de la minimisation PSAS	37
3.8	Courbe de l'équivalent 3D-var au cours de la minimisation PSAS, échelle tronquée à $3 \cdot 10^6$	38
3.9	Effet du pré-conditionnement de la hessienne sur les fonctions objectives 3D-Var et 3D-PSAS.	39
3.10	Effet du conditionnement de la hessienne sur l'équivalent 3D-Var au cours de la minimisation 3D-PSAS.	40
4.1	Courbe des fonctions coût 3D-Var-dual, 3D-Var-dual après pré-conditionnement de sa hessienne (3D-Var-dual-w) et 3D-Var	43
4.2	Schéma PALM du 3D-Var-dual	44
5.1	Les blocs dans PALM-MP.	47

Bibliographie

- [1] T. Lagarde. *Nouvelle approche des méthodes d'assimilation de données : les algorithmes de point selle*. thèse PhD, juin 2000, CERFACS.
- [2] P. Courtier. *Dual formulation of four-dimensional variational assimilation*. Quart. J. Roy. Meteorol. Soc. 1997, 123, pp.2449-2461.
- [3] P. Gauthier, C. Chouinard, B. Brasnett. *Quality control : methodology and applications*. NATO Science Series, IV. Earth and Environmental Sciences, Vol. 26, 2003.
- [4] R. Daley & E. Barker. *NAVDAS Source Book 2000 NRL Atmospheric Variational Data Assimilation System*. Naval Research Laboratory, Atmospheric Dynamics and Prediction Branch, Marine Meteorology Division, Monterey, CA 93943-5502, 2000.
- [5] CERFACS/Global Change/Equipe PALM. *Etude de faisabilité du projet PALM*. Projet MERCATOR : revue de fin de phase A, mai 1998.
- [6] F. Bouttier & P. Courtier. *Data assimilation concepts and methods* ECMWF Meteorological Training Course Lecture Series, March 1999.
- [7] A. Da Silva, J. Pfaendtner, J. Guo, M. Sienkiewicz et S. Cohn. *Assessing the effect of data selection with DAO's Physical-space Statistical Analysis System*. WMO/TD 651, 273-278.
- [8] S. Louvel. *Implementation of dual variational algorithm for assimilation of synthetic altimeter data in the oceanic primitive equation model micom*.
- [9] P. Gauthier and M. Buehner. *Background-error statistics modelling in a 3D variational data assimilation scheme : estimation and impact on the analyses*. Processings of the 1998 Statistical Meetings, August 9-13, 1998, Texas, U.S.A.
- [10] C. Ficher with contributions by L. Berre *The variational computations inside ARPEGE/ALADIN : cycle CY25T1*. Météo-France, CNRM/GMAP, November 22, 2002.
- [11] L. Fillion, M. Buehner, C. Charette, C. Chouinard, P. Gauthier, P. Koclas, J. Morneau, S. Pellerin, R. Sarrazin, J. St-James, G. Verner. *The canadian 3D-Var Analysis Scheme on model's vertical coordinate*. Atmospheric Environment Service, 11 Mai 1999.
- [12] I.M Navon et D.M. Legler, 1987. *Conjugate-gradient methods for large-scale minimization in meteorology*. M.W.R., 115, 1479-1502.
- [13] S. Laroche and P. Gauthier. *A validation of the incremental formulation of 4D-Var variational data assimilation in a nonlinear barotropic flow*. Tellus (1998), 50A, 557-572.

- [14] P. Gauthier, C. Charette, L. Fillion, P. Koclas and S. Laroche. *Implementation of a 3D variational data assimilation system at the Canadian Meteorological Center. PART 1 : the global analysis*. Atmosphere ocean, Vol.XXXVII, No. 2, pp. 103-156. June, 1999.
- [15] Andersson, E., M. Fisher, R. Munro and A. McNally, 2000. *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.R. Meteor. Soc., 126, 1455-1472.
- [16] A. El Akkraoui, N. Chapelon, O, Pannekouke. *Implémentation de nouvelles méthodes d'assimilation variationnelle dans un modèle océanique simplifié* rapport de stage de modélisation, CERFACS, 2003.
- [17] Gilbert, J.C. and C. LeMaréchal, 1989. *Some numerical experiments with variable-storage quasi-Newton algorithms*. Mathematical Programming, 45, p. 407-435.