HAL open science

Assimilation de données pour l'initialisation et l'estimation de paramètres d'un modèle d'évolution de calotte polaire

Bertrand Bonan

▶ **To cite this version:**

Bertrand Bonan. Assimilation de données pour l'initialisation et l'estimation de paramètres d'un modèle d'évolution de calotte polaire. Géophysique [physics.geo-ph]. Université de Grenoble, 2013. Français. NNT : 2013GRENM091 . tel-00930097v2

HAL Id: tel-00930097

https://theses.hal.science/tel-00930097v2

Submitted on 1 Jun 2017

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.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques Appliquées**

Arrêté ministériel : 7 août 2006

Présentée par

Bertrand Bonan

Thèse dirigée par **Maëlle Nodet**
et codirigée par **Catherine Ritz**

préparée au sein **Laboratoire Jean Kuntzmann**
et de **École Doctorale Mathématiques, Sciences et Technologies de l'Information, Informatique**

Assimilation de données pour l'initialisation et l'estimation de paramètres d'un modèle d'évolution de calotte polaire

Thèse soutenue publiquement le **15 Novembre 2013**,
devant le jury composé de :

M. Stéphane Labbé
Professeur, Université Joseph Fourier – Grenoble 1, Président
M. Étienne Mémin
Directeur de recherche, Inria Rennes - Bretagne Atlantique, Rapporteur
Mme Frédérique Rémy
Directeur de recherche CNRS, LEGOS, Rapporteur
M. Marc Bocquet
Professeur, École des Ponts ParisTech, Examinateur
M. Fabien Gillet-Chaulet
Chargé de recherche CNRS, LGGE, Examinateur
Mme Sophie Ricci
Chargée de recherche, CERFACS, Examinateur
Mme Maëlle Nodet
Maître de Conférence, Université Joseph Fourier – Grenoble 1, Directeur de thèse
Mme Catherine Ritz
Directeur de recherche CNRS, LGGE, Co-Directeur de thèse

Remerciements

Je voudrais en premier lieu remercier mes deux directrices de thèse, Maëlle Nodet et Catherine Ritz. Ce manuscrit n'aurait jamais pu voir le jour sans vous. Merci Maëlle pour m'avoir guidé à travers les méandres de l'assimilation de données. Ta suggestion d'utiliser un filtre de Kalman d'ensemble à un moment donné a été plus que déterminante. Merci aussi pour l'attention que tu as portée à l'encadrement de ma thèse, pour ta génerosité dans tes conseils et pour ta clémence quant à mes bourdes verbales répétées (merci bébé Winnie). Merci Catherine de m'avoir fait découvrir le monde merveilleux des calottes polaires (sans doute l'un des rares endroit où je ne pourrais pas me balader en chemisette á carreaux). Merci aussi pour l'ouverture d'esprit que tu as su montré face à cette bête curieuse qu'est l'assimilation. Merci à toutes les deux pour tous les bons moments (et même les moins bons) partagés autour d'une bonne bière le vendredi soir.

Je voudrais remercier Frédérique Remy et Etienne Mémin pour le temps qu'ils ont consacré à lire et rapporter ma thèse en un si court laps de temps. Leurs critiques ont permis d'améliorer grandement la qualité du manuscrit. Je voudrais aussi remercier Sophie Ricci et Fabien Gillet-Chaulet pour leur commentaires constructifs et conseils avisés. Je remercie également Marc Bocquet pour l'intéret qu'il a porté à mes travaux durant ces trois ans ainsi que son expertise précieuse concernant les filtres de Kalman d'ensemble. Merci à Stéphane Labbé d'avoir eu la gentillesse de présider ce jury et pour sa bienveillance tout au long de ma thèse.

Je voudrais remercier l'ensemble de l'équipe MOISE pour son accueil. En particulier j'aimerais remercier Eric Blayo et Arthur Vidard. Mais aussi j'aimerais remercier vivement la véritable âme de l'équipe MOISE, son centre névralgique, Anne Pierson pour sa gentillesse et son efficacité redoutable face aux tâches administratives les plus ardues.

Je voudrais remercier l'équipe EDGe, Olivier Gagliardini, Gaël Durand, Fabien Gillet-Chaulet, Vincent Peyaud et tous les autres, pour leur accueil au LGGE et pour leurs commentaires toujours positifs quant à l'utilité de mes travaux pour les glaciologues.

Je voudrais remercier Nancy Nichols, Mike Baines et l'Université de Reading pour avoir facilité ce mois et demi compliqué à Reading juste avant ma soutenance.

Je voudrais remercier tout le personnel du LJK pour ces trois ans passés chez eux. En particulier je voudrais remercier les thésards, post-docs et ingé pour avoir rendu la tour IRMA et son septième étage si accueillant. Merci à Vincent, Gaëlle, Manel, Fédérico et Romain (mes pauvres co-bureaux du bureau 36), Madison et Lukas (pour mon court

séjour au bureau 34 après l'invasion du bureau 36 par un essaim d'abeilles), Chloé, Thomas, Pierre-Olivier, Pierre-Luc, Matthias, Roland, Bénédicte, Jean-Matthieu, Alexandre, Adrien, Emilie, Morgane et tant d'autres.

Je voudrais remercier aussi tous mes amis qui ont eu la lourde tâche de supporter mes blagues pas drôles et autres fantaisies pour leur soutien. Mention spéciale à Bobo (l'homme de ma vie) et Cécile, Jéjé, Nono, Lolo, Antoine (et les Chimay bleues aux Houches), Rom Pi, Arnaud, Adrien, Jeff, Lab et Grinch.

Pour finir, je voudrais remercier mes parents et ma sœur Bérénice pour leur soutien infaillible tout au long de ces trois ans. Les mots me manquent pour vous dire à quel point vous m'êtes précieux.

Table des matières

Table des figures **9**

Introduction **13**

1 Modélisation et observation de l'évolution des calottes polaires **15**
1.1 Mécanismes physiques en jeu 16
1.1.1 Écoulement de la glace 18
1.1.2 Évolution de la géométrie 26
1.1.3 Bilan de masse 28
1.2 Écoulement de Stokes et ses approximations 29
1.2.1 Rapport d'aspect et ses conséquences 29
1.2.2 Approximation hydrostatique 30
1.2.3 Modèle de Blatter-Pattyn 31
1.2.4 Shallow Shelf – Shelfy Stream Approximation 33
1.2.5 Shallow Ice Approximation 35
1.2.6 Autres modèles 36
1.3 Un modèle d'évolution simplifié : Winnie 37
1.3.1 Équation du bilan de masse et vitesses de bilan 38
1.3.2 Modélisation de la vitesse à la base 38
1.3.3 Calcul du bilan de masse et de la température en surface 38
1.3.4 Discrétisation et résolution numérique 40
1.4 Les observations disponibles 41
1.4.1 Observation de l'état actuel des calottes polaires 42
1.4.2 États passés 46
1.5 Assimilation de données en glaciologie 48
1.5.1 Initialisation de modèle 49
1.5.2 Paléoclimatologie 50

2 Assimilation de données **53**
2.1 Ingrédients de l'assimilation de données 54
2.1.1 Remarques préliminaires 54
2.1.2 Vecteur d'état et vecteur d'état augmenté 56
2.1.3 Modèle dynamique et prévision 57
2.1.4 Observations 58
2.1.5 Les différentes estimations du vecteur d'état 59
2.2 Approche bayésienne 60
2.2.1 Règle de Bayes 60

2.2.2 Assimilation d'une observation . . . 61
2.2.3 Estimation bayésienne . . . 62
2.2.4 Assimilation séquentielle : filtrage bayésien . . . 63
2.2.5 Filtres particulaires . . . 65
2.3 Best Linear Unbiased Estimator (BLUE) . . . 67
2.3.1 Généralités . . . 68
2.3.2 BLUE et assimilation . . . 69
2.4 Méthodes variationnelles . . . 71
2.4.1 3D-Var . . . 72
2.4.2 4D-Var . . . 77
2.5 Filtre de Kalman . . . 80
2.5.1 Filtre de Kalman . . . 81
2.5.2 Filtre de Kalman étendu et limite de cette approche . . . 83
2.5.3 Filtres de Kalman de rang réduit . . . 83
2.6 Filtre de Kalman d'ensemble . . . 86
2.6.1 Echantillonner pour approcher . . . 86
2.6.2 Phase de prévision . . . 87
2.6.3 Phase d'analyse . . . 89
2.6.4 Problèmes liés à l'utilisation de l'EnKF et solutions envisagées . . . 97
2.6.5 Inflation . . . 99
2.6.6 Filtre de Kalman d'ensemble de taille finie . . . 100
2.6.7 Localisation . . . 102
2.7 Approches hybrides . . . 105

3 Un problème simplifié de paléoclimatologie 107
3.1 Description du problème . . . 108
3.2 4D-Var et modèle adjoint . . . 110
3.2.1 Quelques détails sur la construction d'un modèle adjoint discret . . . 111
3.2.2 Construction de l'adjoint discret de Winnie . . . 113
3.2.3 Validation du code adjoint . . . 119
3.3 Expériences jumelles réalisées . . . 121
3.3.1 Supposer un forçage climatique sinusoïdal . . . 122
3.3.2 Retrouver le forçage climatique avec une bonne ébauche . . . 125
3.3.3 Que faire si nous n'avons pas une bonne ébauche . . . 126
3.4 Conclusion . . . 128

4 Initialisation de modèle de calotte polaire par filtre de Kalman d'ensemble 131
4.1 Résumé des travaux . . . 131
4.2 Une approche ETKF pour l'estimation jointe de l'état initial et de paramètres pour un modèle de calotte polaire . . . 133
4.2.1 Introduction . . . 133
4.2.2 Methods . . . 135
4.2.3 Ice sheet model . . . 138
4.2.4 Numerical results . . . 142
4.2.5 Conclusion and further directions . . . 148
4.3 Essais avec un ETKF-N version duale . . . 155

4.4 Conclusion 156

Conclusions et perspectives **159**

A Quelques rappels sur les probabilités **163**
A.1 Variable aléatoire et densité de probabilité 163
A.2 Espérance mathématique 163
A.3 Matrice de covariance 164
A.4 Probabilités conditionnelles 164
A.4.1 Variables à densité jointe 164
A.4.2 Probabilité conditionnelle 164
A.4.3 Espérance conditionnelle 165

B Bibliographie **167**

Table des figures

1.1 Principaux mécanismes intervenant dans l'évolution d'une calotte polaire (figure de C. Ritz). 19
1.2 Schéma d'une coupe d'une calotte polaire selon les axes x et z. 20
1.3 Évolution de l'accumulation, de l'ablation et du bilan de masse en fonction de la température en surface dans Winnie. 39
1.4 Discrétisation de l'axe x employée dans Winnie. 40
1.5 Cartographie de la surface dérivée du DEM de [Bamber *et al.*, 2009]. . . . 42
1.6 Carte d'estimation des erreurs RMS du DEM de [Bamber *et al.*, 2009]. . . 43
1.7 Carte de vitesses de surface issue de [Rignot *et al.*, 2011]. 44
1.8 Carte de dH/dt tirée de [Flament et Rémy, 2012] 45
1.9 Couverture des données de localisaton de l'interface socle-calotte utilisées dans le projet Bedmap2 [Fretwell *et al.*, 2013]. 46
1.10 Carte de la topographie du socle rocheux antarctique issue du projet Bedmap2 [Fretwell *et al.*, 2013]. 47
1.11 Carte de l'incertitude estimée sur la topographie du socle rocheux antarctique issue du projet Bedmap2 [Fretwell *et al.*, 2013]. 48
1.12 Évolution de l'étendue de la calotte Fennoscandienne reconstruite à partir de données géomorphologiques issue de [Kleman *et al.*, 1997]. 51

2.1 Illustration explicitant la dégénérescence du filtre particulaire bootstrap issue de [van Leeuwen, 2009]. 66
2.2 Figure illustrant l'intérêt d'ajouter un terme d'ébauche à la fonction écart aux observations dans le but de la régulariser. 74
2.3 Cycle du 3D-Var. 76
2.4 Cycle du 4D-Var. 79
2.5 Fonctionnement séquentiel du filtre de Kalman. 82
2.6 Fonctionnement séquentiel du filtre de Kalman d'ensemble. 87
2.7 Impact de la localisation de covariance [Sakov et Bertino, 2011]. 104
2.8 Impact de l'analyse locale [Sakov et Bertino, 2011]. 104

3.1 Topographie de la surface de la Terre au dernier maximum glaciaire (estimation). Focus sur l'hémisphère Nord (figure de C. Dumas). 108
3.2 Exemple d'évolution du volume au cours du temps obtenue en utilisant un forçage climatique sinusoïdal. 109
3.3 Évolution de l'épaisseur de la calotte polaire utilisant le forçage climatique sinusoïdal de la figure 3.2. 110
3.4 Exemple de convergence pour le premier ordre du gradient 120

3.5 Exemple de convergence pour le deuxieme ordre du gradient. 121
3.6 Graphe de l'écart aux observations $\mathcal{J}^o$ en fonction de a_1 et a_2 122
3.7 Graphe de la fonction coût $\mathcal{J} = \mathcal{J}^o + \mathcal{J}^b$ en fonction de a_1 et a_2. 123
3.8 Evolution des paramètres en fonction des itérations 124
3.9 Nous cherchons à reconstruire le forçage de référence (en bleu) à partir de l'ébauche sinusoïdale (en vert). Les termes de forçages sont à déterminer tous les 500 ans (entre ces deux termes, nous appliquons la régression linéaire entre ces deux points indiquée à l'équation (3.2)). 125
3.10 Volume observé (obtenu par le forçage climatique de référence) et le volume à l'ébauche (obtenu par l'ébauche sinusoïdale du forçage climatique) 125
3.11 Forçage climatique reconstruit (croix vertes) par méthode adjointe, comparaison avec la valeur de référence (ronds bleus) 126
3.12 Ecart relatif entre le volume reconstruit à l'aide de la méthode adjointe et le volume observé au cours du temps. 126
3.13 Forçage climatique reconstitué (en croix bleu) en prenant pour ébauche $F^b_{\text{clim}}(t_i) = F^b_{\text{clim}}(t_{i-1})$. 127
3.14 Evolution du volume de la calotte reconstruit (en bleu) suite au forçage climatique trouvé dans la figure 3.13 128
3.15 Forçage climatique reconstitué (en croix bleu) en prenant pour ébauche $F^b_{\text{clim}}(t_i)$ calculé à l'aide de la relation (3.39). 129

4.1 Geométrie de la calotte simplifiée 138
4.2 Géométrie de la calotte référence et de l'ébauche 143
4.3 Paramètre de glissement basal $\beta(x)$: référence et ébauche 144
4.4 Vitesse de la glace en surface et ses composantes dus au glissement et à la déformation 145
4.5 Écarts types pour la matrice $\mathbf{\Sigma}$ utilisée pour générer l'ensemble de topographies du socle rocheux de départ 146
4.6 Fonction de corrélation utilisée pour générer l'ensemble de topographies du socle rocheux de départ 147
4.7 Exemple ensemble initial de calottes (50 membres) 148
4.8 Assimilation avec ensemble de taille 1 000 : résultat géométrie de la calotte 149
4.9 Écart entre le socle reconstruit et la référence pour un ensemble de taille 1 000 150
4.10 Assimilation avec ensemble de taille 1 000 : résultat β 150
4.11 Assimilation avec ensemble de taille 1 000 : résultat vitesse de glissement . 151
4.12 Écart entre vitesse de glissement reconstruite et référence pour un ensemble de taille 1 000 151
4.13 Assimilation avec ensemble de taille 100 : résultat géométrie de la calotte . 152
4.14 Géométrie de la calotte reconstruite par LETKF avec inflation pour des tailles d'ensemble de 30, 50 et 100 152
4.15 Écart entre le socle rocheux reconstruit et la référence pour des tailles d'ensemble de 30, 50 et 100 153
4.16 Vitesse de glissement reconstruite par LETKF avec inflation pour des tailles d'ensemble de 30, 50 et 100 153

4.17 Écart entre la vitesse de glissement reconstruite et la référence pour des tailles d'ensemble de 30, 50 et 100 154
4.18 Calotte reconstruite par ETKF-N version duale pour un ensemble de taille 100. 155
4.19 Vitesse de glissement reconstruite par ETKF-N version duale pour un ensemble de taille 100. 156
4.20 Évolution de l'inflation estimée à chaque itération par l'ETKF-N version duale pour un ensemble de taille 100. 157

Introduction

Au cours du quaternaire (période géologique couvrant les événements s'étant déroulés les 2,6 derniers millions d'années), la Terre a été soumise à une alternance de périodes chaudes, semblables à l'époque actuelle, et de périodes froides pendant lesquelles d'énormes calottes de glace couvraient l'Amérique du Nord et le nord de l'Eurasie. De ces temps anciens subsistent encore les calottes formées au Groenland et en Antarctique. Les calottes peuvent être vues comme les témoins des évolutions du climat mais affirmer cela est réducteur par rapport à leur véritable rôle. En effet, elles forment un des éléments du système climatique et interagissent avec les autres composantes. Par exemple, lorsqu'une calotte polaire fond, elle apporte aux océans un flux d'eau douce froide. Ce flux impacte la salinité et la température des masses d'eau et donc la circulation thermohaline. La relation avec l'océan n'est bien sûr pas à sens unique puisque la température de l'eau a un impact direct sur la fonte des plateformes flottantes composant une partie du pourtour des calottes polaires. Au niveau des interactions entre les calottes et l'atmosphère, les calottes sont couvertes de neige. Elles augmentent ainsi considérablement l'albédo de la planète. De plus, leur altitude élevée affecte la circulation atmosphérique. A ce sujet, une rétroaction particulièrement importante peut être constatée. Elle provient du fait que, plus une calotte est haute, plus la température à sa surface est basse ce qui favorise la persistance de la glace.

L'évolution des calottes polaires est régie à la fois par ses interactions avec les autres composantes du système climatique et par une dynamique de l'écoulement complexe (fluide visqueux à rhéologie non-newtonienne). D'autre part, des mécanismes comme le glissement à la base, la température de la glace ou l'isostasie (enfoncement du socle rocheux sous le poids de la calotte) influent lourdement sur cette évolution. Tout ceci fait que les calottes polaires constituent un système complexe à modéliser. Bien qu'il ne soit pas chaotique, de nombreuses boucles de rétroactions existent et des phénomènes de seuils sont constatés. Néanmoins, ces difficultés n'ont pas empêché les glaciologues de modéliser l'évolution des calottes polaires. Les premiers modèles apparus dans les années 1990 se concentraient sur l'évolution des calottes à grande échelle de temps (minimum 10 000 ans) pour des problèmes de paléoclimatologie. L'utilisation des modèles pour des simulations à des échelles de temps plus courtes (quelques années) est apparue beaucoup plus récemment, notamment dans le cadre de l'estimation de l'impact des calottes polaires sur l'élévation du niveau des mers. Notons que, contrairement à d'autres composantes des géosciences, les mêmes modèles sont parfois utilisés pour des études à des échelles de temps bien différentes.

Les modèles utilisés pour les calottes polaires font intervenir des paramètres. Or cer-

tains sont peu ou pas connus, ils nécessitent pourtant d'être correctement spécifiés car ils sont influents sur l'évolution des calottes. L'idée naturelle est de tenter d'estimer ces paramètres. C'est là qu'intervient l'assimilation de données qui consiste à combiner de « façon optimale » le modèle et l'information contenue dans les observations pour mieux connaître ces paramètres. Contrairement aux communautés météorologiques et océanographiques, l'utilisation de l'assimilation de données n'est pas encore classique en glaciologie. Ceci est probablement dû au fait que les observations à notre disposition sont bien moindres que dans ces communautés de recherche. Ainsi des données mesurant l'évolution temporelle de l'altitude de la surface des calottes ne sont apparues que très récemment. Jusqu'à peu, les glaciologues pensaient ne pouvoir observer les calottes qu'à un instant t.

Cette thèse porte sur la mise en place de systèmes d'assimilation performants dans le contexte décrit plus haut. Pour mieux nous concentrer sur ce point, nous avons travaillé durant toute cette thèse avec un modèle d'évolution de calotte simplifié appelé Winnie. Les problèmes inverse abordés et les solutions apportées par cette thèse sont les suivants

— Problème de paléoclimatologie : nous reconstruisons l'évolution d'un paramètre climatique influant sur l'évolution des calottes polaires à échelle de temps longue (20 000 ans) à l'aide d'une approche variationnelle (4D-Var). Celle-ci a nécessité le développement du code adjoint du modèle. Le problème abordé n'a été que très rarement traité auparavant.

— Problème du *spin up* : nous reconstruisons l'état initial, la topographie du socle rocheux et les paramètres de glissement basal par filtre de Kalman d'ensemble. Ce problème de calibration de modèle pour des simulations à échelle de temps courte (pas plus de 100 ans) n'est généralement traité que partiellement (souvent les seuls paramètres de glissement basal sont recherchés). Le filtre de Kalman d'ensemble est utilisé pour la première fois dans ce contexte. Nous faisons aussi attention à prendre en compte les temps d'observations (rarement fait jusqu'à présent) et nous travaillons à l'échelle d'une calotte polaire entière.

Notons que les problèmes inverses étudiés ici couvrent les différentes échelles de temps classiques des modèles de calotte et que nous employons durant cette thèse deux méthodes d'assimilation de données différentes.

Ce travail de thèse est présenté en 5 chapitres. Le chapitre 1 porte sur la modélisation et l'observation des calottes polaires. En particulier, nous décrivons le modèle d'évolution simplifié Winnie utilisé tout au long de cette thèse. Le chapitre 2 décrit les méthodes d'assimilation de données classiques, dont celles employées dans cette thèse. Le chapitre 3 aborde le problème de paléoclimatologie. Une partie de ce travail a fait l'objet d'un acte de conférence contenu dans le recueil *Advanced Data Assimilation for Geosciences*. Le chapitre 4 est consacré quant à lui au problème du *spin up* et à sa résolution par filtre de Kalman d'ensemble. Une partie de ce travail a été soumis sous forme d'article au journal *Nonlinear Processes in Geophysics* dans un special issue consacré à l'utilisation des méthodes d'ensemble en géosciences. Enfin, le dernier chapitre conclut et propose des perspectives.

Chapitre 1

Modélisation et observation de l'évolution des calottes polaires

Depuis des millions d'années, l'évolution des calottes polaires est un des éléments façonnant le visage de la Terre. Son rôle direct dans l'élévation du niveau des mers ainsi que ses interactions avec le climat et les océans impactent l'évolution de la surface de la Terre et de ceux qui y vivent. Songeons par exemple qu'au dernier maximum glaciaire, il y a de cela environ 20 000 ans, le niveau des mers était de 120 m plus bas (estimation), il était ainsi possible aux populations humaines et animales de passer de la Sibérie à l'Alaska à pied. Cette problématique du changement du niveau des mers se fait encore plus sensible à l'heure actuelle du fait des changements climatiques en cours. En effet, il est constaté depuis plusieurs années que le niveau des mers monte et ce phénomène devrait s'accélerer dans les prochaines années. Sachant que près de 40 % de la population mondiale vit à moins de 60 km des côtes, il est nécessaire de prévoir au mieux cette hausse. A l'heure actuelle, on estime que cette hausse est due pour moitié à la fonte des calottes antarctique et groenlandaise. Modéliser et observer cette évolution semble donc primordial à cet égard comme l'a souligné le dernier rapport du Groupe d'experts Intergouvernemental sur l'Evolution du Climat en 2007.

Dans le présent chapitre, nous nous attachons à décrire quels sont les phénomènes principaux intervenant dans l'évolution des calottes et à comment les modéliser correctement (sections 1.1 et 1.2). Nous détaillons en particulier le modèle d'évolution de calotte polaire simplifié mais incluant les principaux processus que nous avons utilisé durant toute cette thèse (section 1.3). Nous voyons aussi quelles sont les observations sur les calottes polaires à notre disposition (section 1.4). Nous distinguons notamment les observations utiles pour décrire l'état actuel des calottes polaires et celles utiles pour reconstruire son passé. Ces sections mettent aussi en lumière le rôle de paramètres mal connus mais très influents sur l'évolution des calottes polaires. Les techniques d'assimilation de données permettent classiquement de fournir des estimations de ces paramètres à partir du modèle d'évolution employé et des observations utilisées. La section 1.5 décrit rapidement l'état actuel des travaux en assimilation de données dans le cadre des calottes polaires.

Le contenu de ce chapitre doit beaucoup à la dernière édition de la « Bible » des glaciologues [Cuffey et Paterson, 2010], au livre de [Greve et Blatter, 2009], aux notes de cours

dispensés durant la *Karthaus summer school on ice sheets and glaciers in the climate system 2010*, au manuscrit de la thèse d'État de [Ritz, 1992] ainsi qu'aux manuscrits de thèses de [Drouet, 2012], [Dumas, 2002], [Morlighem, 2011], [Peyaud, 2006], [Quiquet, 2012] et [Rommelaere, 1997].

1.1 Mécanismes physiques en jeu

Pour décrire rapidement ce qui régit l'évolution d'une calotte polaire, celle-ci est due à la différence entre l'alimentation en glace de la surface de la calotte appelée bilan de masse en surface et le drainage par écoulement de la glace du cœur de la calotte vers ses pourtours où la glace fond ou se met à flotter sur la mer (*ice shelves*). Bien évidemment, modéliser l'évolution est un peu plus compliqué que ce qui vient d'être écrit car de nombreux processus (voir figure 1.1) interviennent et interagissent entre eux (présence de boucles de rétroaction). Nous listons ici en préambule quelques interactions majeures entre les différents mécanismes entrant en jeu afin que le lecteur puisse toucher du doigt la complexité d'un modèle d'évolution de calotte polaire :

— L'écoulement de la glace est modulé par un couplage thermomécanique influençant à la fois la déformation de la glace et le glissement de la calotte à la base.

— L'isostasie (enfoncement du socle rocheux sous le poids de la glace) réagit aux variations d'épaisseur mais décale la surface de la calotte.

— Le niveau des mers modifie l'endroit où la glace se met à flotter (ligne d'échouage).

— Le bilan de masse en surface est affecté par l'altitude de la surface de la calotte.

— L'écoulement gouverné par la gravité est fonction de la pente de la surface.

Notons aussi que le rapport d'aspect (épaisseur de la calotte sur son étendue) est très faible pour une calotte polaire. L'utilisation de ce rapport permet la simplification des équations régissant l'évolution de la calotte (voir section 1.2).

Avant de commencer la description plus en détail des mécanismes cités en introduction à cette section, nous introduisons les notations principales employées pour la physique dans le tableau suivant qui, tel un phare, nous l'espérons, permettra au lecteur de se repérer parmi les notations utilisées s'il vient à se perdre dans les étendues autères que peuvent représenter les équations employées.

Table 1.1 – Notations principales employées dans ce chapitre

Notations	Description	Unité
	Coordonnées cartésiennes	
t	temps	(an)
x, y	coordonnées horizontales	(m)
z	coordonnée verticale	(m)

Variables géométriques		
$S(x,y,t)$	altitude de la surface de la calotte	(m)
$B(x,y,t)$	altitude de la base de la glace	(m)
$B_{soc}(x,y,t)$	altitude de la surface du socle rocheux	(m)
$H(x,y,t)$	épaisseur de la calotte : $H = S - B$	(m)

Vitesses		
$\mathbf{u}(x,y,z,t)$	vitesse d'écoulement de la glace : $\mathbf{u} = (u_x, u_y, u_z)$	$(m.an^{-1})$
$\mathbf{u}_s(x,y,t)$	vitesse de la glace en surface : $\mathbf{u}_s = (u_{s,x}, u_{s,y}, u_{s,z})$	$(m.an^{-1})$
$\mathbf{u}_b(x,y,t)$	vitesse de la glace à la base : $\mathbf{u}_b = (u_{b,x}, u_{b,y}, u_{b,z})$	$(m.an^{-1})$
$u_b(x,y,t)$	norme de la vitesse de la glace à la base	$(m.an^{-1})$
$\overline{\mathbf{U}}(x,y,t)$	vitesse de la glace moyennée sur l'épaisseur : $\overline{\mathbf{U}} = (\overline{U}_x, \overline{U}_y)$	$(m.an^{-1})$

Tenseurs		
$\boldsymbol{\tau}(x,y,z,t)$	tenseur des contraintes de Cauchy	(Pa)
$\boldsymbol{\tau}'(x,y,z,t)$	déviateur des contraintes	(Pa)
$\dot{\boldsymbol{\varepsilon}}(x,y,z,t)$	tenseur des taux de déformation	(an^{-1})
$\tau_{ij}(x,y,z,t)$	contrainte de cisaillement $(i,j = x,y,z$ avec $i \neq j)$	(Pa)
$\sigma_i(x,y,z,t)$	contrainte normale $(i = x,y,z)$: $\tau_{ii} = \sigma_i$	(Pa)
$\sigma'_i(x,y,z,t)$	composante normale du déviateur des contraintes	(Pa)
$\dot{\varepsilon}_{ij}(x,y,z,t)$	composante du tenseur des taux de déformation	(an^{-1})
$p(x,y,z,t)$	pression octaédrale	(Pa)
$\tau(x,y,z,t)$	cission efficace, $2^{\text{ème}}$ invariant du tenseur des contraintes	(Pa)
$\boldsymbol{\tau}_b(x,y,t)$	contrainte tangentielle du frottement à la base	(Pa)
$\tau_b(x,y,t)$	norme du frottement basal	(Pa)

Loi de déformation		
$\eta(x,y,z,t)$	viscosité effective	$(Pa.an^{-1})$
$\overline{\eta}(x,y,z,t)$	viscosité effective moyennée sur l'épaisseur	$(Pa.an^{-1})$
n	exposant de la loi de déformation	-
B_{ATn}	coefficient de la loi de déformation	$(Pa^{-n}.an^{-1})$
E_a	énergie d'activation de la loi du fluage	$(J.mol^{-1})$

Loi de glissement		
$\mathbf{u}_b(x,y,t)$	vitesse de la glace à la base : $\mathbf{u}_b = (u_{b,x}, u_{b,y}, u_{b,z})$	$(m.an^{-1})$
$u_b(x,y,t)$	norme de la vitesse de la glace à la base	$(m.an^{-1})$
$\beta(x,y)$	coefficient de la loi de glissement	$(Pa.an.m^{-1})$

Propriétés thermiques		
$T(x,y,z,t)$	température de la glace	(K)
$T_f(x,y,z,t)$	température de fusion de la glace : $T_f = -0,0074\,\rho_i\,g\,H$	(K)
$\kappa(x,y,z,t)$	conductivité thermique de la glace, dépend de T	$(W.K^{-1}.m^{-1})$
$c_V(x,y,z,t)$	capacité calorifique de la glace, dépend de T	$(J.mol^{-1}.kg^{-1})$

Bilan de masse et climat		
$\dot{b}_m(x,y,t)$	bilan de masse en surface	$(m.an^{-1})$
$Acc(x,y,t)$	accumulation de glace en surface	$(m.an^{-1})$
$Abl(x,y,t)$	ablation (fonte) de glace en surface	$(m.an^{-1})$
$T_s(x,y,t)$	température à la surface	$(^\circ\ C)$
$F_{\text{clim}}(t)$	terme de forçage climatique	$(^\circ\ C)$
λ	gradient latitudinal de température	$(^\circ\ C.m^{-1})$
γ	gradient vertical de température	$(^\circ\ C.m^{-1})$
$\dot{f}(x,y,t)$	bilan de masse à la base	$(m.an^{-1})$
$\ell_w(t)$	niveau de la mer	(m)

Constantes		
g	accélération de la pesanteur à la surface de la Terre	$9,81\ m.s^{-2}$
ρ_i	densité de la glace	$910\ kg.m^{-3}$
ρ_w	densité de l'eau de mer	$1028\ kg.m^{-3}$
R	constante des gaz parfaits	$8,3145\ J.mol^{-1}.K^{-1}$

Nous illustrons aussi les variables géométriques intervenant dans l'évolution des calottes polaires par la figure 1.2. Notez que dans ce document l'axe vertical z est toujours orienté vers le haut [1], l'origine est placée au niveau de la mer actuel.

Suite à cette nécessaire introduction, nous pouvons commencer à étudier chacun des mécanismes en jeu dans l'évolution des calottes polaires. Nous débutons par une étude de l'écoulement de la glace.

1.1.1 Écoulement de la glace

La modélisation de l'écoulement de la glace fait intervenir des quantités comme la vitesse de la glace $\mathbf{u}$, le tenseur des contraintes $\boldsymbol{\tau}$ ou le tenseur des taux de déformations $\dot{\boldsymbol{\varepsilon}}$. Nous faisons quelques rappels à leurs sujets avant d'étudier les équations régissant l'écoulement.

Tenseur des contraintes

Le tenseur des contraintes de Cauchy $\boldsymbol{\tau}$ symétrique est représenté par une matrice dont les composantes sont notées τ_{ij} avec i la direction de la contrainte et j la direction de la normale à la surface sur laquelle s'exerce la contrainte $(i,j = x,y,z)$. Les composantes normales du tenseur des contraintes sont traditionnellement notées $\tau_{ii} = \sigma_i$ $(i = x,y,z)$. $\boldsymbol{\tau}$ admet ainsi la forme suivante

$$\boldsymbol{\tau} = \begin{pmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \sigma_y & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \sigma_z \end{pmatrix} \tag{1.1}$$

1. En glaciologie, il est possible de rencontrer Oz orienté vers le bas pour des raisons liées à l'analyse des forages glaciaires (calcul des profondeurs).

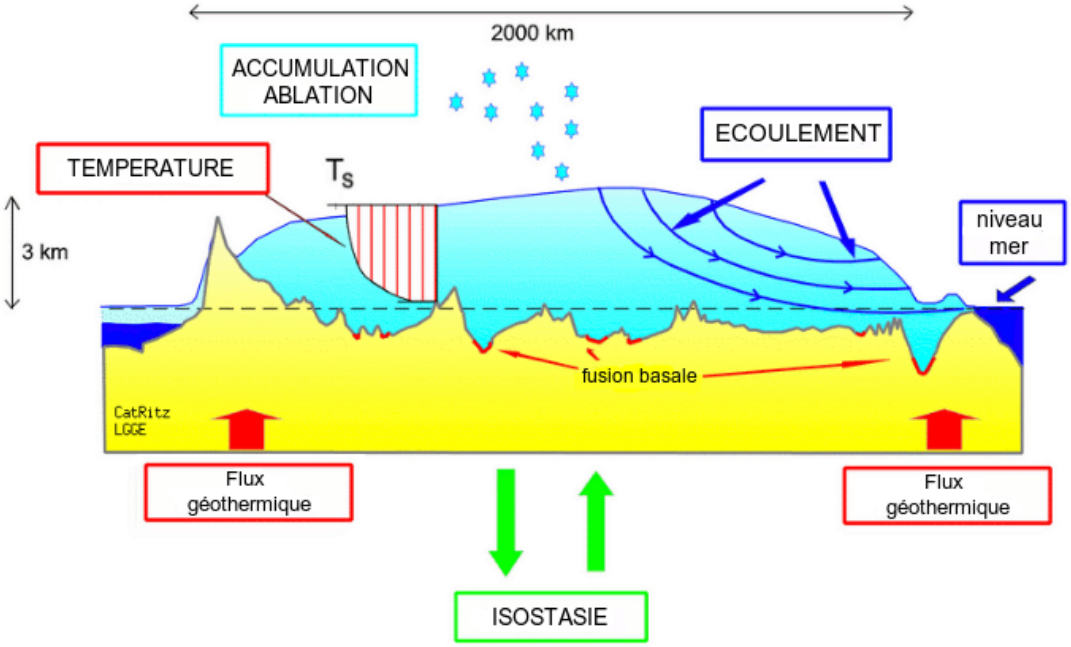

FIGURE 1.1 – Principaux mécanismes intervenant dans l'évolution d'une calotte polaire (figure de C. Ritz). L'accumulation (apport de neige par précipitation principalement) et l'ablation (fonte) de glace composent le flux de glace apporté en surface appelé bilan de masse en surface. Ce potentiel apport de glace est ensuite transporté du cœur de la calotte vers ses extrémités par écoulement. Via la dépendance de la viscosité de la glace à la température et l'impact de la fusion basale sur le glissement, l'écoulement est influencé par la température de la glace qui elle même est fonction du flux géothermique et de l'écoulement lui-même (couplage thermomécanique). Aux extrémités, les éventuelles parties flottantes de la calotte dépendent du niveau des mers. Enfin, le socle rocheux s'enfonce sous le poids de la calotte (isostasie). Pour finir, remarquons que le rapport d'aspect (épaisseur de la calotte sur son étendue) est très faible. Ce point sera utile pour simplifier les équations mécaniques.

Notons enfin p la pression octaédrale définie par la relation

$$p = -\frac{\sigma_x + \sigma_y + \sigma_z}{3} \tag{1.2}$$

$-p$ est le premier invariant du tenseur des contraintes.

Déviateur des contraintes

Définissons maintenant le déviateur des contraintes $\boldsymbol{\tau}'$. Les composantes τ'_{ij} de cette matrice valent

$$\tau'_{ij} = \tau_{ij} + \delta_{ij}\, p \qquad\qquad i, j = x, y, z \tag{1.3}$$

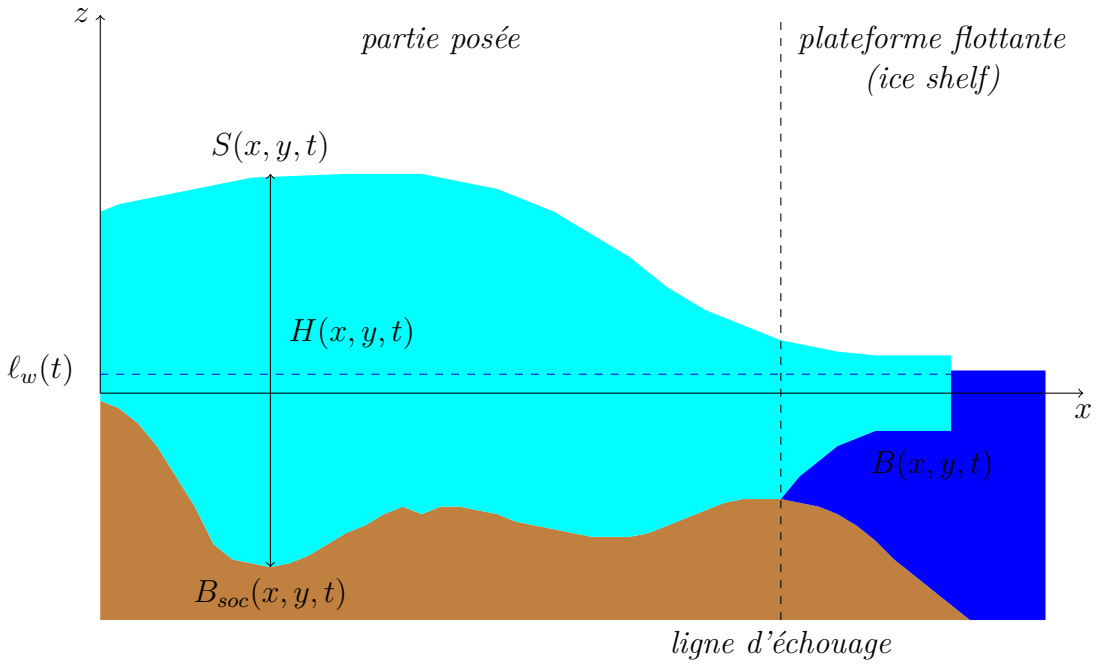

FIGURE 1.2 – Schéma d'une coupe d'une calotte polaire selon les axes x et z. La calotte a pour épaisseur H, l'altitude de sa surface est notée S et l'altitude de sa base B. La partie principale de la calotte repose sur le socle rocheux d'altitude B_{soc} (dans ce cas $B = B_{soc}$) tandis qu'aux extrémités peuvent apparaître des plateformes flottantes appelées *ice shelves* (dans ce cas $B > B_{soc}$). La ligne d'échouage délimite les parties posées et flottantes de la calotte. Enfin ℓ_w correspond au niveau de la mer à un instant donné (le niveau de la mer actuel est pris pour origine).

avec δ_{ij} le symbole de Kronecker (vaut 1 si $i = j$, 0 sinon). Le déviateur des contraintes vaut donc

$$\boldsymbol{\tau}' = \begin{pmatrix} \sigma'_x & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \sigma'_y & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \sigma'_z \end{pmatrix} = \begin{pmatrix} \sigma_x + p & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \sigma_y + p & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \sigma_z + p \end{pmatrix} \tag{1.4}$$

Enfin notons τ la cission efficace (deuxième invariant du tenseur des contraintes) définit par la relation suivante

$$\tau^2 = \frac{1}{2}\left({\sigma'_x}^2 + {\sigma'_y}^2 + {\sigma'_z}^2\right) + \tau_{xy}^2 + \tau_{xz}^2 + \tau_{yz}^2 \tag{1.5}$$

Tenseur des taux de déformation

Nous définissons maintenant le tenseur des taux de déformation noté $\dot{\boldsymbol{\varepsilon}}$. Ses composantes notées $\dot{\varepsilon}_{ij}$ $(i, j = x, y, z)$ sont définies par la relation suivante

$$\dot{\varepsilon}_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial j} + \frac{\partial u_j}{\partial i}\right) \tag{1.6}$$

Dynamique de la glace

A de grandes échelles spatiales et temporelles, la glace peut être considérée comme un fluide incompressible non-newtonnien. L'écoulement de la glace est alors régi par les équations de conservation de la masse (1.7) et de conservation des moments (1.8)

$$\nabla \cdot \mathbf{u} = 0 \tag{1.7}$$

$$\rho_i \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = \nabla \cdot \boldsymbol{\tau} + \rho_i \, \mathbf{g} - 2 \, \rho_i \, \boldsymbol{\Omega} \wedge \mathbf{u} \tag{1.8}$$

L'équation (1.7) appelée usuellement équation de continuité traduit ici l'incompressibilité du fluide. Après analyse dimensionnelle du système (voir [Greve et Blatter, 2009] et [Morlighem, 2011] pour plus de détails), l'accélération, le terme de convection ($\mathbf{u} \cdot \nabla \mathbf{u}$) et la force de Coriolis ($\rho_i \, \boldsymbol{\Omega} \wedge \mathbf{u}$) sont négligeables par rapport aux autres termes de l'équation (1.8). Elle se réécrit

$$\nabla \cdot \boldsymbol{\tau} + \rho_i \, \mathbf{g} = 0 \tag{1.9}$$

Dans un repère cartésien, l'équation (1.9) se réécrit

$$\begin{cases} \dfrac{\partial \sigma_x}{\partial x} + \dfrac{\partial \tau_{xy}}{\partial y} + \dfrac{\partial \tau_{xz}}{\partial z} = 0 \\ \dfrac{\partial \tau_{xy}}{\partial x} + \dfrac{\partial \sigma_y}{\partial y} + \dfrac{\partial \tau_{yz}}{\partial z} = 0 \\ \dfrac{\partial \tau_{xz}}{\partial x} + \dfrac{\partial \tau_{yz}}{\partial y} + \dfrac{\partial \sigma_z}{\partial z} = \rho_i \, g \end{cases} \tag{1.10}$$

Elle est communément appelée ***équation d'équilibre quasi-statique***. Le système précédent ne fait pas intervenir directement les vitesses contrairement à l'équation d'incompressibilité, cependant il est possible de relier le déviateur des contraintes aux vitesses par une loi de déformation de la glace.

Loi de déformation

Nous considérons ici que la glace est incompressible et isotrope [2]. Le tenseur des taux de déformation et le déviateur des contraintes sont reliés grâce à la viscosité effective η.

$$\tau'_{ij} = 2 \, \eta \, \dot{\varepsilon}_{ij} \qquad i, j = x, y, z \tag{1.11}$$

La glace est un fluide non-newtonien, sa viscosité dépend de la cission efficace suivant une loi puissance de type Norton-Hoff communément appelée ***loi de Glen généralisée*** [3] en glaciologie (voir à ce sujet [Duval, 1979] ou [Cuffey et Paterson, 2010]). Celle-ci prend la forme

$$\frac{1}{\eta} = B_{ATn} \, \tau^{n-1} \tag{1.12}$$

avec $B_{AT,n}$ coefficient de la loi de déformation associé à l'exposant n. Traditionnellement, celui-ci prend pour valeur $n = 3$ même si des mesures expérimentales ont montré que sa

2. Cette hypothèse n'est pas forcément valable, voir entre autre [Cuffey et Paterson, 2010], chapitre 3 ou [Gillet-Chaulet, 2006] sur l'anisotropie de la glace

3. On rencontre aussi le terme de *loi isotropique de Nye-Glen.*

valeur réelle pouvait fluctuer en fonction de la contrainte et de la profondeur. Cependant, il est possible d'utiliser d'autres lois. En particulier, nous choisissons ici d'employer dans le modèle simplifié utilisé dans cette thèse une loi polynomiale suggérée par [Lliboutry, 1993] et [Cuffey et Paterson, 2010] tel que

$$\frac{1}{\eta} = B_{AT3}\,\tau^2 + B_{AT1} \tag{1.13}$$

Cette dernière loi offre plusieurs avantages. Elle se rapproche d'une loi newtonienne à faible contrainte (ainsi la viscosité effective n'est pas infinie même s'il n'y a pas de déformation) et de la loi de Glen à forte contrainte.

Les coefficients B_{ATn} traduisent la fluidité de la glace et varient généralement en fonction de la température de la glace (d'où le couplage thermomécanique de l'écoulement déjà évoqué en introduction). Cette dépendance en température est souvent décrite par une loi d'Arrhenius

$$B_{ATn} = B_{ATn,0}\,\exp\left(\frac{E_a}{R}\left(\frac{1}{T_f} - \frac{1}{T}\right)\right) \tag{1.14}$$

avec E_a l'énergie d'activation de la loi du fluage, R la constante des gaz parfaits, T la température de la glace en Kelvin et $T_f = -0,0074\,\rho_i\,g\,H$ température de fusion de la glace. Dans le cas où la calotte est supposée isotherme, les coefficients de la loi de déformation sont des constantes.

Équations de Stokes

Grâce à la loi de déformation (1.11) et aux relations (1.2), (1.3), il est possible de réécrire l'équation d'équilibre quasi-statique (1.10) en fonction des vitesses et de la pression octaédrale. Ce sont les ***équations de Stokes***.

$$\left\{\begin{array}{rcl}
\frac{\partial}{\partial x}\left(2\eta\frac{\partial u_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(\eta\left(\frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y}\right)\right) + \frac{\partial}{\partial z}\left(\eta\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)\right) - \frac{\partial p}{\partial x} & = & 0 \\
\frac{\partial}{\partial x}\left(\eta\left(\frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y}\right)\right) + \frac{\partial}{\partial y}\left(2\eta\frac{\partial u_y}{\partial y}\right) + \frac{\partial}{\partial z}\left(\eta\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right)\right) - \frac{\partial p}{\partial y} & = & 0 \\
\frac{\partial}{\partial x}\left(\eta\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)\right) + \frac{\partial}{\partial y}\left(\eta\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right)\right) + \frac{\partial}{\partial z}\left(2\eta\frac{\partial u_z}{\partial z}\right) - \frac{\partial p}{\partial z} & = & \rho_i\,g
\end{array}\right. \tag{1.15}$$

Ce système est non-linéaire en vitesse du fait de la viscosité non-newtonienne de la glace. En effet, si nous reprenons la loi polynomiale (1.13), η est solution de l'équation

$$2B_{AT3}\,\dot{\varepsilon}^2\,\eta^3 + B_{AT1}\,\eta - 1 = 0 \tag{1.16}$$

avec $\dot{\varepsilon}^2$ deuxième invariant du tenseur des taux de déformation $\dot{\boldsymbol{\varepsilon}}$

$$\dot{\varepsilon}^2 = \left(\frac{\partial u_x}{\partial x}\right)^2 + \left(\frac{\partial u_y}{\partial y}\right)^2 + \left(\frac{\partial u_z}{\partial z}\right)^2 + \frac{1}{2}\left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y}\right)^2 \quad (1.17)$$

Pour obtenir que le système soit bien posé, il faut adjoindre à ces équations l'équation d'incompressibilité (1.7) (d'où 4 inconnues à rechercher reliées entre elles par 4 équations indépendantes) ainsi que les conditions limites à la surface et à la base.

Conditions aux limites pour le tenseur des contraintes

— À la surface, le tenseur des contraintes est soumis à la condition de Dirichlet traduisant la pression de l'atmosphère sur la calotte $\boldsymbol{\tau}|_{z=S} \cdot \mathbf{n} = -p_{atm}\ \mathbf{n}$ avec $\mathbf{n}$ la normale à la surface de la calotte pointée vers l'extérieur et p_{atm} la pression atmosphérique. Généralement, celle-ci est négligeable d'où la condition suivante

$$\boldsymbol{\tau}|_{z=S} \cdot \mathbf{n} = 0$$

Réécrite en terme de déviateurs des contraintes, nous avons

$$\begin{cases} \sigma'_x \dfrac{\partial S}{\partial x} + \tau_{xy}\dfrac{\partial S}{\partial y} - \tau_{xz} - p\dfrac{\partial S}{\partial x} &= 0 \\ \tau_{xy}\dfrac{\partial S}{\partial x} + \sigma'_y \dfrac{\partial S}{\partial y} - \tau_{yz} - p\dfrac{\partial S}{\partial y} &= 0 \\ \tau_{xz}\dfrac{\partial S}{\partial x} + \tau_{yz}\dfrac{\partial S}{\partial y} - \sigma'_z + p &= 0 \end{cases} \quad (1.18)$$

Le système précédent peut s'exprimer en terme de vitesses à l'aide de la loi de déformation (1.11), cette condition limite est alors une condition de Neumann pour les vitesses en surface.

— Dans le cas du fond flottant ($B > B_{soc}$, critère de flottaison : $\rho_i\, H < -\rho_w\,(\ell_w - B_{soc})$), la seule force s'appliquant est la poussée d'Archimède et il n'y a pas de contrainte tangentielle. La condition à la limite à la base se met alors sous la forme

$$\boldsymbol{\tau}|_{z=B} \cdot \mathbf{n} = -\rho_w\, g\,(\ell_w - B)\,\mathbf{n}$$

avec ℓ_w le niveau de la mer à un instant donné et $\mathbf{n}$ la normale à la base de la calotte pointée vers l'extérieur. Comme la condition limite en surface, cette condition peut être réécrite en terme de déviateurs des contraintes

$$\begin{cases} \sigma'_x \dfrac{\partial B}{\partial x} + \tau_{xy}\dfrac{\partial B}{\partial y} - \tau_{xz} - p\dfrac{\partial B}{\partial x} &= -\rho_w\, g\,(\ell_w - B)\,\dfrac{\partial B}{\partial x} \\ \tau_{xy}\dfrac{\partial B}{\partial x} + \sigma'_y \dfrac{\partial B}{\partial y} - \tau_{yz} - p\dfrac{\partial B}{\partial y} &= -\rho_w\, g\,(\ell_w - B)\,\dfrac{\partial B}{\partial y} \\ \tau_{xz}\dfrac{\partial B}{\partial x} + \tau_{yz}\dfrac{\partial B}{\partial y} - \sigma'_z + p &= \rho_w\, g\,(\ell_w - B) \end{cases} \quad (1.19)$$

Une nouvelle fois, le système précédent peut s'exprimer en terme de vitesses, cette condition limite est alors une condition de Neumann pour les vitesses à la base flottante.

— Dans le cas du fond posé ($B = B_{soc}$, critère : $\rho_i\, H \geq -\rho_w\, (\ell_w - B_{soc})$), nous avons la condition basale suivante

$$\boldsymbol{\tau}|_{z=B} \cdot \mathbf{n} = \boldsymbol{\tau}_b - \rho_i\, g\, H\, \mathbf{n}$$

avec $-\rho_i\, g\, H\, \mathbf{n}$ la contribution due au poids de la calotte et $\boldsymbol{\tau}_b$ la composante tangentielle supposée connue du tenseur des contraintes à la base (d'où $\boldsymbol{\tau}_b \cdot \mathbf{n} = 0$). Dans le cas d'une calotte polaire, la composante verticale (en z) des vecteurs tangents orthogonaux au vecteur normal $\mathbf{n}$ est négligeable par rapport aux composantes horizontales (du fait du faible rapport d'aspect épaisseur/étendue). Dans ce cadre, la composante tangentielle du tenseur des contraintes admet alors la forme suivante $(\tau_{b,x}, \tau_{b,y}, 0)^T$. Tout ceci peut bien sûr se réécrire en terme de déviateurs des contraintes

$$\left\{\begin{array}{rcl} \sigma'_x \dfrac{\partial B}{\partial x} + \tau_{xy} \dfrac{\partial B}{\partial y} - \tau_{xz} - p \dfrac{\partial B}{\partial x} & = & \tau_{b,x} - \rho_i\, g\, H \dfrac{\partial B}{\partial x} \\ \tau_{xy} \dfrac{\partial B}{\partial x} + \sigma'_y \dfrac{\partial B}{\partial y} - \tau_{yz} - p \dfrac{\partial B}{\partial y} & = & \tau_{b,y} - \rho_i\, g\, H \dfrac{\partial B}{\partial y} \\ \tau_{xz} \dfrac{\partial B}{\partial x} + \tau_{yz} \dfrac{\partial B}{\partial y} - \sigma'_z + p & = & \rho_i\, g\, H \end{array}\right. \tag{1.20}$$

Encore une fois, le système précédent peut s'exprimer en terme de vitesses (condition de Neumann).

Toutes les conditions limites données sont des conditions de Neumann pour les vitesses. Pour la résolution de l'équation de Stokes, il est aussi nécessaire d'avoir une condition de Dirichlet sur les vitesses à un point donné de la calotte. Généralement, il y a au moins une zone à la base posée de la calotte où il n'y a pas de glissement (par exemple, pour des raisons thermiques) permettant d'assurer l'existence d'un point Dirichlet pour les vitesses.

Glissement basal

Nous avons affirmé plus haut que la composante tangentielle du frottement basal $\boldsymbol{\tau}_b$ était supposément connue dans l'équation (1.20). Estimer sa valeur a fait l'objet de nombre d'études qui généralement cherchent à la relier à la vitesse basale $\mathbf{u}_b$ (voir cependant la remarque dans le cas plastique un peu plus loin).

Le glissement à la base est un élément complexe et potentiellement mal connu de la modélisation des calottes polaires. Il dépend notamment du type de sédiment reposant sur le socle rocheux et de la lubrification de l'interface calotte-socle due à la présence d'eau liquide sous la calotte polaire. L'interface socle-calotte contient tout un système d'hydrologie sous-glaciaire (au sujet de la modélisation de tels systèmes lire [de Fleurian, 2010]). Notons par exemple la présence de nombreux lacs sous-glaciaires en Antarctique. Le plus

connu est sans aucun doute le lac Vostok dont les dimensions approximatives sont de 250 km de longueur pour 50 km de largeur.

À l'échelle de temps étudiée, le socle peut être supposé comme fixé (l'isostasie est un phénomène n'affectant pas la vitesse relative de la glace par rapport au socle). Supposons aussi une fusion basale nulle. Si $\mathbf{n}$ désigne un vecteur normal à la base de la calotte posée, alors la vitesse à la base $\mathbf{u}_b$ obéit à la relation

$$\mathbf{u}_b \cdot \mathbf{n} = 0 \tag{1.21}$$

La vitesse à la base est entièrement composée de la vitesse de glissement tangentielle à l'interface calotte-socle. De nombreuses lois de glissement ont été proposées afin de la relier à τ_b. [Cuffey et Paterson, 2010] note cependant qu'elles sont presque toutes de la forme suivante

$$u_b = A\frac{(\tau_b)^{\frac{1}{m}}}{N^{\theta}} \tag{1.22}$$

avec u_b la norme de la vitesse de glissement, τ_b la norme du frottement basal et N la pression effective traduisant la différence entre le poids de la glace et la pression de l'eau liquide contenue dans la cavité à l'interface socle-calotte. Les paramètres A, m et θ à prendre varient selon les auteurs. D'autre part, il est de plus en plus admis qu'ils peuvent varier spatialement.

Pour résoudre le système d'équations décrivant l'écoulement de Stokes, il nous faut connaître $\boldsymbol{\tau}_b$. Une bonne modélisation doit prendre en compte les faits énoncés précédemment dont notamment la présence d'eau liquide à l'interface calotte-socle. Celle-ci n'est possible que si la température de la glace à la base est au point de fusion T_{fusion}. Une modélisation classique de $\boldsymbol{\tau}_b$ permettant de fixer sa direction est

$$\boldsymbol{\tau}_b = \begin{cases} -C\,|\mathbf{u}_b|^{m-1}\,\mathbf{u}_b & \text{si } T(z=B) = T_{fusion} \\ \mathbf{0} & \text{sinon} \end{cases} \tag{1.23}$$

avec $C > 0$ et m des paramètres variant potentiellement en espace et en temps. $T(z = B)$ désigne la température à la base de la calotte. Notons que cette modélisation permet d'intégrer le cas où le sédiment a un comportement plastique. Dans ce cas, le frottement s'établit à la valeur du seuil de plasticité. L'équation (1.23) peut être utilisée avec une valeur très petite de m ($< 1/100$ suffit) et C prend la valeur du seuil de plasticité.

Dans le cas d'études à court terme, la modélisation précédente peut être remplacée par

$$\boldsymbol{\tau}_b = \begin{cases} -\beta\,\mathbf{u}_b & \text{si } T(z=B) > T_{fusion} \\ \mathbf{0} & \text{sinon} \end{cases} \tag{1.24}$$

avec $\beta(x, y) > 0$ un coefficient de frottement variant en espace et supposé fixe en temps.

Modèle Full Stokes

Les différent paragraphes précédents nous donnent toutes les briques nécessaires afin de considérer le problème complet de l'écoulement de la calotte. Les modèles résolvant ce problème sont appelés ***Full Stokes***, ils se résument de la manière suivante.

Modèle 1.1 (*Full Stokes*)

A un instant donné t, évaluer les 4 variables u_x, u_y, u_z et p dépendant de x, y, z solutions du système comprenant :

- *les équations de Stokes* (1.15).
- *l'équation d'incompressibilité* (1.7).
- *la calcul de la viscosité effective (si loi polynomiale, utiliser* (1.16) *et* (1.17)*).*
- *les conditions limites en surface* (1.18) *traduites en fonction de u_x, u_y, u_z et p.*
- *les conditions limites à la base (pour cas flottant :* (1.19)*, pour cas posé :* (1.20)*) traduites en fonction de u_x, u_y, u_z et p.*
- *la loi de glissement basal* (1.24).

Un tel système est coûteux à résoudre d'autant qu'il est surtout intéressant pour tenir compte des petites échelles spatiales nécessitant un maillage fin en temps et en espace. Cette restriction limite principalement son utilisation à l'étude de processus locaux où il est l'outil idéal.

Notons qu'un modèle Full-Stokes appelé Elmer/ICE [Gagliardini *et al.*, 2013] a été développé depuis une dizaine d'années conjointement par le LGGE de Grenoble et le CSC en Finlande. Ce modèle a pu être utilisé pour étudier des bassins versants (cas du glacier Variegated en Alaska [Jay-Allemand *et al.*, 2011]) voire la totalité d'une calotte (cas du Groenland [Gillet-Chaulet *et al.*, 2012]) sur de courtes échelles de temps (le siècle).

Calottes et températures

Nous avons vu plus tôt que la température de la glace joue un grand rôle dans l'évaluation de la viscosité effective (voir équation (1.14)). Elle intervient aussi au niveau de la loi de glissement (voir équations (1.23) et (1.24)). Il faut donc connaître généralement la température T de la glace dans la calotte. Elle est obtenue en résolvant l'équation suivante.

$$\rho_i \, c_V \left(\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right) = \frac{\partial}{\partial x} \left(\kappa \frac{\partial T}{\partial x} \right) + \frac{\partial}{\partial y} \left(\kappa \frac{\partial T}{\partial y} \right) + \frac{\partial}{\partial z} \left(\kappa \frac{\partial T}{\partial z} \right) + Q \tag{1.25}$$

avec $\kappa = \kappa(T)$ la conductivité thermique de la glace, $c_V = c_V(T)$ sa capacité calorifique et Q un terme source de chaleur. Ce terme fait notamment intervenir le produit des taux de déformations et de contraintes pour la chaleur produite par la déformation de la glace.

Nous ne rentrons pas dans les détails sur la résolution de cette équation (conditions limites, termes sources, ...). Nous renvoyons au chapitre 9 de [Cuffey et Paterson, 2010] pour qui serait interessé par plus d'informations sur ce point. Notons enfin que nous faisons l'hypothèse d'une calotte isotherme dans le modèle d'évolution employé durant cette thèse rendant du coup la résolution de l'équation plus haut superflue.

1.1.2 Évolution de la géométrie

Jusqu'à présent, nous n'avons considéré que l'écoulement de la glace sans évoquer l'évolution de la géométrie de la calotte polaire. L'évolution de la surface libre ou de

l'épaisseur est gérée par la conservation de la masse. Dans le cas de la surface, il suffit d'étudier les conditions limites sur les vitesses. Dans le cas de l'épaisseur, il faut intégrer l'équation d'incompressibilité (1.7) (forme locale de la conservation de la masse) verticalement sur l'épaisseur pour obtenir l'équation dite du bilan de masse. Enfin, la topographie du socle rocheux évolue au cours du temps du fait de l'isostasie.

Conditions limites sur les vitesses

Considérons la surface de la calotte. Elle peut être décrite par l'équation implicite suivante

$$F_S(x, y, z, t) = z - S(x, y, t) = 0 \tag{1.26}$$

Notons $\mathbf{n}$ le vecteur normal (unitaire) de cette surface, il est orienté vers l'extérieur de la calotte

$$\mathbf{n} = \frac{\nabla F_S}{|\nabla F_S|} = \frac{1}{N_S} \begin{pmatrix} -\frac{\partial S}{\partial x} \\ -\frac{\partial S}{\partial y} \\ 1 \end{pmatrix} \qquad N_S = \left(1 + \left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2\right)^{1/2} \tag{1.27}$$

Si nous notons $\mathbf{v}$ la vitesse de la surface (différente de la vitesse de la glace en surface $\mathbf{u}_s$) alors la dérivée temporelle de F_S suivant le mouvement de la surface doit être nulle d'où

$$\frac{\partial F_S}{\partial t} + \nabla F_S \cdot \mathbf{v} = 0 \tag{1.28}$$

Introduisons maintenant $a_S^\perp = (\mathbf{v} - \mathbf{u}_s) \cdot \mathbf{n}$ le flux de glace perpendiculaire à la surface traversant cette dernière. Nous avons alors

$$\frac{\partial F_S}{\partial t} + \nabla F_S \cdot \mathbf{u}_s = \frac{\partial F_S}{\partial t} + \nabla F_S \cdot \mathbf{v} - \nabla F_S \cdot (\mathbf{v} - \mathbf{u}_s) = -N_S\, a_S^\perp \tag{1.29}$$

En posant $N_S\, a_S^\perp = \dot{b}_m$, nous avons finalement

$$\frac{\partial S}{\partial t} + u_{s,x}\, \frac{\partial S}{\partial x} + u_{s,y}\, \frac{\partial S}{\partial y} - u_{s,z} = \dot{b}_m \tag{1.30}$$

Le terme $\dot{b}_m$ est appelé bilan de masse en surface. Il dépend des conditions climatiques influençant les précipitations de neige ou la fonte (pour plus de détails voir 1.1.3).

De la même manière peut être obtenue la relation suivante à la base de la calotte

$$\frac{\partial B}{\partial t} + u_{b,x}\, \frac{\partial B}{\partial x} + u_{b,y}\, \frac{\partial S}{\partial y} - u_{b,z} = \dot{f} \tag{1.31}$$

avec $\dot{f}$ le bilan de masse à la base. Ce terme dépend de la fusion (ou du regel) de la glace sous l'influence du flux géothermique, si cette glace est posée ; ou de la chaleur apportée par les océans si cette glace est flottante (pour plus de détails voir 1.1.3).

Les modèles Full Stokes ont directement les 3 composantes de la vitesse, il peuvent calculer directement dS/dt et dB/dt grâce aux relations (1.30) et (1.31).

Équation du bilan de masse

Intégrons maintenant l'équation (1.7) sur l'épaisseur de la calotte. En utilisant les conditions aux limites (1.30) et (1.31), nous obtenons la relation suivante dite ***équation du bilan de masse***

$$\frac{\partial H}{\partial t} = \dot{b}_m - \dot{f} - \frac{\partial (H\overline{U}_x)}{\partial x} - \frac{\partial (H\overline{U}_y)}{\partial y} \tag{1.32}$$

avec $\overline{\mathbf{U}} = (\overline{U}_x, \overline{U}_y)$ la vitesse de la glace moyennée sur l'épaisseur de la calotte. La dépendance temporelle de l'écoulement est ainsi explicitée.

Isostasie

Le poids de la calotte polaire entraîne à long terme (temps de relaxation d'environ 3000 ans) un enfoncement de la lithosphère et un fluage de l'asthénosphère. Cet enfoncement représente environ 1/3 de l'épaisseur de la calotte. Ce phénomène appelé ***isostasie*** ne peut être négligé dans le cas d'études paléoclimatiques. Nous renvoyons à [Le Meur, 1996] pour de plus amples informations à ce sujet. Notons cependant un fait caractéristique, conséquence directe du temps de réaction du socle par rapport à l'évolution d'une calotte polaire. Le socle rocheux reprend de l'altitude pendant et surtout après la fonte totale d'une calotte polaire, c'est le rebond post-glaciaire.

1.1.3 Bilan de masse

Nous venons de voir que la géométrie de la calotte évolue au cours du temps du fait notamment de la conservation de la masse. Elle se traduit par l'intervention de deux termes appelés bilan de masse en surface $\dot{b}_m$ et à la base $\dot{f}$. Nous donnons dans la suite quelques détails sur ces deux termes.

Bilan de masse en surface

Le terme $\dot{b}_m$ appelé bilan de masse en surface est le flux de glace (positif ou négatif) apporté par l'extérieur à la surface de la calotte. Nous décomposons ce terme en deux termes, un positif nommé accumulation Acc et l'autre négatif appelé ablation Abl.

$$\dot{b}_m = Acc + Abl \tag{1.33}$$

L'accumulation regroupe tous les phénomènes apportant de la glace au système en surface (précipitations de neige, regel de l'eau, ...). L'ablation englobe tous les mécanismes extérieurs retirant de la glace en surface (fonte, transport par le vent, ...).

Le bilan de masse en surface dépend de nombreux phénomènes extérieurs dont les conditions météorologiques et la température en surface. Il est possible de le simuler en utilisant notamment des modèles de bilan énergétique impliquant les radiations reçues en surface (consulter à ce sujet [Mölg, 2010] et références à l'intérieur). Cette approche valable pour les glaciers nécessite cependant de connaître un certain nombre de paramètres climatiques.

Dans le cas des calottes polaires, le bilan de masse peut être évalué en couplant le modèle d'évolution avec un modèle climatique, par exemple un modèle de climat de complexité intermédiaire pour les grandes échelles de temps appelé aussi EMICS (voir par exemple [Philippon *et al.*, 2006] pour le couplage entre le modèle de calotte polaire GRISLI et le modèle de climat et biosphère CLIMBER).

L'autre solution est d'utiliser comme forçage un modèle climatique global (GCM) ou régional (RGM). Citons dans ce cas les 2 RGM souvent utilisés en glaciologie : MAR (pour Modèle Atmosphérique Régional) et RACMO (pour Regional Atmospheric Climate MOdel). Nous renvoyons aux articles de [Fettweis, 2007] et [Ettema *et al.*, 2009] pour une idée de l'évolution du bilan de masse en surface pour la calotte groenlandaise sur les 30 dernières années et à l'article de [van de Berg *et al.*, 2006] pour le cas de l'Antarctique.

Dans le cadre de cette thèse, le calcul du bilan de masse en surface se fait avec un modèle très simple dépendant de la température en surface (voir 1.3.3).

Bilan de masse à la base

Le terme $\dot{f}$ appelé bilan de masse à la base est le flux de glace (positif ou négatif) à la base de la calotte.

Dans le cas de la partie posée sur le socle rocheux, $\dot{f}$ se résume principalement à de la fusion basale ($\dot{f} < 0$) du fait d'une température de la glace supérieure au point de fusion (notons que la pression due au poids de la calotte joue un rôle non négligeable dans ce cas). Cependant, si de l'eau est présente à la base, il peut y avoir aussi du regel. $\dot{f}$ fait partie des variables calculées dans la résolution thermique.

Dans le cas des plateformes flottantes, la fusion ou le regel sont liés aux échanges avec l'océan. Pour des raisons de circulation océanique, la fusion se trouve près de la ligne d'échouage et au bord de la plateforme, le regel quant à lui se produit plutôt au milieu.

1.2 Écoulement de Stokes et ses approximations

L'objectif de cette thèse est de développer une méthodologie d'assimilation de données pour les calottes à grande échelle spatiale (Groenland et Antarctique) et potentiellement à grande échelle de temps (voir le chapitre 3 consacré aux expériences en paléoclimatologie). Il est alors judicieux d'utiliser des modèles moins coûteux en temps de calcul que les modèles Full Stokes. Il nous faut donc approcher ces modèles. La suite de cette section est consacrée à ce sujet et décrit plusieurs alternatives connues au problème Full-Stokes. Elles supposent toutes un rapport d'aspect petit et que l'approximation hydrostatique soit valide.

1.2.1 Rapport d'aspect et ses conséquences

Toutes les approximations que nous allons étudier reposent sur le fait que le rapport d'aspect de la calotte δ est petit. Notons $[H]$ et $[L]$ les dimensions caractéristiques verti-

cales et horizontales d'une calotte polaire. Typiquement, le rapport d'aspect d'une calotte polaire est de l'ordre de

$$\delta = \frac{[H]}{[L]} \approx 10^{-3} \tag{1.34}$$

Si nous considérons δ comme quantité négligeable, un certain nombre d'approximations en découle notamment sur les vitesses. Notons $[U]$ et $[W]$ les valeurs caractéristiques horizontales et verticales des vitesses de la glace. L'équation d'incompressibilité (1.7) permet de relier les quantités caractéristiques par la relation

$$\frac{[U]}{[L]} + \frac{[W]}{[H]} \sim 0 \Longrightarrow [W] \sim [U]\frac{[H]}{[L]} \sim [U]\,\delta \tag{1.35}$$

Cela implique que

$$\frac{\partial u_z}{\partial x} \sim \delta^2\,\frac{\partial u_x}{\partial z} \qquad\qquad \frac{\partial u_z}{\partial y} \sim \delta^2\,\frac{\partial u_y}{\partial z} \tag{1.36}$$

Les composantes $\dot{\varepsilon}_{xz}$ et $\dot{\varepsilon}_{yz}$ du tenseur des taux de déformation peuvent alors se simplifier sous la forme suivante

$$\dot{\varepsilon}_{xz} \approx \frac{1}{2}\frac{\partial u_x}{\partial z} \qquad\qquad \dot{\varepsilon}_{yz} \approx \frac{1}{2}\frac{\partial u_y}{\partial z} \tag{1.37}$$

1.2.2 Approximation hydrostatique

Toutes les approximations que nous allons étudier supposent que l'approximation hydrostatique soit valable sur l'axe vertical. Cela veut dire que la composante en z de l'équation d'équilibre quasi-statique (1.10) se réécrit

$$\frac{\partial \sigma_z}{\partial z} \cong \rho_i g \tag{1.38}$$

et que les composantes en z des conditions aux limites (1.18), (1.19) et (1.20) sur le tenseur des contraintes se réécrivent

$$\left\{\begin{array}{lcll} \sigma_z|_{z=S} & = & 0 & \\ \sigma_z|_{z=B} & = & -\rho_w\,g\,(\ell_w - B) & \text{cas flottant} \\ \sigma_z|_{z=B} & = & -\rho_i\,g\,H & \text{cas posé} \end{array}\right. \tag{1.39}$$

Les équations précédentes peuvent se réécrire en terme de vitesses et pression octaédrale. Ainsi l'équation (1.38) peut se mettre sous la forme

$$\frac{\partial p}{\partial z} - \frac{\partial}{\partial z}\left(2\eta\frac{\partial u_z}{\partial z}\right) = -\rho_i\,g \tag{1.40}$$

et les conditions aux limites (1.39) sous la forme

$$\left\{\begin{array}{lcll} p|_{z=S} - \left(2\eta\dfrac{\partial u_z}{\partial z}\right)\Big|_{z=S} & = & 0 & \\ p|_{z=B} - \left(2\eta\dfrac{\partial u_z}{\partial z}\right)\Big|_{z=B} & = & \rho_w\,g\,(\ell_w - B) & \text{cas flottant} \\ p|_{z=B} - \left(2\eta\dfrac{\partial u_z}{\partial z}\right)\Big|_{z=B} & = & \rho_i\,g\,H & \text{cas posé} \end{array}\right. \tag{1.41}$$

Intégrons l'équation (1.40) d'une altitude z quelconque à la surface S, la pression octaédrale admet alors pour valeur

$$p = \rho_i\, g\,(S-z) + 2\eta\frac{\partial u_z}{\partial z} \tag{1.42}$$

$$= \rho_i\, g\,(S-z) - 2\eta\frac{\partial u_x}{\partial x} - 2\eta\frac{\partial u_y}{\partial y} \tag{1.43}$$

La dernière ligne est obtenue en utilisant l'équation d'incompressibilité (1.7).

Remarquons que l'utilisation de la seule approximation hydrostatique permet de supprimer la recherche d'une variable (en l'occurence la pression octaédrale p) par rapport au problème Full-Stokes.

1.2.3 Modèle de Blatter-Pattyn

En utilisant simplement l'hypothèse hydrostatique et un rapport d'aspect petit, le problème Full Stokes se simplifie largement. Le modèle obtenu est connu sous le nom de modèle de Blatter-Pattyn car introduit par [Blatter, 1995] puis réécrit en terme de vitesses par [Pattyn, 2003]. Celui-ci approche bien le modèle Full Stokes jusqu'à l'ordre 2 en terme de δ quel que soit le type de frottement rencontré à la base [Schoof et Hindmarsh, 2010]. Nous donnons à présent les principales clés permettant à l'établissement de ce modèle.

En utilisant le fait que le rapport d'aspect de la calotte soit petit, les composantes horizontales de l'équation de Stokes (1.15) se mettent sous la forme

$$\left\{\begin{array}{rcl} \frac{\partial}{\partial x}\left(2\eta\frac{\partial u_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(\eta\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\right) + \frac{\partial}{\partial z}\left(\eta\left(\frac{\partial u_x}{\partial z}+\cancel{\frac{\partial u_z}{\partial x}}\right)\right) - \frac{\partial p}{\partial x} & = & 0 \\ \frac{\partial}{\partial x}\left(\eta\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\right) + \frac{\partial}{\partial y}\left(2\eta\frac{\partial u_y}{\partial y}\right) + \frac{\partial}{\partial z}\left(\eta\left(\frac{\partial u_y}{\partial z}+\cancel{\frac{\partial u_z}{\partial y}}\right)\right) - \frac{\partial p}{\partial y} & = & 0 \end{array}\right. \tag{1.44}$$

En injectant la formule (1.43) pour p dans les équations précédentes, nous obtenons

$$\left\{\begin{array}{rcl} \frac{\partial}{\partial x}\left(2\eta\left(2\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}\right)\right) + \frac{\partial}{\partial y}\left(\eta\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\right) + \frac{\partial}{\partial z}\left(\eta\frac{\partial u_x}{\partial z}\right) & = & \rho_i\, g\frac{\partial S}{\partial x} \\ \frac{\partial}{\partial x}\left(\eta\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\right) + \frac{\partial}{\partial y}\left(2\eta\left(\frac{\partial u_x}{\partial x}+2\frac{\partial u_y}{\partial y}\right)\right) + \frac{\partial}{\partial z}\left(\eta\frac{\partial u_y}{\partial z}\right) & = & \rho_i\, g\frac{\partial S}{\partial y} \end{array}\right. \tag{1.45}$$

Nous n'avons plus que 2 équations elliptiques 3D à 2 inconnues u_x et u_y valables sur toute l'épaisseur de la calotte à résoudre. Dans le cas de la loi polynomiale (1.13) utilisée pour la viscosité, cette dernière est solution de l'équation

$$4\,B_{AT3}\,\dot{\varepsilon}^2\,\eta^3 + B_{AT1}\,\eta - 1 = 0 \tag{1.46}$$

avec

$$\dot{\varepsilon}^2 = \left(\frac{\partial u_x}{\partial x}\right)^2 + \left(\frac{\partial u_y}{\partial y}\right)^2 + \frac{\partial u_x}{\partial x}\frac{\partial u_y}{\partial y} + \frac{1}{4}\left(\frac{\partial u_x}{\partial y}+\frac{\partial u_y}{\partial x}\right) + \frac{1}{4}\left(\frac{\partial u_x}{\partial z}\right)^2 + \frac{1}{4}\left(\frac{\partial u_y}{\partial z}\right)^2 \tag{1.47}$$

Les composantes horizontales des conditions aux limites (1.18), (1.19) et (1.20) sur le tenseur des contraintes se simplifient de la manière suivante :

— surface

$$\left\{\begin{array}{rcl} \left(2\left(2\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}\right)\right)\Big|_{z=S}\frac{\partial S}{\partial x}+\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\Big|_{z=S}\frac{\partial S}{\partial y}-\frac{\partial u_x}{\partial z}\Big|_{z=S} & = & 0 \\ \left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\Big|_{z=S}\frac{\partial S}{\partial x}+\left(2\left(\frac{\partial u_x}{\partial x}+2\frac{\partial u_y}{\partial y}\right)\right)\Big|_{z=S}\frac{\partial S}{\partial y}-\frac{\partial u_y}{\partial z}\Big|_{z=S} & = & 0 \end{array}\right. \tag{1.48}$$

— base flottante

$$\left\{\begin{array}{rcl} \left(2\left(2\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\frac{\partial B}{\partial x}+\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\Big|_{z=B}\frac{\partial B}{\partial y}-\frac{\partial u_x}{\partial z}\Big|_{z=B} & = & 0 \\ \left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\Big|_{z=B}\frac{\partial B}{\partial x}+\left(2\left(\frac{\partial u_x}{\partial x}+2\frac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\frac{\partial B}{\partial y}-\frac{\partial u_y}{\partial z}\Big|_{z=B} & = & 0 \end{array}\right. \tag{1.49}$$

— base posée

$$\left\{\begin{array}{rcl} \left(2\eta\left(2\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\frac{\partial B}{\partial x}+\left(\eta\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\right)\Big|_{z=B}\frac{\partial B}{\partial y}-\left(\eta\frac{\partial u_x}{\partial z}\right)\Big|_{z=B} & = & \tau_{b,x} \\ \left(\eta\left(\frac{\partial u_y}{\partial x}+\frac{\partial u_x}{\partial y}\right)\right)\Big|_{z=B}\frac{\partial B}{\partial x}+\left(2\eta\left(\frac{\partial u_x}{\partial x}+2\frac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\frac{\partial B}{\partial y}-\left(\eta\frac{\partial u_y}{\partial z}\right)\Big|_{z=B} & = & \tau_{b,y} \end{array}\right. \tag{1.50}$$

Le modèle de Blatter-Pattyn se résume finalement de la manière suivante

Modèle 1.2 (*Blatter-Pattyn*)

1. *A chaque pas de temps, calcul des vitesses horizontales u_x et u_y dépendantes de x, y, z sur toute la calotte par résolution du système comprenant*
   — *les équations* (1.45).
   — *la calcul de la viscosité effective (si loi polynomiale, utiliser* (1.46) *et* (1.47)*).*
   — *les conditions limites* (1.48), (1.49) *et* (1.50).
   — *la loi de glissement basal* (1.24).
2. *A chaque pas de temps, calcul des vitesses moyennées sur l'épaisseur $\overline{U}_x$ et $\overline{U}_y$.*
3. *Evolution temporelle gérée par l'équation du bilan de masse* (1.32).

Par rapport au modèle Full Stokes, la pression octaédrale p et la vitesse verticale u_z s'expriment en fonction de u_x et u_y grâce à la relation (1.43) et par l'équation suivante

$$u_z = u_z(z=B) - \int_B^z \left(\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}\right) dz' \tag{1.51}$$

1.2.4 Shallow Shelf – Shelfy Stream Approximation

Cette approximation n'a pas été conçue pour traiter une calotte polaire dans sa globalité mais seulement certaines zones d'intérêt. La shallow shelf approximation (SSA) fut introduite en premier par [Morland, 1987] pour les ice-shelves non confinés puis par [MacAyeal, 1989] pour les fleuves de glace (d'où le nom de *Shelfy Stream Approximation*). En plus des hypothèses classiques déjà énoncées, la SSA fait l'hypothèse suivante

Hypothèses 1.1 (*SSA*)

Les vitesses horizontales ne dépendent pas de l'axe vertical Oz.

$$\frac{\partial u_x}{\partial z} \approx 0 \qquad \frac{\partial u_y}{\partial z} \approx 0 \tag{1.52}$$

Cela veut donc dire que les vitesses horizontales ne dépendent pas de la profondeur contrairement au modèle de Blatter-Pattyn. Cette hypothèse n'est valide que si le frottement basal est faible (cas des fleuves de glace) ou nul (cas des ice shelves). Une dérivation élégante de la SSA par adimensionalisation des équations est disponible dans [MacAyeal, 1989] pour les fleuves de glace et dans [Rommelaere, 1997] pour les plateformes flottantes. Nous donnons ici seulement les points clés de la dérivation de la SSA.

Reprenons l'équation principale (1.45) pour les vitesses horizontales dans le modèle de Blatter-Pattyn. Suite à l'hypothèse effectuée précédemment, nous avons

$$\left\{\begin{array}{rcl} \dfrac{\partial}{\partial x}\left(2\eta\left(2\dfrac{\partial u_x}{\partial x}+\dfrac{\partial u_y}{\partial y}\right)\right)+\dfrac{\partial}{\partial y}\left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right) & = & \rho_i\, g\dfrac{\partial S}{\partial x} \\ \dfrac{\partial}{\partial x}\left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)+\dfrac{\partial}{\partial y}\left(2\eta\left(\dfrac{\partial u_x}{\partial x}+2\dfrac{\partial u_y}{\partial y}\right)\right) & = & \rho_i\, g\dfrac{\partial S}{\partial y} \end{array}\right. \tag{1.53}$$

L'idée de la SSA est d'intégrer ces équations sur l'épaisseur de la calotte car les vitesses horizontales ne dépendent pas de la profondeur. Pour cela, il est nécessaire de connaître les conditions limites à la surface et à la base. Suite à l'hypothèse 1.1, les conditions (1.48), (1.49) et (1.50) se réécrivent de la manière suivante :

— surface

$$\left\{\begin{array}{rcl} \left(2\eta\left(2\dfrac{\partial u_x}{\partial x}+\dfrac{\partial u_y}{\partial y}\right)\right)\Big|_{z=S}\dfrac{\partial S}{\partial x}+\left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)\Big|_{z=S}\dfrac{\partial S}{\partial y} & = & 0 \\ \left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)\Big|_{z=S}\dfrac{\partial S}{\partial x}+\left(2\eta\left(\dfrac{\partial u_x}{\partial x}+2\dfrac{\partial u_y}{\partial y}\right)\right)\Big|_{z=S}\dfrac{\partial S}{\partial y} & = & 0 \end{array}\right. \tag{1.54}$$

— base flottante

$$\left\{\begin{array}{rcl} \left(2\eta\left(2\dfrac{\partial u_x}{\partial x}+\dfrac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial x}+\left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial y} & = & 0 \\ \left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial x}+\left(2\eta\left(\dfrac{\partial u_x}{\partial x}+2\dfrac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial y} & = & 0 \end{array}\right. \tag{1.55}$$

— base posée

$$\left\{\begin{array}{rcl} \left(2\eta\left(2\dfrac{\partial u_x}{\partial x}+\dfrac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial x}+\left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial y} & = & \tau_{b,x} \\ \left(\eta\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial x}+\left(2\eta\left(\dfrac{\partial u_x}{\partial x}+2\dfrac{\partial u_y}{\partial y}\right)\right)\Big|_{z=B}\dfrac{\partial B}{\partial y} & = & \tau_{b,y} \end{array}\right. \tag{1.56}$$

Nous intégrons maintenant (1.53) avec les conditions limites fournies plus haut. En utilisant la viscosité effective moyenne $\overline{\eta}$ définie par

$$\overline{\eta}(x,y,t)=\frac{1}{H}\int_B^S \eta(x,y,z,t)dz \tag{1.57}$$

l'approximation SSA revient à résoudre le système suivant

$$\left\{\begin{array}{rcl} \dfrac{\partial}{\partial x}\left(2H\overline{\eta}\left(2\dfrac{\partial u_x}{\partial x}+\dfrac{\partial u_y}{\partial y}\right)\right)+\dfrac{\partial}{\partial y}\left(H\overline{\eta}\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right) & = & \rho_i\, g\, H\dfrac{\partial S}{\partial x}-\tau_{b,x} \\ \dfrac{\partial}{\partial x}\left(H\overline{\eta}\left(\dfrac{\partial u_y}{\partial x}+\dfrac{\partial u_x}{\partial y}\right)\right)+\dfrac{\partial}{\partial y}\left(2H\overline{\eta}\left(\dfrac{\partial u_x}{\partial x}+2\dfrac{\partial u_y}{\partial y}\right)\right) & = & \rho_i\, g\, H\dfrac{\partial S}{\partial y}-\tau_{b,y} \end{array}\right. \tag{1.58}$$

($\tau_{b,x}=\tau_{b,y}=0$ si la glace flotte) avec une viscosité effective moyenne solution de l'équation

$$4\,B_{AT3}\,\dot{\varepsilon}^2\,\overline{\eta}^3+B_{AT1}\,\overline{\eta}-1=0 \tag{1.59}$$

avec

$$\dot{\varepsilon}^2=\left(\frac{\partial u_x}{\partial x}\right)^2+\left(\frac{\partial u_y}{\partial y}\right)^2+\frac{\partial u_x}{\partial x}\frac{\partial u_y}{\partial y}+\frac{1}{4}\left(\frac{\partial u_x}{\partial y}+\frac{\partial u_y}{\partial x}\right) \tag{1.60}$$

Les relations sur la viscosité effective moyenne sont valables si la calotte est supposée isotherme, sinon il faut faire ce calcul à toutes les profondeurs puis intégrer la viscosité verticalement.

La SSA se résume finalement à

Modèle 1.3 (*SSA*)

1. *A chaque pas de temps, calcul des vitesses horizontales u_x et u_y dépendantes de x, y constantes en profondeur par résolution du système comprenant*
   — *les équations* (1.58).
   — *la calcul de la viscosité effective moyenne.*
   — *les conditions limites* (1.54), (1.55) *et* (1.56).
   — *la loi de glissement basal* (1.24).
2. *A chaque pas de temps, calcul des vitesses moyennées sur l'épaisseur $\overline{U}_x$ et $\overline{U}_y$.*
3. *Evolution temporelle gérée par l'équation du bilan de masse* (1.32).

Par rapport au modèle de Blatter-Pattyn, l'équation principale n'est plus à résoudre pour chaque profondeur mais seulement une fois (cas 2D horizontal au lieu de 3D).

1.2.5 Shallow Ice Approximation

Formalisée par [Hutter, 1983] puis [Morland, 1984] mais utilisée implicitement depuis plus longtemps (voir [Böðvarsson, 1955]), l'approximation Shallow Ice (SIA) est la plus ancienne approximation faite sur l'écoulement d'une calotte polaire. Elle est aussi la plus simple et la moins coûteuse. Cependant l'approximation supplémentaire effectuée est importante.

Hypothèses 1.2 (*SIA*)

Les contraintes de cisaillement vertical (τ_{xz} et τ_{yz}) prédominent sur les autres contraintes dans le déviateur des contraintes.

Grâce à cette hypothèse, il est possible d'évaluer directement les composantes u_x et u_y des vitesses comme nous allons le voir. Une dérivation propre de la SIA par adimensionalisation des équations est disponible dans [Greve, 1997]. Nous donnons ici uniquement les points clés de cette dérivation.

En reprenant l'hypothèse de la SIA, nous avons que

$$\sigma_x = \sigma_y = \sigma_z = -p \tag{1.61}$$

p correspond alors à la pression hydrostatique. L'équation d'équilibre quasi-statique (1.10) se simplifie de la manière suivante

$$\left\{\begin{array}{rcl} \cancel{\frac{\partial \sigma'_x}{\partial x}} + \cancel{\frac{\partial \tau_{xy}}{\partial y}} + \frac{\partial \tau_{xz}}{\partial z} - \frac{\partial p}{\partial x} & = & 0 \\ \cancel{\frac{\partial \tau_{xy}}{\partial x}} + \cancel{\frac{\partial \sigma'_y}{\partial y}} + \frac{\partial \tau_{yz}}{\partial z} - \frac{\partial p}{\partial y} & = & 0 \\ \cancel{\frac{\partial \tau_{xz}}{\partial x}} + \cancel{\frac{\partial \tau_{yz}}{\partial y}} + \cancel{\frac{\partial \sigma'_z}{\partial z}} - \frac{\partial p}{\partial z} & = & \rho_i\, g \end{array}\right. \tag{1.62}$$

D'autre part, considérons la condition limite à la surface (1.18). Sous l'hypothèse 1.2, cette condition limite se réécrit

$$\left\{\begin{array}{rcl} \tau_{xz} + p\frac{\partial S}{\partial x}\Big|_{z=S} & = & 0 \\ \tau_{yz} + p\frac{\partial S}{\partial y}\Big|_{z=S} & = & 0 \\ \tau_{xz}\frac{\partial S}{\partial x} + \tau_{yz}\frac{\partial S}{\partial y} + p\Big|_{z=S} & = & 0 \end{array}\right. \quad \Longrightarrow \quad \left\{\begin{array}{rcl} p\,|_{z=S} & = & 0 \\ \tau_{xz}\,|_{z=S} & = & 0 \\ \tau_{yz}\,|_{z=S} & = & 0 \end{array}\right. \tag{1.63}$$

En intégrant (1.62) sur l'épaisseur de la calotte et en utilisant les conditions limites (1.63), nous obtenons

$$\left\{\begin{array}{rcl} p & = & \rho_i\, g\,(S-z) \\ \tau_{xz} & = & -\rho_i\, g\frac{\partial S}{\partial x}\,(S-z) \\ \tau_{yz} & = & -\rho_i\, g\frac{\partial S}{\partial y}\,(S-z) \end{array}\right. \tag{1.64}$$

Le carré de la cission efficace définie en (1.5) vaut alors

$$\tau^2 = \tau_{xz}^2 + \tau_{yz}^2 = (\rho_i\, g)^2\, (S - z)^2\, (\nabla S \cdot \nabla S) \tag{1.65}$$

Il nous reste encore à évaluer les composantes horizontales des vitesses. En combinant (1.37), (1.65), (1.11) et (1.13), nous avons pour $i = x, y$

$$\frac{\partial u_i}{\partial z} = -\Big(B_{AT1}\, \rho_i\, g\, (S - z) + B_{AT3}\, (\rho_i\, g)^3\, (S - z)^3\, (\nabla S \cdot \nabla S)\Big)\, \frac{\partial S}{\partial i} \tag{1.66}$$

En intégrant l'égalité précédente entre le fond de la calotte B et une altitude z, les vitesses horizontales valent

$$u_i(z) = u_{B,i} - \Big[\rho_i\, g \int_B^z B_{AT1}\, (S - z')\, dz' + (\rho_i\, g)^3\, (\nabla S \cdot \nabla S) \int_B^z B_{AT3}\, (S - z')^3\, dz'\Big]\, \frac{\partial S}{\partial i} \tag{1.67}$$

et les vitesses horizontales moyennées sur l'épaisseur

$$\overline{U}_i = u_{B,i} - \Big[\rho_i\, g \int_B^S \int_B^z B_{AT1}\, (S - z')\, dz'\, dz + (\rho_i\, g)^3\, (\nabla S \cdot \nabla S) \int_B^S \int_B^z B_{AT3}\, (S - z')^3\, dz'\, dz\Big]\, \frac{\partial S}{\partial i} \tag{1.68}$$

Suite à ces résultats, un modèle SIA se résume de la manière suivante

Modèle 1.4 (*SSA*)

1. *Calcul des vitesses moyennées sur l'épaisseur* $\overline{U}_x$ *et* $\overline{U}_y$ *par la formule* (1.68).
2. *Evolution temporelle gérée par l'équation du bilan de masse* (1.32).

Contrairement à l'approche SSA et si les coefficients B_{AT1} et B_{AT3} sont supposés constants (cas calotte isotherme), les vitesses horizontales et les vitesses horizontales moyennées admettent une expression analytique. Notons enfin qu'il faut fournir pour le modèle SIA les vitesses de glissement à la base.

Quelques remarques pour finir sur la SIA, la partie déformation de l'équation des vitesses est locale et ne dépend que de H et de la pente de la surface. D'autre part, la direction de l'écoulement ne dépend pas de la profondeur et suit la plus grande pente à la surface, cela justifie l'usage de modèle en ligne d'écoulement SIA comme le modèle d'évolution simplifié utilisé durant cette thèse.

1.2.6 Autres modèles

Il existe d'autres modèles basés sur les approximations classiques du rapport d'aspect petit et de l'hypothèse hydrostatique sur l'axe vertical (voir [Hindmarsh, 2004]). Ce sont des modèles d'ordre supérieur ou *Higher order models* (HOM). Le plus connu de ses modèles est celui de Blatter-Pattyn décrit plus tôt. Signalons néanmoins l'existence

d'un autre modèle HOM proposé par [Schoof et Hindmarsh, 2010]. Celui-ci propose de résoudre une équation proche de la SSA dans le but de calculer les vitesses horizontales à la base puis de calculer les vitesses horizontales grâce à une formule établie dépendant de la vitesse à la base (voir aussi [Cornford *et al.*, 2013] pour une implémentation de ce modèle). Ce modèle représente une alternative à coût moindre au modèle de Blatter-Pattyn (plus de résolution d'équations sur la calotte entière).

Les modèles dits hybrides SIA+SSA représente une autre alternative intéressante (voir notamment [Bueler et Brown, 2009]). L'idée est d'utiliser l'équation de la SSA comme loi de glissement à la base puis d'utiliser la vitesse à la base trouvée dans la formule établie pour le calcul des vitesses moyennes comme dans la SIA. Le coût numérique de ces modèles est alors légèrement supérieur à celui de la SSA.

1.3 Un modèle d'évolution simplifié : Winnie

Nous présentons dans cette section le modèle simplifié que nous avons utilisé tout au long de cette thèse pour nos expériences d'assimilation. Ce modèle appelé Winnie est une version ligne d'écoulement isotherme du modèle d'évolution 3D GRISLI (lire à ce sujet [Ritz *et al.*, 2001], [Peyaud *et al.*, 2007] et [Quiquet *et al.*, 2012]). Nous considérons ici la version de Winnie utilisant l'approximation SIA avec une loi de glissement donnée à la base. Nous listons ici les hypothèses effectuées.

Hypothèses 1.3 (*Winnie SIA*)

1. *Approximation SIA.*
2. *Le flux de glace suit une ligne d'écoulement.*
3. *La calotte n'a pas de plateforme flottante.*
4. *La calotte est isotherme.*
5. *La topographie du socle rocheux est fixe.*
6. *Les vitesses de glissement à la base sont données par une loi de frottement linéaire avec un coefficient dépendant de l'espace.*
7. *Le bilan de masse à la base est nul.*
8. *Le bilan de masse à la surface est une fonction de la température de la surface.*
9. *La température de la surface est une fonction dépendant de paramètres climatiques et de l'altitude de la surface de la calotte.*

Bien que la liste des hypothèses soit longue, Winnie contient néanmoins les points principaux de la physique d'une calotte à grande échelle et reste encore réaliste. En effet, selon l'approximation SIA, le flux de glace a pour direction la pente à la surface la plus abrupte, il peut alors être bien représenté par une ligne d'écoulement dans cette direction. D'autre part, la vitesse de glissement reste dépendante de la géométrie de la calotte comme dans les modèles plus sophistiqués. Elle est également fonction des propriétés de l'interface glace-socle via son coefficient et nous cherchons justement à développer une méthodologie efficace afin de cartographier ce coefficient. Nous décrivons dans la suite de cette section

les équations régissant l'évolution du modèle suite aux hypothèses précédentes ainsi que son implémentation pratique.

1.3.1 Équation du bilan de masse et vitesses de bilan

L'évolution temporelle de la calotte est régie par l'équation du bilan de masse (eq. (1.32)) ici simplifiée pour une ligne d'écoulement. Elle vaut

$$\frac{\partial H}{\partial t} = \dot{b}_m - \frac{\partial (H\,\overline{U}_x)}{\partial x} \tag{1.69}$$

Notons que nous supposons que le bilan de masse à la base $\dot{f}$ est nul. Nous supposons qu'à chaque extrémité de l'axe x, l'épaisseur de la calotte est nulle (conditions limites de Dirichlet).

Pour pouvoir résoudre l'équation précédente, il faut pouvoir calculer la vitesse horizontale moyenne $\overline{U}_x$. Nous supposons la calotte isotherme, l'équation (1.68) pour les vitesses en SIA se simplifie grandement. Ainsi, nous avons

$$\overline{U}_x = u_{b,x} - B_{AT3}\frac{(\rho_i\,g)^3}{5}H^4\left(\frac{\partial S}{\partial x}\right)^3 - B_{AT1}\frac{\rho_i\,g}{3}H^2\,\frac{\partial S}{\partial x} \tag{1.70}$$

avec $u_{b,x}$ vitesse de glissement à la base et $S = H + B_{soc}$ l'altitude de surface (pas de plateformes flottantes considérées ici). Nous supposons aussi que la topographie du socle rocheux B_{soc} est fixe. Il s'agit d'un paramètre du modèle (que nous pourrons rechercher par assimilation de données).

1.3.2 Modélisation de la vitesse à la base

La vitesse de glissement à la base est donnée par une loi de frottement linéaire où le frottement à la base τ_b est donné par la SIA

$$\tau_b = \rho_i\,g\,H\,\frac{\partial S}{\partial x} = -\beta\,u_{b,x} \tag{1.71}$$

avec β coefficient de frottement basal positif. Comme ce paramètre peut prendre des valeurs de différentes magnitudes, nous préférons plutôt utiliser $\alpha = \log_{10}(\beta)$ (que ce soit pour la modélisation ou pour des expériences d'assimilation).

1.3.3 Calcul du bilan de masse et de la température en surface

Le bilan de masse $\dot{b}_m$ est la somme de l'accumulation $Acc > 0$ (due aux précipitations de neige) et de l'ablation $Abl < 0$ (due à la fonte) de glace en surface. L'accumulation est calculée à l'aide de la paramétrisation suivante

$$Acc = Acc_0\,e^{c_1 T_s} \tag{1.72}$$

L'hypothèse sous-jacente est que la précipitation de neige est gouvernée par la quantité d'eau dans l'air et donc par la pression de vapeur saturante qui est une exponentielle de la température en surface T_s. Pour l'ablation, nous avons

$$Abl = \begin{cases} Abl_0 \left(\dfrac{T_s - T_{\text{no melt}}}{T_{\text{no melt}}} \right)^2 & \text{si } T_s > T_{\text{no melt}} \\ 0 & \text{sinon} \end{cases} \tag{1.73}$$

Dans nos expériences, $T_{\text{no melt}}$ vaut $-6°\,C$, il s'agit de la moyenne annuelle. Notons aussi que cette modélisation de l'ablation préserve la non-linéarité de l'ablation face à la température en surface. Pour voir comment évolue le bilan de masse en surface en fonction de T_s, nous renvoyons à la figure 1.3.

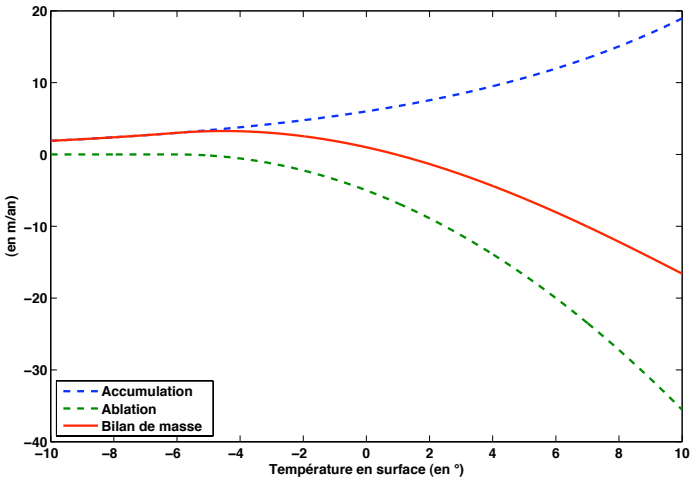

FIGURE 1.3 – Évolution de l'accumulation, de l'ablation et du bilan de masse en fonction de la température en surface dans Winnie. Ici, le bilan de masse est positif pour $T_s < 1°$ et négatif sinon. Le bilan de masse maximum vaut autour de 3,3 m/an pour T_s autour de $-4,5°$.

La température en surface T_s dépend de l'altitude de la surface de la calotte S, de la distance selon l'axe x (représentant la latitude, l'origine est au nord) et d'un terme de forçage climatique F_{clim} potentiellement à rechercher. Ainsi, T_s suit la formule suivante

$$T_s(x,t) = F_{\text{clim}}(t) + \lambda\, x + \gamma\, S(x,t) \tag{1.74}$$

avec λ gradient latitudinal ($1°\,C$ par degré de latitude) et γ gradient vertical. Dans toutes les expériences effectuées durant cette thèse, les paramètres λ et γ valent

$$\lambda = \frac{1}{111000}^{\circ}\, C.m^{-1} \qquad \gamma = -0,0063°\, C.m^{-1} \tag{1.75}$$

Cette paramétrisation de la température en surface prend en compte la rétroaction altitude - T_s fondamentale dans l'évolution des calottes telles que le Groenland ou les calottes passées (Laurentides, Fennoscandie).

1.3.4 Discrétisation et résolution numérique

Nous avons maintenant tous les éléments en main afin de simuler l'évolution d'une calotte le long d'une ligne d'écoulement à l'aide de Winnie. Il ne nous reste plus qu'à détailler certaines phases de l'algorithme de Winnie.

Discrétisation par différences finies

Présentons tout d'abord la discrétisation utilisée dans Winnie. Winnie est un modèle utilisant les différences finies pour la discrétisation en temps et en espace. Notons Δt le pas de temps de la résolution numérique, Δx le pas d'espace de la grille régulière employée et L la longueur maximale de l'axe x. La grille est constituée de n_x nœuds majeurs notés x_i et de $n_x - 1$ nœuds mineurs notés $x_{i+\frac{1}{2}}$ définis tels que

$$x_i = \frac{iL}{\Delta x} \qquad i = 0, \ldots, n_x - 1 \qquad \text{et} \qquad x_{i+\frac{1}{2}} = \frac{(i+\frac{1}{2})L}{\Delta x} \qquad i = 0, \ldots, n_x - 2 \tag{1.76}$$

Sur les nœuds majeurs sont calculés l'épaisseur de la calotte, l'altitude de la surface, la température en surface et le bilan de masse. Sur les nœuds mineurs sont calculées la vitesse à la base et la vitesse de bilan.

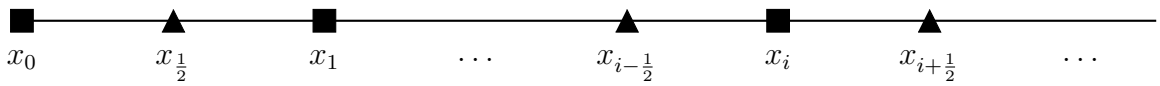

FIGURE 1.4 – Discrétisation de l'axe x employée dans Winnie. Les carrés désignent les nœuds majeurs où sont calculés H, S, T_s et le bilan de masse. Les triangles désignent les nœuds mineurs où sont calculées les vitesses de bilan et à la base.

Fonctionnement synthétique de Winnie

Nous résumons de manière rapide le fonctionnement de Winnie à chaque pas de temps.

Algorithme 1.1 (*Winnie*)

A chaque pas de temps $t_k = t_{k-1} + \Delta t$

1. *Calcul de la température en surface T_s avec l'altitude de la surface $S(t_{k-1})$ sur les nœuds majeurs.*
2. *Calcul du bilan de masse sur les nœuds majeurs.*
3. *Calcul des vitesses sur les nœuds mineurs utilisant $S(t_{k-1})$ et $H(t_{k-1})$.*
4. *Résolution numérique de l'équation du bilan de masse pour obtenir $H(t_k)$.*
5. *Si certaines nouvelles épaisseurs sont négatives, les remettre à 0.*
6. *Calcul de l'altitude de la surface $S(t_k)$.*

Calcul numérique des vitesses

Contrairement aux autres variables, les vitesses de base et de bilan se calculent sur les nœuds mineurs. Or les équations (1.70) et (1.71) permettant de les calculer utilisent l'épaisseur H disponible uniquement sur les nœuds majeurs. Le calcul de la pente à la surface aux nœuds mineurs est lui naturel. Finalement au nœud mineur $x_{i+\frac{1}{2}}$, l'épaisseur de la calotte et la pente de la surface sont approchées par les quantités suivantes

$$H_{i+\frac{1}{2}} \approx \frac{H_i + H_{i+1}}{2} \qquad \left.\frac{\partial S}{\partial x}\right|_{i+\frac{1}{2}} \approx \frac{S_{i+1} - S_i}{\Delta x} \tag{1.77}$$

Résolution numérique de l'équation du bilan de masse

Le schéma numérique employé pour la résolution numérique de l'équation (1.69) est un schéma amont semi-implicite de telle sorte que les épaisseurs H_i^k aux points x_i et au temps t_k suivent les relations suivantes

$$\begin{cases} \dfrac{H_i^k - H_i^{k-1}}{\Delta t} + \dfrac{1}{\Delta x}\left[\left(c^+_{i+\frac{1}{2}} H_i^k - c^-_{i+\frac{1}{2}} H_{i+1}^k\right) - \left(c^+_{i-\frac{1}{2}} H_{i-1}^k - c^-_{i-\frac{1}{2}} H_i^k\right)\right] = \dot{b}_{m,i} \\ H_0^k = 0 \\ H_{n_x-1}^k = 0 \end{cases} \tag{1.78}$$

avec

$$c^+_{i+\frac{1}{2}} = \max\left(\overline{U}_{x,i+\frac{1}{2}}, 0\right) \qquad c^-_{i+\frac{1}{2}} = \max\left(-\overline{U}_{x,i+\frac{1}{2}}, 0\right) \tag{1.79}$$

Le schéma utilisé est ainsi explicite pour les vitesses mais implicite pour l'épaisseur. Enfin notons que ce schéma n'est pas inconditionnellement stable du fait de la non-linéarité de D coefficient de diffusion défini de la manière suivante

$$D = B_{AT3}\frac{(\rho_i\, g)^3}{5} H^5 \left(\frac{\partial S}{\partial x}\right)^2 + B_{AT1}\frac{\rho_i\, g}{3} H^2 + \frac{\rho_i\, g}{\beta} H^2 \tag{1.80}$$

Le critère de stabilité du schéma employé est de la forme

$$\Delta t < \frac{(\Delta x)^2}{aD} \tag{1.81}$$

avec a un réel fixé. Une démonstration de ce critère de stabilité est disponible dans [Hindmarsh, 2001] pour des discrétisations proches de celle utilisée ici.

1.4 Les observations disponibles

L'objectif de cette section est de montrer quels types d'observations sur les calottes polaires sont disponibles et avec quelle précision. Cependant, les applications détaillées par la suite dans les chapitres 3 et 4 sont des expériences synthétiques. C'est pourquoi nous nous contentons d'un survol rapide des vraies données accessibles.

1.4.1 Observation de l'état actuel des calottes polaires

Les calottes polaires antarctiques et groenlandaises sont des objets de grandes dimensions soumis à des conditions climatiques difficiles limitant les observations de terrain. Comme nous le verrons par la suite, la plupart des observations disponibles à l'échelle de la calotte a pour origine les satellites.

Altitude de la surface

L'un des enjeux majeurs de l'observation des calottes polaires est la production de modèles numériques de terrain (DEM pour *Digital elevation models*) consistant à cartographier au mieux la topographie de la surface de la calotte. Ces DEM sont pour la plupart constituées de données obtenues par altimétrie radar satellitaire (ERS-1, Envisat, SPOT, ...). Ces données peuvent être combinées avec des mesures satellites faites par altimétrie laser (ICESat) ou par photoclinométrie (AVHRR, voir [Scambos et Haran, 2002]) par exemple. Les résolutions des derniers DEM sont de l'ordre du kilomètre. Ces données étant cruciales pour la modélisation des calottes polaires, un effort est fait au niveau international pour produire des bases de données incluant les DEM. Nous renvoyons à [Bamber *et al.*, 2009] (voir figure 1.5) ou [Fretwell *et al.*, 2013] pour des exemples de DEM en Antarctique et à [Bamber *et al.*, 2013] pour un exemple de DEM au Groenland. Les utilisations des DEM sont nombreuses, citons par exemple [Flament *et al.*, 2013] évaluant à l'aide de DEM la masse d'eau drainée à partir d'un lac sous-glaciaire.

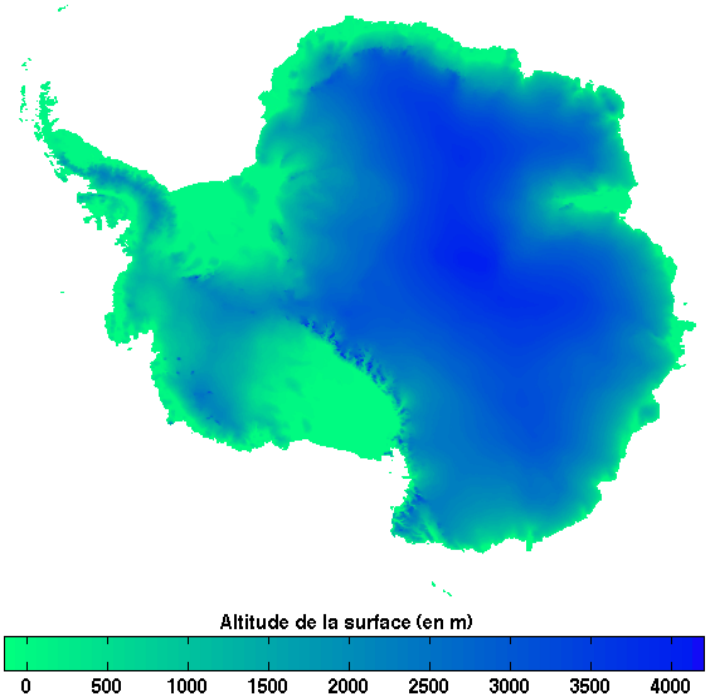

FIGURE 1.5 – Cartographie de la surface dérivée du DEM de [Bamber *et al.*, 2009] agrégeant des mesures par altimétrie radar (ERS-1) et des mesures par altimétrie laser (ICESat). La résolution de ce DEM est de 1 km.

Les DEM peuvent être validés par comparaison avec des données radar acquises par survol de la surface en avion. Les résultats estiment en général l'erreur commise moyenne autour de 3 m (voir figure 1.6 pour une carte d'erreur estimée du DEM de [Bamber *et al.*, 2009]).

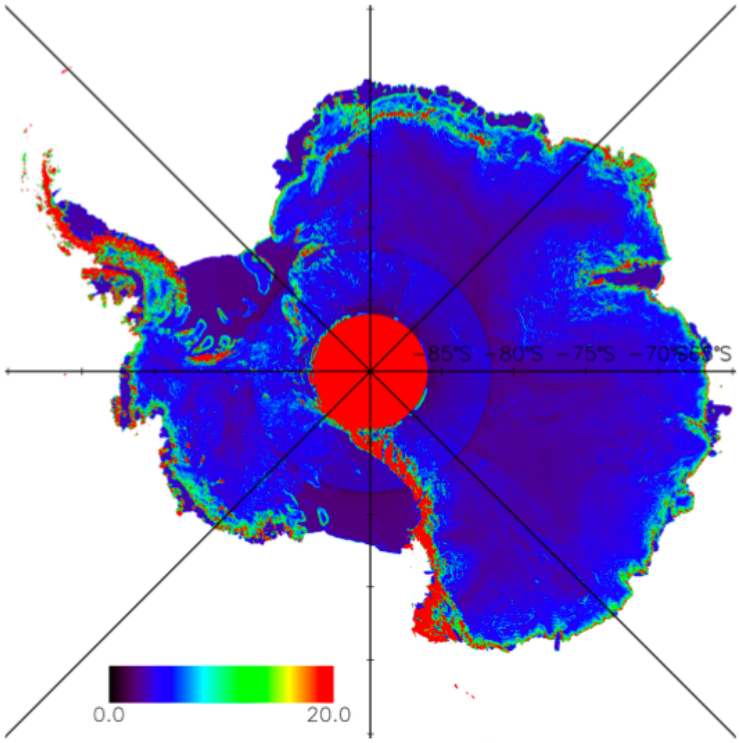

FIGURE 1.6 – Carte d'estimation des erreurs RMS du DEM de [Bamber *et al.*, 2009] obtenue par régression basée sur des données radar par avion extraite de [Griggs et Bamber, 2009]. La gamme d'erreur est de l'ordre de 1 m pour les ice-shelves et entre 2 et 6 m majoritairement pour la glace posée. L'erreur RMS peut cependant atteindre 130 m dans les régions montagneuses. Néanmoins, plus de 98 % de la surface couverte par les données satellite a une erreur estimée inférieure à 20 m.

Vitesses de surface

Les observations dédiées à la dynamique de la glace et à sa traduction en surface se développent à partir des années 1990 grâce notamment à l'apport d'images optiques mais surtout de l'imagerie radar (interférométrie). Ainsi des jeux de données d'observations de vitesses de surface ont pu au fil du temps être constitués à l'échelle de la calotte polaire (voir [Rignot *et al.*, 2011] pour l'Antarctique, cf. figure 1.7 et [Joughin *et al.*, 2010] pour le Groenland). Pour l'instant, les cartes de vitesses proposées sont des mosaïques, toutes les régions n'ayant pas forcément été mesurées au même moment. Cependant, cet aspect s'améliore et des instantanés de champs de vitesses à diverses époques commencent à être disponible (voir [Joughin *et al.*, 2010]). La résolution de ces jeux de données est de l'ordre du kilomètre et l'erreur estimée des données acquises est généralement inférieure

à 10 m.an^{-1}. En comparant l'amplitude de la vitesse et l'incertitude, nous voyons que l'information est surtout pertinente vers les côtes.

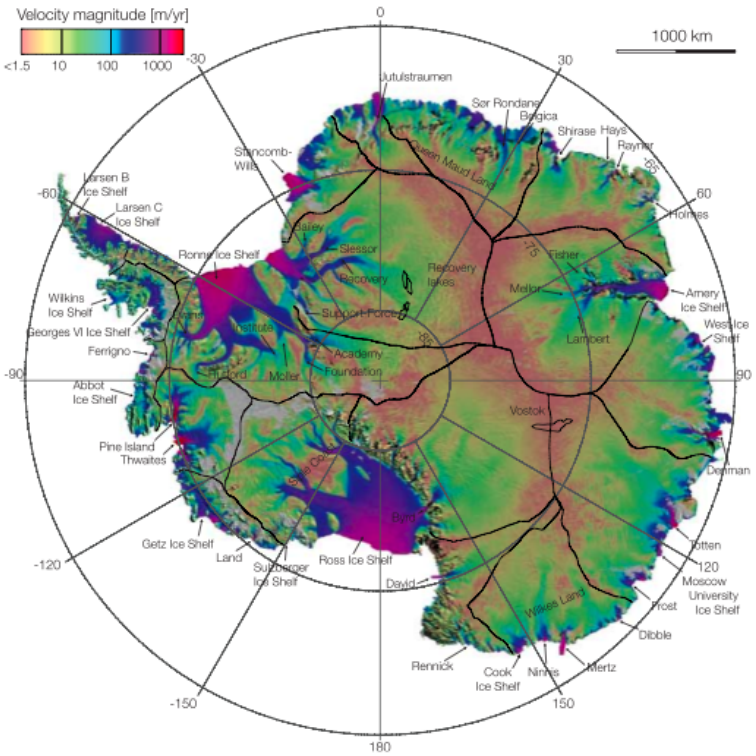

FIGURE 1.7 – Carte de vitesses de surface issue de [Rignot *et al.*, 2011] obtenue par interférométrie radar InSAR à l'aide des satellites ALOS PALSAR, RADARSAT-1/2, Envisat ASAR et ERS-1/2. Les lignes noires épaisses délimitent les principaux bassins versants, alors que les plus fines représentent le contour des lacs sous-glaciaires majeurs. La gamme d'erreur commise sur les vitesses de surface est estimée entre 1 et 17 m.an^{-1}. En comparant l'amplitude de la vitesse et l'incertitude, nous voyons que l'information est surtout pertinente vers les côtes.

Variations d'épaisseur ou d'altitude de surface

Les campagnes répétées de mesures d'altitude de surface acquises notamment par les satellites ICESat et Envisat durant les années 2000 ont permis de souligner les tendances temporelles des observations de la surface et des jeux de variations d'épaisseur ou d'élévation de surface dS/dt ont pu être constitués (voir [Pritchard *et al.*, 2009] pour le Groenland et l'Antarctique, données ICESat et [Flament et Rémy, 2012] pour l'Antarctique, données Envisat, cf. figure 1.8). Notons aussi que [Flament et Rémy, 2012] sont parvenus à construire une carte d'accélération de l'élevation d^2S/dt^2 pour l'Antarctique de l'Ouest qui subit actuellement des modifications dynamiques importantes dans le secteur de la mer d'Amundsen.

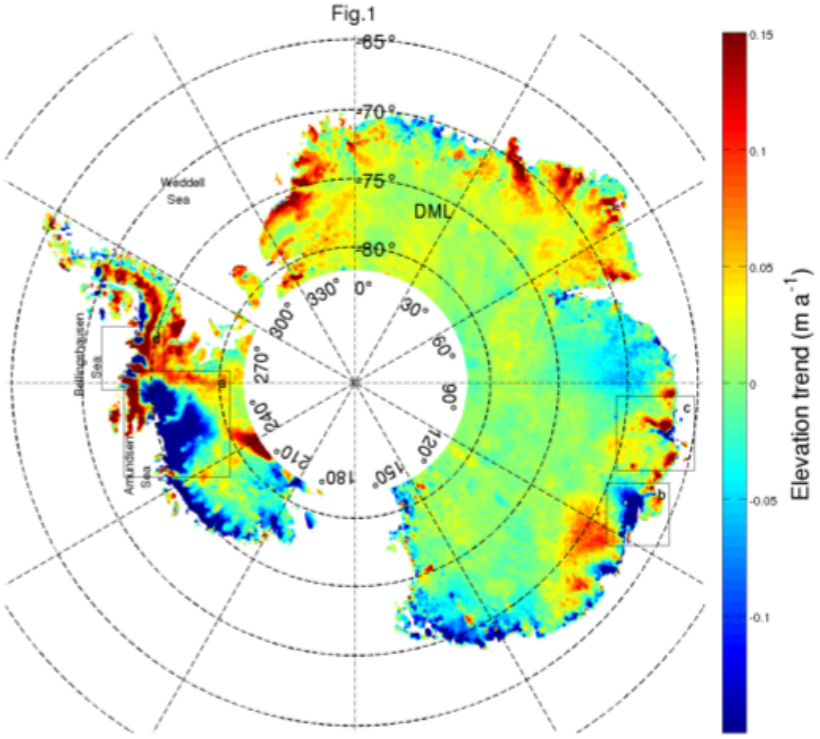

FIGURE 1.8 – Carte de changement d'élévation de surface dH/dt par altimétrie radar obtenue avec le satellite ENVISAT issue de [Flament et Rémy, 2012]. L'amincissement de la calotte est prépondérant en périphérie de la calotte.

Épaisseur et interface socle-calotte

Jusqu'à présent, nous avons abordé les observations de la surface de la calotte grâce aux capteurs radar et laser embarqués par satellite. Ceux-ci ont pour inconvénient de ne pouvoir traiter que la surface ou la sub-surface d'une calotte. Il est possible de pénétrer plus profondément dans la couche de glace en utilisant des radars à plus basse fréquence. Cependant ces derniers ne sont pas embarqués dans les satellites mais placés à bord d'avions chargés de survoler les zones choisies d'où une limitation des zones étudiées.

Ces campagnes d'écho radar ont notamment permis aux glaciologues de localiser plus précisement l'interface socle-calotte aux endroits mesurés. Cependant elles ne couvrent pas entièrement les calottes (voir figure 1.9 pour une carte de couverture des données disponibles). Afin de fournir des DEM de socle rocheux, ces données ont été combinées notamment avec des DEM d'altitude de surface par krigeage mais sans considération pour les flux de glace (Bedmap2 pour l'Antarctique [Fretwell *et al.*, 2013], cf. figure 1.10 et [Bamber *et al.*, 2013] pour le Groenland).

Les erreurs commises estimées sont de l'ordre de 20 m aux points de mesure radar jusqu'à plusieurs centaines de mètres pour les endroits les plus éloignés des points de mesure (voir 1.11 pour la carte d'incertitude dans le cas Bedmap2).

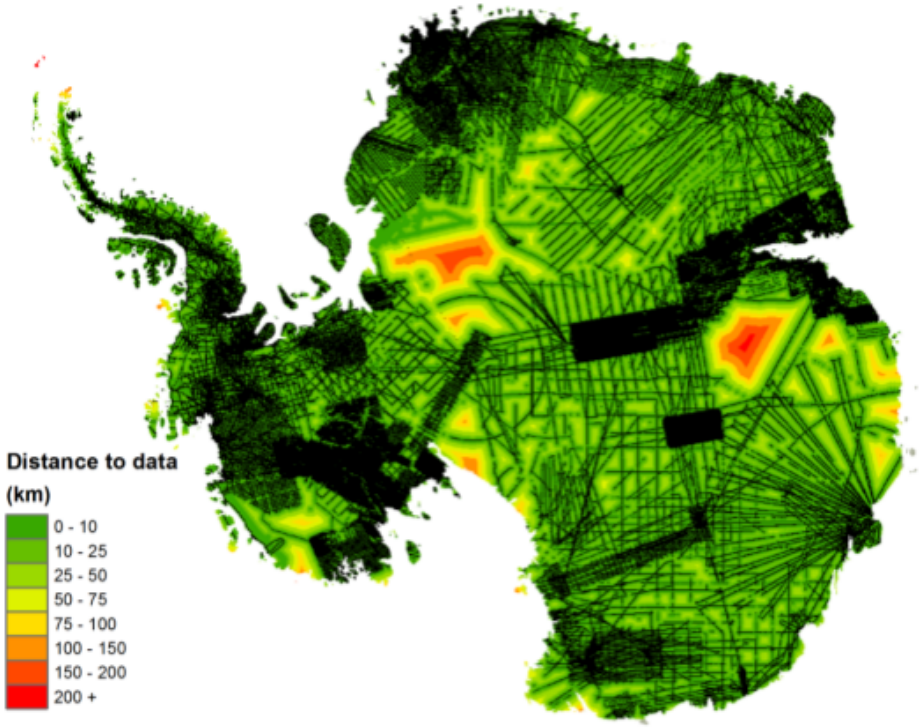

FIGURE 1.9 – Couverture des données de localisaton de l'interface socle-calotte utilisées dans le projet Bedmap2 [Fretwell *et al.*, 2013]. Les traits noirs indiquent les campagnes de survol et les zones en rouge les régions les plus éloignées des campagnes de mesure.

Une autre quantité d'intérêt pour la modélisation des calottes polaires est la mesure du glissement à la base. Cependant, les observations pour ces conditions sont très limitées. En effet, il est quasiment impossible d'aller faire des mesures sous la glace, sauf dans de rares points de forages profonds où le socle a été atteint. D'autre part, la nature du socle n'est pas connue : s'agit-il de sédiments, de roche, de débris de roche ? La présence d'eau liquide, déterminante pour le glissement, est aussi liée à des paramètres mal connus comme le flux géothermique. Nous ne disposons donc d'aucune mesure pour la loi de frottement et les coefficients associés.

1.4.2 États passés

Observer les calottes polaires ne permet pas uniquement d'avoir plus d'informations sur leur présent mais aussi sur leur passé. Différentes données sont disponibles afin de mieux appréhender leur histoire.

Isochrones

Les campagnes de survol par avions ne permettent pas seulement de localiser où se trouve le socle rocheux sous la calotte. Ils permettent aussi de détecter les strates internes de la calotte [Fujita *et al.*, 1999]. Considérées comme des *isochrones*, elles représentent les anciennes surfaces de la calotte. Leurs mesures donnent potentiellement des informations

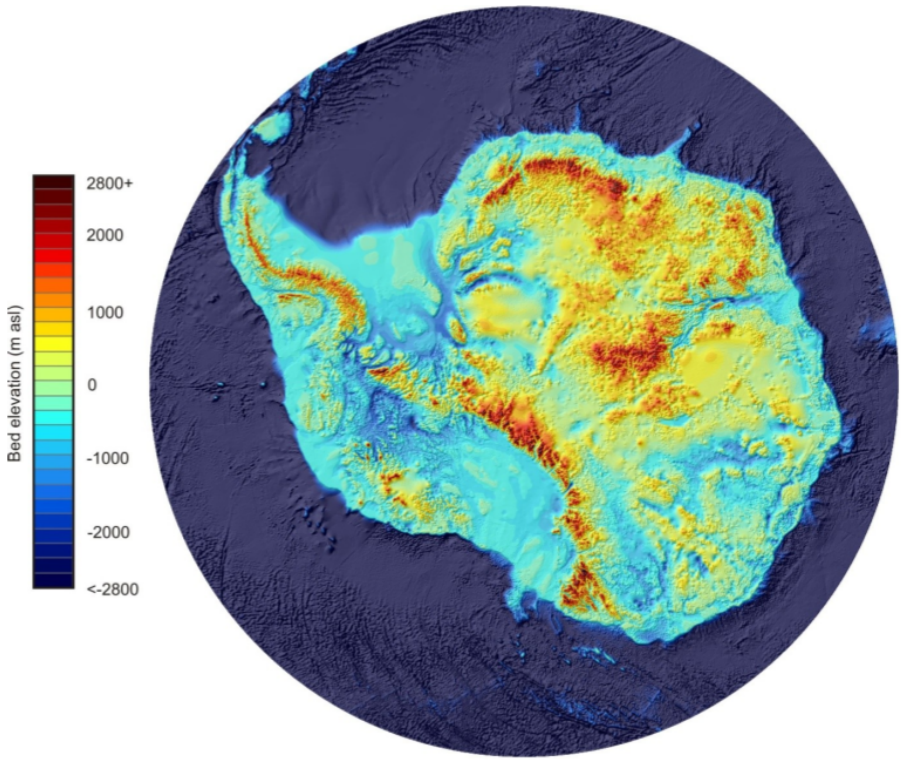

FIGURE 1.10 – Carte de la topographie du socle rocheux antarctique issue du projet Bedmap2 [Fretwell *et al.*, 2013] obtenue par l'utilisation de campagnes d'écho radar et de techniques de krigeage. Le relief de l'Antarctique est très chahuté avec des changements brutaux d'altitude et des chaînes de montagnes. Une carte de l'incertitude commise est disponible figure 1.11.

quant à l'écoulement des calottes dans le passé (consulter à ce sujet [Parrenin *et al.*, 2006] et [Parrenin et Hindmarsh, 2007]).

Carottes de glace

L'étude des carottes de glace constitue un pan entier de la glaciologie. Il serait présomptueux de résumer toutes les informations contenues dans une carotte de glace ici. Nous renvoyons à [Fisher, 2010] et [Cuffey et Paterson, 2010] pour de plus amples informations à ce sujet. Cependant notons que les carottes de glace constituent une grande source d'information sur le climat passé grâce notamment à l'identification de marqueurs temporels (éruptions volcaniques par exemple) et à la mesure de la concentration de certains gaz piégés dans des bulles d'air dans la glace.

Volume total des calottes polaires

Grâce à l'étude de carottes de sédiments prélevées au fond des océans et de leurs marqueurs biostratigraphiques (ex. : foraminifères benthiques), il est possible de reconstruire le volume total de glace des calottes polaires sur de gigantesques échelles de temps (plusieurs centaines de milliers d'années) (voir [Paillard et Parrenin, 2004]).

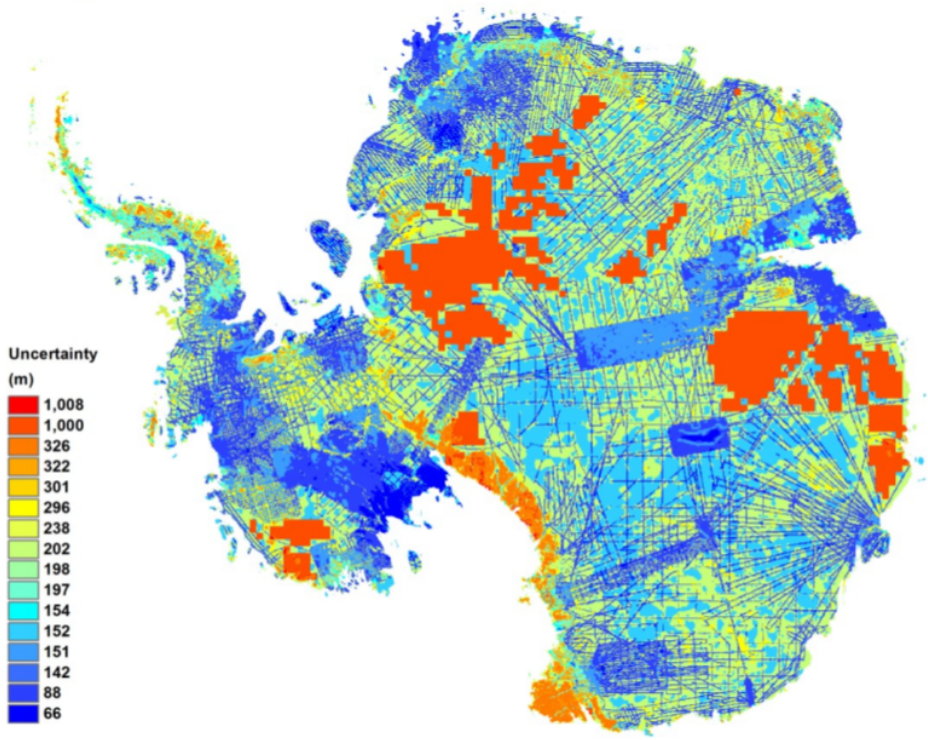

FIGURE 1.11 – Carte de l'incertitude estimée sur la topographie du socle rocheux antarctique issue du projet Bedmap2 [Fretwell *et al.*, 2013]. Les traits bleu foncé représentent les campagnes de survol effectuées (incertitude estimée de 20 m), les zones en orange-rouge indiquent les régions les plus éloignées de ces campagnes. L'erreur commise dans ces zones est grande (estimation de l'ordre de 1000 m). Autant dire qu'à ces endroits, nous ne savons pas où se trouve l'interface socle-calotte.

Données géomorphologiques

Les calottes disparues (Fennoscandie et Laurentides) ont laissé des traces de leur passage (moraines par exemple). Les différentes données géomorphologiques datées correctement peuvent permettre de reconstituer une partie de l'histoire de ces calottes (voir notamment [Kleman *et al.*, 1997], cf. figure 1.12).

1.5 Assimilation de données en glaciologie

Nous venons de voir que les glaciologues disposent d'un certain nombre d'observations sur les calottes polaires. D'autre part, des modèles perfectionnés permettant la simulation de leur évolution existent. Cependant, ces modèles dépendent de paramètres mal (voire pas) connus comme le frottement basal, la topographie du socle rocheux ou le forçage climatique. Les techniques d'assimilation de données peuvent permettre de mieux identifier ces paramètres. Dans la suite de cette section, nous donnons un aperçu non-exhaustif de l'utilisation de l'assimilation de données pour les calottes polaires. Nous insistons particulièrement sur les problèmes d'identification du frottement basal, de la topographie du

socle rocheux ou de paramètres climatiques. Les méthodes d'assimilation classiques ne sont pas décrites ici, elles font l'objet du chapitre 2.

1.5.1 Initialisation de modèle

L'un des problèmes majeurs rencontrés par les modélisateurs de calotte polaire est comment initialiser correctement leurs modèles afin d'obtenir les meilleures simulations d'évolution possibles. Pour cela, ils ont besoin d'identifier correctement le frottement basal et la topographie du socle rocheux dans leurs zones d'étude. Les travaux en assimilation de données sur ce sujet se décomposent en 2 cas : identifier uniquement le frottement basal ou le rechercher conjointement avec la topographie du socle. L'objectif d'une partie de cette thèse (chapitre 4) se place clairement dans cette seconde optique.

Identification du frottement basal

Plusieurs méthodes ont été employées pour identifier le frottement basal. Elles utilisent toutes comme observations les vitesses de la glace à la surface. Cependant, aucun des travaux cités par la suite ne prend en compte la dépendance temporelle des observations.

La première méthode fut introduite par [MacAyeal, 1992] et [MacAyeal, 1993]. Il s'agit d'une méthode adjointe (appelée méthode de contrôle en glaciologie) afin d'identifier le coefficient de frottement β (voir eq. (1.24)) à la base d'un fleuve de glace modélisée à l'aide de la seule équation de la SSA (eq. (1.58)). Son utilisation est facilitée par le caractère auto-adjoint de l'équation employée ce qui explique en partie son relatif succès dans la communauté. Cette approche a été ensuite appliquée à divers fleuves de glace ou bassins versants (consulter notamment [Rommelaere et MacAyeal, 1997] et [Vieli et Payne, 2003]). Cette méthode a aussi été appliquée de manière approchée pour un modèle Full-Stokes par [Morlighem *et al.*, 2010]. Ce dernier utilise le modèle comme approximation de l'adjoint car les équations de Stokes résolues classiquement à l'aide d'une itération de Picard ne sont auto-adjointes que dans le cas où la viscosité est linéaire (loi de Glen $n = 1$, cf. eq. (1.12)). Cette méthode a aussi été employée par [Gillet-Chaulet *et al.*, 2012] sur la calotte groenlandaise dans sa globalité.

Il est aussi possible de construire un adjoint complet à partir de modèles simplifiés. Ainsi [Heimbach et Bugnion, 2009] développe l'adjoint du modèle couplé SIA (pour les parties posées) – SSA (pour les parties flottantes) appelé SICOPOLIS, l'adjoint est obtenu par dérivation automatique et est utilisé pour des expériences d'analyses de sensibilité. De son côté, [Goldberg et Sergienko, 2011] dérive de manière théorique le modèle adjoint à partir du modèle d'ordre supérieur proposé par [Goldberg, 2011].

D'autres méthodes ont été utilisées plus récemment. Les méthodes du BLUE et de l'interpolation optimale ont été aussi utilisées pour les calottes polaires (voir [Arthern, 2003] pour l'évaluation du bilan de masse, [Arthern et Hindmarsh, 2006] pour l'évaluation des variations d'épaisseur à l'échelle de la calotte groenlandaise et [Berliner *et al.*, 2008] dans le cas du frottement basal). Une méthode introduite par [Chaabane et Jaoua, 1999] a également été employée notamment par [Arthern et Gudmundsson, 2010]. Appelée méthode de Robin, elle consiste à minimiser une fonction coût mesurant l'écart aux observations

par une descente de gradient sur β par itérations tirant parti des résolutions du modèle Full Stokes avec successivement une condition de Dirichlet en surface (vitesses observées imposées) et la condition de Neumann naturelle du modèle. La solution vérifie à convergence les deux conditions et produit le champ de β voulu. Cette méthode a aussi été employée par [Gillet-Chaulet *et al.*, 2012] sur la calotte groenlandaise dans sa globalité et comparée à la méthode adjointe approchée.

Enfin citons le travail récent de [Pollard et DeConto, 2012] utilisant une simulation à échelle de temps long avec ajustement tous les 5000 ans des coefficients de frottement avec un terme de rappel aux observations. Cette méthode a été employée à l'échelle de la calotte antarctique.

Frottement basal et topographie du socle

Contrairement à l'identification du frottement basal seul, les méthodes cherchant à reconstruire conjointement le frottement basal et la topographie du socle rocheux sont plus restreintes. Citons cependant [Raymond-Pralong et Gudmundsson, 2011] minimisant le BLUE directement par gradient approché (appelé dans l'article méthode de Bayes), [van Pelt *et al.*, 2013] utilisant une itération de Picard pour réduire l'écart aux observations ou [Goldberg et Heimbach, 2013] mettant en place une méthode adjointe.

Notons enfin qu'aucune dépendance temporelle des observations n'est prise en compte dans tous les travaux cités plus haut à l'exception de [Goldberg et Heimbach, 2013].

1.5.2 Paléoclimatologie

Datation de carottes

Les méthodes inverses rencontrées en paléoclimatologie sont principalement consacrées à la datation de carottes de glace. Des méthodes de Monte-Carlo ont été employées à cet effet (consulter notamment [Parrenin *et al.*, 2001], [Grinsted et Dahl-Jensen, 2002] et [Parrenin *et al.*, 2004]). Notons aussi l'utilisation d'un 4D-Var pour la datation croisée de plusieurs carottes dans [Lemieux-Dudon *et al.*, 2010].

Reconstruction du climat passé

La question de la reconstitution du climat passé par des observations de volume de glace (problème qui va nous occuper au chapitre 3) n'a été abordée seulement par [Bintanja *et al.*, 2005a] et [Bintanja *et al.*, 2005b]. Les deux papiers utilisent une méthode simple d'assimilation très proche du nudging aussi appelé relaxation newtonienne (consiste juste à mettre un terme de rappel aux observations dans le second membre du modèle d'évolution).

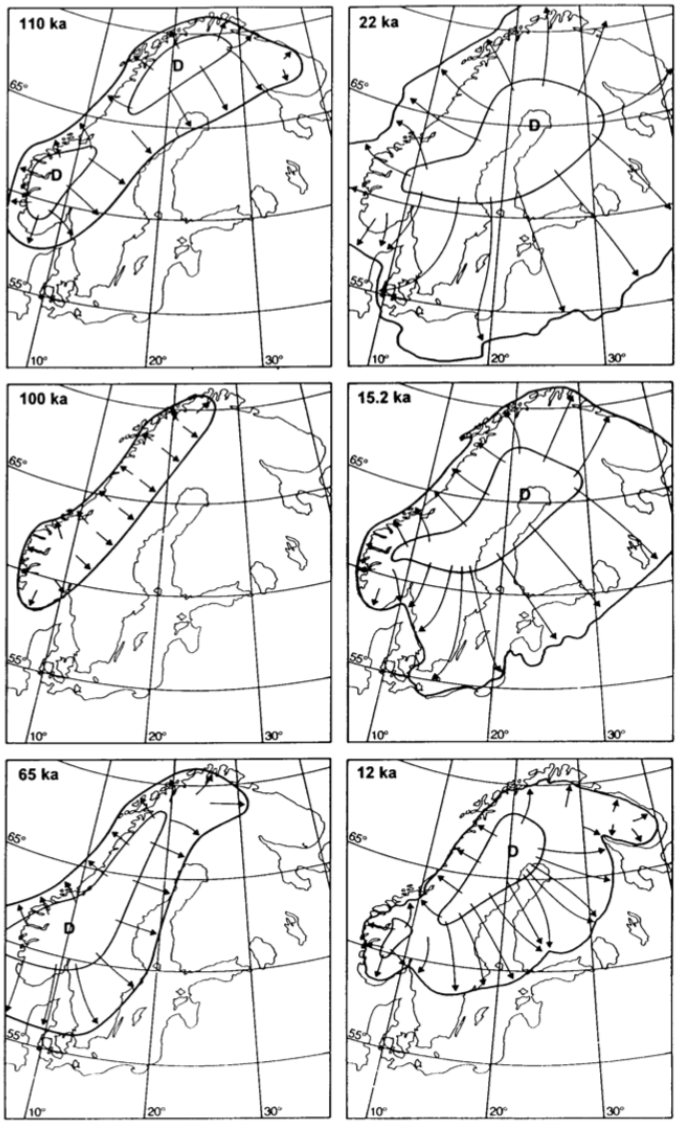

FIGURE 1.12 – Évolution de l'étendue de la calotte Fennoscandienne reconstruite à partir de données géomorphologiques issue de [Kleman *et al.*, 1997].

Chapitre 2

Assimilation de données

Résumer en peu de mots le concept principal de l'assimilation de données peut sembler compliqué tant cette discipline foisonne de méthodes différentes mais ayant peu ou prou le même but. Cependant, du fait de la multiplication de ses champs d'action (météorologie, océanographie, hydrologie, agriculture, ...), une définition claire semble indispensable. Nous choisissons de reproduire ici la définition donnée par [Le Dimet et Blum, 2002] dans le cadre de l'assimilation de données pour l'étude d'un fluide géophysique : « *L'assimilation de données est l'ensemble des techniques qui permettent de combiner, de façon optimale (en un sens à définir), l'information mathématique contenue dans les équations et l'information physique provenant des observations en vue de reconstituer l'état de l'écoulement.* ».

Écartons dans un premier temps l'information contenue dans les équations d'évolution (pas de prise en compte du temps) et notons $\mathbf{x}$ l'état de l'écoulement (nous détaillerons de manière plus précise dans le paragraphe 2.1 les notations utilisées dans ce chapitre ainsi que ce qui est sous-entendu par le mot *état*) décrit par les équations et $\mathbf{y}$ les observations fournies. En combinant ces deux connaissances, l'assimilation de données peut fournir différents résultats :

— une nouvelle estimation de $\mathbf{x}$ avec une estimation de l'incertitude commise prenant compte des observations.

— l'estimation de la loi de probabilité auquel obéit $\mathbf{x}$ lorsque $\mathbf{y}$ est connu (en général, elle est notée $p(\mathbf{x}|\mathbf{y})$).

Ces 2 produits sont bien évidemment liés mais l'information contenue dans $p(\mathbf{x}|\mathbf{y})$ est plus complète qu'une estimation de $\mathbf{x}$ et de son incertitude. Connaître $p(\mathbf{x}|\mathbf{y})$ est utile pour donner une estimation de $\mathbf{x}$ (par exemple, l'espérance mathématique avec comme estimation de l'incertitude la matrice de covariance). Cependant certaines questions légitimes peuvent être soulevées : quelle estimation de $\mathbf{x}$ donner à partir de $p(\mathbf{x}|\mathbf{y})$ (maximum, moyenne, modes, ...) ? L'incertitude estimée est-elle représentative de $p(\mathbf{x}|\mathbf{y})$? Ces questions rendent déjà l'assimilation de données moins simple.

Ajoutons maintenant cet élément essentiel qu'est le temps pour l'assimilation de données. Plus exactement, nous prenons en compte l'évolution dans le temps de l'état du système, des observations ainsi que les équations d'évolutions régissant ces modifications temporelles. Trois objectifs différents mais complémentaires s'offrent désormais à nous :

— le **filtrage** : l'assimilation utilise les observations passées et présentes afin de donner la meilleure estimation (en un sens qui reste encore à définir) de l'état présent du système étudié. Seul le présent est connu de manière optimale. C'est en général un des objectifs des centres opérationnels utilisant l'assimilation de données. Son complément naturel est
— la **prévision** : l'assimilation utilise les observations passées et présentes afin d'estimer au mieux l'état futur du système étudié. En général, c'est le prolongement direct du problème de filtrage.
— le **lissage** : l'assimilation utilise les observations passées et présentes afin de construire la meilleure estimation de la trajectoire du système étudié au cours du temps. Il permet aussi la réanalyse des données (c'est la réanalyse des états reconstruits par filtrage sur une fenêtre de temps plus longue en utilisant les observations collectées sur cette fenêtre et un modèle d'évolution généralement plus moderne que celui utilisé lors des analyses produites par filtrage).

Le but du présent chapitre est de décrire les principales méthodes utilisées à l'heure actuelle en assimilation de données avec un intérêt tout particulier pour le filtre de Kalman d'ensemble (voir 2.6) et le 4D-Var (voir 2.4) qui sont les méthodes utilisées durant cette thèse. Nous détaillons les notations employées dans ce chapitre dans la section 2.1. Ce chapitre doit beaucoup aux notes de cours de [Blayo et Nodet, 2012], [Bocquet, 2013], [Bouttier et Courtier, 1999], [Snyder, 2012] ainsi qu'aux états de l'art effectués dans le manuscrit d'Habilitation à diriger les recherches de [Desroziers, 2007] et dans les manuscrits de thèses de [Broquet, 2007], [Daget, 2008], [Neveu, 2011], [Nodet, 2005], [Papadakis, 2007], [Ricci, 2004], [Souopgui, 2010] et [Vidard, 2001].

2.1 Ingrédients de l'assimilation de données

2.1.1 Remarques préliminaires

Nous introduisons ici les notations utilisées dans ce chapitre. Nous reprenons dans la mesure du possible les notations standards de l'assimilation de données telles qu'indiquées dans [Ide *et al.*, 1997]. Nous prenons les convention suivantes :

— Les scalaires sont notés en italique minuscule (ex. : t, α, ...)
— Les vecteurs sont notés en gras minuscule (ex. : $\mathbf{x}$, $\mathbf{y}$, ...)
— Les opérateurs linéaires et matrices sont notés en gras majuscule (ex. : $\mathbf{M}$, $\mathbf{P}$, $\mathbf{R}$, ...)
— Les opérateurs non-linéaires, fonctions coût sont notés en style calligraphique (ex. : $\mathcal{M}$, $\mathcal{H}$, ...).

Nous résumons dans le tableau suivant les principaux symboles utilisés dans ce chapitre. Ce tableau détaille notamment la signification des indices et exposants utilisés. Cependant, il n'a pas pour but d'expliciter ce qui se cache derrière chaque symbole. Son utilité est plus de l'ordre de l'aide mémoire afin que le lecteur puisse retrouver au même endroit un résumé des nombreux symboles rencontrés dans ce chapitre sur l'assimilation.

TABLE 2.1 – Notations principales employées dans ce chapitre

Symbole	Définition	Remarques
	Vecteurs	
$\mathbf{x}$	vecteur d'état (augmenté ou non)	
$\overline{\mathbf{x}}$	moyenne ensemble de $\mathbf{x}$	
$\mathbf{w}$	vecteur espace d'ensemble	
$\mathbf{y}$	vecteur équivalent observation	
$\overline{\mathbf{y}}$	moyenne ensemble de $\mathbf{y}$	
$\mathbf{y}^o$	vecteur d'observation	contient les mesures effectuées
$\boldsymbol{\varepsilon}$	vecteur erreur	excepté pour l'erreur modèle
$\boldsymbol{\eta}$	erreur modèle	
	Opérateurs non-linéaires	
$\mathcal{H}$	Opérateur d'observation	
$\mathcal{J}$	Fonction coût	
$\mathcal{M}$	Modèle dynamique	
	Matrices et opérateurs linéaires	
$\mathbb{E}$	Espérance mathématique	
$\mathbf{\Pi}$	Projection de l'état continu au discret	
$\mathbf{B}$	Covariance d'erreur d'ébauche	
$\mathbf{H}$	Opérateur d'observation tangent	
$\mathbf{I}_n$	Matrice identité de taille $n \times n$	
$\mathbf{K}$	Matrice de gain	
$\mathbf{M}$	Modèle tangent	
$\mathbf{P}$	Covariance d'erreur estimation de $\mathbf{x}$	en général, $\mathbf{P}^f$ ou $\mathbf{P}^a$
$\mathbf{P}_e$	Estimation de $\mathbf{P}$ par un ensemble	
$\mathbf{Q}$	Covariance d'erreur modèle	
$\mathbf{R}$	Covariance d'erreur d'observations	
$\mathbf{X}$	Matrice anomalies ensemble de $\mathbf{x}$	
$\mathbf{Y}$	Matrice anomalies ensemble de $\mathbf{y}$	
	Scalaires	
t	temps	
$n_{\mathbf{x}}$	taille du vecteur d'état	
$n_{\mathbf{y}}$	taille du vecteur d'observation	
N_e	taille de l'ensemble	
	Indices	
$(\cdot)_k$	indice temporel pour l'instant t_k	
	Exposants	
$(\cdot)^a$	analyse	résultat de l'assimilation
$(\cdot)^b$	ébauche	b pour background
$(\cdot)^f$	prévision	f pour forecast

$(\cdot)^o$	observation	
$(\cdot)^t$	vérité	t pour true
$(\cdot)^c$	continu	
$(\cdot)^{(i)}$	$i$ème élément de l'ensemble	
$(\cdot)^T$	transposée	
$(\cdot)^{-1}$	inverse ou inverse généralisé	

2.1.2 Vecteur d'état et vecteur d'état augmenté

Le vecteur d'état noté $\mathbf{x}(t)$, de taille $n_{\mathbf{x}}$, contient l'état du système étudié à un instant t. Il s'agit plus exactement d'une discrétisation (par différences finies, éléments finies, méthode spectrale, ...) des variables continues du système. Lorsqu'il n'y a pas d'ambiguité, $\mathbf{x}(t)$ est tout simplement noté $\mathbf{x}$. Dans le cas d'une étude en temps discret, on note $\mathbf{x}_k$ le vecteur d'état à l'instant t_k.

Cependant, il n'est pas nécessaire de conserver toutes les variables d'un modèle pour connaître l'état du système étudié. Ainsi, le vecteur d'état contient uniquement l'ensemble des variables pronostiques du modèle étudié, ce sont les variables dont la dérivée temporelle intervient explicitement dans les équations du modèle. Les autres variables du modèle (ou variables diagnostiques) sont entièrement déterminées par le vecteur d'état. Illustrons cette définition par l'exemple suivant

Exemple 2.1 (*Variables pronostiques et diagnostiques*)

Considérons un modèle d'évolution de calotte polaire utilisant l'approximation shallow ice (voir section 1.2.5) avec un socle rocheux supposé fixé. La seule variable pronostique de ce système est l'épaisseur de la calotte H. En effet, les vitesses d'écoulement nécessaires à la résolution de l'équation bilan (1.32) *se calculent par la relation* (1.68) *à partir de H et de différents paramètres. Ce sont des variables diagnostiques.*

L'assimilation de données a généralement pour but d'estimer au mieux le vecteur d'état. Cependant, il peut être nécessaire d'estimer certains paramètres (qui par définition ne sont pas modifiés par le modèle) influents mais mal connus intervenant dans le modèle (c'est le cas dans notre thèse). L'assimilation de données est tout à fait envisageable dans ce contexte (on parle alors d'estimation de paramètres), il suffit simplement d'ajouter dans le vecteur d'état les paramètres cherchés. $\mathbf{x}$ est alors appelé vecteur d'état augmenté. L'assimilation de données cherche alors à estimer au mieux $\mathbf{x}$ (les variables et les paramètres intéressants). Reprenons le contexte de l'exemple précédent pour illustrer le contenu de $\mathbf{x}$.

Exemple 2.2 (*Vecteur d'état augmenté*)

L'évolution d'une calotte polaire dépend notamment du socle rocheux dont la position B_{soc} est mal connue dans certaines zones. Il est possible d'ajouter B_{soc} au vecteur d'état et le vecteur d'état augmenté nouvellement constitué comprend l'épaisseur H et la topographie du socle B_{soc}

Enfin, l'appelation vecteur de contrôle désigne l'ensemble des quantités recherchées par notre système d'assimilation.

2.1.3 Modèle dynamique et prévision

Pour simuler l'évolution de notre système physique est utilisé un modèle numérique, en général non-linéaire, noté $\mathcal{M}$ construit à partir de la discrétisation des équations continues modélisant l'évolution du système. Il est possible à partir de ce modèle d'effectuer une prévision de l'état du système à l'instant t_{k+1} notée $\mathbf{x}_{k+1}^f$ (le f signifie forecast) à partir de l'état $\mathbf{x}_k$ connu à l'instant précédent t_k ($t_k \leq t_{k+1}$). Nous avons la relation suivante

$$\mathbf{x}_{k+1}^f = \mathcal{M}\left(t_k, t_{k+1}\right)\left(\mathbf{x}_k\right) \tag{2.1}$$

Cette prévision résulte uniquement de la connaissance de $\mathbf{x}_k$ et de $\mathcal{M}\left(t_k, t_{k+1}\right)$. Pour simplifier les notations, nous posons $\mathcal{M}\left(t_k, t_{k+1}\right) = \mathcal{M}_k$.

Soit $\mathbf{x}_k^t$ la vraie (t pour true) valeur du vecteur d'état à l'instant t. En théorie, nous espérons que le modèle utilisé soit parfait. Dans ce cas, nous avons l'égalité suivante $\mathbf{x}_{k+1}^t = \mathcal{M}_k\left(\mathbf{x}_k^t\right)$. Cependant, ce cas n'est jamais rencontré en pratique et nous avons plutôt la relation suivante

$$\mathbf{x}_{k+1}^t = \mathcal{M}_k\left(\mathbf{x}_k^t\right) + \boldsymbol{\eta}_k \tag{2.2}$$

avec $\boldsymbol{\eta}_k$ l'erreur due au modèle.

Les sources d'erreur modèle peuvent être nombreuses mais, en général, il est possible de les classer en 3 catégories que voici :

— *Discrétisation du système*

Pour mieux comprendre où intervient cette source, soulignons tout d'abord que le vecteur $\mathbf{x}_k^t$ est une version discrétisée de l'état du système. Notons maintenant $\mathbf{x}_k^c$ la version continue vraie de l'état du système. Il est possible de relier $\mathbf{x}_k^c$ et $\mathbf{x}_k^t$ par l'opérateur $\mathbf{\Pi}$ qui permet la projection du vecteur continu sur la discrétisation employée. Nous avons ainsi

$$\mathbf{x}_k^t = \mathbf{\Pi}\,\mathbf{x}_k^c \tag{2.3}$$

Posons maintenant $\mathcal{M}_k^c$ le modèle d'évolution continu (donc exact) permettant de relier $\mathbf{x}_{k+1}^c$ à $\mathbf{x}_c^k$ par la relation suivante

$$\mathbf{x}_{k+1}^c = \mathcal{M}_k^c\left(\mathbf{x}_k^c\right) \tag{2.4}$$

Nous pouvons alors réécrire l'erreur modèle de la manière suivante

$$\boldsymbol{\eta}_k = \mathbf{x}_{k+1}^t - \mathcal{M}_k(\mathbf{x}_k^t) = \mathbf{\Pi}\,\mathcal{M}_k^c\left(\mathbf{x}_k^c\right) - \mathcal{M}_k\left(\mathbf{\Pi}\,\mathbf{x}_k^c\right) \tag{2.5}$$

Celle-ci dépend bien de la discrétisation employée pour le modèle.

— *Hypothèses simplifiant la physique*

Cette source d'erreur est inhérente à la modélisation de la dynamique de systèmes physiques complexes. La modélisation des calottes polaire ne fait pas exception à cette affirmation comme le souligne l'exemple suivant.

Exemple 2.3 (*Hypothèses simplifiant la physique*)

L'approximation Shallow ice (voir section 1.2.5) utilisée tout au long de cette thèse approche le problème Full Stokes (voir le modèle 1.1) et introduit des erreurs quant à la modélisation de l'évolution des calottes polaires en simplifiant le problème.

— *Mauvais ajustement des paramètres*

Cette dernière catégorie est liée à la catégorie précédente puisque la paramétrisation d'un système physique peut apparaître pour simplifier la modélisation d'un système. Cependant, certains paramètres du modèle peuvent être incontournables comme l'illustre notre exemple fil rouge sur les calottes polaires.

Exemple 2.4 (*Ajustement des paramètres*)

Si l'échelle de temps de l'étude est courte (100 ans par exemple), la topographie du socle rocheux B_{soc} peut être supposée comme fixe. C'est alors un paramètre du modèle et s'il est mal spécifié, des erreurs apparaissent.

Notons que si le paramètre étudié est recherché (donc inclus dans le vecteur d'état augmenté), il doit être exclus comme source d'erreur modèle (puisque dans ce cas, l'erreur ne provient pas de $\mathcal{M}_k$ mais de $\mathbf{x}_k$).

Comme nous venons de le voir, les sources d'erreur modèle sont variées et généralement s'entremêlent souvent de manière inextricable. Pour simplifier le problème, l'erreur modèle $\boldsymbol{\eta}_k$ est considérée comme une variable aléatoire de moyenne

$$\overline{\boldsymbol{\eta}}_k = \mathbb{E}\left[\boldsymbol{\eta}_k\right] \tag{2.6}$$

et de matrice de covariance

$$\mathbf{Q}_k = \mathbb{E}\left[\left(\boldsymbol{\eta}_k - \overline{\boldsymbol{\eta}}_k\right)\left(\boldsymbol{\eta}_k - \overline{\boldsymbol{\eta}}_k\right)^T\right] \tag{2.7}$$

$\mathbb{E}$ désigne l'espérance mathématique, sa définition est rappelée dans l'annexe A.

2.1.4 Observations

Posons $\mathbf{y}_k^o$ le vecteur de taille $n_{\mathbf{y}_k}$ (la taille peut varier au cours du temps !) contenant les observations effectuées à l'instant t_k. Les sources de ces observations sont multiples (données in-situ, radars, satellites, ...) et les quantités observées ne sont pas forcément des variables du modèle ou localisées aux points de grille utilisés. Pour relier le vecteur d'état à l'espace des observations, il faut définir un opérateur d'observation discret $\mathcal{H}_k$ tel que

$$\mathbf{y}_k^o = \mathcal{H}_k\left(\mathbf{x}_k^t\right) + \boldsymbol{\varepsilon}_k^o \tag{2.8}$$

avec $\boldsymbol{\varepsilon}_k^o$ l'erreur d'observation à l'instant t_k. Notons qu'en général, $n_{\mathbf{y}} \ll n_{\mathbf{x}}$.

Regardons maintenant de plus près ce qui se cache derrière l'erreur d'observation. Notons $\mathcal{H}_k^c$ la version continue (donc a priori vraie) de l'opérateur d'observation, nous avons

$$\mathbf{y}_k^o = \mathcal{H}_k^c\left(\mathbf{x}_k^c\right) + \boldsymbol{\varepsilon}_k^m \tag{2.9}$$

où $\boldsymbol{\varepsilon}_k^m$ est l'erreur de mesure dépendant de l'instrument de mesure utilisé. Elle est supposée généralement indépendante du modèle et de sa discrétisation. L'équation (2.8) peut maintenant se réécrire

$$\mathbf{y}_k^o = \mathcal{H}_k\left(\mathbf{x}_k^t\right) + \boldsymbol{\varepsilon}_k^m + \boldsymbol{\varepsilon}_k^r \tag{2.10}$$

avec $\boldsymbol{\varepsilon}_k^r$ l'erreur de représentativité valant

$$\boldsymbol{\varepsilon}_k^r = \mathbf{y}_k^o - \mathcal{H}_k\left(\mathbf{x}_k^t\right) - \boldsymbol{\varepsilon}_k^m = \mathcal{H}_k^c\left(\mathbf{x}_k^c\right) - \mathcal{H}_k\left(\mathbf{\Pi}\mathbf{x}_k^c\right) \tag{2.11}$$

Contrairement à l'erreur de mesure, l'erreur de représentativité dépend du modèle et est généralement beaucoup plus difficile à évaluer.

Pour simplifier les chose, l'erreur d'observation est considérée comme une variable aléatoire de moyenne donnée

$$\overline{\boldsymbol{\varepsilon}}_k^o = \mathbb{E}\left[\boldsymbol{\varepsilon}_k^o\right] \tag{2.12}$$

et de matrice de covariance connue

$$\mathbf{R}_k = \mathbb{E}\left[\left(\boldsymbol{\varepsilon}_k^o - \overline{\boldsymbol{\varepsilon}}_k^o\right)\left(\boldsymbol{\varepsilon}_k^o - \overline{\boldsymbol{\varepsilon}}_k^o\right)^T\right] \tag{2.13}$$

En général, les observations sont supposées débiaisées donc que $\overline{\boldsymbol{\varepsilon}}_k^o = 0$.

2.1.5 Les différentes estimations du vecteur d'état

Ébauche

Généralement, nous disposons avant assimilation d'une estimation du vecteur d'état appelé ébauche et notée $\mathbf{x}^b$ (b pour background). L'ébauche est issue de connaissances a priori du système (climatologie, processus d'assimilation antérieur, prévision du modèle, ...). L'erreur d'ébauche notée $\boldsymbol{\varepsilon}^b$ vaut

$$\boldsymbol{\varepsilon}^b = \mathbf{x}^b - \mathbf{x}^t \tag{2.14}$$

et est supposée être une variable aléatoire de moyenne donnée

$$\overline{\boldsymbol{\varepsilon}}^b = \mathbb{E}\left[\boldsymbol{\varepsilon}^b\right] \tag{2.15}$$

et de matrice de covariance connue

$$\mathbf{B} = \mathbb{E}\left[\left(\boldsymbol{\varepsilon}^b - \overline{\boldsymbol{\varepsilon}}^b\right)\left(\boldsymbol{\varepsilon}^b - \overline{\boldsymbol{\varepsilon}}^b\right)^T\right] \tag{2.16}$$

Prévision

Le vecteur prévision noté $\mathbf{x}^f$ (f pour forecast) est le produit de la simulation de l'évolution du modèle étudié. L'erreur de prévision notée $\boldsymbol{\varepsilon}^f$ vaut

$$\boldsymbol{\varepsilon}^f = \mathbf{x}^f - \mathbf{x}^t \tag{2.17}$$

et est supposée être une variable aléatoire de moyenne

$$\overline{\boldsymbol{\varepsilon}}^f = \mathbb{E}\left[\boldsymbol{\varepsilon}^f\right] \tag{2.18}$$

et de matrice de covariance

$$\mathbf{P}^f = \mathbb{E}\left[\left(\boldsymbol{\varepsilon}^f - \overline{\boldsymbol{\varepsilon}}^f\right)\left(\boldsymbol{\varepsilon}^f - \overline{\boldsymbol{\varepsilon}}^f\right)^T\right] \tag{2.19}$$

Analyse

Le vecteur analyse noté $\mathbf{x}^a$ est l'estimation du vecteur d'état après assimilation de données. L'erreur d'analyse notée $\boldsymbol{\varepsilon}^a$ vaut

$$\boldsymbol{\varepsilon}^a = \mathbf{x}^a - \mathbf{x}^t \tag{2.20}$$

et est supposée être une variable aléatoire de moyenne

$$\overline{\boldsymbol{\varepsilon}}^a = \mathbb{E}\left[\boldsymbol{\varepsilon}^a\right] \tag{2.21}$$

et de matrice de covariance

$$\mathbf{P}^a = \mathbb{E}\left[\left(\boldsymbol{\varepsilon}^a - \overline{\boldsymbol{\varepsilon}}^a\right)\left(\boldsymbol{\varepsilon}^a - \overline{\boldsymbol{\varepsilon}}^a\right)^T\right] \tag{2.22}$$

2.2 Approche bayésienne

Comme indiqué en introduction de ce chapitre, l'assimilation de données peut fournir différents résultats, soit une analyse $\mathbf{x}^a$ du vecteur d'état $\mathbf{x}$ avec une estimation de l'incertitude commise sur $\mathbf{x}^a$ (en général, cette estimation prend la forme d'une matrice de covariance d'erreur notée $\mathbf{P}^a$), soit une estimation de la loi de probabilité à laquelle obéit $\mathbf{x}$ lorsque la loi de probabilité de l'observation est connue. Cependant, la plupart des méthodes employées en assimilation de données sont profondément liées au théorème de Bayes. Nous nous concentrons donc dans cette section à la présentation de ce théorème ainsi qu'à son utilisation potentielle dans les différents cas de l'assimilation de données susmentionnés.

2.2.1 Règle de Bayes

Le présent paragraphe a pour but de rappeler la règle de Bayes[1]. Pour cela, nous effectuons en amont un bref rappel sur les probabilités conditionnelles. Considérons un thésard adepte du loto, si tant est que cet espèce existe, et notons l'événement A « *Le*

1. On rencontre aussi l'appelation théorème de Bayes pour la règle de Bayes.

thésard gagne la cagnotte au prochain tirage du loto » et l'événement B « *Le thésard paie au moins une bouteille de champagne à ses amis* ». Depuis la mise en place en 2008 de la nouvelle version du loto de la Française des Jeux, la probabilité de l'événement A est

$$P(A) = \frac{1}{19\,068\,840} \approx 5.2\,10^{-8}$$

Nous supposons aussi que si le thésard gagne la calotte au loto, alors il y a 95 % de chance qu'il paie au moins une bouteille de champagne à ses amis pour fêter l'événement. Cela correspond à dire que la probabilité de l'événement B connaissant l'événement A notée $P(B|A)$ vaut 0.95. Maintenant, pour connaitre la probabilité d'avoir les 2 évenements, on utilise la formule des probabilités conditionnelles et dans ce cas

$$P(A \cap B) = P(B|A)P(A) \approx 5.0\,10^{-8}$$

Notons qu'il est donc probable que si un thésard de votre connaissance vous offre du champagne, cela ne soit pas dû à la cagnotte du loto [2].

Il était possible aussi de calculer $P(A \cap B)$ en utilisant $P(A|B)$ et $P(B)$. En effet, la formule des probabilités conditionnelles fonctionne dans les deux sens

$$P(A \cap B) = P(A|B)P(B) \tag{2.23}$$
$$P(A \cap B) = P(B|A)P(A) \tag{2.24}$$

En utilisant les formules précédentes, nous trouvons

$$P(A|B) = \frac{P(B|A)P(B)}{P(A)} \tag{2.25}$$

La formule précédente est appelée **règle de Bayes** et peut être utilisée en assimilation de données. Nous allons voir durant les quelques lignes suivantes comment.

2.2.2 Assimilation d'une observation

Considérons le vecteur d'état $\mathbf{x}$ comme une variable aléatoire ayant une densité de probabilité a priori $p_{\mathbf{X}}(\mathbf{x})$ et le vecteur observation $\mathbf{y}$ relié à $\mathbf{x}$ par la relation suivante

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon} \tag{2.26}$$

avec $\boldsymbol{\varepsilon}$ une autre variable aléatoire en général indépendante de $\mathbf{x}$. Puisque $\boldsymbol{\varepsilon}$ et $\mathbf{x}$ sont des variables aléatoires ($\mathcal{H}$ est un opérateur déterministe non nécessairement linéaire), $\mathbf{y}$ est aussi une variable aléatoire.

Le but de l'assimilation de données est d'améliorer nos connaissances sur $\mathbf{x}$ grâce à l'utilisation des observations de $\mathbf{y}$. Un des moyens pour cela est d'évaluer la probabilité

2. D'autres exemples sympathiques d'illustration des probabilités conditionnelles se trouvent facilement dans la littérature, on citera par exemple [van Leeuwen, 2009] avec une histoire de lunettes et de présence féminine lors d'un colloque d'assimilation ...

conditionnelle de $\mathbf{x}$ connaissant $\mathbf{y}$ notée $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$. Supposons que la densité a priori $p_{\mathbf{X}}(\mathbf{x})$ et la relation permettant de relier $\mathbf{y}$ à $\mathbf{x}$ soient connues. La loi de probabilité de $\boldsymbol{\varepsilon}$ est aussi supposée connue, cela implique que la probabilité conditionnelle de $\mathbf{y}$ connaissant $\mathbf{x}$ notée $p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})$ l'est aussi. Nous allons voir par la suite que la connaissance de ces 2 densités est suffisante pour permettre l'évaluation de $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$. Pour cela, il suffit d'appliquer la règle de Bayes

$$p_{\mathbf{X}}(\mathbf{x}|\mathbf{y}) = \frac{p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})}{p_{\mathbf{Y}}(\mathbf{y})} \tag{2.27}$$

D'autre part, $p_{\mathbf{Y}}(\mathbf{y})$ se calcule facilement grâce à la formule suivante

$$p_{\mathbf{Y}}(\mathbf{y}) = \int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \tag{2.28}$$

Finalement, $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$ peut se mettre sous la forme suivante

$$p_{\mathbf{X}}(\mathbf{x}|\mathbf{y}) = \frac{p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})}{\displaystyle\int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}} \tag{2.29}$$

Il suffit de connaitre $p_{\mathbf{X}}(\mathbf{x})$ et $p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})$ pour obtenir l'information recherchée.

En pratique, nous disposons d'une réalisation de la variable aléatoire $\mathbf{y}$ notée $\mathbf{y}^o$ (c'est le vecteur contenant les mesures effectuées). La quantité d'intérêt devient ainsi $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y} = \mathbf{y}^o)$ se calculant de la manière suivante

$$p_{\mathbf{X}}(\mathbf{x}|\mathbf{y} = \mathbf{y}^o) = \frac{p_{\mathbf{Y}}(\mathbf{y} = \mathbf{y}^o|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})}{\displaystyle\int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y} = \mathbf{y}^o|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}} \tag{2.30}$$

Notons qu'il n'est pas nécessaire de calculer le dénominateur pour connaitre l'allure de la densité de probabilité recherchée. En effet, $\displaystyle\int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y} = \mathbf{y}^o|\mathbf{x})p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$ est une constante (on évalue l'intégrale en $\mathbf{y}^o$). La relation de proportionnalité (notée par la suite $\propto$) suivante est donc valable

$$p_{\mathbf{X}}(\mathbf{x}|\mathbf{y} = \mathbf{y}^o) \propto p_{\mathbf{Y}}(\mathbf{y} = \mathbf{y}^o|\mathbf{x})p_{\mathbf{X}}(\mathbf{x}) \tag{2.31}$$

Cette relation est un des élément sous-jacents de l'assimilation de données, elle peut servir soit à évaluer la loi de probabilité du vecteur $\mathbf{x}_k$ à un instant t_k connaissant les observations effectuées soit à donner une nouvelle estimation du vecteur d'état $\mathbf{x}_k$ grâce aux observations fournies.

2.2.3 Estimation bayésienne

Nous avons déclaré dans l'introduction que l'assimilation de donnée pouvait servir soit à calculer la densité $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$ soit à donner une estimation $\mathbf{x}^a$ prenant en compte les observations fournies. Cependant lorsque la densité $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$ est connue, il est aisé de construire des estimations basées sur la connaissance de cette densité. C'est l'estimation bayésienne. Dans la suite du paragraphe, nous donnons quelques estimateurs classiques s'appuyant sur la connaissance complète ou partielle de la densité $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$.

Estimateur de variance minimale

L'estimateur de variance minimale (en anglais : Minimum Variance Estimator (MV) ou Minimum Mean Square Error Estimator (MMSE)) noté $\mathbf{x}^a_{MV}$ minimise parmi tous les estimateurs de $\mathbf{x}$ possibles, notés indifféremment $\mathbf{x}^a$, l'espérance mathématique de $\|\mathbf{x}-\mathbf{x}^a\|^2$ connaissant $\mathbf{y}$. Cette espérance est notée $\mathbb{E}\left[\|\mathbf{x}-\mathbf{x}^a\|^2|\mathbf{y}\right]$ (nous effectuons un rappel sur le calcul de l'espérance d'une variable aléatoire conditionnelle dans l'annexe A).

Remarquons tout d'abord l'égalité suivante

$$\|\mathbf{x}-\mathbf{x}^a\|^2 = (\mathbf{x}-\mathbf{x}^a)^T(\mathbf{x}-\mathbf{x}^a) = \mathrm{Tr}\left((\mathbf{x}-\mathbf{x}^a)(\mathbf{x}-\mathbf{x}^a)^T\right) \tag{2.32}$$

avec Tr désignant la trace d'une matrice. Par linéarité de l'espérance mathématique, rechercher $\mathbb{E}\left[\|\mathbf{x}-\mathbf{x}^a\|^2|\mathbf{y}\right]$ revient à rechercher

$$\mathbb{E}\left[\|\mathbf{x}-\mathbf{x}^a\|^2 \,|\, \mathbf{y}\right] = \mathrm{Tr}\left(\mathbb{E}\left[(\mathbf{x}-\mathbf{x}^a)(\mathbf{x}-\mathbf{x}^a)^T \,|\, \mathbf{y}\right]\right) \tag{2.33}$$

$\mathbf{x}^a_{MV}$ doit donc minimiser la trace de la covariance de l'estimation commise. $\mathbf{x}^a_{MV}$ minimise ainsi la fonction coût suivante

$$\mathcal{J}_{MV}(\mathbf{x}^a) = \int_{\mathbb{R}^{n_\mathbf{x}}} (\mathbf{x}-\mathbf{x}^a)^T(\mathbf{x}-\mathbf{x}^a)\, p_\mathbf{X}(\mathbf{x}|\mathbf{y})\, d\mathbf{x} \tag{2.34}$$

Sous certaines hypothèses de régularités non-explicitées ici, le minimum de $\mathcal{J}_{MV}$ est unique et vaut

$$\mathbf{x}^a_{MV}(\mathbf{y}) = \int_{\mathbb{R}^{n_\mathbf{x}}} \mathbf{x}\, p_\mathbf{X}(\mathbf{x}|\mathbf{y})\, d\mathbf{x} = \mathbb{E}\left[\mathbf{x}|\mathbf{y}\right] \tag{2.35}$$

Il est à noter que cet estimateur est sans biais ($\mathbb{E}[\mathbf{x}^a_{MV}] = \mathbb{E}\left[\mathbf{x}\right]$) ce qui est toujours appréciable pour un estimateur statistique.

Estimateur du maximum a posteriori

L'estimateur du maximum a posteriori (en anglais : Maximum A Posteriori (MAP)) noté $\mathbf{x}^a_{MAP}$ maximise la densité a posteriori $p_\mathbf{X}(\mathbf{x}|\mathbf{y})$. $\mathbf{x}^a_{MAP}$ peut être aussi trouvé en maximisant le logarithme de cette densité. En utilisant la règle de Bayes, ce logarithme se réécrit

$$\ln\left(p_\mathbf{X}(\mathbf{x}|\mathbf{y})\right) = \ln\left(p_\mathbf{Y}(\mathbf{y}|\mathbf{x})p_\mathbf{X}(\mathbf{x})\right) - \ln\left(p_\mathbf{Y}(\mathbf{y})\right) \tag{2.36}$$

En pratique, nous disposons d'une réalisation $\mathbf{y}^o$ de la variable aléatoire $\mathbf{y}$, la densité $p_\mathbf{Y}(\mathbf{y}=\mathbf{y}^o)$ est alors une constante. Dans ce cas, $\mathbf{x}^a_{MAP}$ est le maximum de $p_\mathbf{Y}(\mathbf{y}=\mathbf{y}^o|\mathbf{x})p_\mathbf{X}(\mathbf{x})$ ou de $\ln\left(p_\mathbf{Y}(\mathbf{y}=\mathbf{y}^o|\mathbf{x})p_\mathbf{X}(\mathbf{x})\right)$. Notons que, contrairement à l'estimateur du minimum de variance, l'estimateur du maximum a posteriori n'est pas garanti sans biais.

2.2.4 Assimilation séquentielle : filtrage bayésien

Jusqu'à présent, nous nous sommes concentrés sur la prise en compte des observations dans l'estimation du vecteur d'état $\mathbf{x}$ ou de sa loi de probabilité.

Principe du filtrage

Nous nous plaçons désormais dans le cadre du filtrage. Notons $\mathbf{y}_{0:k}$ l'ensemble des observations (variables aléatoires) entre t_0 et t_k, soit

$$\mathbf{y}_{0:k} = \{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_k\} \tag{2.37}$$

L'objectif du ***filtrage bayésien*** est d'évaluer à chaque instant t_k la probabilité conditionnelle du vecteur d'état $\mathbf{x}_k$ connaissant les observations passées $\mathbf{y}_{0:k}$, notée $p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})$. L'assimilation des données va être effectuée ici de manière séquentielle, les observations sont prises en compte au fur et à mesure donc $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1})$ est calculé à partir de $p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})$. Ce calcul est effectué en 2 étapes nommées phase de prévision et phase d'analyse.

Phase de prévision

La phase de prévision consiste à évaluer la distribution $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})$ connaissant $p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})$. Ces deux distributions sont reliées de la manière suivante (résultat issu d'une formule sur les probabilités conditionnelles, voir annexe A)

$$p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k}) = \int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{x}_k) p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})\, d\mathbf{x}_k \tag{2.38}$$

et nécessite la connaissance de la probabilité de $\mathbf{x}_{k+1}$ connaissant $\mathbf{x}_k$, c'est typiquement là qu'intervient le modèle. En effet, ces deux états sont reliés classiquement par la relation suivante inspirée de l'équation (2.2).

$$\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k) + \boldsymbol{\eta}_k \tag{2.39}$$

$\boldsymbol{\eta}_k$ est un vecteur aléatoire supposé blanc (décorrélé en temps) et de densité de probabilité $p_{\mathbf{X}}(\boldsymbol{\eta}_k) = p_{\mathbf{X}}(\mathbf{x}_{k+1} - \mathcal{M}_k(\mathbf{x}_k))$. La densité de probabilité de $\mathbf{x}_{k+1}$ connaissant $\mathbf{x}_k$ se résume alors à $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{x}_k) = p_{\mathbf{X}}(\boldsymbol{\eta}_k)$. L'égalité (2.38) peut ainsi se réécrire pour permettre l'évaluation de $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})$.

$$p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k}) = \int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{X}}(\mathbf{x}_{k+1} - \mathcal{M}_k(\mathbf{x}_k))\, p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})\, d\mathbf{x}_k \tag{2.40}$$

Dans le cas particulier où le modèle est parfait,

$$p_{\mathbf{X}}(\mathbf{x}_{k+1} - \mathcal{M}_k(\mathbf{x}_k)) = \delta(\mathbf{x}_{k+1} - \mathcal{M}_k(\mathbf{x}_k)) \tag{2.41}$$

avec δ la distribution de Dirac (vaut 1 si $\mathbf{x}_{k+1} = \mathcal{M}_k(\mathbf{x}_k)$, 0 sinon). Dans ce cas, l'évaluation de $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})$ par l'équation (2.40) s'en trouve grandement facilitée.

Phase d'analyse

La phase d'analyse consiste à évaluer la distribution $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1})$ alors que la densité $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})$ est connue. Pour cela, il suffit d'utiliser l'égalité (2.29) et alors nous obtenons la relation suivante

$$p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1}) = \frac{p_{\mathbf{Y}}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}, \mathbf{y}_{0:k})\, p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})}{\displaystyle\int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1}, \mathbf{y}_{0:k})\, p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})\, d\mathbf{x}_{k+1}} \tag{2.42}$$

Les observations sont en général supposées indépendantes les unes des autres en temps. La formule précédente se simplifie alors sous la forme

$$p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1}) = \frac{p_{\mathbf{Y}}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})\, p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})}{\displaystyle\int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y}_{k+1}|\mathbf{x}_{k+1})\, p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})\, d\mathbf{x}_{k+1}} \tag{2.43}$$

2.2.5 Filtres particulaires

Nous venons de décrire comment obtenir d'un point de vue théorique la solution au problème du filtrage bayésien. Cependant, cette solution n'est souvent pas calculable directement en pratique. En effet, comme la densité $p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})$ est une fonction de $\mathbf{x}_k$, vecteur de $\mathbb{R}^{n_{\mathbf{x}}}$, le stockage de la densité voulue est impossible en pratique car $n_{\mathbf{x}}$ est souvent très grand (de l'ordre du million pour des systèmes opérationnels comme en météorologie).

Pour pouvoir résoudre de manière pratique la solution du filtrage bayésien vu plus haut, il faut tenter un calcul approché. Une manière de faire est d'utiliser une approche Monte Carlo consistant à approcher les densités recherchées $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})$ et $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1})$ par un échantillon discret. Lorsque le nombre de membres (appelés aussi particules) N_e de l'échantillon tend vers l'infini, la solution approchée tend normalement vers la solution exacte du filtre bayésien.

Plusieurs versions de cette approche existent, ce sont les ***filtres particulaires***. Nous présentons ici la version la plus simple appelée filtre particulaire bootstrap introduite par [Gordon *et al.*, 1993]. Notons que ce dernier a été utilisé avec succès pour des problèmes de petite taille ($n_{\mathbf{x}} < 10$) notamment par la communauté du climat. (voir [Goosse *et al.*, 2010] ainsi que [Dubinkina *et al.*, 2011] et [Dubinkina et Goosse, 2013]).

Echantillonnage

La solution du filtrage bayésien $p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k})$ est à tout instant t_k approchée par un ensemble représentatif de N_e états du système $\left\{\mathbf{x}_k^{(1)}, \mathbf{x}_k^{(2)}, \ldots, \mathbf{x}_k^{(N_e)}\right\}$ pondéré par un ensemble de poids $\left\{w_k^{(1)}, w_k^{(2)}, \ldots, w_k^{(N_e)}\right\}$ dont la somme fait 1. Ainsi la densité recherchée est approchée par la relation suivante

$$p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}_{0:k}) \approx \sum_{i=1}^{N_e} w_k^{(i)} \delta\left(\mathbf{x}_k - \mathbf{x}_k^{(i)}\right) \tag{2.44}$$

Phase de prévision

À l'étape de prévision, les particules sont propagées par le modèle

$$\mathbf{x}_{k+1}^{(i)} = \mathcal{M}_k\left(\mathbf{x}_k^{(i)}\right) \tag{2.45}$$

et les poids restent inchangés. La densité $p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k})$ est alors approchée par

$$p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k}) \approx \sum_{i=1}^{N_e} w_k^{(i)} \delta\left(\mathbf{x}_{k+1} - \mathbf{x}_{k+1}^{(i)}\right) = \sum_{i=1}^{N_e} w_k^{(i)} \delta\left(\mathbf{x}_{k+1} - \mathcal{M}_k\left(\mathbf{x}_k^{(i)}\right)\right) \tag{2.46}$$

Si nous nous plaçons dans le cas où le modèle est supposé parfait, la précédente relation approche l'égalité (2.40).

Phase d'analyse

L'étape d'analyse consiste à calculer les nouveaux poids $w_{k+1}^{(i)}$ à partir des anciens poids $w_k^{(i)}$ en utilisant la règle de Bayes. Nous possédons une réalisation $\mathbf{y}_{k+1}^o$ du vecteur d'observation à l'instant t_{k+1}. La relation de proportionnalité suivante est donc valable

$$p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k+1}) \propto p_{\mathbf{Y}}(\mathbf{y}_{k+1} = \mathbf{y}_{k+1}^o|\mathbf{x}_{k+1})p_{\mathbf{X}}(\mathbf{x}_{k+1}|\mathbf{y}_{0:k}) \tag{2.47}$$

La mise à jour des poids se fait alors avec la relation suivante

$$w_{k+1}^{(i)} \propto w_k^{(i)}\ p_{\mathbf{Y}}(\mathbf{y}_{k+1} = \mathbf{y}_{k+1}^o|\mathbf{x}_{k+1} = \mathbf{x}_{k+1}^{(i)}) \qquad i = 1, \ldots, N_e \tag{2.48}$$

Dégénérescence du filtre particulaire bootstrap

En pratique, les résultats obtenus par ce filtre sont souvent décevants. Il est souvent constaté qu'après quelques phases d'analyse la densité de probabilité recherchée est approchée uniquement par une seule particule. Le poids des autres particules devient quant à lui négligeable. Notons que le phénomène décrit (et illustré par la figure 2.1) peut être constaté même pour des systèmes de très petite taille. Il est appelé ***dégénérescence du filtre***.

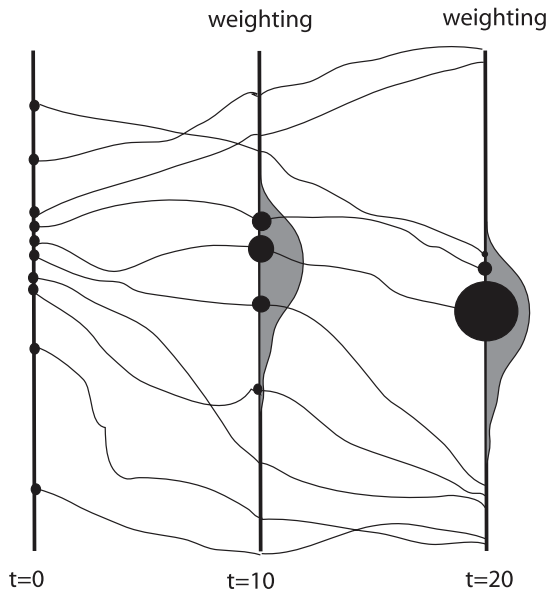

FIGURE 2.1 – Illustration explicitant la dégénérescence du filtre particulaire bootstrap issue de [van Leeuwen, 2009]. Au temps $t = 20$, la densité de probabilité recherchée est presque entièrement approchée par une seule particule, le poids des autres particules étant devenu négligeable.

Une stratégie connue permettant d'éviter cette divergence consiste à rééchantilloner l'ensemble des particules après la phase d'analyse afin d'obtenir un nouvel ensemble de particules ayant toutes des poids proches (il existe différentes stratégies, nous renvoyons à [Doucet *et al.*, 2000] et à [van Leeuwen, 2009] pour de plus amples informations à ce sujet). Cette stratégie appliquée au filtre particulaire bootstrap a été utilisée avec succès pour des problèmes en grande dimension (par exemple [van Leeuwen, 2003] ou [Zhou *et al.*, 2006]). Cependant, le succès du filtre bootstrap avec rééchantillonage pour les applications en grande dimension est à nuancer comme le souligne de manière approfondie [Bengtsson *et al.*, 2008], [Bickel *et al.*, 2008] et [Snyder *et al.*, 2008] (le nombre de particules nécessaires au bon fonctionnement de ce filtre dépend exponentiellement du nombre d'observations indépendantes fournies lors de la phase d'analyse sous certaines hypothèses). Leurs remarques sont consistantes avec les 2 applications citées plus tôt.

Depuis quelques années, de nouvelles stratégies sont développées afin de rendre le filtre particulaire plus accessible pour les systèmes de grande taille. Citons notamment les travaux de [Morzfeld et Chorin, 2012] (voir aussi [Chorin *et al.*, 2010]) et [Avenel *et al.*, 2013]. Nous renvoyons aussi à l'excellente revue sur les filtres particulaires utilisés en géosciences faite par [van Leeuwen, 2009].

Une autre voie à envisager est l'hybridation du filtre particulaire avec d'autres méthodes d'assimilation comme le filtre de Kalman d'ensemble. Nous renvoyons à la section 2.7 pour plus d'informations à ce sujet.

2.3 Best Linear Unbiased Estimator (BLUE)

L'approche bayésienne vue précédemment permet de donner une estimation de la loi de probabilité du vecteur d'état connaissant les observations $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$ ou une estimation du vecteur $\mathbf{x}$ tirée de cette loi. Elle nécessite la connaissance de la densité de probabilité a priori sur le vecteur d'état $p_{\mathbf{X}}(\mathbf{x})$ et celle de la probabilité conditionnelle du vecteur observation $\mathbf{y}$ connaissant $\mathbf{x}$, $p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})$. Lorsque ces densités sont mal connues, c'est à dire lorsque que nous ne connaissons plus que les moments de premier (l'espérance mathématique) et second ordre (la matrice de covariance) des lois de probabilités, il est encore possible de donner une estimation de $\mathbf{x}$ connaissant $\mathbf{y}$ ainsi qu'une estimation de l'incertitude commise par notre estimateur sous forme de matrice de covariance d'erreur. L'estimateur le plus connu dans le cas précédemment décrit est le Best Linear Unbiased Estimator ou BLUE. Soulignons d'ores et déjà que le BLUE est équivalent à l'estimateur du maximum a posteriori et à l'estimateur de minimum de variance (voir section 2.2.3 pour leurs définitions respectives) dans le cas gaussien. Cette équivalence n'est valable que dans ce cas, si nous nous en éloignons le BLUE n'est plus qu'un estimateur parmi d'autres et sa qualité est variable selon les cas.

La suite de cette section est consacrée à la description du BLUE et à son utilisation potentielle en assimilation.

2.3.1 Généralités

Plaçons nous tout d'abord dans le cas suivant : nous cherchons à donner une estimation du vecteur aléatoire $\mathbf{x}$ en utilisant le vecteur aléatoire $\mathbf{y}$. Nous supposons connaître uniquement les quantités suivante :

— les moments de premier et second ordre de la loi de probabilité $p_{\mathbf{X}}(\mathbf{x})$ soit la moyenne $\mathbb{E}[\mathbf{x}]$ et la matrice de covariance $\mathbf{Cov}(\mathbf{x})$.

— les moments de premier et second ordre de la loi de probabilité $p_{\mathbf{Y}}(\mathbf{y})$ soit la moyenne $\mathbb{E}[\mathbf{y}]$ et la matrice de covariance $\mathbf{Cov}(\mathbf{y})$.

— la covariance entre $\mathbf{x}$ et $\mathbf{y}$ notée $\mathbf{Cov}(\mathbf{x}, \mathbf{y})$.

Nous nous restreignons aux estimateurs de $\mathbf{x}$ notés $\mathbf{x}^a$ de la forme suivante

$$\mathbf{x}^a = \mathbf{K}\mathbf{y} + \mathbf{b} \tag{2.49}$$

avec une matrice $\mathbf{K}$ et un vecteur $\mathbf{b}$ déterministes à fixer.

$\mathbf{x}^a$ est une fonction affine du vecteur aléatoire d'observation $\mathbf{y}$ (d'où le Linear de BLUE). Nous allons maintenant chercher $\mathbf{x}^a_{BLUE}$ le meilleur estimateur sous cette forme (d'où le Best de BLUE). Le meilleur est à prendre au sens de variance minimale, soit nous cherchons $\mathbf{x}^a_{BLUE}$ minimisant l'espérance mathématique de $\|\mathbf{x} - \mathbf{x}^a\|^2$ (nous avons déjà utilisé ce genre de procédé pour calculer l'estimateur de variance minimale durant l'approche bayésienne).

Posons la quantité

$$\theta^2 = \|\mathbf{x} - \mathbf{x}^a\|^2 = \mathbf{x}^T\mathbf{x} + \mathbf{y}^T\mathbf{K}^T\mathbf{K}\mathbf{y} + \mathbf{b}^T\mathbf{b} + 2\left(\mathbf{b}^T\mathbf{K}\mathbf{y} - \mathbf{x}^T\mathbf{b} - \mathbf{x}^T\mathbf{K}\mathbf{y}\right)$$

Cette quantité est quadratique en $\mathbf{K}$ et en $\mathbf{b}$ donc $\mathbb{E}[\theta^2]$ est dérivable en $\mathbf{K}$ et en $\mathbf{b}$, les dérivées partielles valent

$$\frac{\partial \mathbb{E}[\theta^2]}{\partial \mathbf{K}} = \mathbb{E}\left[\frac{\partial \theta^2}{\partial \mathbf{K}}\right] = 2\mathbf{K}\mathbb{E}\left[\mathbf{y}\mathbf{y}^T\right] + 2\mathbf{b}\left(\mathbb{E}[\mathbf{y}]\right)^T - 2\mathbb{E}[\mathbf{x}\mathbf{y}^T] \tag{2.50}$$

$$\frac{\partial \mathbb{E}[\theta^2]}{\partial \mathbf{b}} = \mathbb{E}\left[\frac{\partial \theta^2}{\partial \mathbf{b}}\right] = 2\mathbf{b} + 2\mathbf{K}\mathbb{E}[\mathbf{y}] - 2\mathbb{E}[\mathbf{x}] \tag{2.51}$$

Pour obtenir l'estimateur BLUE, il suffit de chercher $\mathbf{K}$ et $\mathbf{b}$ tel que les 2 dérivées partielles s'annulent (dans ce cas, la quantité $\mathbb{E}[\theta^2]$ est minimale). Le résultat obtenu est

$$\mathbf{K} = \mathbf{Cov}(\mathbf{x}, \mathbf{y})\left(\mathbf{Cov}(\mathbf{y})\right)^{-1} \tag{2.52}$$

$$\mathbf{b} = \mathbb{E}[\mathbf{x}] - \mathbf{K}\,\mathbb{E}[\mathbf{y}] \tag{2.53}$$

où par définition

$$\begin{aligned}\mathbf{Cov}(\mathbf{x}, \mathbf{y}) &= \mathbb{E}\left[\mathbf{x}\mathbf{y}^T\right] - \mathbb{E}[\mathbf{x}]\left(\mathbb{E}[\mathbf{y}]\right)^T \\ \mathbf{Cov}(\mathbf{y}) &= \mathbb{E}\left[\mathbf{y}\mathbf{y}^T\right] - \mathbb{E}[\mathbf{y}]\left(\mathbb{E}[\mathbf{y}]\right)^T\end{aligned}$$

Notons que cet estimateur est sans biais ($\mathbb{E}\left[\mathbf{x}^a_{BLUE}\right] = \mathbb{E}[\mathbf{x}]$ d'où le Unbiased de BLUE).

Évaluons maintenant, la matrice de covariance de l'erreur commise sur l'estimation $\mathbf{P}^a_{BLUE}$. Nous avons déjà que

$$\begin{aligned}
\left(\mathbf{x} - \mathbf{x}^a_{BLUE}\right)\left(\mathbf{x} - \mathbf{x}^a_{BLUE}\right)^T &= \left(\mathbf{x} - \mathbb{E}[\mathbf{x}] - \mathbf{K}(\mathbf{y} - \mathbb{E}[\mathbf{y}])\right)\left(\mathbf{x} - \mathbb{E}[\mathbf{x}] - \mathbf{K}(\mathbf{y} - \mathbb{E}[\mathbf{y}])\right)^T \\
&= \left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)\left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)^T - \left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)\left(\mathbf{y} - \mathbb{E}[\mathbf{y}]\right)^T \mathbf{K}^T \\
&\quad - \mathbf{K}\left(\mathbf{y} - \mathbb{E}[\mathbf{y}]\right)\left(\mathbf{x} - \mathbb{E}[\mathbf{x}]\right)^T \\
&\quad + \mathbf{K}\left(\mathbf{y} - \mathbb{E}[\mathbf{y}]\right)\left(\mathbf{y} - \mathbb{E}[\mathbf{y}]\right)^T \mathbf{K}^T
\end{aligned}$$

$\mathbf{P}^a_{BLUE}$ correspond à l'espérance de la quantité précédente. En utilisant la valeur de $\mathbf{K}$ calculée plus haut, $\mathbf{P}^a_{BLUE}$ vaut

$$\mathbf{P}^a_{BLUE} = \mathbf{Cov}(\mathbf{x}) - \mathbf{Cov}(\mathbf{x}, \mathbf{y})\left(\mathbf{Cov}(\mathbf{y})\right)^{-1}\left(\mathbf{Cov}(\mathbf{x}, \mathbf{y})\right)^T \tag{2.54}$$

$\mathbf{P}^a_{BLUE}$ ne dépend pas de $\mathbf{y}$.

Enfin, dans le cas particulier gaussien, il est possible de relier l'estimateur BLUE aux estimateurs de variance minimale et de maximum a posteriori vus dans l'approche bayésienne (voir section 2.2.3) grâce à la proposition suivante.

Proposition 2.1

Si $\mathbf{x}$ et $\mathbf{y}$ sont des variables aléatoires suivant une loi normale, alors la densité $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$ est une gaussienne de moyenne $\mathbb{E}[\mathbf{x}|\mathbf{y}]$ valant $\mathbf{x}^a_{BLUE}$ et de matrice de covariance $\mathbf{P}_{\mathbf{x}|\mathbf{y}}$ valant $\mathbf{P}^a_{BLUE}$.

D'après la proposition précédente, dans le cas gaussien, l'estimateur BLUE correspond à l'espérance de $\mathbf{x}$ connaissant $\mathbf{y}$ qui est l'estimateur de variance minimale. Il est aussi égal à l'estimateur du maximum a posteriori car dans ce cas, le maximum de $p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})$ est atteint pour $\mathbf{x}$ valant $\mathbb{E}[\mathbf{x}|\mathbf{y}]$. Il y a donc équivalence entre les 3 estimateurs dans le cas gaussien décrit dans la proposition. Cette propriété nous sera utile par la suite. Notons bien que cette équivalence n'est vraie que dans le cas gaussien. Dans le cas général, l'estimateur du maximum a posteriori et l'estimateur de variance minimale ne sont pas égaux.

2.3.2 BLUE et assimilation

Jusqu'à présent, nous n'avons pas fait intervenir l'opérateur d'observation et n'avons fait aucune hypothèse. Repositionnons nous maintenant dans le cadre classique de l'assimilation de données. Avant l'assimilation, nous disposons des quantités suivantes :

— une ébauche $\mathbf{x}^b$ du vecteur d'état $\mathbf{x}$ tel que $\mathbf{x}^b = \mathbf{x} + \boldsymbol{\varepsilon}^b$.

— une réalisation $\mathbf{y}^o$ du vecteur d'observation $\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}^o$.

Hypothèses 2.1 (*BLUE cas linéaire*)

1. *L'erreur d'ébauche* $\boldsymbol{\varepsilon}^b$ *est non-triviale, sans biais* ($\mathbb{E}[\boldsymbol{\varepsilon}^b] = \mathbf{0}$) *et de matrice de covariance* $\mathbf{B}$ *connue (matrice définie positive).*

2. *L'erreur d'observation* $\boldsymbol{\varepsilon}^o$ *est non-triviale, sans biais* ($\mathbb{E}[\boldsymbol{\varepsilon}^o] = \mathbf{0}$) *et de matrice de covariance* $\mathbf{R}$ *connue (matrice définie positive).*

3. *Les erreurs d'ébauche et d'observation sont indépendantes, ce qui veut dire que*

$$\mathbb{E}[\boldsymbol{\varepsilon}^o(\boldsymbol{\varepsilon}^b)^T] = \mathbf{0} \qquad \mathbb{E}[\boldsymbol{\varepsilon}^b(\boldsymbol{\varepsilon}^o)^T] = \mathbf{0} \tag{2.55}$$

4. *L'opérateur d'observation est linéaire autour de* $\mathbf{x}^b$*, c'est à dire que pour tout* $\mathbf{x}$ *proche de* $\mathbf{x}^b$

$$\mathcal{H}(\mathbf{x}) - \mathcal{H}\left(\mathbf{x}^b\right) = \mathbf{H}\left(\mathbf{x} - \mathbf{x}^b\right) \tag{2.56}$$

avec $\mathbf{H}$ *linéarisation de l'opérateur d'observation* $\mathbf{H}$ *autour de* $\mathbf{x}^b$*.*

Que vaut maintenant le BLUE dans ce cas ? Suite aux hypothèses précédentes, nous admettons que la densité a priori sur $\mathbf{x}$ a pour moyenne $\mathbb{E}[\mathbf{x}] = \mathbf{x}^b$ et pour matrice de covariance $\mathbf{Cov}(\mathbf{x}) = \mathbf{B}$. Il nous reste encore à calculer $\mathbb{E}[\mathbf{y}]$, $\mathbf{Cov}(\mathbf{x}, \mathbf{y})$ et $\mathbf{Cov}(\mathbf{y})$. Ainsi

$$\begin{aligned}
\mathbb{E}[\mathbf{y}] = \mathbb{E}[\mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}^o] &= \mathbb{E}[\mathcal{H}(\mathbf{x}^b) + \mathbf{H}(\mathbf{x} - \mathbf{x}^b) + \boldsymbol{\varepsilon}^o] \\
&= \mathcal{H}(\mathbf{x}^b) - \mathbf{H}\,\mathbb{E}[\boldsymbol{\varepsilon}^b] + \mathbb{E}[\boldsymbol{\varepsilon}^o] \\
&= \mathcal{H}(\mathbf{x}^b)
\end{aligned}$$

$$\begin{aligned}
\mathbf{Cov}(\mathbf{x}, \mathbf{y}) = \mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}^o) &= \mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x}^b) - \mathbf{H}\boldsymbol{\varepsilon}^b + \boldsymbol{\varepsilon}^o) \\
&= \mathbf{Cov}(\mathbf{x}, -\mathbf{H}\boldsymbol{\varepsilon}^b) \\
&= \mathbf{Cov}(\boldsymbol{\varepsilon}^b, \mathbf{H}\boldsymbol{\varepsilon}^b) \\
&= \mathbf{B}\mathbf{H}^T
\end{aligned}$$

$$\begin{aligned}
\mathbf{Cov}(\mathbf{y}) &= \mathbf{Cov}(\mathcal{H}(\mathbf{x}^b) - \mathbf{H}\boldsymbol{\varepsilon}^b + \boldsymbol{\varepsilon}^o) \\
&= \mathbf{Cov}(-\mathbf{H}\boldsymbol{\varepsilon}^b + \boldsymbol{\varepsilon}^o) \\
&= \mathbf{Cov}(\mathbf{H}\boldsymbol{\varepsilon}^b) + \mathbf{Cov}(\boldsymbol{\varepsilon}^o) \\
&= \mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}
\end{aligned}$$

Le BLUE s'écrit dans ce cas

Algorithme 2.1 (*BLUE cas linéaire*)

1. *L'analyse est obtenue à partir des relations*

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)\right) \tag{2.57}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1} \tag{2.58}$$

La matrice $\mathbf{K}$ *est appelée matrice de gain.*

2. *La matrice de covariance d'erreur associée vaut*

$$\mathbf{P}^a = (\mathbf{I}_{n_\mathbf{x}} - \mathbf{KH})\,\mathbf{B} \tag{2.59}$$

Dans le cas gaussien, nous supposons en plus que la densité de probabilité a priori $p_\mathbf{X}(\mathbf{x})$ est normale de moyenne $\mathbf{x}^b$ et de matrice de covariance $\mathbf{B}$ et que la loi de $\boldsymbol{\varepsilon}^o$ est aussi normale de moyenne nulle et de matrice de covariance $\mathbf{R}$. Sous ces hypothèses, la loi de $\mathbf{x}$ connaissant $\mathbf{y}$ est aussi normale de moyenne $\mathbf{x}^a$ et de matrice de covariance $\mathbf{P}^a$. Dans ce cas, l'estimation BLUE permet de déterminer entièrement la loi de probabilité recherchée par l'approche bayésienne. Cette proporiété est exploitée notamment par le filtre de Kalman qui est un filtre gaussien. Les propriétés essentielles du BLUE permettant de dériver le filtre de Kalman sont entièrement détaillées et démontrées dans la thèse de [Arnaud, 2004]. Nous renvoyons à sa lecture pour plus d'informations à ce sujet.

Dans le cas où la linéarisation de l'opérateur d'observation $\mathcal{H}$ n'est pas pertinente (opérateur trop non-linéaire par exemple), l'estimation BLUE reste pertinente (les autres hypothèses formulées restent valables). Cependant, l'espérance mathématique $\mathbb{E}[\mathbf{y}]$ est plus difficile à obtenir puisque

$$\mathbb{E}[\mathbf{y}] = \mathbb{E}[\mathcal{H}(\mathbf{x})]$$

et les covariances liées à $\mathbf{y}$ valent

$$\mathbf{Cov}(\mathbf{x},\mathbf{y}) = \mathbf{Cov}(\mathbf{x},\mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}^o) = \mathbf{Cov}(\mathbf{x},\mathcal{H}(\mathbf{x}))$$

$$\mathbf{Cov}(\mathbf{y}) = \mathbf{Cov}(\mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}^o) = \mathbf{Cov}(\mathcal{H}(\mathbf{x})) + \mathbf{R}$$

Le BLUE se réécrit alors

Algorithme 2.2 (*BLUE cas non-linéaire*)

1. *L'analyse est obtenue à partir des relations*

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)\right) \tag{2.60}$$

$$\mathbf{K} = \mathbf{Cov}(\mathbf{x},\mathcal{H}(\mathbf{x}))\left(\mathbf{Cov}(\mathcal{H}(\mathbf{x})) + \mathbf{R}\right)^{-1} \tag{2.61}$$

2. *La matrice de covariance d'erreur associée vaut*

$$\mathbf{P}^a = \mathbf{Cov}(\mathbf{x}) - \mathbf{K}\left(\mathbf{Cov}(\mathbf{x},\mathcal{H}(\mathbf{x}))\right)^T \tag{2.62}$$

Cette version nous sera utile notamment pour comprendre le fonctionnement du filtre de Kalman d'ensemble.

2.4 Méthodes variationnelles

Contrairement aux autres méthodes d'assimilation de données présentées dans ce manuscrit, les ***méthodes variationnelles*** ne sont pas principalement basées sur la théorie de l'estimation statistique mais sur la théorie de l'optimisation et du contrôle optimal.

Elles consistent toutes en la construction d'une estimation minimisant une fonctionnelle.

Introduites dès 1955 ([Sasaki, 1955], [Sasaki, 1958]), leur utilisation dans le cadre opérationnel à partir des années 90 par les services de prévision météorologique (en 1991 au NMC [3] [Parrish et Derber, 1992], en 1996 au Centre européen de prévision météorologique à moyen terme [Courtier *et al.*, 1998], [Andersson *et al.*, 1998], en 1997 à Météo France) a assuré leur succès durant de nombreuses années dans la communauté de l'assimilation de données. Les méthodes variationnelles sont toujours utilisées par les centre opérationnels dans ce cadre conjointement avec les filtres de Kalman d'ensemble (cette technique est détaillée à la section 2.6).

Nous présentons dans la suite les 2 principaux réprésentants de l'assimilation variationnelle, l'assimilation 3D-Var et l'assimilation 4D-Var. Notons pour finir cette rapide introduction que, même si les statistiques ne forment pas les bases de ces méthodes, elles restent partie prenante. Enfin, les méthodes variationnelles sous certaines hypothèses peuvent être rattachées au BLUE ou aux approches bayésiennes.

2.4.1 3D-Var

Introduite par [Sasaki, 1955], le ***3D-Var*** (pour 3 Dimensional VARiationnal assimilation) a pour habitude de traiter des problèmes tridimensionnels (pas de dimension temporelle impliquée). Par abus de langage, cette dénomination est conservée pour les problèmes unidimensionels et bidimensionnels [4].

Le problème du 3D-Var se présente de la manière suivante. A un instant donné, nous disposons d'une ébauche $\mathbf{x}^b$ du vecteur d'état avec la matrice de covariance d'erreur d'ébauche $\mathbf{B}$ qui lui est associée ainsi que d'observations $\mathbf{y}^o$ avec la matrice de covariance d'erreur d'observation $\mathbf{R}$. L'objectif de l'assimilation 3D-Var est de trouver une estimation $\mathbf{x}^a$ minimisant la fonction coût suivante

$$\mathcal{J}(\mathbf{x}) = \mathcal{J}^b(\mathbf{x}) + \mathcal{J}^o(\mathbf{x}) \tag{2.63}$$

avec $\mathcal{J}^b$ représentant l'écart par rapport à l'ébauche

$$\mathcal{J}^b(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x} - \mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x} - \mathbf{x}^b\right) \tag{2.64}$$

et $\mathcal{J}^o$ l'écart par rapport aux observations

$$\mathcal{J}^o(\mathbf{x}) = \frac{1}{2}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x})\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x})\right) \tag{2.65}$$

Minimiser $\mathcal{J}$ représente donc un compromis (au sens des moindres carrés) entre être proche de l'ébauche et être proche des observations. Ces deux termes sont pondérés par l'inverse des matrices $\mathbf{B}$ et $\mathbf{R}$ (de manière simpliste, si une observation est sûre, alors sa variance

3. devenu par la suite le National Center for Environmental Prediction

4. Dans le cas d'un modèle de calotte polaire utilisant l'approximation *shallow ice*, le problème étudié n'est que bidimensionnel, les variables d'état du modèle (ici l'épaisseur de la calotte) ne dépendent que de x et y. On devrait parler dans ce cas de 2D-Var, cependant l'expression consacrée reste 3D-Var.

est faible et son poids dans la fonction coût est plus important). Bien que l'approche soit philosophiquement différente de celle du BLUE, les deux méthodes fournissent les mêmes résultats dans le cas linéaire.

Fondements : équivalence avec le BLUE dans le cas linéaire

Nous nous replaçons sous les hypothèses du BLUE dans le cas linéaire (Hypothèses 2.1). La fonction $\mathcal{J}^o$ peut se réécrire

$$\mathcal{J}^o(\mathbf{x}) = \frac{1}{2}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b) - \mathbf{H}\left(\mathbf{x} - \mathbf{x}^b\right)\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b) - \mathbf{H}\left(\mathbf{x} - \mathbf{x}^b\right)\right)$$

La fonction coût $\mathcal{J}$ est alors quadratique en $\mathbf{x}$. Comme $\mathbf{B}$ et $\mathbf{R}$ sont supposées définies positives, elle admet un unique minimum atteint lorsque son gradient $\mathcal{J}$ s'annule. Hors celui-ci vaut

$$\nabla\mathcal{J}(\mathbf{x}) = \mathbf{B}^{-1}\left(\mathbf{x} - \mathbf{x}^b\right) - \mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b) - \mathbf{H}\left(\mathbf{x} - \mathbf{x}^b\right)\right)$$

Résoudre $\nabla\mathcal{J}(\mathbf{x}^a) = \mathbf{0}$ conduit à obtenir l'égalité suivante

$$\mathbf{x}^a = \mathbf{x}^b + \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x}^b)\right)$$

Or, avec un peu de calcul matriciel, l'égalité suivante est obtenue

$$\begin{aligned}
& \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1} \\
= & \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1} \\
= & \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{B}^{-1}\right)\mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1} \\
= & \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1} \qquad (2.66)
\end{aligned}$$

Dès lors, nous reconnaissons bien la formule du BLUE pour $\mathbf{x}^a$. D'autre part, il est aussi possible de relier la matrice de covariance d'erreur de l'estimation $\mathbf{P}^a$ et la hessienne de la fonction coût au minimum $\mathbf{x}^a$. Pour cela, nous avons besoin de la formule de Sherman-Morrisson-Woodbury

$$(\mathbf{A} + \mathbf{U}\mathbf{C}\mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1} + \mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)\mathbf{V}\mathbf{A}^{-1} \qquad (2.67)$$

reliant $\mathbf{A}$, $\mathbf{U}$, $\mathbf{C}$ et $\mathbf{V}$ des matrices de tailles correctement spécifiées.

La hessienne de $\mathcal{J}$ vaut

$$\mathbf{Hess}(\mathbf{x}) = \mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} \qquad (2.68)$$

Or la matrice $\mathbf{P}^a$ du BLUE se réécrit de la manière suivante en utilisant la formule de Sherman-Morrison-Woodbury (2.67).

$$\mathbf{P}^a = \mathbf{B} - \mathbf{B}\mathbf{H}^T\left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{B} = \left(\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}$$

Donc

$$\mathbf{P}^a = \left(\mathbf{Hess}(\mathbf{x}^a)\right)^{-1} \qquad (2.69)$$

Cette relation nous permet de donner une autre manière d'évaluer la matrice de covariance d'erreur de l'estimation $\mathbf{P}^a$ et aussi d'affirmer que la précision de l'estimation est liée à la convexité de la fonction coût. Ainsi, plus la fonction coût est convexe, plus l'estimation est de bonne qualité. Cette relation nous servira par la suite à estimer $\mathbf{P}^a$ dans le cadre général du 3D-Var et du 4D-Var.

Bases du 3D-Var

Si l'approche 3D-Var est équivalente au BLUE dans le cas d'un opérateur d'observation linéaire, pourquoi privilégier le 3D-Var par rapport au BLUE ? En fait, l'un des inconvénients majeurs du BLUE est le calcul de la matrice de gain $\mathbf{K}$ (nous détaillons cette problématique dans la section consacrée au filtre de Kalman, cf. 2.5.2). Le 3D-Var contourne efficacement ce problème en minimisant numériquement la fonction coût $\mathcal{J}$.

Avant d'évoquer la résolution plus pratique de cette minimisation, nous effectuons quelques remarques sur l'ébauche. Elle est issue de la connaissance a priori du système étudié, cela peut être par exemple une prévision à un instant précédent. Le terme d'ébauche peut être aussi vu comme une régularisation de Tikhonov de l'écart aux observations $\mathcal{J}^o$. En effet, il est fort probable que $\mathcal{J}^o$ n'admette pas de minimum global unique ($n_{\mathbf{y}}$ la taille du vecteur d'observation est bien souvent plus petite que $n_{\mathbf{x}}$ la taille du vecteur d'état). Ajouter $\mathcal{J}^b$ à $\mathcal{J}^o$ permet en général d'assurer l'unicité du minimum global de la fonction coût $\mathcal{J}$ (illustrée à la figure 2.2) ou tout du moins d'améliorer la convexité de la fonction, ce qui est toujours agréable lorsque nous voulons minimiser cette quantité.

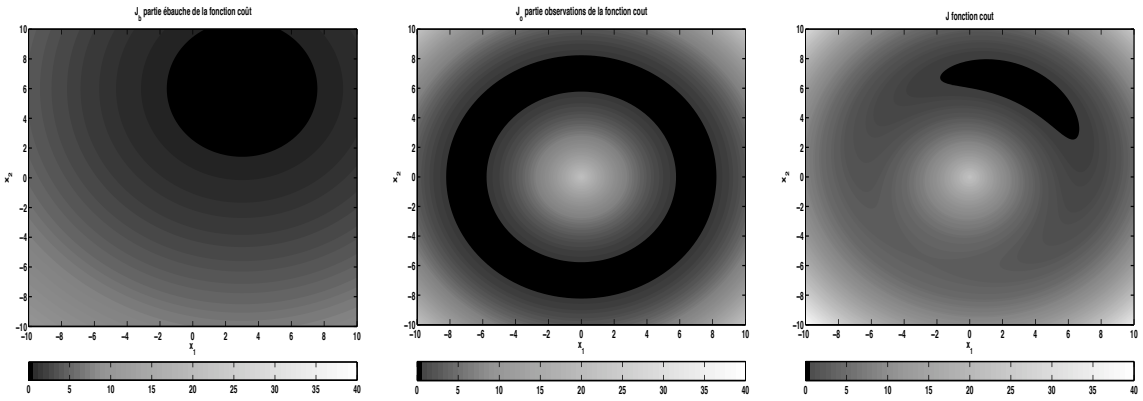

FIGURE 2.2 – A gauche, tracé du terme d'ébauche $\mathcal{J}^b(\mathbf{x}) = \frac{1}{30}\left((x_1-3)^2+(x_2-6)^2\right)$. Au centre, tracé du terme dû aux observations $\mathcal{J}^o(\mathbf{x}) = \frac{1}{2}\left(7-\|\mathbf{x}\|\right)^2$ avec $\|.\|$ la norme euclidienne. $\mathcal{J}^o$ a clairement une infinité de minima. Enfin à droite, tracé de la fonction coût somme des deux précédente, cette dernière n'a plus qu'un seul minimum global.

Comme $n_{\mathbf{x}}$ est grand, la minimisation de $\mathcal{J}$ est généralement effectuée avec l'aide du gradient. Contrairement à ce qui précède, nous ne supposons plus que l'opérateur d'observation est linéaire autour de $\mathbf{x}^b$. Nous supposons cependant que $\mathcal{H}$ non-linéaire est assez régulier pour admettre l'existence de $\mathbf{H}$ opérateur ***linéaire tangent*** de $\mathcal{H}$ en $\mathbf{x}$. Cet opérateur varie en fonction de $\mathbf{x}$. Le gradient de la fonction coût vaut alors

$$\nabla\mathcal{J}(\mathbf{x}) = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right) - \left(\mathbf{H}(\mathbf{x})\right)^T\mathbf{R}^{-1}\left(\mathbf{y}^o-\mathcal{H}(\mathbf{x})\right) \tag{2.70}$$

et fait intervenir $\mathbf{H}^T$ l'opérateur ***adjoint*** de $\mathcal{H}$ au point $\mathbf{x}$ obtenu à partir de $\mathbf{H}$. Notons que si $\mathcal{H}$ est non-linéaire, $\mathcal{J}$ peut admettre plusieurs minima locaux ce qui rend la

recherche du minimum global plus difficile. En général, les méthodes utilisées pour minimiser la fonction coût font intervenir le gradient de la fonction coût et ne peuvent accéder qu'aux minima locaux (gradient nul à ce point) sans savoir s'il s'agit du minimum global ou non.

La qualité de l'estimation se calcule toujours par $\mathbf{P}^a = (\mathbf{Hess}(\mathbf{x}^a))^{-1}$ avec

$$\mathbf{Hess}(\mathbf{x}) = \mathbf{B}^{-1} + (\mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}\mathbf{H}(\mathbf{x}) \tag{2.71}$$

Contrairement au cas du BLUE, la hessienne dépend de $\mathbf{x}$.

Au final, l'algorithme du 3D-Var peut être résumé de la manière suivante

Algorithme 2.3 (*3D-Var*)

1. *Initialisation de l'algorithme à* $\mathbf{x}_{j=0} = \mathbf{x}^b$
2. *Calcul du gradient en* $\mathbf{x}_j$ *par la relation (2.70)*
3. *Appel d'une méthode d'optimisation pour trouver* $\mathbf{x}_{j+1}$
4. $j \longleftarrow j+1$
5. *Si* $\|\nabla\mathcal{J}(\mathbf{x}_j)\| < \epsilon$*, arrêt. Sinon retour à 2.*

Utilisation classique du 3D-Var et variantes

Jusqu'à présent, nous n'avons pas parlé de l'aspect temporel des observations dans cette section. Pourtant, le 3D-Var a été utilisé de manière opérationnelle dans les centres de prévision météorologique. Nous allons expliquer rapidement son fonctionnement dans ce cadre.

L'assimilation est effectué sur des cycles (fenêtres temporelles de 12h par exemple). Pour un cycle donné, toutes les observations effectuées sur la fenêtre d'assimilation sont regroupées à l'instant t_c où l'assimilation est effectuée (généralement le début ou le milieu de la fenêtre temporelle). La méthode ne prend pas en compte l'origine temporelle des observations, elles sont alors comparées à l'état du modèle à un instant donné (n'étant pas celui des observations). La fonction coût se réécrit alors avec les termes suivants

$$\mathcal{J}^b(\mathbf{x}) = \frac{1}{2}\left(\mathbf{x}-\mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}^b\right)$$

$$\mathcal{J}^o(\mathbf{x}) = \frac{1}{2}\sum_{k=0}^{N_t-1}\left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}(t_c))\right)^T \mathbf{R}_k^{-1}\left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}(t_c))\right) \tag{2.72}$$

avec $\mathbf{y}_k^o$ une observation sur la fenêtre considérée, $\mathbf{R}_k$ sa matrice de covariance d'erreur et $\mathcal{H}_k$ l'opérateur d'observation associé. La figure 2.3 explicite le fonctionnement du 3D-Var en cycles.

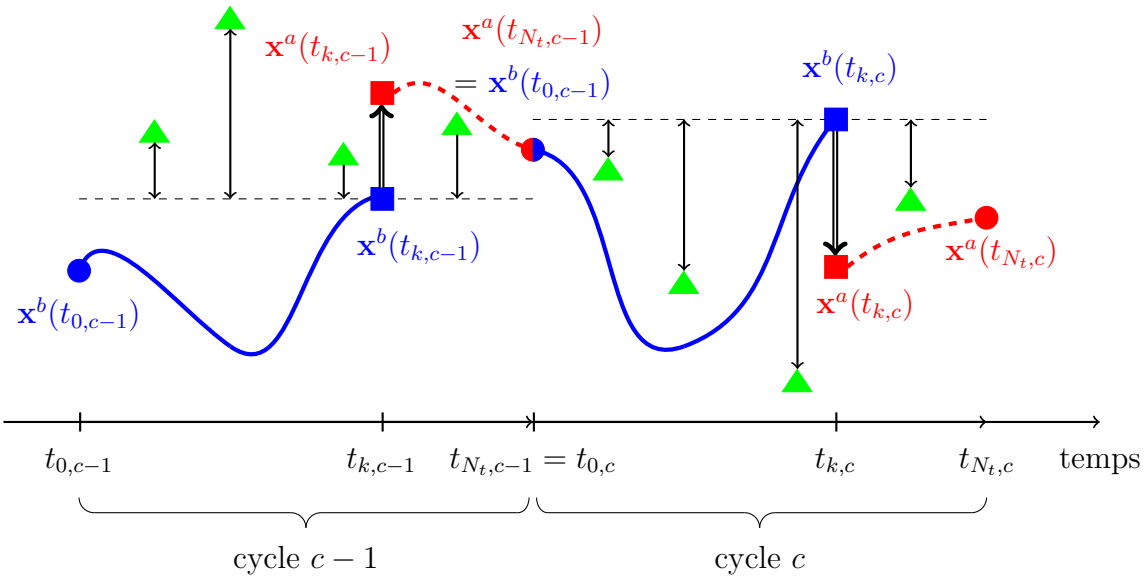

FIGURE 2.3 – Cycle du 3D-Var. A chaque cycle c, le modèle d'évolution propage l'ébauche du système au début de la fenêtre $\mathbf{x}^b(t_{0,c})$ jusquà l'instant $t_{k,c}$ pour produire $\mathbf{x}^b(t_{k,c})$ (courbe bleue). L'assimilation consiste ensuite à utiliser les observations (triangles verts) pour produire une analyse $\mathbf{x}^a(t_{k,c})$. Pour cela, il faut calculer la distance entre les observations (triangles verts) et leurs équivalents modèle $\mathcal{H}_k(\mathbf{x})$. Après la minimisation, $\mathbf{x}^a(t_{k,c})$ est propagé par le modèle jusqu'à la fin de la fenêtre (courbe rouge pointillée). L'état analysé $\mathbf{x}^a(t_{N_t,c})$ peut être utilisé comme ébauche au cycle suivant.

Du fait de la taille du problème et des non-linéarités des opérateurs d'observation (d'où une minimisation moins efficace), un certain nombre de variantes du 3D-Var a été inventé. Par exemple, le formalisme ***PSAS*** (Physical Space Assimilation System [5]) permet de minimiser la fonction coût uniquement dans l'espace dual des observations (de taille $n_{\mathbf{y}}$ beaucoup plus petite que la taille du vecteur d'état). En effet, le PSAS est basé sur un résultat de dualité lagrangienne : si l'opérateur d'observation $\mathcal{H}$ est linéaire, alors le 3D-Var admet une formulation duale (voir [Lions, 1968]) Le 3D-PSAS a été introduit par [Bennett et Thorburn, 1992] puis amélioré par [Amodei, 1995] (les 2 sous le nom de méthode des représenteurs).

Il est aussi possible de mieux prendre en compte l'origine temporelle des observations (sans faire intervenir le modèle plusieurs fois cependant) grâce au ***3D-FGAT*** (First Guess at Appropriate Time). Celui-ci sera exposé plus en détail après la description du 4D-Var, il constitue un pont entre le 3D-Var et le 4D-Var.

5. Bien que cette appellation puisse induire le lecteur en erreur (comme si le vecteur d'état corrigé dans le 3D-Var classique n'était pas physique ...), elle reste populaire. Selon [Bouttier et Courtier, 1999], le nom aurait été rendu populaire du fait d'une prononciation de l'acronyme PSAS proche de la prononciation anglaise de pizzas...

2.4.2 4D-Var

Au lieu de construire l'état optimal du système à un instant donné comme dans l'approche 3D-Var, nous cherchons la trajectoire optimale sur une fenêtre de temps donnée. Cette amélioration du 3D-Var introduite par [Le Dimet, 1980] ajoute la connaissance de l'évolution du système (le modèle) au système d'assimilation. Elle est appelée 4D-Var (pour 4 Dimensional VARiationnal assimilation) et connaît de nombreuses applications réalistes. Ainsi, le 4D-Var a remplacé le 3D-Var dans les systèmes de prévision météorologique du Centre européen de prévision météorologique à moyen terme en 1997 et à Météo France en 2000.

Bases du 4D-Var

Le 4D-Var consiste classiquement en l'estimation combinée de l'état initial du système et de certains paramètres inclus dans le vecteur $\mathbf{x}_0$. La fonction coût $\mathcal{J}$ à minimiser devient alors

$$\mathcal{J}(\mathbf{x}_0) = \mathcal{J}^b(\mathbf{x}_0) + \mathcal{J}^o(\mathbf{x}_0) \tag{2.73}$$

$\mathcal{J}^b$ représente l'écart par rapport à l'ébauche à l'état initial

$$\mathcal{J}^b(\mathbf{x}_0) = \frac{1}{2}\left(\mathbf{x}_0 - \mathbf{x}^b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}^b\right) \tag{2.74}$$

et $\mathcal{J}^o$ l'écart par rapport aux observations

$$\mathcal{J}^o(\mathbf{x}_0) = \frac{1}{2}\sum_{k=0}^{N_t-1}\left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k)\right)^T \mathbf{R}_k^{-1}\left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k)\right) \tag{2.75}$$

$\mathcal{J}^o$ fait intervenir l'état du système à différents instants.

$$\left\{\begin{array}{lcl} \mathbf{x}_{k+1} & = & \mathcal{M}_k(\mathbf{x}_k) \\ \mathbf{x}_0 & & \text{connu} \end{array}\right. \tag{2.76}$$

Si les opérateurs adjoints du modèle $\mathcal{M}_k$ notés $\mathbf{M}_k^T$ et les opérateurs adjoints des opérateurs d'observations $\mathcal{H}_k$ notés $\mathbf{H}_k^T$ sont connus, alors le gradient de la fonction coût utilisé pour la minimisation vaut

$$\begin{aligned}\nabla\mathcal{J}(\mathbf{x}_0) = \mathbf{B}^{-1}&\left(\mathbf{x}_0 - \mathbf{x}^b\right) \\ &- \sum_{k=0}^{N_t-1}\left(\mathbf{M}_0(\mathbf{x}_0)\right)^T \ldots \left(\mathbf{M}_{k-1}(\mathbf{x}_{k-1})\right)^T \left(\mathbf{H}_k(\mathbf{x}_k)\right)^T \mathbf{R}_k^{-1}\left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k)\right)\end{aligned} \tag{2.77}$$

Les opérateurs adjoints dépendent de la trajectoire du système.

Le calcul du gradient est généralement effectué par méthode adjointe. Celle-ci provient directement de la théorie du contrôle optimal des équations aux dérivées partielles [Lions, 1968]. Notons que cette technique a été introduite en météorologie par [Le Dimet et Talagrand, 1986].

La méthode adjointe simplifie le calcul du gradient. En effet, il peut être réécrit sous la forme suivante

$$\nabla \mathcal{J}(\mathbf{x}_0) = \mathbf{p}_0^* + \mathbf{B}^{-1}\left(\mathbf{x}_0 - \mathbf{x}^b\right) \tag{2.78}$$

avec l'ensemble des états adjoints $\mathbf{p}_k^*$ (ce sont des vecteurs de taille $n_{\mathbf{x}}$) définis de la manière suivante

$$\left\{ \begin{array}{rcl} \mathbf{p}_{N_t}^* & = & 0 \\ \mathbf{p}_k^* & = & (\mathbf{M}_k(\mathbf{x}_k))^T \mathbf{p}_{k+1}^* - (\mathbf{H}_k(\mathbf{x}_k))^T \mathbf{R}_k^{-1} (\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k)) \end{array} \right. \tag{2.79}$$

Remarquons que l'ensemble des états adjoints se calculent dans le sens contraire au sens du calcul des états directs du modèle (passage de t_{k+1} à t_k au lieu du passage classique de t_k à t_{k+1}).

L'algorithme du 4D-Var peut finalement se résumer de la manière suivante

Algorithme 2.4 (*4D-Var*)

1. *Initialisation de l'algorithme à* $\mathbf{x}_{0,j=0} = \mathbf{x}^b$
2. *Run du modèle direct par (2.76) en partant de* $\mathbf{x}_0 = \mathbf{x}_{0,j}$ *et stockage des innovations* $\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k)$
3. *Run du modèle adjoint associé par (2.79)*
4. *Calcul du gradient par la relation (2.78)*
5. *Appel d'une méthode d'optimisation pour trouver* $\mathbf{x}_{0,j+1}$
6. $j \longleftarrow j+1$
7. *Si* $\|\nabla \mathcal{J}(\mathbf{x}_{0,j})\| < \epsilon$, *arrêt. Sinon retour à 2.*

Dans le cas où le modèle est linéaire et parfait (erreur modèle nulle) et les opérateurs d'observations sont linéaires, l'état final de la trajectoire optimale construite par le 4D-Var est indentique au résultat obtenu par le filtre de Kalman (décrit à la section 2.5.1) au bout de la fenêtre d'assimilation. Cette équivalence est facilement démontrable par récurrence (voir par exemple [Li et Navon, 2001]). Enfin, la qualité de l'estimation à l'instant initial peut se calculer par $\mathbf{P}^a = (\mathbf{Hess}(\mathbf{x}_0^a))^{-1}$ avec

$$\mathbf{Hess}(\mathbf{x}_0) = \mathbf{B}^{-1} + \sum_{k=0}^{N_t-1} \mathbf{M}_0^T \ldots \mathbf{M}_{k-1}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k^T \mathbf{M}_{k-1}^T \ldots \mathbf{M}_0^T \tag{2.80}$$

Notons cependant que nous ne connaissons pas avec ce calcul la qualité de l'estimation sur toute la trajectoire reconstruite mais seulement au début de la fenêtre d'assimilation utilisée.

Utilisation pratique et variantes

Dans un cadre opérationnel (comme pour les systèmes de prévision météorologique), l'assimilation est effectuée par cycles. Pour un cycle donné, le 4D-Var utilise toutes les observations de la fenêtre de temps utilisée pour estimer au mieux l'état initial (au début

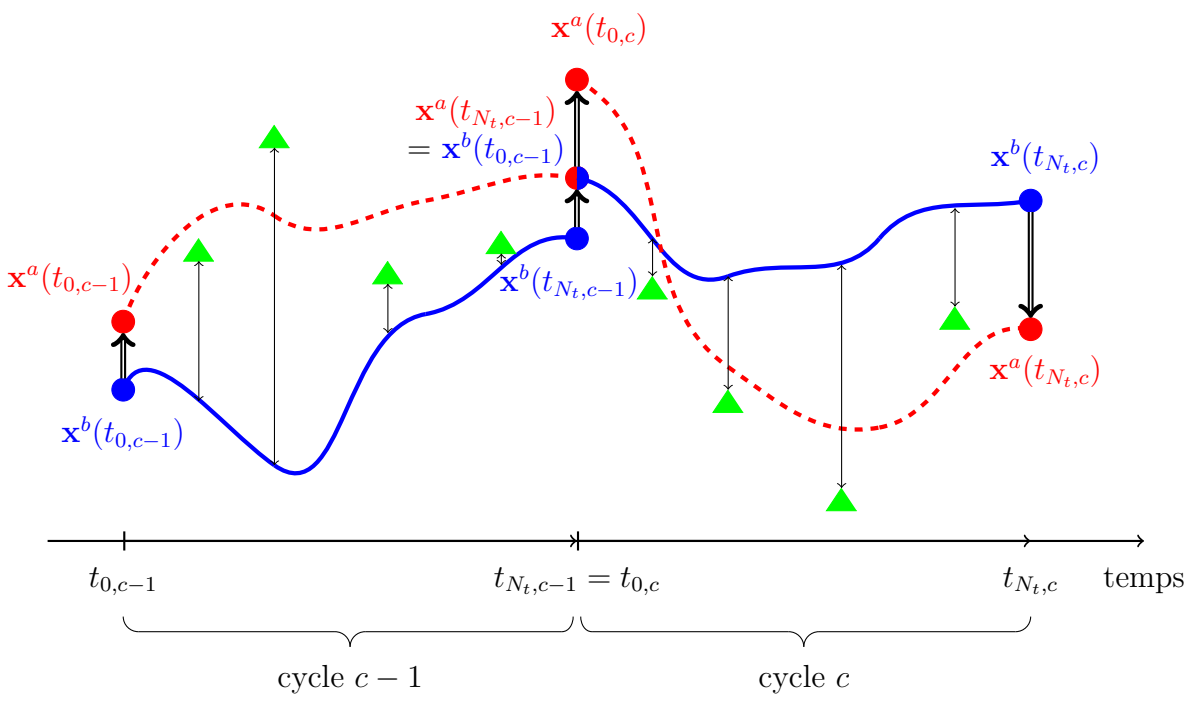

FIGURE 2.4 – Cycle du 4D-Var. A chaque cycle c, le modèle d'évolution propage l'ébauche du système au début de la fenêtre $\mathbf{x}^b(t_{0,c})$ jusqu'à l'instant final $t_{N_t,c}$ (courbe bleue). L'assimilation consiste ensuite à utiliser les observations (triangles verts) pour corriger l'état initial et produire une analyse $\mathbf{x}^a(t_{0,c})$. Pour cela, il faut calculer la distance entre les observations (triangles verts) et leurs équivalents modèle $\mathcal{H}_k(\mathbf{x}_k)$. Après la minimisation, $\mathbf{x}^a(t_{0,c})$ est propagé par le modèle jusqu'à la fin de la fenêtre, c'est la trajectoire optimale (courbe rouge pointillée). L'état analysé $\mathbf{x}^a(t_{N_t,c})$ (état final de la trajectoire reconstruite au cycle précédent) peut être utilisé comme ébauche au cycle suivant.

de la fenêtre étudiée). La trajectoire optimale consiste alors en la propagation par le modèle de cet état initial optimisé. La figure 2.4 illustre les cycles d'assimilation.

L'une des grandes questions en assimilation de données variationnelle est l'implémentation des codes adjoints. Il peut être fastidieux de les coder à la main (c'est néanmoins faisable), pour cela il « suffit » d'utiliser les règles de dérivation du code adjoint détaillées dans [Giering et Kaminski, 1998]. Le développeur peut aussi envisager l'utilisation d'un logiciel de différentiation automatique (consulter entre autre [Griewank et Walther, 2008] pour plus d'informations à ce sujet). Parmi les plus connus citons TAPENADE développé par Inria à Sophia Antipolis [Hascoet et Pascual, 2013].

Le calcul de la hessienne peut aussi se faire par méthode adjointe, au second ordre cette fois-ci, cf. notamment [Wang *et al.*, 1992] et [Le Dimet *et al.*, 2002] à ce sujet.

Comme dans le cas du 3D-Var, la question de la taille du problème reste valable pour le 4D-Var. Des variantes ont donc été trouvées pour contourner cet obstacle. Nous listons ici les plus connues :

— ***3D-FGAT*** (First Guess at Appropriate Time) : c'est une variante du 4D-Var dans laquelle les opérateurs adjoints $\mathbf{M}_k^T$ sont remplacés par l'identité. Le calcul du gradient ne nécessite plus que les adjoints des opérateurs d'observation. Le 3D-FGAT peut être aussi vu comme un 3D-Var amélioré car il prend en compte les temps d'observations contrairement au 3D-Var classique. Le 3D-FGAT constitue alors un intermédiaire très intéressant entre 3D-Var et 4D-Var.

— ***4D-PSAS*** : le formalisme PSAS décrit pour le 3D-Var est généralisable au 4D-Var (introduit par [Da Silva *et al.*, 1995], voir aussi [Courtier, 1997]).

— Variante incrémentale du 4D-Var (voir [Courtier *et al.*, 1994]).

— Stratégies de réduction d'ordre (consulter [Blayo *et al.*, 1998], [Durbiano, 2001] et [Robert, 2004] à ce propos). Ce point sera plus largement abordé dans le cas des filtres de Kalman.

Prise en compte de l'erreur modèle : 4D-Var à contrainte faible

La présentation précédente du 4D-Var supposait le modèle parfait (pas d'erreur commise par le modèle), il est alors dit « à contrainte forte » [Sasaki, 1970]. Cependant, cette hypothèse est généralement fausse. Le modèle est généralement de la forme

$$\left\{ \begin{array}{lcl} \mathbf{x}_{k+1} & = & \mathcal{M}_k(\mathbf{x}_k) + \boldsymbol{\eta}_k \\ \mathbf{x}_0 & & \text{connu} \end{array} \right. \tag{2.81}$$

Pour mieux prendre en compte les inexactitude, les vecteurs $\boldsymbol{\eta}_k$ peuvent être ajoutés au vecteur de contrôle afin de corriger les erreurs du modèle. Il faut alors introduire un terme correctif dans la fonction coût à minimiser :

$$\mathcal{J}(\mathbf{x}_0, \boldsymbol{\eta}_0, \ldots, \boldsymbol{\eta}_{N_t-1}) = \mathcal{J}^b(\mathbf{x}_0) + \mathcal{J}^o(\mathbf{x}_0) + \frac{1}{2} \sum_{k=0}^{N_t-2} \boldsymbol{\eta}_k^T \mathbf{Q}_k^{-1} \boldsymbol{\eta}_k \tag{2.82}$$

avec $\mathbf{Q}_k$ la matrice de covariance d'erreur modèle connue à l'instant t_k.

Cette version du 4D-Var est appelée 4D-Var ***à contrainte faible***. Remarquons que le vecteur de contrôle a une taille sensiblement augmentée par rapport au 4D-Var à contrainte forte (on ajoute $(N_t - 1) \times n_{\mathbf{x}}$ variables au vecteur de contrôle qui à la base a pour taille $n_{\mathbf{x}}$). D'autre part il est nécessaire de proposer une modélisation des $\mathbf{Q}_k$. Tout ceci rend son utilisation complexe. Cependant en utilisant des stratégies de réduction d'ordre sur les $\boldsymbol{\eta}_k$, il est possible de se ramener à un vecteur de contrôle de taille acceptable (voir [Durbiano, 2001], [Vidard, 2001] ou [Di Lorenzo *et al.*, 2007] pour des exemples d'utilisation du 4D-Var à contrainte faible avec stratégies de réduction d'ordre).

2.5 Filtre de Kalman

Le principal inconvénient de l'approche variationnelle 4D-Var est qu'elle ne donne pas directement d'estimation quant à la qualité de la trajectoire retrouvée (elle donne

seulement une estimation de la qualité de l'état initial retrouvé). Nous revenons alors à une approche plus statistique appelée filtre de Kalman et basée principalement sur le BLUE. Dans la suite de cette section, nous évoquons ce filtre ainsi que quelques unes de ses variantes, conçues pour alléger son fonctionnement.

2.5.1 Filtre de Kalman

Introduit par [Kalman, 1960], le filtre de Kalman généralise le BLUE au cas dépendant du temps, avec les hypothèses suivantes :

Hypothèses 2.2 (*Kalman*)

1. *Les hypothèses 2.1 utilisées pour le BLUE sont généralisées en temps.*
2. *A un instant donné t_k, l'ébauche du vecteur d'état est donné par la prévision $\mathbf{x}_k^f$, sa matrice de covariance d'erreur associée est notée $\mathbf{P}_k^f$*
3. *Les erreurs modèle $\boldsymbol{\eta}_k$ sont non-triviales, sans biais ($\mathbb{E}[\boldsymbol{\eta}_k] = \mathbf{0}$) et de matrices de covariance $\mathbf{Q}_k$ connues (matrices définies positives).*
4. *Les erreurs modèle et d'analyse sont indépendantes, ce qui veut dire que*
$$\mathbb{E}[\boldsymbol{\varepsilon}_k^a(\boldsymbol{\eta}_k)^T] = \mathbb{E}[\boldsymbol{\eta}_k(\boldsymbol{\varepsilon}_k^a)^T] = \mathbf{0} \tag{2.83}$$
5. *Le modèle de prévision est linéaire autour de $\mathbf{x}_k^a$, c'est à dire que pour tout $\mathbf{x}$ proche de $\mathbf{x}_k^a$*
$$\mathcal{M}_k(\mathbf{x}) - \mathcal{M}\left(\mathbf{x}_k^a\right) = \mathbf{M}_k\left(\mathbf{x} - \mathbf{x}_k^a\right) \tag{2.84}$$
avec $\mathbf{M}_k$ linéarisation de l'opérateur d'observation $\mathbf{M}_k$ autour de $\mathbf{x}_k^a$.

Comme pour le filtre bayésien, l'estimation de la valeur du vecteur d'état à un instant donné t_k s'effectue en 2 temps : la phase de prévision et la phase d'analyse (voir figure 2.5 pour illustration). Durant ces deux phases, nous évaluons aussi la matrice de covariance d'erreur de l'estimation effectuée.

Décrivons maintenant les deux phases du filtre de Kalman. La phase de prévision consiste à propager l'état analysé de l'instant précédent à l'instant courant. Ainsi $\mathbf{x}_{k+1}^f = \mathcal{M}_k(\mathbf{x}_k^a)$. L'erreur d'estimation commise vaut alors

$$\begin{aligned}
\boldsymbol{\varepsilon}_{k+1}^f &= \mathbf{x}_{k+1}^f - \mathbf{x}_{k+1} \\
&= \mathcal{M}_k(\mathbf{x}_k^a) - \mathcal{M}_k(\mathbf{x}_k) - \boldsymbol{\eta}_k \\
&= \mathbf{M}_k\boldsymbol{\varepsilon}_k^a - \boldsymbol{\eta}_k
\end{aligned}$$

et la matrice de covariance d'erreur de prévision

$$\begin{aligned}
\mathbf{P}_{k+1}^f &= \mathbb{E}\left[\boldsymbol{\varepsilon}_{k+1}^f\left(\boldsymbol{\varepsilon}_{k+1}^f\right)^T\right] \\
&= \mathbb{E}\left[\left(\mathbf{M}_k\boldsymbol{\varepsilon}_k^a - \boldsymbol{\eta}_k\right)\left(\mathbf{M}_k\boldsymbol{\varepsilon}_k^a - \boldsymbol{\eta}_k\right)^T\right] \\
&= \mathbf{M}_k\mathbb{E}\left[\boldsymbol{\varepsilon}_k^a\left(\boldsymbol{\varepsilon}_k^a\right)^T\right]\mathbf{M}_k^T + \mathbb{E}\left[\boldsymbol{\eta}_k\left(\boldsymbol{\eta}_k\right)^T\right] \\
&= \mathbf{M}_k\mathbf{P}_k^a\mathbf{M}_k^T + \mathbf{Q}_k
\end{aligned}$$

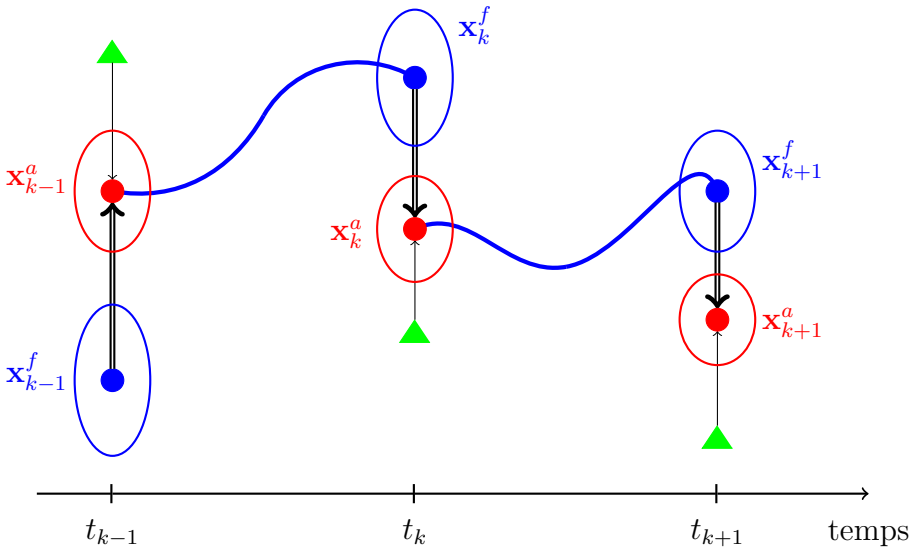

FIGURE 2.5 – Fonctionnement séquentiel du filtre de Kalman. A un instant donné t_k, nous disposons de $\mathbf{x}_k^f$ (point bleu) une prévision faite grâce au modèle à partir de l'état analysé au temps précédent, de sa matrice de covariance d'erreur $\mathbf{P}_k^f$ (symbolisée par l'ellipse bleue) et d'une observation $\mathbf{y}_k^o$ (triangle vert). La prévision est corrigée à l'aide de l'observation pour produire l'analyse $\mathbf{x}_k^a$ (point rouge) avec pour matrice de covariance d'erreur $\mathbf{P}_k^a$ (symbolisée par l'ellipse rouge).

La phase d'analyse quant à elle consiste à mettre à jour l'estimation grâce au BLUE en prenant pour ébauche la prévision. L'algorithme du filtre de Kalman se résume finalement à l'algorithme suivant.

Algorithme 2.5 (*Filtre de Kalman*)

Phase d'analyse :

1. *Calcul de la matrice de gain de Kalman*

$$\mathbf{K}_k = \mathbf{P}_k^f \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \tag{2.85}$$

2. *Calcul de l'état analysé*

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k \left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^f)\right) \tag{2.86}$$

3. *Calcul de la matrice de covariance d'erreur d'analyse*

$$\mathbf{P}_k^a = \left(\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}_k \mathbf{H}_k\right) \mathbf{P}_k^f \tag{2.87}$$

Phase de prévision :

1. *Calcul de l'état prévu*

$$\mathbf{x}_{k+1}^f = \mathcal{M}_k(\mathbf{x}_k^a) \tag{2.88}$$

2. *Calcul de la matrice de covariance d'erreur de prévision*

$$\mathbf{P}_{k+1}^f = \mathbf{M}_k \mathbf{P}_k^a \mathbf{M}_k^T + \mathbf{Q}_k \tag{2.89}$$

Dans le cas gaussien (les erreurs modèles, d'observation et d'ébauche sont gaussiennes), le filtre de Kalman permet de déterminer entièrement l'évolution de la loi de probabilité $p_{\mathbf{X}}(\mathbf{x}_k|\mathbf{y}^o_{0:k})$, c'est un filtre gaussien.

Si le modèle est en plus parfait ($\mathbf{Q}_k = \mathbf{0}$), le résultat obtenu par le filtre de Kalman au bout de la fenêtre d'assimilation est équivalent à l'état final de la trajectoire optimale construite par le 4D-Var. L'équivalence est facilement démontrable par récurrence (voir par exemple [Li et Navon, 2001]).

2.5.2 Filtre de Kalman étendu et limite de cette approche

Le filtre de Kalman vu précédemment peut être généralisé aux modèles et opérateurs d'observations non-linéaires. Pour cela, il suffit de prendre dans l'algorithme 2.5 $\mathbf{M}_k$ l'opérateur linéaire tangent de $\mathcal{M}_k$ en $\mathbf{x}^a_k$ et $\mathbf{H}_k$ l'opérateur linéaire tangent de $\mathcal{H}_k$ en $\mathbf{x}^f_k$. Par contre, l'état analysé n'est plus optimal. Cette version généralisée du filtre de Kalman est appelée ***filtre de Kalman étendu*** (EKF pour *Extended Kalman Filter*).

Le filtre de Kalman et l'EKF sont utilisable en pratique pour des systèmes d'assimilation de faible taille (inférieure à 100 ou 200). Pour des systèmes de plus grande taille, leur utilisation est plus compliquée. En effet, ces algorithmes nécessitent le stockage des matrices de covariance d'erreur $\mathbf{P}^f_k$ et $\mathbf{P}^a_k$ de tailles $n_{\mathbf{x}} \times n_{\mathbf{x}}$. $n_{\mathbf{x}}$ pouvant valoir plusieurs millions pour les systèmes opérationnels, le stockage est impossible dans ce cas et les calculs sont limités. Pour contourner cet obstacle, plusieurs stratégies existent. Nous allons maintenant évoquer l'une d'entre elle : la réduction de rang.

2.5.3 Filtres de Kalman de rang réduit

La littérature en assimilation de données regorge depuis les années 70 d'un certain nombre d'adaptations du filtre de Kalman. L'utilisation de filtres de rang réduit en assimilation de données pour les géosciences date quant à elle des années 90. Dans la suite, nous détaillons rapidement 2 filtres de rang réduit populaires : le filtre RRSQRT et le filtre SEEK. Il en existe d'autres (voir [Hoteit, 2001], [Cohn et Todling, 1996] et références mentionnées dans ces deux documents pour plus d'informations à ce sujet).

Filtre RRSQRT

Définit par [Verlaan et Heemink, 1995], le ***filtre RRSQRT*** (acronyme de *Reduced Rank SQuare RooT*) est basé sur la stratégie suivante : approcher les matrices $\mathbf{P}_k$ par le produit $\mathbf{S}_k\mathbf{S}^T_k$ avec $\mathbf{S}_k$ matrice racine carrée (d'où le *SQuare RooT*) approchée de $\mathbf{P}_k$. $\mathbf{S}_k$ contient les n_{mod} premiers modes de la matrice de covariance d'erreur (d'où le *Reduced Rank*). n_{mod} est pris généralement à une valeur beaucoup plus petite que $n_{\mathbf{x}}$. La principale subtilité de ce filtre est la phase de réduction nécessaire avant la phase de prévision afin de pouvoir tenir compte de l'erreur modèle (détails fournis dans l'algorithme 2.6).

Algorithme 2.6 (*Filtre RRSQRT*)

Initialisation

1. *ébauche du système à l'état initial* $\mathbf{x}_0^f$ *de matrice de covariance d'erreur* $\mathbf{P}_0^f$.
2. *Selection des* n_{mod} *premiers modes de* $\mathbf{P}_0^f \approx \mathbf{S}_0^f \left(\mathbf{S}_0^f\right)^T$.
3. *Selection des* $n_q < n_{mod}$ *premiers modes des matrices de covariances d'erreur modèle* $\mathbf{Q}_k \approx \widetilde{\mathbf{Q}}_k^{1/2} \left(\widetilde{\mathbf{Q}}_k^{1/2}\right)^T$.

Phase d'analyse

1. *Calcul de la matrice de gain*
   (a) *Projection des* n_{mod} *premiers modes de la matrice de covariance d'erreur de prévision dans l'espace des observations* $\mathbf{\Psi} = \left(\mathbf{H}_k \mathbf{S}_k^f\right)^T$
   (b) *Calcul de la matrice de gain approchée*
$$\mathbf{K}_k^* = \mathbf{S}_k^f \mathbf{\Psi} \left(\mathbf{\Psi}^T \mathbf{\Psi} + \mathbf{R}_k\right)^{-1}$$
2. *Calcul de l'état analysé*
$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* \left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^f)\right)$$
3. *Calcul approché des* n_{mod} *premiers modes de la matrice de covariance d'erreur d'analyse*
$$\mathbf{S}_k^a = \mathbf{S}_k^f \left(\mathbf{I}_{n_{mod}} - \mathbf{\Psi} \left(\mathbf{\Psi}^T \mathbf{\Psi} + \mathbf{R}_k\right)^{-1} \mathbf{\Psi}^T\right)^{1/2}$$
   La matrice racine carrée est facile à calculer car de taille $n_{mod} \times n_{mod}$.

Phase de réduction

1. *Diagonalisation de* $(\mathbf{S}_k^a)^T \mathbf{S}_k^a$ *(de taille* $n_{mod} \times n_{mod}$*),* $\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T = (\mathbf{S}_k^a)^T \mathbf{S}_k^a$.
2. *Réduction de* $\mathbf{S}_k^a$ *à ses* $n_{mod} - n_q$ *premiers modes.* $\widetilde{\mathbf{V}}$ *contient les* $n_{mod} - n_q$ *vecteurs propres associés aux* $n_{mod} - n_q$ *plus grandes valeurs propres de* $\mathbf{\Lambda}$ *d'où calcul de* $\widetilde{\mathbf{S}}_k^a = \mathbf{S}_k^a \widetilde{\mathbf{V}}$.

Phase de prévision

1. *Calcul de l'état prévu*
$$\mathbf{x}_{k+1}^f = \mathcal{M}_k(\mathbf{x}_k^a)$$
2. *Évaluation des* n_{mod} *premiers modes de la matrice de covariance d'erreur de prévision par concaténation des* $n_{mod} - n_q$ *premiers modes de* $\widetilde{\mathbf{S}}_k^a$ *propagés par le modèle linéaire tangent et des* n_q *premiers modes de* $\mathbf{Q}_k$ *contenus dans* $\widetilde{\mathbf{Q}}_k^{1/2}$
$$\mathbf{S}_{k+1}^f = \left[\mathbf{M}_k \widetilde{\mathbf{S}}_k^a \mathbf{M}_k^T \ , \ \widetilde{\mathbf{Q}}_k^{1/2}\right]$$

L'idée d'approcher les matrices de covariance $\mathbf{P}_k$ par $\mathbf{S}_k \mathbf{S}_k^T$ avec des matrices de rang réduit sera réutilisée abondamment par les filtres de Kalman d'ensemble déterministes

(cf. section 2.6.3).

Filtre SEEK

Le ***filtre SEEK*** (pour *Singular Evolutive Extended Kalman*) a été introduit par [Pham *et al.*, 1996]. Il repose sur la supposition que le rang des matrices de covariance d'erreur $\mathbf{P}_k$ stagne ou décroit au cours du temps (cette propriété est vraie dans le cas où le modèle utilisé est parfait, elle est généralement fausse dans le cas général). Sous cette hypothèse, si le rang r de la matrice de covariance initiale est faible par rapport à la taille de l'espace d'état $n_{\mathbf{x}}$, il le reste au cours du temps. L'idée du SEEK est donc d'approcher les matrices $\mathbf{P}_k$ par le produit $\mathbf{L}_k\mathbf{U}_k\mathbf{L}_k^T$ avec $\mathbf{L}_k$ la matrice contenant les directions principales des erreurs (de taille $n_{\mathbf{x}} \times r$) et $\mathbf{U}_k$ (de taille $r \times r$) la matrice de covariance réduite.

Au départ, la matrice $\mathbf{U}_k$ est diagonale mais elle évolue par la suite. L'un des avantages du SEEK est que $\mathbf{L}_k$ est modifé uniquement durant la phase de prévision. En effet, sous ce formalisme, il est possible de montrer que la phase d'analyse ne modifie pas l'espace vectoriel engendré par les directions principales des erreurs (même si ces directions changent elles-même). Enfin, l'erreur modèle est prise en compte en projetant la matrice de covariance d'erreur modèle $\mathbf{Q}_k$ sur l'espace engendré par les vecteurs colonnes de $\mathbf{L}_k$. L'opérateur de projection est $\mathbf{L}_k \left(\mathbf{L}_k^T\ \mathbf{L}_k\right)^{-1} \mathbf{L}_k^T$. Au final, l'algorithme du SEEK se formule de la manière suivante

Algorithme 2.7 (*Filtre SEEK*)

Initialisation

1. *ébauche du système à l'état initial* $\mathbf{x}_0^f$ *de matrice de covariance d'erreur* $\mathbf{P}_0^f$.
2. *Diagonalisation de* $\mathbf{P}_0^f \approx \mathbf{L}_0\mathbf{U}_0^f\mathbf{L}_0^T$. $\mathbf{U}_0^f$ *est diagonale et contient les* r *plus grandes valeurs propres de* $\mathbf{P}_0^f$.

Phase d'analyse

1. *Calcul de la matrice de covariance d'erreur réduite*
$$(\mathbf{U}_k^a)^{-1} = \left(\mathbf{U}_k^f\right)^{-1} + \mathbf{L}_k^T\ \mathbf{H}_k^T\ \mathbf{R}_k^{-1}\ \mathbf{H}_k\ \mathbf{L}_k$$
2. *Calcul de la matrice de gain approchée*
$$\mathbf{K}_k^* = \mathbf{L}_k\ \mathbf{U}_k^a\ \mathbf{L}_k^T\ \mathbf{H}_k^T\ \mathbf{R}_k^{-1}$$
3. *Calcul de l'état analysé*
$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* \left(\mathbf{y}_k^o - \mathcal{H}_k(\mathbf{x}_k^f)\right)$$

Phase de prévision

1. *Calcul de l'état prévu*
$$\mathbf{x}_{k+1}^f = \mathcal{M}_k(\mathbf{x}_k^a)$$
2. *Mise à jour des vecteurs engendrant l'espace des directions principales* $\mathbf{L}_{k+1} = \mathbf{M}_{k+1}\mathbf{L}_k$

3. Calcul de la matrice de covariance d'erreur réduite

$$\mathbf{U}_{k+1}^{f} = \mathbf{U}_{k}^{a} + \left(\mathbf{L}_{k+1}^{T}\mathbf{L}_{k+1}\right)^{-1} \mathbf{L}_{k+1}^{T}\, \mathbf{Q}_{k}\, \mathbf{L}_{k+1} \left(\mathbf{L}_{k+1}^{T}\mathbf{L}_{k+1}\right)^{-1}$$

Ce filtre sert de base au filtre SEIK qui est la première version déterministe d'un filtre de Kalman d'ensemble. Nous détaillons dans la section 2.6.3 ce filtre.

Les 2 filtres de rang réduit que nous venons de détailler sont bien sûr plus facilement utilisables que le filtre de Kalman étendu. Cependant, ils possèdent encore un inconvénient majeur : ils linéarisent le modèle et l'opérateur d'observation (comme l'EKF) pour estimer les matrices de covariances d'erreur. Cela peut créer des problèmes d'instabilité sur l'évalution de ces matrices (voir entre autre [Evensen, 1992], [Gauthier *et al.*, 1993], [Hoteit, 2001], et [Kushner, 1967]). Pour contourner ce problème, il faut pouvoir évaluer les matrices $\mathbf{P}_k$ en utilisant le modèle et les opérateurs d'observations non-linéaires. C'est l'objectif du filtre de Kalman d'ensemble que nous allons décrire dans la suite.

2.6 Filtre de Kalman d'ensemble

Comme nous l'avons vu précédemment, utiliser le filtre de Kalman ou ses variantes peut se révéler plus complexe en pratique notamment en terme d'évaluation de la matrice de covariance d'erreur d'estimation. [Evensen, 1994] introduit une approche de type Monte-Carlo permettant d'approcher les matrices $\mathbf{P}$ par un ensemble. Cette méthode appelée ***filtre de Kalman d'ensemble*** (EnKF pour *Ensemble Kalman Filter*) reçoit depuis les années 2000 l'attention d'une grande partie de la communauté assimilation. Rappelons en préambule que l'ensemble de l'EnKF ne converge vers la distribution du filtrage bayésien que dans le cas linéaire gaussien pour une taille d'ensemble infinie. Dans le cas non-linéaire, il converge vers une distribution qui n'est pas celle du filtrage bayésien (lire à ce sujet [Le Gland *et al.*, 2011] et [Mandel *et al.*, 2011]). Cependant, son succès en géosciences a eu tendance à éclipser ce fait.

Le fonctionnement de l'EnKF est séquentiel avec une phase de prévision et une phase d'analyse (voir la figure 2.6 pour un fonctionnement schématisé). Dans la suite, nous nous proposons de détailler rapidement les différentes versions existantes de l'EnKF. Nous évoquons ensuite les problèmes intrinsèques à cette méthode d'assimilation et les corrrectifs proposés pour contrebalancer leurs effets. La suite de cette section est entre autres basée sur [Evensen, 2007].

2.6.1 Echantillonner pour approcher

L'idée de base de l'EnKF introduite par [Evensen, 1994] est d'approcher l'estimation (prévision ou analyse) $\mathbf{x}_k$ et sa matrice de covariance d'erreur associée $\mathbf{P}_k$ par un ensemble de N_e états représentatifs du système

$$\left\{\mathbf{x}_k^{(1)}, \mathbf{x}_k^{(2)}, \ldots, \mathbf{x}_k^{(N_e)}\right\} \tag{2.90}$$

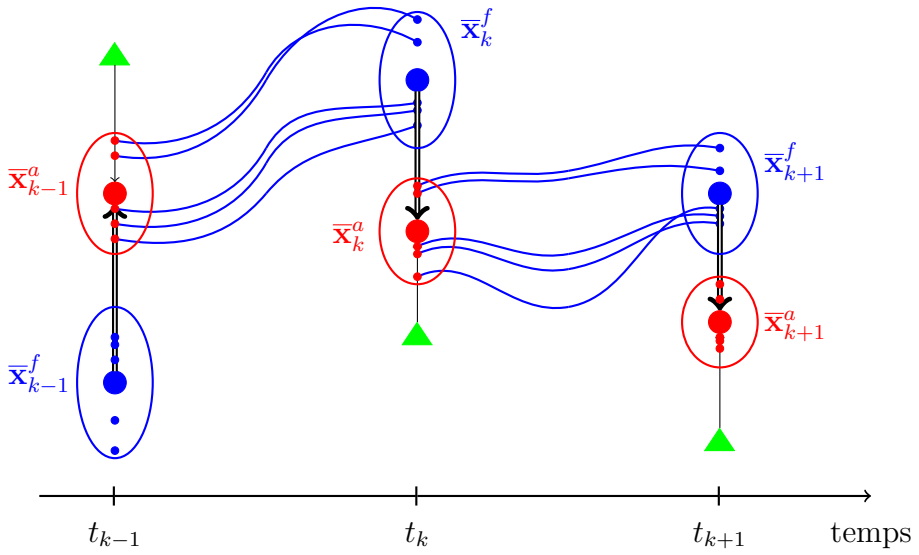

FIGURE 2.6 – Fonctionnement séquentiel du filtre de Kalman d'ensemble. A un instant donné t_k, nous disposons d'un ensemble d'états prévus (petits points bleus) de moyenne $\overline{\mathbf{x}}_k^f$ (gros point bleu) et de matrice de covariance d'erreur estimée $\mathbf{P}_{k,e}^f$ (symbolisée par l'ellipse bleue), et d'une observation $\mathbf{y}_k^o$ (triangle vert). L'ensemble prévu est issue de la propagation par le modèle de l'ensemble analysé au temps précédent. L'ensemble prévu est corrigé à l'aide de l'observation pour produire un nouvel ensemble d'analyses (petits points rouges) de moyenne $\overline{\mathbf{x}}_k^a$ (gros point rouge) et de matrice de covariance d'erreur estimée $\mathbf{P}_{k,e}^a$ (symbolisée par l'ellipse rouge).

N_e est en général bien plus petit que $n_{\mathbf{x}}$ (sinon l'approche est inutilisable). L'estimation $\mathbf{x}_k$ est approchée par la moyenne de l'ensemble

$$\overline{\mathbf{x}}_k = \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{x}_k^{(i)} \tag{2.91}$$

Notons $\mathbf{X}_k$ la matrice contenant dans ses colonnes les anomalies de l'ensemble

$$\mathbf{X}_k = \left[\mathbf{x}_k^{(1)} - \overline{\mathbf{x}}_k\ ,\ \mathbf{x}_k^{(2)} - \overline{\mathbf{x}}_k\ ,\ \ldots\ ,\ \mathbf{x}_k^{(N_e)} - \overline{\mathbf{x}}_k\right] \tag{2.92}$$

Cette matrice de taille $n_x \times N_e$ est au plus de rang $N_e - 1$ (car la somme des colonnes par construction est nulle). De son côté, la matrice $\mathbf{P}_k$ est évaluée par la matrice de covariance estimée de l'ensemble valant

$$\mathbf{P}_{k,e} = \frac{1}{N_e - 1}\sum_{i=1}^{N_e}\left(\mathbf{x}_k^{(i)} - \overline{\mathbf{x}}_k\right)\left(\mathbf{x}_k^{(i)} - \overline{\mathbf{x}}_k\right)^T = \frac{1}{N_e - 1}\mathbf{X}_k\mathbf{X}_k^T \tag{2.93}$$

Ceci est le fondement de l'EnKF. Il ne reste plus qu'à voir comment réaliser la phase de prévision et la phase d'analyse sous ce cadre.

2.6.2 Phase de prévision

Commençons d'abord par étudier la phase de prévision. Dans le cas du modèle parfait, cette phase est évidente. Il suffit de propager l'ensemble des états analysés à t_k par le

modèle jusqu'à l'instant d'assimilation suivant (approche Monte-Carlo). L'ensemble prévu est alors défini par la relation suivante pour chaque membre [6]

$$\mathbf{x}_{k+1}^{f\,(i)} = \mathcal{M}_k\left(\mathbf{x}_k^{a(i)}\right) \qquad i = 1, \ldots, N_e \tag{2.94}$$

Dans le cas où le modèle n'est pas parfait, plusieurs possibilités s'offrent à nous. Si nous connaissons (où avons une idée de la forme de) la loi de probabilité de l'erreur modèle $\boldsymbol{\eta}_k$, nous pouvons générer de manière aléatoire un ensemble de vecteurs d'erreur modèle $\left\{\boldsymbol{\eta}_k^{(1)}, \ldots, \boldsymbol{\eta}_k^{(N_e)}\right\}$. L'ensemble prévu est alors redéfini par

$$\mathbf{x}_{k+1}^{f\,(i)} = \mathcal{M}_k\left(\mathbf{x}_k^{a(i)}\right) + \boldsymbol{\eta}_k^{(i)} \qquad i = 1, \ldots, N_e \tag{2.95}$$

Maintenant, si nous ne connaissons que la matrice de covariance d'erreur modèle $\mathbf{Q}_k$, l'approche précédente est encore utilisable. Dans ce cas, nous supposons que l'erreur modèle suit une loi normale de moyenne nulle et de matrice de covariance $\mathbf{Q}_k$. Un ensemble de vecteurs d'erreur modèle suivant cette loi normale peut alors être généré. L'ensemble $\left\{\boldsymbol{\eta}_k^{(1)}, \ldots, \boldsymbol{\eta}_k^{(N_e)}\right\}$ a ainsi pour matrice de covariance estimée $\mathbf{Q}_{k,e}$, cette quantité tend vers la matrice $\mathbf{Q}_k$ lorsque N_e tend vers l'infini.

L'inconvénient de la pratique juste décrite est que la convergence de cette méthode est lente (ce fait réinterviendra dans la phase d'analyse). Nous ne prenons alors pas en compte exactement $\mathbf{Q}_k$. Une autre approche appelée *tirage exact au second ordre sous contrainte linéaire* (en anglais *Second-order-exact Sampling with linear contraint*) proposée par [Pham, 2001] permet de résoudre en partie ce problème. Elle consiste en théorie à calculer l'ensemble $\left\{\boldsymbol{\eta}_k^{(1)}, \ldots, \boldsymbol{\eta}_k^{(N_e)}\right\}$ de manière aléatoire ou déterministe tel que

$$\frac{1}{N_e}\sum_{i=1}^{N_e} \boldsymbol{\eta}_k^{(i)} = \mathbf{0} \qquad\qquad \frac{1}{N_e}\sum_{i=1}^{N_e} \boldsymbol{\eta}_k^{(i)}\left(\boldsymbol{\eta}_k^{(i)}\right)^T = \mathbf{Q}_k$$

sous la contrainte

$$\sum_{i=1}^{N_e} \boldsymbol{\eta}_k^{(i)}\left(\mathbf{x}_k^{f\,(i)} - \overline{\mathbf{x}}_k^f\right)^T = \mathbf{0}$$

Cette technique est cependant difficile à mettre en œuvre en pratique (lire les pages consacrées à ce sujet dans [Hoteit, 2001]) car nos connaissances sur les matrices $\mathbf{Q}_k$ et/ou l'erreur modèle sont généralement limitées. Cette remarque est aussi valable pour les autres méthodes décrites précédement.

Une autre possibilité pour mieux prendre en compte l'erreur modèle est de la contrôler (comme dans le 4D-Var à contrainte faible), il suffit alors de l'intégrer dans le vecteur d'état augmenté $\mathbf{x}$ comme un paramètre dont l'évolution temporelle serait inconnue.

Enfin, nous pouvons choisir de ne pas prendre en compte l'erreur modèle et de se contenter d'utiliser la relation (2.94) comme phase de prévision. Ceci a l'avantage d'être simple mais peut mener à sous-estimer les variances des estimations effectuées. Une procédure corrective appelée inflation a été introduite historiquement dans le but de contrebalancer la non-prise en compte de l'erreur modèle. Nous détaillons ses fondements en section 2.6.5.

6. On préfère parler de membre de l'ensemble dans le cadre de l'EnKF plutôt que de particule.

2.6.3 Phase d'analyse

La phase de prévision décrite plus haut s'avère facile à mettre en œuvre. Qu'en est-il de la phase d'analyse ? L'objectif de cette phase est d'approcher la phase d'analyse du filtre de Kalman dans le cas où l'opérateur d'observation est linéaire ($\mathcal{H}_k = \mathbf{H}_k$). L'idée la plus naturelle est de remplacer $\mathbf{P}_k^f$ par $\mathbf{P}_{k,e}^f$ dans le calcul de la matrice de gain. Dans la suite du paragraphe, nous omettons l'indice temporel k pour éviter une surchage en notation. La matrice de gain approchée vaut maintenant dans le cas linéaire

$$\mathbf{K}^* = \mathbf{P}_e^f \mathbf{H} \left(\mathbf{H} \mathbf{P}_e^f \mathbf{H}^T + \mathbf{R} \right)^{-1} \tag{2.96}$$

[Evensen, 1994] propose de construire l'ensemble analysé par la relation suivante

$$\mathbf{x}^{a(i)} = \mathbf{x}^{f(i)} + \mathbf{K}^* \left(\mathbf{y}^o - \mathbf{H} \mathbf{x}^{f(i)} \right) \tag{2.97}$$

L'ensemble construit de cette manière a pour moyenne $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{K}^* \left(\mathbf{y}^o - \mathbf{H} \overline{\mathbf{x}}^f \right)$ ce qui approche correctement l'état analysé construit par le filtre de Kalman.

Notons maintenant $\mathbf{X}^a$ la matrice contenant les anomalies de l'ensemble analysé, $\mathbf{X}^a$ est égal dans le cas linéaire à $(\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^* \mathbf{H}) \mathbf{X}^f$. Dans ce cas, la matrice de covariance estimée $\mathbf{P}_e^a$ vaut

$$\mathbf{P}_e^a = (\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^* \mathbf{H}) \mathbf{P}_e^f (\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^* \mathbf{H})^T \tag{2.98}$$

$\mathbf{P}_e^a$ ne tend donc pas vers $\mathbf{P}^a$ version Kalman lorsque la taille de l'ensemble tend vers l'infini. La méthode proposée par [Evensen, 1994] n'est donc pas satisfaisante, elle sous-estime clairement les variances de l'erreur d'estimation. La phase d'analyse doit donc être plus ingénieuse. Dans la suite de la section, nous exposons les différentes manières d'effectuer la phase d'analyse. Nous distinguons 2 versions : l'EnKF version stochastique et les versions déterministes.

EnKF version stochastique

Proposée indépendamment par [Burgers *et al.*, 1998] et [Houtekamer et Mitchell, 1998], cette version corrige les problèmes de la version de [Evensen, 1994] en bruitant les observations. Nous revenons au cas général, l'opérateur d'observation $\mathcal{H}$ n'est pas forcément linéaire. La phase d'analyse de l'EnKF version stochastique se déroule de la manière suivante

Algorithme 2.8 (*phase analyse EnKF version stochastique*)

1. *Projeter l'ensemble prévu dans l'espace des observations*
   — $\mathbf{y}^{f(i)} = \mathcal{H}\left(\mathbf{x}^{f(i)}\right)$ *pour* $i = 1, \ldots, N_e$.
   — $\overline{\mathbf{y}}^f$ *est la moyenne de cet ensemble et* $\mathbf{Y}^f$ *est la matrice contenant les anomalies de l'ensemble prévu dans l'espace des observations.*
2. *Simuler un ensemble* $\left\{ \boldsymbol{\varepsilon}^{o(1)}, \ldots, \boldsymbol{\varepsilon}^{o(N_e)} \right\}$ *de vecteurs aléatoires générés à partir de la loi normale de moyenne nulle et de matrice de covariance* $\mathbf{R}$. *Cet ensemble est de moyenne estimée* $\overline{\boldsymbol{\varepsilon}}^o$ *(ne vaut pas forcément 0) et de matrice de covariance*

estimée $\mathbf{R}_e$.

3. *Calcul de la matrice de gain approchée*

$$\mathbf{K}^* = \frac{1}{N_e - 1}\mathbf{X}^f \left(\mathbf{Y}^f\right)^T \left(\frac{1}{N_e - 1}\mathbf{Y}^f \left(\mathbf{Y}^f\right)^T + \mathbf{R}_e\right)^{-1} \tag{2.99}$$

4. *Calcul de l'ensemble analysé*

$$\mathbf{x}^{a(i)} = \mathbf{x}^{f(i)} + \mathbf{K}^* \left(\mathbf{y}^o + \boldsymbol{\varepsilon}^{o(i)} - \mathbf{y}^{f(i)}\right) \tag{2.100}$$

Vérifions maintenant si la phase d'analyse tend vers un résultat satisfaisant lorsque la taille de l'ensemble N_e tend vers l'infini. Bien sûr, nous avons les limites suivantes

$$\mathbf{P}_e^f \underset{N_e \to +\infty}{\longrightarrow} \mathbf{P}^f \qquad \mathbf{R}_e \underset{N_e \to +\infty}{\longrightarrow} \mathbf{R}$$

Réutilisons les notations $\mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x}))$ et $\mathbf{Cov}(\mathcal{H}(\mathbf{x}))$ déjà employées pour l'estimation BLUE (cf. section 2.3.2) et remarquons que

$$\frac{1}{N_e - 1}\mathbf{X}^f \left(\mathbf{Y}^f\right)^T \underset{N_e \to +\infty}{\longrightarrow} \mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x})) \qquad \frac{1}{N_e - 1}\mathbf{Y}^f \left(\mathbf{Y}^f\right)^T \underset{N_e \to +\infty}{\longrightarrow} \mathbf{Cov}(\mathcal{H}(\mathbf{x}))$$

Du coup, la matrice de gain admet pour limite

$$\mathbf{K}^* \underset{N_e \to +\infty}{\longrightarrow} \mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x})) \left(\mathbf{Cov}(\mathcal{H}(\mathbf{x})) + \mathbf{R}\right)^{-1}$$

La moyenne de l'ensemble analysé vaut $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{K}^* \left(\mathbf{y}^o + \overline{\boldsymbol{\varepsilon}}^o - \overline{\mathbf{y}}^f\right)$. Par construction, la moyenne de l'ensemble des bruits appliqués aux observations tend vers $\mathbf{0}$ quand N_e augmente. De même, $\overline{\mathbf{y}}^f$ tend vers $\mathbb{E}\left[\mathcal{H}(\mathbf{x})\right]$ et $\overline{\mathbf{x}}^f$ vers $\mathbf{x}^f$. Finalement, la moyenne de l'ensemble analysé admet pour limite

$$\overline{\mathbf{x}}^a \underset{N_e \to +\infty}{\longrightarrow} \mathbf{x}^f + \mathbf{K}\left(\mathbf{y}^o - \mathbb{E}\left[\mathcal{H}(\mathbf{x})\right]\right)$$

La moyenne de l'ensemble analysé approche donc l'estimateur BLUE dans le cas non-linéaire. D'autre part, [Burgers *et al.*, 1998] montre que

$$\mathbf{P}_e^a = \mathbf{P}_e^f - \frac{1}{N_e - 1}\mathbf{K}^*\mathbf{Y}^f \left(\mathbf{X}^f\right)^T + O(N_e^{-1/2})$$

Par passage à la limite,

$$\mathbf{P}_e^a \underset{N_e \to +\infty}{\longrightarrow} \mathbf{P}^f - \mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x})) \left(\mathbf{Cov}(\mathcal{H}(\mathbf{x})) + \mathbf{R}\right)^{-1} \left(\mathbf{Cov}(\mathbf{x}, \mathcal{H}(\mathbf{x}))\right)^T$$

La phase d'analyse du filtre de Kalman d'ensemble dans cette version approche donc le BLUE dans le cas non-linéaire, cela explique aussi pourquoi l'EnKF gère mieux les non-linéarités qu'un filtre de Kalman étendu. Notons que si l'opérateur d'observation est linéaire, nous retrouvons le BLUE dans le cas linéaire et alors l'EnKF approche le filtre de Kalman par une méthode Monte-Carlo.

Cette version de l'EnKF est appelée ***stochastique*** (synonyme d'aléatoire) car il nous faut ajouter du bruit simulé aléatoirement aux observations. Ainsi, le filtre ne prend en compte qu'une version approchée $\mathbf{R}_e$ de la matrice de covariance d'erreur d'observation $\mathbf{R}$. Il est possible de remplacer la loi normale générant ces vecteurs aléatoires si plus d'information est accessible sur la forme de l'erreur d'observation. Cependant cette version de l'EnKF n'est pas forcément des plus performantes du fait de la lenteur de la convergence des méthodes de Monte-Carlo. D'autres versions, déterministes cette fois-ci, prennent en compte directement la matrice $\mathbf{R}$ et se révèlent plus performantes en pratique (voir notamment [Whitaker et Hamill, 2002] et [Sakov et Oke, 2008b] pour des expériences comparatives sur les performances obtenues par différentes versions de l'EnKF).

Versions déterministes

Comme dans le cas du filtre de Kalman, il est possible d'approcher la matrice de covariance approchée $\mathbf{P}_e$ par un produit de matrices racines carrés en utilisant la relation (2.93). La phase d'analyse consiste alors en deux étapes pour produire l'ensemble analysé :

— mise à jour des anomalies $\mathbf{X}^a$.
— calcul de la moyenne $\overline{\mathbf{x}}^a$ et ajout de la moyenne à chaque colonne de de $\mathbf{X}^a$.

Les versions déterministes appelées généralement ***square root*** (pour racine carré) suivent toutes ce schéma. Nous nous restreignons dans le reste de ce paragraphe au cas où l'opérateur d'observation est linéaire (la projection des anomalies de l'ensemble prévu dans l'espace des observations $\mathbf{Y}^f$ vaut $\mathbf{H}\mathbf{X}^f$). La question qui se pose maintenant est celle du calcul de $\mathbf{X}^a$. Nous donnons ici une première variante appelée *Ensemble Transform Kalman Filter* (abrégée ETKF) introduite en premier lieu par [Bishop *et al.*, 2001]. Pour rappel, la matrice de covariance est calculée dans le filtre de Kalman par les relations (2.85) et (2.87). Si nous remplaçons les matrices $\mathbf{P}$ par $\frac{1}{N_e-1}\mathbf{X}\mathbf{X}^T$ dans les relations précédentes, le produit $\mathbf{X}^a\mathbf{X}^{aT}$ doit s'écrire sous la forme

$$\begin{aligned}
\mathbf{X}^a\mathbf{X}^{aT} &= \mathbf{X}^f\mathbf{X}^{fT} - \frac{1}{N_e-1}\mathbf{X}^f\mathbf{X}^{fT}\mathbf{H}^T\left(\frac{1}{N_e-1}\mathbf{H}\mathbf{X}^f\mathbf{X}^{fT}\mathbf{H}^T+\mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X}^f\mathbf{X}^{fT} \\
&= \mathbf{X}^f\left(\mathbf{I}_{N_e} - \frac{1}{N_e-1}\mathbf{X}^{fT}\mathbf{H}^T\left(\frac{1}{N_e-1}\mathbf{H}\mathbf{X}^f\mathbf{X}^{fT}\mathbf{H}^T+\mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X}^f\right)\mathbf{X}^{fT}
\end{aligned}$$

Cette relation est satisfaite si $\mathbf{X}^a = \mathbf{X}^f\mathbf{T}$ avec $\mathbf{T}$ matrice de transformation telle que

$$\mathbf{T}\mathbf{T}^T = \mathbf{I}_{N_e} - \frac{1}{N_e-1}\mathbf{X}^{fT}\mathbf{H}^T\left(\frac{1}{N_e-1}\mathbf{H}\mathbf{X}^f\mathbf{X}^{fT}\mathbf{H}^T+\mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X}^f \tag{2.101}$$

En utilisant la formule de Sherman-Morrison-Woodbury (2.67), la relation précédente est équivalente à

$$\mathbf{T}\mathbf{T}^T = \left(\mathbf{I}_{N_e} + \frac{1}{N_e-1}\mathbf{X}^{fT}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{X}^f\right)^{-1} \tag{2.102}$$

Nous choisissons $\mathbf{T}$ de taille $N_e \times N_e$, cette matrice est une racine carré de ce qui précède.

Remarque 2.1 (*Non-unicité dans le choix de* T)

Il n'y a pas d'unicité dans le choix de $\mathbf{T}$. *En effet, si nous remplaçons* $\mathbf{T}$ *par* $\mathbf{TU}$ *avec* $\mathbf{U}$ *une matrice orthogonale quelconque* ($\mathbf{U}\mathbf{U}^T = \mathbf{I}_{N_e}$), *la relation précédente reste inchangée.*

Il n'y a pas d'implémentation unique des filtres de Kalman d'ensemble déterministes (plusieurs ensembles différents peuvent avoir la même moyenne et la même matrice de covariance estimée). Cependant, ils sont formellement équivalents au sens qu'ils approchent correctement les quantités voulues (voir [Tippett *et al.*, 2003]). En outre, chacune des variantes proposées revient à chercher $\mathbf{X}^a$ sous la forme $\mathbf{X}^f\mathbf{T}$ ou $\tilde{\mathbf{T}}\mathbf{X}^f$ avec $\mathbf{T}$ et $\tilde{\mathbf{T}}$ matrices de transformation à déterminer (cf. [Sakov et Oke, 2008b]).

Nous avons déjà cité rapidement le filtre ETKF que nous verrons plus en détail dans le paragraphe suivant. Nous listons maintenant les autres principales variantes de ces filtres de Kalman d'ensemble déterministes afin que le lecteur puisse se repérer dans la jungle des abbréviations existantes dans le domaine :

— EnSRF (pour *Ensemble Square Root Filter*) [Whitaker et Hamill, 2002]

L'idée de base est de calculer une matrice de gain $\mathbf{K}^*$ tel que

$$\mathbf{P}_e^a = (\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^*\mathbf{H})\,\mathbf{P}_e^f = (\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^*\mathbf{H})\,\mathbf{P}_e^f\,(\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^*\mathbf{H})^T$$

Ceci est possible en utilisant la formule développée par [Andrews, 1968]. La matrice d'anomalies est alors mise à jour selon la relation $\mathbf{X}^a = (\mathbf{I}_{n_\mathbf{x}} - \mathbf{K}^*\mathbf{H})\,\mathbf{X}^f$.

— EAKF (pour *Ensemble Adjustment Kalman Filter*) [Anderson, 2001]

L'EAKF utilise deux décompositions en valeurs singulières :

$$\mathbf{P}_e^f = \mathbf{L}\,\mathbf{\Sigma}\,\mathbf{L}^T \qquad \text{et} \qquad \mathbf{I}_{N_e} + \frac{1}{N_e - 1}\mathbf{X}^{f^T}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\mathbf{X}^f = \mathbf{C}\,\mathbf{\Gamma}\,\mathbf{C}^T$$

$\mathbf{X}^a = \tilde{\mathbf{T}}\mathbf{X}^f$ est ensuite calculé avec $\tilde{\mathbf{T}} = \mathbf{X}^f\mathbf{C}\,\mathbf{\Gamma}^{-1/2}\,\mathbf{\Sigma}^{-1/2}\,\mathbf{L}$.

— SEIK (pour *Singular Evolutive Interpolated Kalman*) [Pham, 1996] et [Pham *et al.*, 1998]

Premier filtre de Kalman d'ensemble déterministe inventé, l'idée principale est de réutiliser la phase d'analyse proposée pour le filtre SEEK (voir algorithme 2.7). Nous reprenons les notations utilisées pour le SEEK avec la relation suivante le rang r vaut $N_e - 1$. Pour utiliser la phase d'analyse du SEEK, il faut évaluer les directions principales $\mathbf{L}$ et la matrice de covariance réduite $\mathbf{U}^f$. Ces matrices sont calculées en utilisant la matrice suivante

$$\mathbf{W} = \begin{bmatrix} & \mathbf{I}_r & \\ 0 & \cdots & 0 \end{bmatrix} - \frac{1}{r+1}\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}$$

$$\mathbf{L} = \left[\mathbf{x}^{f(1)} \dots \mathbf{x}^{f(r+1)}\right] \mathbf{W} \qquad \mathbf{U}^f = (r+1)\mathbf{W}^T\mathbf{W}$$

Il est possible d'intégrer l'erreur modèle dans $\mathbf{U}^f$ comme le fait le SEEK. La phase d'analyse nous produit une mise à jour de la matrice de covariance réduite $\mathbf{U}^a$. La matrice d'anomalies $\mathbf{X}^a$ vaut alors $\sqrt{r+1}\,\mathbf{L}\,\mathbf{U}^{a\,1/2}$ avec $\mathbf{U}^{a\,1/2}$ calculée à partir d'une décomposition (possiblement aléatoire) de $\mathbf{U}^a$. Notons enfin que le SEIK est implémentable sous la forme d'un ETKF comme le mentionne [Nerger *et al.*, 2012].

— ESSE (pour *Error Subspace Statistical Estimation*) [Lermusiaux et Robinson, 1999]

L'idée est de considérer la décomposition des matrices de covariance estimées en éléments propres. Nous décomposons $\mathbf{P}_e^f = \mathbf{L}^f\mathbf{\Sigma}^f\mathbf{L}^{f^T}$ et les matrices $\mathbf{L}$ et $\mathbf{\Sigma}$ sont mises à jour par la phase d'analyse. Pour cela, il est nécessaire d'effectuer la décomposition en éléments propres de la matrice

$$\tilde{\mathbf{\Sigma}} = \mathbf{S}\mathbf{\Sigma}^a\mathbf{S}^T = \mathbf{\Sigma}^f - \mathbf{\Sigma}^f\mathbf{L}^{f^T}\mathbf{H}^T\left(\mathbf{H}\mathbf{L}^f\mathbf{\Sigma}^f\mathbf{L}^{f^T}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{L}^f\mathbf{\Sigma}^f$$

$\mathbf{\Sigma}^a$ est obtenu durant la décomposition précédente et $\mathbf{L}^a = \mathbf{L}^f\mathbf{S}$. La matrice d'anomalies vaut alors $\mathbf{X}^a = \sqrt{N_e - 1}\,\mathbf{L}^f\,\mathbf{S}\,\mathbf{\Sigma}^{a\,1/2}$.

— DEnKF (pour *Deterministic Ensemble Kalman Filter*) [Sakov et Oke, 2008a]

L'idée de base est d'utiliser les deux premiers termes de la décomposition en série de Taylor de

$$\left(\mathbf{I}_{N_e} - \frac{1}{N_e-1}\mathbf{X}^{f^T}\mathbf{H}^T\left(\frac{1}{N_e-1}\mathbf{H}\mathbf{X}^f\mathbf{X}^{f^T}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X}^f\right)^{1/2}$$
$$\approx \mathbf{I}_{N_e} - \frac{1}{2(N_e-1)}\mathbf{X}^{f^T}\mathbf{H}^T\left(\frac{1}{N_e-1}\mathbf{H}\mathbf{X}^f\mathbf{X}^{f^T}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{X}^f$$

La matrice d'anomalies vaut alors $\mathbf{X}^a = \mathbf{X}^f\mathbf{T}$ avec la matrice de transformation $\mathbf{T}$ valant l'approximation effectuée plus haut. Bien que ne convergeant pas vers la bonne solution lorsque la taille de l'ensemble tend vers l'infini, les quelques expériences effectuées sur des modèles jouets semblent donner des résultats très proches des autres filtres déterministes.

Tous les filtres indiqués approchent de manière satisfaisante le même problème, ils sont fondamentalement équivalents. Néanmoins, certains sont plus coûteux en temps de calcul que d'autres (voir à ce sujet [Tippett *et al.*, 2003]). Nous choisissons dans cette thèse d'employer une version de l'ETKF. Nous l'étudions plus en détail dans le paragraphe suivant. Comme montré par [Tippett *et al.*, 2003], le coût algorithmique de la phase d'analyse d'un ETKF est en $O(N_e^3 + n_{\mathbf{x}}N_e^2 + n_{\mathbf{y}}N_e^2)$, il est donc linéaire en fonction de la taille du vecteur d'état et du nombre d'observation. Cette propriété est très appréciable car ces deux quantités sont les principaux facteurs limitants de l'utilisation des filtres de Kalman et leurs dérivés.

Ensemble Transform Kalman filter

Revenons au cas général (l'opérateur d'observation n'est pas nécessairement linéaire). Notons $\mathbf{Y}^f$ la matrice d'anomalies de l'ensemble prévu projeté dans l'espace des observations. L'idée de l'ETKF est de mettre à jour $\mathbf{X}^a = \mathbf{X}^f\mathbf{T}$ avec la matrice de transformation $\mathbf{T}$. En remplaçant $\mathbf{HX}^f$ par $\mathbf{Y}^f$ dans les équations (2.101) et (2.102), $\mathbf{T}$ obéit à la contrainte suivante

$$\mathbf{T}\mathbf{T}^T = \mathbf{I}_{N_e} - \frac{1}{N_e - 1}\mathbf{Y}^{f^T}\left(\frac{1}{N_e - 1}\mathbf{Y}^f\mathbf{Y}^{f^T} + \mathbf{R}\right)^{-1}\mathbf{Y}^f \tag{2.103}$$

La formule précédente est équivalente à (formule de Sherman-Morrison-Woodbury (2.67))

$$\mathbf{T}\mathbf{T}^T = \left(\mathbf{I}_{N_e} + \frac{1}{N_e - 1}\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f\right)^{-1} \tag{2.104}$$

Dans la première version de l'ETKF, [Bishop *et al.*, 2001] propose d'effectuer la décomposition en éléments propres de $\mathbf{I}_{N_e} + \frac{1}{N_e - 1}\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f = \mathbf{C}\,\mathbf{\Gamma}\,\mathbf{C}^T$ et de prendre $\mathbf{T} = \mathbf{C}\,\mathbf{\Gamma}^{-1/2}$. De son côté, [Evensen, 2004] effectue une décomposition en éléments propres directement sur $\mathbf{I}_{N_e} - \frac{1}{N_e - 1}\mathbf{Y}^{f^T}\left(\frac{1}{N_e - 1}\mathbf{Y}^f\mathbf{Y}^{f^T} + \mathbf{R}\right)^{-1}\mathbf{Y}^f = \mathbf{C}\,\mathbf{\Gamma}^{-1}\,\mathbf{C}^T$ et prend $\mathbf{T} = \mathbf{C}\,\mathbf{\Gamma}^{-1/2}$. Cette manière est souvent plus coûteuse que la méthode précédente. Elle nécessite d'inverser une matrice de taille $n_\mathbf{y} \times n_\mathbf{y}$ alors que l'autre proposition ne nécessite généralement que l'inversion d'une matrice de taille $N_e \times N_e$ (si la matrice $\mathbf{R}$ est prise diagonale).

Ces versions de l'ETKF (dites *one-sided*) souffrent de deux maux. Le premier est qu'elles ne préservent pas la moyenne calculée $\overline{\mathbf{x}}^a$, [Wang et Bishop, 2003] montrent que la somme des colonnes de la matrice d'anomalie $\mathbf{X}^a$ ainsi construite n'est pas nulle (elles sont orthogonales les unes par rapport aux autres dans l'espace des observations). [Sakov et Oke, 2008b] et [Livings *et al.*, 2008] soulignent l'intêret de préserver la moyenne calculée respectivement d'un point de vue pratique et d'un point de vue théorique.

La seconde déficience est une répartition des variances déséquilibrée. En effet, l'ETKF one-sided a la tendance de mal répartir les membres de l'ensemble autour de leur moyenne. Ainsi [Leeuwenburgh *et al.*, 2005] montre dans le cas d'un modèle scalaire que l'assimilation d'une seule observation réelle conduit à ce que toutes les anomalies de l'ensemble analysé sauf une valent zéro. Ce déséquilibre dans la répartition des variances entre les membres de l'ensemble conduit à la production de valeurs atypiques pour certains membres (en anglais : *outliers*) comme souligné par [Lawson et Hansen, 2004]. Ce défaut a été reconnu par [Evensen, 2004] et [Leeuwenburgh *et al.*, 2005]. Pour résoudre le problème, ces derniers proposent à la place de prendre comme matrice de transformation $\mathbf{T} = \mathbf{C}\mathbf{\Gamma}^{-1/2}\mathbf{U}$ avec $\mathbf{U}$ une matrice orthogonale aléatoire. Cette solution ne préserve toujours pas la moyenne de l'ensemble mais améliore les performances de l'ETKF en répartissant mieux les variances entre les membres de l'ensemble analysé.

Il est possible de prendre une matrice de transformation préservant la moyenne : L'ETKF symétrique (connu aussi sous le nom de *spherical simplex* ETKF) proposé par

[Wang *et al.*, 2004] et [Ott *et al.*, 2004] est une solution aux problèmes énoncés précédement. Pour cela, il suffit de poser $\mathbf{T} = \mathbf{C}\,\mathbf{\Gamma}^{-1/2}\,\mathbf{C}^T\,\mathbf{U}$ avec $\mathbf{U}$ une matrice orthogonale préservant la moyenne. $\mathbf{U}$ est choisie arbitrairement (en général on prend l'identité). Dans cette thèse, nous choisissons cette version de l'ETKF. Sa phase d'analyse se résume de la manière suivante

Algorithme 2.9 (*phase analyse ETKF*)

1. *Projeter l'ensemble prévu dans l'espace des observations*
   — $\mathbf{y}^{f(i)} = \mathcal{H}\left(\mathbf{x}^{f(i)}\right)$ *pour* $i = 1, \ldots, N_e$.
   — $\overline{\mathbf{y}}^f$ *est la moyenne de cet ensemble et* $\mathbf{Y}^f$ *est la matrice contenant les anomalies de l'ensemble.*
2. *Décomposition en valeurs singulières de la matrice symétrique de taille* $N_e \times N_e$
$$(N_e - 1)\mathbf{I}_{N_e} + \left(\mathbf{Y}^f\right)^T \mathbf{R}^{-1}\mathbf{Y}^f = \mathbf{C}\,\mathbf{\Gamma}\,\mathbf{C}^T$$
3. *Calcul de la matrice d'anomalies de l'ensemble analysé*
$$\mathbf{X}^a = \sqrt{N_e - 1}\,\mathbf{X}^f\,\mathbf{C}\,\mathbf{\Gamma}^{-1/2}\,\mathbf{C}^T\mathbf{U}$$
4. *Calcul de la moyenne de l'ensemble analysé*
$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{C}\,\mathbf{\Gamma}^{-1}\,\mathbf{C}^T\left(\mathbf{Y}^f\right)^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right)$$
5. *Calcul de l'ensemble analysé.*
$$\mathbf{x}^{a(i)} = \overline{\mathbf{x}}^a + \mathbf{X}^{a(i)}$$
   $\mathbf{X}^{a(i)}$ *correspond à la* i*ème colonne de* $\mathbf{X}^a$.

Équivalence variationnelle de l'ETKF

En plus des avantages déjà soulignés, l'ETKF a le bon goût de pouvoir se réécrire sous forme variationnelle ([Hunt *et al.*, 2007], [Harlim et Hunt, 2007]). Traditionnellement, la fonction coût à minimiser est celle du 3D-Var (cf. 2.4.1). Nous essayons de minimiser cette fonction coût adaptée à la phase d'analyse en prenant pour ébauche $\mathbf{x}^b$ la moyenne de l'ensemble prévu $\overline{\mathbf{x}}^f$ et pour matrice de covariance d'éauche $\mathbf{B}$ l'estimation de la matrice de covariance d'erreur de prévision $\mathbf{P}_e^f$. La fonction coût à minimiser se réécrit alors

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x})\right)^T\mathbf{R}^{-1}\left(\mathbf{y}^o - \mathcal{H}(\mathbf{x})\right) + \frac{1}{2}\left(\mathbf{x} - \overline{\mathbf{x}}^f\right)^T\left(\mathbf{P}_e^f\right)^{-1}\left(\mathbf{x} - \overline{\mathbf{x}}^f\right) \qquad (2.105)$$

Notons que, par construction, le rang de la matrice $\mathbf{P}_e^f$ est au plus $N_e - 1$. Dans le cas usuel où $N_e \ll n_{\mathbf{x}}$, la matrice $\mathbf{P}_e^f$ n'est donc pas inversible (la pseudo-inverse existe par contre). L'hypothèse principale de l'ETKF est de supposer que $\mathbf{x} - \overline{\mathbf{x}}^f$ vit dans le sous-espace $\mathcal{V}$ généré par les perturbations $\mathbf{x}^{f(i)} - \overline{\mathbf{x}}^f$. Le vecteur d'état $\mathbf{x}$ se réécrit sous la

forme suivante

$$\mathbf{x} = \overline{\mathbf{x}}^f + \sum_{k=1}^{N_e} w_k \left(\mathbf{x}^{f^{(i)}} - \overline{\mathbf{x}}^f \right) = \overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}$$

avec $\mathbf{w} = (w_1, \ldots, w_{N_e})^T$ le vecteur coordonnées dans le sous-espace d'ensemble $\mathcal{V}$. Notons que $\mathbf{x} - \overline{\mathbf{x}}^f = \mathbf{X}^f \mathbf{w}$. La fonction coût peut se mettre sous la forme suivante

$$\mathcal{J}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right) = \frac{1}{2}\left(\mathbf{y}^o - \mathcal{H}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right)\right)^T \mathbf{R}^{-1} \left(\mathbf{y}^o - \mathcal{H}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right)\right) + \frac{N_e - 1}{2} \mathbf{w}^T \mathbf{X}^{f^T} \left(\mathbf{X}^f \mathbf{X}^{f^T}\right)^{-1} \mathbf{X}^f \mathbf{w}$$

Remarquons à présent que $\mathcal{J}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right)$ est invariante selon la translation $\mathbf{w}^* = \mathbf{w} + \alpha \mathbf{v}$ avec α un réel et $\mathbf{v} = (1, \ldots, 1)^T$. En effet,

$$\mathbf{X}^f \mathbf{w}^* = \sum_{k=1}^{N_e} (w_k + \alpha) \left(\mathbf{x}^{f^{(i)}} - \overline{\mathbf{x}}^f \right) = \sum_{k=1}^{N_e} w_k \left(\mathbf{x}^{f^{(i)}} - \overline{\mathbf{x}}^f \right) + \alpha \sum_{k=1}^{N_e} \left(\mathbf{x}^{f^{(i)}} - \overline{\mathbf{x}}^f \right) = \mathbf{X}^f \mathbf{w}$$

En fait, le nombre de degrés de liberté dus à cette invariance est donné par la dimension du noyau de $\mathbf{X}^f$ (vue comme une application linéaire de $\mathbb{R}^{N_e}$ vers $\mathcal{V}$). D'après ce qui précède, il y a au moins un degré de liberté. Tout ceci implique la non-unicité de la minimisation de $\mathcal{J}$ sous les hypothèses précédentes. Pour résoudre ce problème, nous choisissons de minimiser la fonction coût $\mathcal{J}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right)$ sur les $\mathbf{w}$ ayant une projection orthogonale nulle sur le noyau de $\mathbf{X}^f$ soit

$$\left(\mathbf{I}_{N_e} - \mathbf{X}^{f^T} \left(\mathbf{X}^f \mathbf{X}^{f^T}\right)^{-1} \mathbf{X}^f \right) \mathbf{w} = 0$$

Néanmoins, effectuer une minimisation sous contrainte n'est pas forcément agréable. Pour éviter cela, nous ajoutons un terme de jauge $\mathcal{G}(\mathbf{w})$ à la fonction coût $\mathcal{J}$

$$\mathcal{G}(\mathbf{w}) = \frac{N_e - 1}{2} \mathbf{w}^T \left(\mathbf{I}_N - \mathbf{X}^{f^T} \left(\mathbf{X}^f \mathbf{X}^{f^T}\right)^{-1} \mathbf{X}^f \right) \mathbf{w}$$

ce qui nous donne la nouvelle fonction coût $\widetilde{\mathcal{J}}(\mathbf{w}) = \mathcal{J}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right) + \mathcal{G}(\mathbf{w})$. Elle vaut

$$\widetilde{\mathcal{J}}(\mathbf{w}) = \frac{1}{2}\left(\mathbf{y}^o - \mathcal{H}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right)\right)^T \mathbf{R}^{-1} \left(\mathbf{y}^o - \mathcal{H}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}\right)\right) + \frac{N_e - 1}{2} \mathbf{w}^T \mathbf{w}$$

Notons que si $\mathbf{w}^a$ minimise $\widetilde{\mathcal{J}}$, alors $\overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}^a$ minimise la fonction coût $\mathcal{J}$.

Afin de reformuler la phase d'analyse de manière plus explicite, linéarisons $\mathcal{H}$ autour de $\overline{\mathbf{x}}^f$. Si $\mathcal{H}$ est linéaire, le minimum exact de $\widetilde{\mathcal{J}}$ sera atteint. Sinon, si l'étendue de l'ensemble prévu n'est pas trop large, la linéarisation devrait être une approximation acceptable, similaire à l'approximation considérant que $\mathbf{x} - \overline{\mathbf{x}}^f$ vit dans le sous-espace $\mathcal{V}$ généré par les perturbations $\mathbf{x}^{f^{(i)}} - \overline{\mathbf{x}}^f$.

Pour cela, nous avons simplement besoin d'évaluer $\mathcal{H}$ dans un sous-espace restreint (d'après l'hypothèse de départ de ETKF). La manière la plus simple de linéariser $\mathcal{H}$

consiste à appliquer cet opérateur à chacun des membres de l'ensemble $\left\{\mathbf{x}^{f^{(i)}}\right\}$. Nous projetons l'ensemble prévu dans l'espace des observations. Le nouvel ensemble construit a pour moyenne $\overline{\mathbf{y}}^f$ et pour matrice d'anomalies $\mathbf{Y}^f$. L'approximation linéaire est effectuée de la manière suivante

$$\mathcal{H}\left(\overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w}\right) \approx \overline{\mathbf{y}}^f + \mathbf{Y}^f\mathbf{w}$$

Cela donne une nouvelle fonction coût quadratique approchée

$$\widetilde{\mathcal{J}}^*(\mathbf{w}) = \frac{1}{2}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right) + \frac{N_e - 1}{2}\mathbf{w}^T\mathbf{w} \tag{2.106}$$

Son minimum vaut $\mathbf{w}^a$ avec

$$\begin{aligned}
\mathbf{w}^a &= \widetilde{\mathbf{P}}^a\mathbf{Y}^{f^T}\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right) \\
\widetilde{\mathbf{P}}^a &= \left((N_e - 1)\,\mathbf{I}_{N_e} + \mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f\right)^{-1}
\end{aligned}$$

Dans l'espace d'état du modèle, la moyenne de l'ensemble analysé et sa matrice de covariance valent

$$\begin{aligned}
\overline{\mathbf{x}}^a &= \overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w}^a \\
\mathbf{P}^a &= \mathbf{X}^f\widetilde{\mathbf{P}}^a\mathbf{X}^{f^T}
\end{aligned}$$

Pour construire l'ensemble analysé, il suffit ensuite de prendre la matrice d'anomalies $\mathbf{X}^a$ sous la forme $\mathbf{X}^f\mathbf{T}$ avec la matrice de transformation $\mathbf{T}$ choisie telle que $\mathbf{T}\mathbf{T}^T = (N_e - 1)\widetilde{\mathbf{P}}^a$. Nous retrouvons bien le principe de l'ETKF tel que détaillé plus haut.

Pour conclure cette sous-section, nous avons vu que les filtres de Kalman d'ensemble, bien qu'ayant plusieurs déclinaisons possibles, étaient formellement équivalents. Notre choix du filtre ETKF symétrique comme filtre de Kalman d'ensemble déterministe (car plus efficace que la version stochastique) se justifie par plusieurs éléments. Le premier est son coût algorithmique relativement bas. D'autre part, sa formulation séduisante permet de bien comprendre les liens de la phase d'analyse avec le BLUE mais aussi les approches variationnelles.

2.6.4 Problèmes liés à l'utilisation de l'EnKF et solutions envisagées

Le filtre de Kalman d'ensemble possède un certain nombre d'avantages dans sa mise en place. Néanmoins, estimer les matrices de covariance d'erreurs de prévisions et d'analyse à l'aide d'ensembles de taille finie (N_e dépasse rarement 100) introduit un certain nombre d'erreurs d'échantillonnage. Quelles sont leurs implications ?

Pour avoir une première idée de leur impact, plaçons nous dans le cadre suivant (tiré de [Mallat *et al.*, 1998], [Furrer et Bengtsson, 2007], [Raynaud *et al.*, 2009] et [Bocquet, 2011]). Considérons un ensemble de vecteur d'états $\mathbf{x}^{(i)}$ de $\mathbb{R}^{n_\mathbf{x}}$ indépendants de taille N_e fixée

suivant une loi normale multidimensionnelle de matrice de covariance $\mathbf{P}$ inconnue. En géneral, $\mathbf{P}$ diffère de la matrice de covariance estimée $\mathbf{P}_e$. En utilisant le théorème d'Isserlis [7], il est possible de montrer que

$$\mathbb{E}\left([\mathbf{P}_e - \mathbf{P}]_{rs}^2\right) = \frac{1}{N_e}\left([\mathbf{P}]_{rr}[\mathbf{P}]_{ss} + [\mathbf{P}]_{rs}^2\right) \tag{2.107}$$

avec $r, s = 1, \ldots, n_{\mathbf{x}}$. $[\mathbf{P}]_{rs}$ correspond au coefficient de la matrice $\mathbf{P}$ ligne r, colonne s. Une idée de la preuve est disponible dans [Mallat *et al.*, 1998]. En particulier, la moyenne de l'erreur des variances estimées de $\mathbf{P}$ vaut (preuve détaillée cf. [Raynaud *et al.*, 2009])

$$\mathbb{E}\left([\mathbf{P}_e - \mathbf{P}]_{rr}^2\right) = \frac{2}{N_e}[\mathbf{P}]_{rr}^2 \tag{2.108}$$

Dans le cas des covariances à longue distance ($r \neq s$), $[\mathbf{P}]_{rs}$ est souvent négligeable dans la plupart des systèmes géophysiques. Cependant, les erreurs d'estimations sur $[\mathbf{P}]_{rs}$

$$\mathbb{E}\left([\mathbf{P}_e - \mathbf{P}]_{rs}^2\right) \sim \frac{1}{N_e}[\mathbf{P}]_{rr}[\mathbf{P}]_{ss} \tag{2.109}$$

restent loin d'être négligeables pour des ensembles de petite taille.

Si nous revenons au cas général, les erreurs d'échantillonnage entraînent l'apparition de covariance d'erreur artificielles (comme l'illustre l'équation (2.109)) mais aussi une sous-estimation des variances lors de la phase d'analyse. Cette sous-estimation, sous certaines hypothèses, a été demontrée par [van Leeuwen, 1999], [Furrer et Bengtsson, 2007] et [Sacher et Bartello, 2008] entre autres. Au bout de plusieurs itérations, cette sous-estimation peut conduire à la divergence du filtre. Ce risque était déjà souligné par [Julier et Uhlmann, 1997]. Le fait d'avoir seulement des observations spatiallement éparses peut aussi conduire aux mêmes symptômes dont la divergence du filtre (consulter à ce sujet [Harlim et Majda, 2010] et [Gottwald et Majda, 2013]).

Au fil des ans, différentes stratégies ont été développées pour réduire l'influence des erreurs d'échantillonnage. Les plus connues sont l'inflation et la localisation. Comme un médecin ne pouvant lutter contre la maladie, ces méthodes agissent non contre la source du problème mais luttent contre les symptômes (sous-évaluation des variances et apparition de covariances artificielles). Dans le cas de l'EnKF stochastique, une autre méthode appelée configuration multi-ensemble a été introduite par [Houtekamer et Mitchell, 1998], [Houtekamer et Mitchell, 2001]. Elle semble donner de bons résultats (voir notamment [Mitchell et Houtekamer, 2009] ou [Sacher et Bartello, 2008]), cependant les justifications empiriques accompagnant ces bons résultats laissent circonspect l'auteur de ces lignes (voir aussi critiques dans [van Leeuwen, 1999] et [Whitaker et Hamill, 2002]).

Dans la suite de cette section sur le filtre de Kalman d'ensemble, nous proposons un rapide descriptif de quelques méthodes permettant de lutter efficacement contre les erreurs d'échantillonnage. Nous décrivons notamment l'inflation et la localisation.

7. Ce théorème est aussi connu sous le nom de théorème de Wick. Il porte sur le calcul des moments d'ordre supérieur d'une loi normale multivariée.

2.6.5 Inflation

Introduite par [Anderson et Anderson, 1999] puis par [Hamill *et al.*, 2001], son objectif consiste à compenser la sous-estimation des variances de l'ensemble analysé par l'inflation des anomalies de l'ensemble, grâce à un coefficient réel ρ plus grand que 1. L'ensemble est alors corrigé de la manière suivante

$$\mathbf{x}^{(i)} = \rho\left(\mathbf{x}^{(i)} - \overline{\mathbf{x}}\right) + \overline{\mathbf{x}} \qquad i = 1, \ldots, N_e \tag{2.110}$$

Le nouvel ensemble est de même moyenne mais sa variance est augmentée afin de compenser la trop grande réduction de la variance durant la phase d'analyse. Ce correctif est usuellement appliqué juste avant la phase d'analyse. Dans le cas particulier de l'ETKF symétrique, il peut être judicieux de remplacer la procédure précédente par la modification de l'étape 2 de l'algorithme 2.9 que voici :

Décomposition en valeurs singulières de la matrice symétrique de taille $N_e \times N_e$

$$\frac{N_e - 1}{\rho^2}\mathbf{I}_{N_e} + \left(\mathbf{Y}^f\right)^T \mathbf{R}^{-1}\mathbf{Y}^f = \mathbf{C}\,\mathbf{\Gamma}\,\mathbf{C}^T \tag{2.111}$$

Cette manière de procéder moins coûteuse algorithmiquement (en $O(N_e)$ au lieu de $O(n_{\mathbf{x}}\,N_e)$) revient à enfler la matrice d'anomalies de l'ensemble prévu projeté dans l'espace des observations ($\mathbf{Y}^f \longleftarrow \rho\mathbf{Y}^f$ au lieu de $\mathbf{X}^f \longleftarrow \rho\mathbf{X}^f$ dans le premier cas). Lorsque l'opérateur d'observation est linéaire, les deux méthodes sont équivalentes. Notons enfin qu'une idée équivalente à l'inflation avait été proposée plus tôt par [Pham *et al.*, 1996] dans le cadre du filtre SEEK sous le nom de *forgetting factor*.

La technique de l'inflation a été introduite originellement pour compenser la non-prise en compte de l'erreur modèle dans la phase de prévision. En effet, enfler la matrice d'anomalie augmente la variance donc l'incertitude de l'estimation. Il est donc possible de remplacer la phase d'inflation (dite multiplicative) par l'ajout de perturbations stochastiques aux membres de l'ensemble (voir notamment [Mitchell et Houtekamer, 2000] et [Corazza *et al.*, 2003]). Dans ce cas, l'inflation est dite additive. Notons cependant que cette pratique reste moins développée que l'inflation multiplicative.

Estimation on-line du paramètre d'inflation

Le terme d'inflation ρ ou ρ^2 a été pendant longtemps réglé de manière manuelle (en gros, la valeur prise était celle qui donnait le plus satisfaction). Cependant, depuis ces dernières années se sont développées des techniques permettant l'évaluation de ce paramètre. Elles se décomposent en 2 catégories. Une première solution est d'estimer le paramètre d'inflation en utilisant une technique du maximum de vraisemblance inspirée par les travaux de [Dee, 1995]. Pour schématiser, cela consiste à trouver ρ_{MV} tel que celui-ci maximise la densité de probabilité $p(\rho|\mathbf{y}^o)$. Des détails de cette estimation (ainsi que des approximations faites) sont accessibles notamment dans [Mitchell et Houtekamer, 2000], [Anderson, 2007] et [Brankart *et al.*, 2010].

Une autre solution proposée par [Li *et al.*, 2009] est d'estimer ρ en utilisant les diagnostics Desroziers (voir [Desroziers *et al.*, 2005]). [Miyoshi, 2011] montre que les deux méthodes citées ont des résultats comparables sur un exemple scalaire.

Les deux méthodes citées contrebalancent normalement les erreurs d'échantillonnage et la non-prise en compte de l'erreur modèle lors de la phase de prévision. Cependant, il est impossible de distinguer ces deux sources d'erreurs (mon ensemble a t-il une taille vraiment petite ou mon modèle commet il beaucoup d'erreur ?) par les méthodes évoquées plus haut. Dans la suite, nous étudions une version modifiée du filtre de Kalman d'ensemble prenant en compte uniquement les erreurs d'échantillonnage. Nous verrons qu'elle reste profondemment liée à la procédure d'inflation.

2.6.6 Filtre de Kalman d'ensemble de taille finie

Introduit par [Bocquet, 2011], le filtre de Kalman d'ensemble de taille finie (EnFK-N) reformule l'étape d'analyse de l'EnKF. Plaçons nous dans l'approximation gaussienne où le filtre de Kalman est optimal. Ainsi nous supposons que que la densité a priori (avant analyse) $p(\mathbf{x}|\mathbf{x}^b, \mathbf{B})$ du vecteur d'état est une gaussienne de moyenne $\mathbf{x}^b$ et de matrice de covariance $\mathbf{B}$ supposées connues. Nous avons alors

$$p(\mathbf{x}|\mathbf{x}^b, \mathbf{B}) \propto \exp(-\mathcal{J}^b(\mathbf{x}))$$

avec $\mathcal{J}^b$ la partie de la fonction coût due à l'ébauche définie dans le 3D-Var. Dans le cas de l'EnKF, il est supposé implicitement que $\mathbf{x}^b = \overline{\mathbf{x}}^f$ et $\mathbf{B} = \mathbf{P}_e^f$. En pratique, nous ne connaissons qu'un échantillon de la loi $p(\mathbf{x}|\mathbf{x}^b, \mathbf{B})$ et $\overline{\mathbf{x}}^f$ et $\mathbf{P}_e^f$ diffèrent généralement des vraies valeurs $\mathbf{x}^b$ et $\mathbf{B}$. L'idée de l'EnKF-N est de considérer la densité $p\left(\mathbf{x}|\mathbf{x}^{f^{(1)}}, \ldots, \mathbf{x}^{f^{(N_e)}}\right)$ à la place de $p(\mathbf{x}|\overline{\mathbf{x}}^f, \mathbf{P}_e^f)$ pour ébauche et de trouver une nouvelle écriture de $\mathcal{J}^b(\mathbf{x})$. Si nous restons dans le cadre gaussien, nous avons

$$p\left(\mathbf{x}\left|\mathbf{x}^{f^{(1)}}, \ldots, \mathbf{x}^{f^{(N_e)}}\right.\right) \propto \int \exp(-\mathcal{L}(\mathbf{x}, \mathbf{x}^b, \mathbf{B}))\, p(\mathbf{x}^b, \mathbf{B})\, d\mathbf{x}^b\, d\mathbf{B} \tag{2.112}$$

avec $\mathcal{L}$ une certaine fonction connue. Pour évaluer l'intégrale, il faut faire une hypothèse sur $p(\mathbf{x}^b, \mathbf{B})$. Généralement, l'hypothèse la plus courante est celle de Jeffreys (c'est l'a priori apportant le moins d'information au sens statistique). Cependant, celui-ci n'est jamais satisfait en pratique. [Bocquet, 2011] relaxe la condition en prenant l'a priori de Jeffreys en supposant que $\mathbf{x}^b$ et $\mathbf{B}$ sont indépendants (voir à ce sujet [Jeffreys, 1961]). C'est à partir de ces hypothèses que [Bocquet, 2011] va construire la phase d'analyse de l'EnKF-N.

Version primale

[Bocquet, 2011] propose une implémentation notée ETKF-N basée sur l'EnKF-N et sur la supposition que $\mathbf{x} - \overline{\mathbf{x}}^f$ vit dans le sous-espace $\mathcal{V}$ généré par les perturbations $\mathbf{x}^{f^{(i)}} - \overline{\mathbf{x}}^f$. La fonction coût à minimiser n'est plus la formule (2.106) mais la suivante

$$\widetilde{\mathcal{J}}^*(\mathbf{w}) = \frac{1}{2}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right) + \frac{N_e}{2}\ln\left(\varepsilon_{N_e} + \mathbf{w}^T\mathbf{w}\right) \tag{2.113}$$

avec ε_{N_e} valant $1 + \frac{1}{N_e}$. Si nous supposons par contre $\mathbf{x}^b = \overline{\mathbf{x}}^f$, alors $\varepsilon_{N_e} = 1$. Son gradient vaut

$$\nabla\widetilde{\mathcal{J}}^*(\mathbf{w}) = -\mathbf{Y}^{f^T}\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right) + N_e\,\frac{\mathbf{w}}{\varepsilon_{N_e} + \mathbf{w}^T\mathbf{w}} \tag{2.114}$$

Enfin, la Hessienne nécessaire pour construire la matrice d'anomalies de l'ensemble analysé vaut

$$\widetilde{\mathbf{Hess}}^*(\mathbf{w}) = \mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + \frac{N_e}{(\varepsilon_{N_e} + \mathbf{w}^T\mathbf{w})^2}\left((\varepsilon_{N_e} + \mathbf{w}^T\mathbf{w})\mathbf{I}_{N_e} - 2\,\mathbf{w}\mathbf{w}^T\right) \tag{2.115}$$

La phase d'analyse de l'ETKF-N se résume donc à

Algorithme 2.10 (*phase analyse ETKF-N*)

1. *Projeter l'ensemble prévu dans l'espace des observations*
   — $\mathbf{y}^{f(i)} = \mathcal{H}\left(\mathbf{x}^{f(i)}\right)$ *pour* $i = 1, \ldots, N_e$.
   — $\overline{\mathbf{y}}^f$ *est la moyenne de cet ensemble et* $\mathbf{Y}^f$ *est la matrice contenant les anomalies de l'ensemble.*
2. *Trouver le vecteur de taille* N_e $\mathbf{w}^a$ *minimisant la fonction coût définie en* (2.113)
3. *Calcul de la moyenne de l'ensemble analysé* $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w}^a$
4. *Décomposition en valeurs singulières de la matrice symétrique de taille* $N_e \times N_e$

$$\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + \frac{N_e}{(\varepsilon_{N_e} + \mathbf{w}^{a^T}\mathbf{w}^a)^2}\left((\varepsilon_{N_e} + \mathbf{w}^{a^T}\mathbf{w}^a)\mathbf{I}_{N_e} - 2\,\mathbf{w}^a\mathbf{w}^{a^T}\right) = \mathbf{C}\,\mathbf{\Gamma}\,\mathbf{C}^T$$

5. *Calcul de la matrice d'anomalies de l'ensemble analysé*

$$\mathbf{X}^a = \sqrt{N_e - 1}\,\mathbf{X}^f\,\mathbf{C}\,\mathbf{\Gamma}^{-1/2}\,\mathbf{C}^T\mathbf{U}$$

6. *Calcul de l'ensemble analysé.*

$$\mathbf{x}^{a(i)} = \overline{\mathbf{x}}^a + \mathbf{X}^{a(i)}$$

   $\mathbf{X}^{a(i)}$ *correspond à la* i*ème colonne de* $\mathbf{X}^a$.

Notons enfin que bien que la fonction coût admette un unique minimum global, elle n'admet pas en général de minimum local unique. La minimisation ne permet donc pas d'accéder forcément au minimum unique.

Version duale et lien avec l'inflation

[Bocquet et Sakov, 2012] permet de résoudre le problème précédent grâce à un résultat de dualité établi dans cet article. Minimiser (2.113) revient à trouver le réel (et non plus un vecteur) ζ^a minimisant sur l'intervalle $\left]0, \frac{N_e}{\varepsilon_{N_e}}\right]$ la fonction suivante

$$\mathcal{D}(\zeta) = \frac{1}{2}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right)^T\left(\mathbf{R} + \mathbf{Y}^f\zeta^{-1}\mathbf{Y}^{f^T}\right)^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right) + \frac{\varepsilon_{N_e}\zeta}{2} + \frac{N_e}{2}\ln\frac{N_e}{\zeta} - \frac{N_e}{2} \tag{2.116}$$

Le minimum $\mathbf{w}^a$ de (2.113) vaut alors

$$\mathbf{w}^a = \left(\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + \zeta^a\,\mathbf{I}_{N_e}\right)^{-1}\mathbf{Y}^{f^T}\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right) \tag{2.117}$$

et la moyenne de l'ensemble analysé vaut $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X}^f \mathbf{w}^a$. Dans le cas ETKF avec inflation, le terme ζ^a est remplacé par $\frac{N_e-1}{\rho^2}$. L'ETKF-N met à jour la moyenne de l'ensemble de la même manière qu'un ETKF avec inflation. Par contre, les matrices d'anomalies diffèrent légèrement car

$$\zeta^a \, \mathbf{I}_{N_e} \neq \frac{N_e}{(\varepsilon_{N_e} + \mathbf{w}^{aT}\mathbf{w}^a)^2} \left((\varepsilon_{N_e} + \mathbf{w}^{aT}\mathbf{w}^a)\mathbf{I}_{N_e} - 2\,\mathbf{w}^a\mathbf{w}^{aT} \right)$$

L'algorithme se réécrit alors

Algorithme 2.11 (*phase analyse ETKF-N version duale*)

1. *Projeter l'ensemble prévu dans l'espace des observations*
   — $\mathbf{y}^{f(i)} = \mathcal{H}\left(\mathbf{x}^{f(i)}\right)$ *pour* $i = 1, \ldots, N_e$.
   — $\overline{\mathbf{y}}^f$ *est la moyenne de cet ensemble et* $\mathbf{Y}^f$ *est la matrice contenant les anomalies de l'ensemble.*
2. *Trouver le réel* ζ^a *minimisant* (2.116) *sur l'intervalle* $\left]0, \frac{N_e}{\varepsilon_{N_e}}\right]$
3. *Calculer le vecteur de taille* N_e

$$\mathbf{w}^a = \left(\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + \zeta^a \, \mathbf{I}_{N_e}\right)^{-1} \mathbf{Y}^{f^T}\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right)$$

4. *Calcul de la moyenne de l'ensemble analysé* $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w}^a$
5. *Décomposition en valeurs singulières de la matrice symétrique de taille* $N_e \times N_e$

$$\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + \frac{N_e}{(\varepsilon_{N_e} + \mathbf{w}^{aT}\mathbf{w}^a)^2} \left((\varepsilon_{N_e} + \mathbf{w}^{aT}\mathbf{w}^a)\mathbf{I}_{N_e} - 2\,\mathbf{w}^a\mathbf{w}^{aT} \right) = \mathbf{C}\,\mathbf{\Gamma}\,\mathbf{C}^T$$

6. *Calcul de la matrice d'anomalies de l'ensemble analysé*

$$\mathbf{X}^a = \sqrt{N_e - 1}\,\mathbf{X}^f\,\mathbf{C}\,\mathbf{\Gamma}^{-1/2}\,\mathbf{C}^T\mathbf{U}$$

7. *Calcul de l'ensemble analysé.*

$$\mathbf{x}^{a(i)} = \overline{\mathbf{x}}^a + \mathbf{X}^{a(i)}$$

$\mathbf{X}^{a(i)}$ *correspond à la* i*ème colonne de* $\mathbf{X}^a$.

Réaffirmons enfin que l'ETKF-N bien que proche de l'inflation ne règle que la sous-estimation des variances dues aux erreurs d'échantillonnage et non à la non-prise en compte de l'erreur modèle.

2.6.7 Localisation

Les deux dernières sous-sections présentaient des méthodes permettant de contrebalancer l'effet des erreurs d'échantillonnage sur les variances. Cependant, elles ne règlent pas le problème de l'apparition de covariances artificielles à longue distance. Ce problème a été résolu par l'utilisation de la localisation. Ce terme recouvre en fait deux

raisonnements différents : la localisation de la covariance et l'analyse locale. Cependant, [Sakov et Bertino, 2011] montre que les deux approches conduisent à des résultats similaires sous certaines conditions.

Localisation de la matrice de covariance

Introduite par [Houtekamer et Mitchell, 2001] et [Hamill *et al.*, 2001], la localisation de covariance a pour but de faire disparaître les covariances à longue distance (qu'elles soient artificielles ou non) au sein des matrices de covariance d'erreurs prévues et analysées estimées. Pour cela, elle utilise un produit d'Hadamard [8] de l'estimation de la covariance d'erreur de prévision estimée avec une fonction de corrélation à support local. Le produit d'Hadamard de 2 matrices $\mathbf{A}$ et $\mathbf{B}$ noté $\mathbf{A} \circ \mathbf{B}$ vaut une matrice $\mathbf{C}$ de même dimension que les deux autres et ses coefficients valent

$$[\mathbf{C}]_{rs} = [\mathbf{A}]_{rs} [\mathbf{B}]_{rs}$$

En pratique, le produit d'Hadamard intervient dans le calcul de la matrice de gain

$$\mathbf{K}^* = \left[\left(\rho \circ \mathbf{P}_e^f\right) \mathbf{H}^T\right] \left[\mathbf{H}_k \left(\rho \circ \mathbf{P}_e^f\right) \mathbf{H}^T + \mathbf{R}\right]^{-1}$$

La notation $\rho \circ \mathbf{B}$ correspond au produit d'Hadamard de la matrice de corrélation $\mathbf{A}$ avec la matrice de covariance $\mathbf{B}$. Les éléments $[\mathbf{A}]_{rs}$ sont obtenus grâce à la fonction de corrélation ρ à support local de la manière suivante. Soient les points p_r et p_s de l'espace correspondants respectivement à la localisation correspondant à la ligne r de la matrice $\mathbf{B}$ et à la colonne s de cette même matrice,

$$[\mathbf{A}]_{rs} = \rho\left(\|p_r - p_s\|\right)$$

La fonction suivante définie dans [Gaspari et Cohn, 1999] sur $\mathbb{R}^+$ est généralement usitée

$$\rho(r,c) = \begin{cases} -\frac{1}{4}\left(\frac{r}{c}\right)^5 + \frac{1}{2}\left(\frac{r}{c}\right)^4 + \frac{5}{8}\left(\frac{r}{c}\right)^3 - \frac{5}{3}\left(\frac{r}{c}\right)^2 + 1 & 0 \leq r \leq c \\ \frac{1}{12}\left(\frac{r}{c}\right)^5 - \frac{1}{2}\left(\frac{r}{c}\right)^4 + \frac{5}{8}\left(\frac{r}{c}\right)^3 + \frac{5}{3}\left(\frac{r}{c}\right)^2 - 5\left(\frac{r}{c}\right) + 4 - \frac{2}{3}\left(\frac{r}{c}\right)^{-1} & c < r \leq 2c \\ 0 & r > 2c \end{cases}$$

Cette fonction à support local a le bon goût d'approcher une gaussienne. Elle dépend d'un paramètre c réglable selon les besoins de l'utilisateur. Un exemple de l'effet du produit d'Hadamard sur $\mathbf{P}_e^f$ est visible figure 2.7. D'autre part, ρ est relativement régulière et si l'opérateur d'observation est simple (ici il est linéaire donc pas de problème), le calcul de la matrice de gain se simplifie

$$\mathbf{K}^* = \left[\rho \circ \left(\mathbf{P}_e^f \mathbf{H}^T\right)\right] \left[\rho \circ \left(\mathbf{H}\mathbf{P}_e^f \mathbf{H}^T\right) + \mathbf{R}\right]^{-1}$$

L'inconvénient principal de la localisation de la covariance est qu'elle nécessite le calcul des matrices de covariance estimées. La version du filtre de Kalman d'ensemble que nous avons choisie (soit l'ETKF) est incompatible avec la localisation de la covariance. Il nous faut donc utiliser une autre méthode : l'analyse locale.

8. Le produit d'Hadamard est aussi connu sous le nom de produit de Schur.

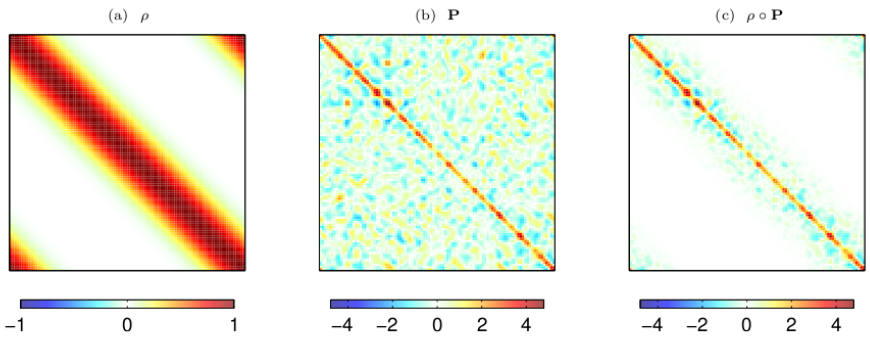

Fig. 1 Example of effect of CL on the state error covariance. See Appendix for details.

FIGURE 2.7 – Illustration de l'impact de la localisation de covariance [Sakov et Bertino, 2011].

Analyse locale

Proposé entre autres par [Anderson, 2001], [Evensen, 2003] et [Ott *et al.*, 2004], l'analyse locale consiste pour un point de grille donné à ne selectionner que les observations proches de ce point.

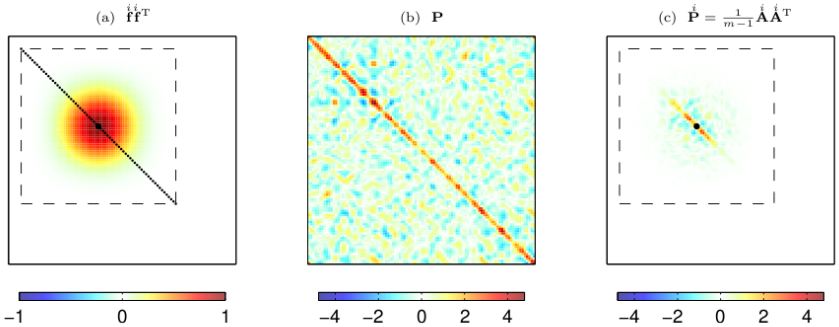

Fig. 2 Example effect of LA on the state error covariance used for for updating the ith ($i = 40$) element of the state vector (see Appendix for details). The dashed rectangle shows the boundaries of the local domain.

FIGURE 2.8 – Illustration de l'impact de l'analyse locale [Sakov et Bertino, 2011].

[Hunt *et al.*, 2007] souligne que cela revient à localiser la matrice $\mathbf{R}$ et baptiste l'ETKF avec localisation LETKF (le L voulant dire *Local*). Il note aussi que cette méthode peut conduire à des discontinuités dans l'analyse lorsqu'une observation quitte par l'exemple l'ensemble des observations utilisées localement. Pour remédier à cet effet, il suffit de remplacer dans la phase d'analyse $\mathbf{R}$ par $\rho^{-1} \circ \mathbf{R}$ (les variances des observations proches de la frontière de la zone d'observation concernée sont augmentées) comme le propose [Hunt *et al.*, 2007]. C'est cette version de la localisation que nous utilisons avec l'ETKF.

La localisation a quelques inconvénients : elle peut faire disparaître des covariances à longue distance réellement existantes et elle nécessite un réglage manuel. Cependant, il existe des tentatives de localisation adaptative (voir à ce sujet [Bishop et Hodyss, 2009]). Malgré tout, cette procédure a prouvé de nombreuses fois son efficacité et son utilisation

avec le filtre de Kalman d'ensemble est devenue standard.

2.7 Approches hybrides

Si nous effectuons un bilan rapide des méthodes abordées dans ce chapitre, remarquons que le filtrage particulaire et l'approche 4D-Var ont l'avantage d'être parfaitement défini mathématiquement. L'approche particulaire est basé sur le filtrage bayésien et le 4D-Var sur le contrôle optimal. De son côté, le filtre de Kalman d'ensemble n'est parfaitement défini que dans le cas linéaire gaussien. Notons d'ailleurs que l'équivalence avec le 4D-Var n'est valable que dans le cas d'un problème de lissage linéaire gaussien. Hors usuellement, le filtre de Kalman d'ensemble est utilisé dans le cadre d'un problème de filtrage dans un contexte non-linéaire. D'autre part, sa phase d'analyse approche l'estimation BLUE par un ensemble qui, selon les cas, peut être un bon estimateur mais qui n'est pas forcément optimal (au sens minimum de variance ou maximum a posteriori). Cependant, dans nombre de cas, cette approche a donné de bons résultats. Il peut être donc avantageux de combiner cette approche avec les autres méthodes décrites avant, ce sont des ***approches hybrides***.

Différentes approches hybrides existent. Des travaux portent sur la combinaison d'une méthode d'ensemble au 4D-Var, on parle de 4D-EnVar (citons dans cette optique l'existence des filtres de Kalman d'ensemble itératifs, consulter notamment [Sakov *et al.*, 2012] et [Bocquet et Sakov, 2012]).

Il peut être aussi intéressant d'allier les avantages pratiques du filtre de Kalman d'ensemble avec la rigueur mathématique du filtre particulaire (voir [Papadakis *et al.*, 2010], [Beyou *et al.*, 2013] et [Nakano et Ueno, 2012]). C'est l'objectif du Weighted Ensemble Kalman Filter (WEnKF) introduit par [Papadakis *et al.*, 2010]. L'application du WEnKF à des problèmes de grande dimension montre une amélioration sensible des résultats obtenus par rapport à un simple filtre de Kalman d'ensemble. Le WEnKF a aussi été appliqué avec succès à l'assimilation d'images [Beyou *et al.*, 2013]. Le papier indiqué précédemment souligne aussi que la précision des méthodes d'ensemble ou du filtre particulaire dépend fortemement du nombre de membre de l'ensemble.

Chapitre 3

Un problème simplifié de paléoclimatologie

Est il possible de reconstruire le climat du passé de la Terre connaissant l'histoire des calottes polaires ? Cette question anime de nombreux glaciologues depuis bien des années. De nombreux travaux en témoignent. A ce sujet, l'étude de carottes de glace prélevées au Groenland et en Antarctique est un outil formidable contenant nombre d'informations. Par exemple, l'identification des isotopes de l'oxygène (rapport oxygène 18 sur oxygène 16 classiquement) contenus dans les bulles d'air emprisonnées dans la glace permet d'évaluer comment a varié la température ambiante à la surface de la calotte dans le passé. Les carottes de glace nous fournissent aussi d'autres informations en terme d'évolution de concentration de dioxyde de carbone, de méthane, ... Ces carottes peuvent nous permettre de remonter très loin dans le temps (jusqu'à 800 000 ans pour le forage EPICA à Dôme C en Antarctique [EPICA community members, 2004] et plus de 135 000 ans pour le forage NEEM au Groenland [NEEM community members, 2013]).

Les informations contenues dans les carottes de glaces sont précieuses sur le climat. Cependant, elles ne donnent pas toute l'histoire de l'évolution des calottes polaires. A quoi ressemblaient elles il y a 20 000 ans ? Pensons notamment aux calottes Fennoscandienne et des Laurentides existant à cette époque et disparues depuis (voir figure 3.1). L'apport des modèles d'évolution de calotte est utile pour mieux évaluer l'impact des scénarii climatiques sur ces objets et leur répercution en terme de niveau des mers. Ainsi, selon les dernières estimations, le niveau des mers était 120 m plus bas il y a 20 000 ans que maintenant. Cette estimation peut être obtenue grâce aux carottes de sédiments marins. En effet, ces dernières permettent de reconstruire l'évolution du volume de glace sur Terre sur plus de 500 000 ans. Dans ce contexte, l'assimilation de donnée peut apporter des informations supplémentaires sur l'évolution des paramètres climatiques grâce aux modèle et aux observations disponibles. Néanmoins, elle n'a que très rarement été utilisée dans ce contexte (voir à ce sujet la section 1.5.2).

L'objectif du présent chapitre est de montrer l'apport de nos travaux à la question de la reconstruction de paramètres climatiques à partir de données sur l'évolution du volume de glace. Nous mettons en place une méthode adjointe sur un cas synthétique (expériences jumelles) qui bien que simplifié permet de mettre en lumière les difficultés de l'utilisation

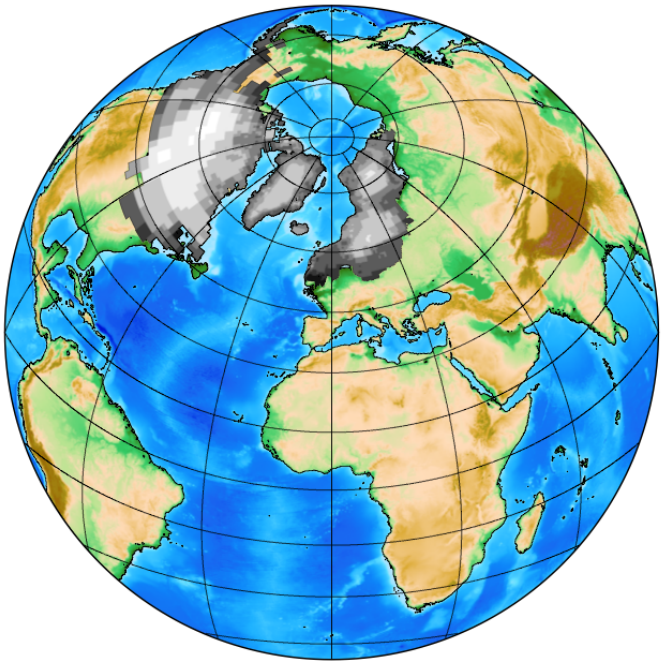

FIGURE 3.1 – Topographie de la surface de la Terre au dernier maximum glaciaire (estimation). Focus sur l'hémisphère Nord (figure de C. Dumas). Outre la présence de la calotte groenlandaise, remarquons l'existence de calottes en Amérique du Nord (calotte des Laurentides) et en Europe du Nord (Fennoscandie).

de l'assimilation de données dans ce contexte. Une partie des travaux décrits ici ont fait l'objet de la publication d'un acte de conférence [Bonan *et al.*, 2012].

3.1 Description du problème

Le problème étudié ici est la reconstruction du paramètre climatique F_{clim} intervenant dans le calcul de la température en surface (voir équation (1.74)) dans notre modèle d'évolution de calotte polaire simplifié Winnie. La fenêtre de temps des expériences est 20 000 ans et les cas considérés ont les propriétés simplificatrices suivantes :

— Socle rocheux plat fixe : $B_{soc}(x) = 0$.
— Épaisseur de la calotte à l'instant initial nulle : $H(x, t = 0) = 0$.
— Pas de frottement basal.

Winnie est donc en pure SIA. Résumons ensuite le problème d'assimilation posé en répondant aux questions suivantes :

— « *Que contrôlons nous ?* » : le seul paramètre climatique F_{clim} variant en temps (l'état initial de la calotte est connu). Nous proposons deux manières de le modéliser :

1. forçage sinusoïdal
$$F_{\text{clim}} = a_0 + a_1 \sin\left(\frac{2\pi t}{a_2}\right) \tag{3.1}$$
Dans ce cas, nous cherchons le vecteur $T_{\text{clim}} = [a_0, a_1, a_2]^T$.

2. interpolation linéaire par morceaux : nous cherchons le vecteur
$$T_{\text{clim}} = [F_{\text{clim}}(t_0 = 0), F_{\text{clim}}(t_1), \ldots, F_{\text{clim}}(t_K = 20\,000)]^T$$
Nous linéarisons ce forçage entre 2 t_k, c'est à dire que nous supposons qu'entre 2 t_k le forçage climatique vaut
$$F_{\text{clim}}(t) = F_{\text{clim}}(t_k) + \frac{t - t_k}{t_{k+1} - t_k}\left(F_{\text{clim}}(t_{k+1}) - F_{\text{clim}}(t_k)\right) \qquad t \in [t_k, t_{k+1}] \tag{3.2}$$

— « *Qu'observons nous ?* » : dans un premier temps, l'évolution de l'épaisseur de la calotte puis l'évolution du volume au cours du temps. Pour une ligne d'écoulement, nous définissons le volume Vol d'une calotte comme l'intégrale de l'épaisseur le long de la ligne d'écoulement.
$$\text{Vol}(t) = \int H(x,t)\,dx \tag{3.3}$$

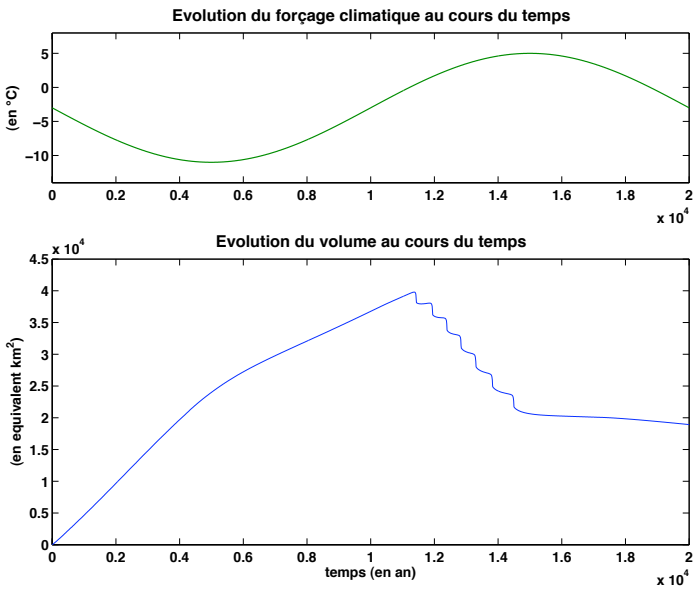

FIGURE 3.2 – Exemple d'évolution du volume au cours du temps (courbe bleue en bas) obtenue en utilisant le forçage climatique sinusoïdal (courbe verte en haut). La forme d'escalier visible lors de la fonte de la calotte est due à la discrétisation spatiale lâche utilisée (si la discrétisation était plus fine, ce phénomène serait toujours présent mais il faudrait zoomer sur la courbe pour l'apercevoir). L'évolution de la calotte en elle-même est visible 3.3.

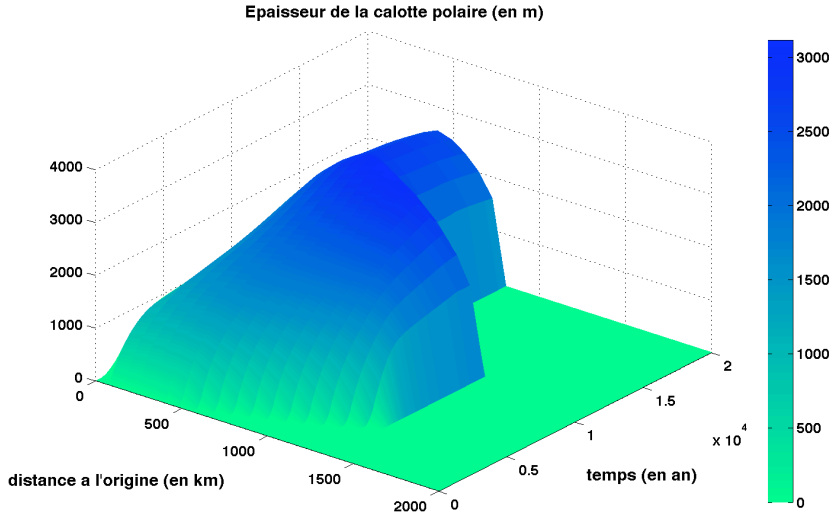

FIGURE 3.3 – Évolution de l'épaisseur de la calotte polaire utilisant le forçage climatique sinusoïdal de la figure 3.2.

— « *Quel système d'assimilation utilisons nous ?* » : assimilation de données variationnelle. Dans le cas où nous assimilons le volume de glace, nous cherchons à minimiser la fonction coût suivante

$$\mathcal{J}(T_{\text{clim}}) = \mathcal{J}^b(T_{\text{clim}}) + \frac{1}{2}\int_{t_0}^{t_K} \|\text{Vol}(t) - \text{Vol}^o(t)\|^2 \, dt \tag{3.4}$$

avec $\text{Vol}^o(t)$ l'évolution du volume observée obtenue à partir d'un run de modèle référence (expérience jumelle). Les observations ne sont pas corrélées en temps. Le terme d'ébauche $\mathcal{J}^b$ est classiquement un terme de rappel par rapport à l'ébauche du vecteur T_{clim} noté T^b_{clim} pondéré par une matrice de covariance d'erreur $\mathbf{B}$.

$$\mathcal{J}^b(T_{\text{clim}}) = \frac{1}{2}\left\|T_{\text{clim}} - T^b_{\text{clim}}\right\|^2_{\mathbf{B}} = \frac{1}{2}\left(T_{\text{clim}} - T^b_{\text{clim}}\right)^T \mathbf{B}^{-1}\left(T_{\text{clim}} - T^b_{\text{clim}}\right) \tag{3.5}$$

Dans le cas où le forçage climatique suit une interpolation linéaire par morceaux, la matrice $\mathbf{B}$ est telle que la covariance $[\mathbf{B}]_{ij}$ de l'erreur d'ébauche entre les temps t_i et t_j est de la forme exponentielle suivante

$$[\mathbf{B}]_{ij} = c_1 \, e^{-c_2|t_i - t_j|} \tag{3.6}$$

Les coefficients c_1 et c_2 ne dépendent pas de i et j. Il ne reste plus qu'à voir comment minimiser la fonction coût.

3.2 4D-Var et modèle adjoint

Nous choisissons de minimiser $\mathcal{J}$ en utilisant un algorithme de descente quasi-Newton (nous utilisons la fonction Matlab pour la fonction `fminunc`, voir la documentation Matlab

pour plus de détails). Cet algorithme fait intervenir le gradient $\nabla\mathcal{J}$ de la fonction coût. Ce dernier est calculé par méthode adjointe. Le modèle adjoint a été dérivé manuellement. Nous détaillons dans cette section le code adjoint de Winnie et les tests de validation complets de cette dérivation.

3.2.1 Quelques détails sur la construction d'un modèle adjoint discret

Plus de détails sur la construction du modèle adjoint discret peuvent être trouvés dans [Giering et Kaminski, 1998].

Adjoint continu

Nous cherchons à minimiser la fonction coût $\mathcal{J}(\mathbf{x})$ définie de manière générale par

$$\mathcal{J}(\mathbf{x}) = \|\mathcal{F}(\mathbf{x}) - F^o\|^2 \tag{3.7}$$

en utilisant le gradient de $\mathcal{J}$ en un point $\mathbf{x}_0$. En utilisant le développement de Taylor (en supposant que la fonction coût soit assez régulière), nous avons

$$\mathcal{J}(\mathbf{x}) = \mathcal{J}(\mathbf{x}_0) + (\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}), \mathbf{x} - \mathbf{x}_0) + o(\|\mathbf{x} - \mathbf{x}_0\|) \tag{3.8}$$

avec $(\cdot\,,\,\cdot)$ désignant le produit scalaire classique. L'égalité précédent peut se réécrire

$$\delta\mathcal{J} = \mathcal{J}(\mathbf{x}) - \mathcal{J}(\mathbf{x}_0) = (\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0) = (\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_0), \delta\mathbf{x}) \tag{3.9}$$

Supposons maintenant que $\mathcal{F}$ soit suffisamment régulier, alors pour chaque ensemble de variables de contrôle $\mathbf{x}_0$ une variation de $\mathcal{F}(\mathbf{x})$ peut être approchée au premier ordre par

$$\delta\mathcal{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0)\delta\mathbf{x} \tag{3.10}$$

$\mathbf{F}(\mathbf{x}_0)$ désigne le jacobien de $\mathcal{F}$ en $\mathbf{x}_0$. En utilisant l'équation (3.7), nous avons

$$\delta\mathcal{J} = (\mathcal{F}(\mathbf{x}_0) - F^o, \mathbf{F}(\mathbf{x}_0)\delta\mathbf{x}) \tag{3.11}$$

En utilisant la définition de l'opérateur adjoint $\mathbf{A}^*$ permettant de relier deux produits scalaires

$$(\mathbf{v}, \mathbf{A}\mathbf{w}) = (\mathbf{A}^*\mathbf{v}, \mathbf{w}) \tag{3.12}$$

nous avons finalement

$$\boxed{\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_0) = \mathbf{F}^*(\mathbf{x}_0)\,(\mathcal{F}(\mathbf{x}_0) - F^o)} \tag{3.13}$$

L'opérateur linéaire $\mathbf{F}(\mathbf{x}_0)$ représente le modèle linéaire tangent. Son adjoint $\mathbf{F}^*(\mathbf{x}_0)$ représente le modèle adjoint, celui que nous voulons construire pour notre modèle d'évolution de calotte polaire.

Passage de l'adjoint continu à l'adjoint discret

Un modèle complexe $\mathcal{F}$ peut être décomposé en $L \in \mathbb{N}$ étapes.

$$\mathcal{F} = \mathcal{F}^L \circ \cdots \circ \mathcal{F}^1 \tag{3.14}$$

Définissons les variables Z^l de la manière suivante

$$Z^l = \mathcal{F}^l \circ \cdots \circ \mathcal{F}^1(\mathbf{x}) \tag{3.15}$$

Définissons aussi les variables **tangentes** δZ^l

$$\delta Z^l = \frac{\partial}{\partial \mathbf{x}} \left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^1\right)\Big|_{\mathbf{x}=\mathbf{x}_0} \delta \mathbf{x} \tag{3.16}$$

où $\delta \mathbf{x}$ est une petite variation de $\mathbf{x}$. Le lien entre deux variables successives est donné par

$$\delta Z^l = \frac{\partial \mathcal{F}^l}{\partial Z^{l-1}}\Big|_{Z^{l-1}=Z_0^{l-1}} \delta Z^{l-1} \tag{3.17}$$

avec $Z_0^l = \mathcal{F}^l \circ \cdots \circ \mathcal{F}^1(X_0)$.

Définissons maintenant les variables **adjointes** $\delta^* Z^l$

$$\delta^* Z^l = \nabla_{Z^l} \left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^{l+1}\right)\Big|_{Z^l=Z_0^l} \tag{3.18}$$

Nous utilisons cette définition pour en déduire le modèle adjoint, autrement dit l'équation qui lie deux variables adjointes successives. $\mathcal{F}$ est la composée de

$$\mathcal{F} = \left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^{l+1}\right) \circ \left(\mathcal{F}^l \circ \cdots \circ \mathcal{F}^1\right)$$

Lorsque cette composée de deux fonctions est dérivée, cela donne

$$\begin{aligned} \frac{\partial \mathcal{F}}{\partial \mathbf{x}} &= \frac{d\left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^{l+1}\right)}{d\left(\mathcal{F}^l \circ \cdots \circ \mathcal{F}^1(\mathbf{x})\right)} \cdot \frac{d\left(\mathcal{F}^l \circ \cdots \circ \mathcal{F}^1(\mathbf{x})\right)}{d\mathbf{x}} \\ &= \frac{d\left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^{l+1}\right)}{dZ^l} \cdot \frac{dZ^l}{d\mathbf{x}} \end{aligned} \tag{3.19}$$

De sorte que nous avons, en utilisant la définition des variables tangentes puis adjointes

$$\begin{aligned} \delta \mathcal{F} &= \frac{\partial \mathcal{F}}{\partial \mathbf{x}} \cdot \delta \mathbf{x} \\ &= \frac{d\left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^{l+1}\right)}{dZ^l} \cdot \frac{dZ^l}{d\mathbf{x}} \cdot \delta \mathbf{x} \\ &= \frac{d\left(\mathcal{F}^L \circ \cdots \circ \mathcal{F}^{l+1}\right)}{dZ^l} \cdot \delta Z^l \\ &= \delta^* Z^l \cdot \delta Z^l \end{aligned} \tag{3.20}$$

Ceci vaut pour tout l, donc en particulier en écrivant ceci pour $l = 0$

$$\delta\mathcal{F} = (\nabla_{\mathbf{x}}\mathcal{F}, \delta\mathbf{x}) = \left(\delta^* Z^l, \delta Z^l\right) = \left(\delta^* Z^{l-1}, \delta Z^{l-1}\right) \tag{3.21}$$

Ceci nous permet d'en déduire $\delta^* Z^{l-1}$ en fonction de $\delta^* Z^l$:

$$\begin{aligned}\delta\mathcal{F} = \left(\delta^* Z^{l-1}, \delta Z^{l-1}\right) &= \left(\delta^* Z^l, \delta Z^l\right) \\ &= \left(\delta^* Z^l, \left.\frac{\partial \mathcal{F}^l}{\partial Z^{l-1}}\right|_{Z^{l-1}=Z_0^{l-1}} \delta Z^{l-1}\right) \\ &= \left(\left[\left.\frac{\partial \mathcal{F}^l}{\partial Z^{l-1}}\right|_{Z^{l-1}=Z_0^{l-1}}\right]^T \delta^* Z^l, \delta Z^{l-1}\right)\end{aligned} \tag{3.22}$$

ce qui donne finalement

$$\boxed{\left\{\begin{array}{rcll} Z^l &=& \mathcal{F}^l(Z^{l-1}) & \leftarrow \text{ modèle direct} \\ \delta Z^l &=& \left.\dfrac{\partial \mathcal{F}^l}{\partial Z^{l-1}}\right|_{Z^{l-1}=Z_0^{l-1}} \delta Z^{l-1} & \leftarrow \text{ modèle tangent} \\ \delta^* Z^{l-1} &=& \left[\left.\dfrac{\partial F^l}{\partial Z^{l-1}}\right|_{Z^{l-1}=Z_0^{l-1}}\right]^T \delta^* Z^l & \leftarrow \text{ modèle adjoint} \end{array}\right.} \tag{3.23}$$

Remarquons enfin que

$$\delta\mathcal{F} = (\nabla_{\mathbf{x}}\mathcal{F}, \delta\mathbf{x}) = (\delta^*\mathbf{x}, \delta\mathbf{x}) = (\delta^* Z_0, \delta\mathbf{x}) \tag{3.24}$$

donc $\nabla_{\mathbf{x}}\mathcal{F} = \delta^* Z_0$ est évalué en dernier.

3.2.2 Construction de l'adjoint discret de Winnie

Après avoir expliqué de manière théorique comment obtenir l'adjoint d'un modèle direct, il nous faut passer à la construction de l'adjoint de notre modèle d'évolution de calotte polaire. Nous rappelons tout d'abord que le code de Winnie peut se mettre sous la forme suivante :

FOR $l = 1, ..., n_t$
- $F^l_{\text{clim}} \leftarrow$ Modele_Fclim(t^l, T_{clim})
- $T^l_s \leftarrow$ Modele_Ts$(H^{l-1}, F^l_{\text{clim}})$
- $Acc^l \leftarrow$ Modele_Acc(T^l_s)
- $Abl^l \leftarrow$ Modele_Abl(T^l_s)
- $\dot{b}^l_m \leftarrow Acc^l + Abl^l$
- $\overline{U}^l_x \leftarrow$ Modele_Ux(H^{l-1})
- $H^l \leftarrow$ Solve_equation_bilan$(\dot{b}^l_m, \overline{U}^l_x H^{l-1})$

END

l en exposant indique que la variable concernée est calculée à t_l le temps d'itération du modèle. Nous détaillons dans la suite l'écriture du code adjoint pour chaque bloc simple puis le code adjoint de l'algorithme principal.

Adjoint du forçage climatique

Dans le cas où le forçage climatique est calculé à un instant t_l par la formule (3.1), les variables actives de ce bloc sont les 3 variables de contrôle a_0, a_1 et a_2 ainsi que F^l_{clim}. Le tangent de ce bloc de calcul vaut

$$\delta F^l_{\text{clim}} = \delta a_0 + \sin\left(\frac{2\pi t^i}{a_2}\right)\delta a_1 - a_1\frac{2\pi t_l}{a_2^2}\cos\left(\frac{2\pi t^i}{a_2}\right)\delta a_2$$

Mis sous forme matricielle, cela donne

$$\begin{pmatrix} \delta F^l_{\text{clim}} \\ \delta a_0 \\ \delta a_1 \\ \delta a_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & \sin\left(\frac{2\pi t^i}{a_2}\right) & -a_1\frac{2\pi t_l}{a_2^2}\cos\left(\frac{2\pi t^i}{a_2}\right) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta F^l_{\text{clim}} \\ \delta a_0 \\ \delta a_1 \\ \delta a_2 \end{pmatrix}$$

Pour obtenir le code adjoint de cette routine, il suffit de prendre la transposée de la matrice précédente

$$\begin{pmatrix} \delta^* F^l_{\text{clim}} \\ \delta^* a_0 \\ \delta^* a_1 \\ \delta^* a_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \sin\left(\frac{2\pi t^i}{a_2}\right) & 0 & 1 & 0 \\ -a_1\frac{2\pi t_l}{a_2^2}\cos\left(\frac{2\pi t^i}{a_2}\right) & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta^* F^l_{\text{clim}} \\ \delta^* a_0 \\ \delta^* a_1 \\ \delta^* a_2 \end{pmatrix}$$

Le code adjoint de ce bloc s'écrit donc :

PROCEDURE **ad_Fclim**$\left(\delta^* F^l_{\text{clim}}, \delta^* a_0, \delta^* a_1, \delta^* a_2\right)$

$\delta^* a_0 \leftarrow \delta^* a_0 + \delta^* F^l_{\text{clim}}$

$\delta^* a_1 \leftarrow \delta^* a_1 + \sin\left(\frac{2\pi t^i}{a_2}\right)\delta^* F^l_{\text{clim}}$

$\delta^* a_2 \leftarrow \delta^* a_2 - a_1\frac{2\pi t^i}{a_2^2}\cos\left(\frac{2\pi t^i}{a_2}\right)\delta^* F^l_{\text{clim}}$

$\delta^* F^l_{\text{clim}} \leftarrow 0$

END

Dans le cas du forçage avec interpolation linéaire par morceaux (3.2), les variables actives de ce bloc sont les F^k_{clim} et F^l_{clim}. Le tangent de ce bloc vaut alors

$$\delta F^l_{\text{clim}} = \frac{t_{k+1} - t_l}{t_{k+1} - t_k}\delta F^k_{\text{clim}} + \frac{t_l - t_k}{t_{k+1} - t_k}\delta F^{k+1}_{\text{clim}} \qquad t_l \in]t_k, t_{k+1}[$$

Mis sous forme matricielle, nous avons

$$\begin{pmatrix} \delta F^l_{\text{clim}} \\ \delta F^k_{\text{clim}} \\ \delta F^{k+1}_{\text{clim}} \end{pmatrix} = \begin{pmatrix} 0 & \frac{t_{k+1}-t_l}{t_{k+1}-t_k} & \frac{t_l-t_k}{t_{k+1}-t_k} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta F^l_{\text{clim}} \\ \delta F^k_{\text{clim}} \\ \delta F^{k+1}_{\text{clim}} \end{pmatrix}$$

Par transposition de la matrice, nous avons le code adjoint

$$\begin{pmatrix} \delta^* F^l_{\text{clim}} \\ \delta^* F^k_{\text{clim}} \\ \delta^* F^{k+1}_{\text{clim}} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ \frac{t_{k+1}-t_l}{t_{k+1}-t_k} & 1 & 0 \\ \frac{t_l-t_k}{t_{k+1}-t_k} & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta^* F^l_{\text{clim}} \\ \delta^* F^k_{\text{clim}} \\ \delta^* F^{k+1}_{\text{clim}} \end{pmatrix}$$

Le code adjoint de ce bloc s'écrit alors :

$$
\begin{array}{l}
\textsc{Procedure}\ \mathbf{ad_Fclim}\Big(\delta^* F^l_{\text{clim}}, \delta^* F^1_{\text{clim}}, \ldots, F^K_{\text{clim}}\Big) \\
\quad \textsc{If}\ t_l \in]t_k, t_{k+1}[\\
\qquad \begin{array}{lcl}
\delta^* F^k_{\text{clim}} & \leftarrow & \delta^* F^k_{\text{clim}} + \frac{t_{k+1}-t_l}{t_{k+1}-t_k} \delta^* F^l_{\text{clim}} \\
\delta^* F^{k+1}_{\text{clim}} & \leftarrow & \delta^* F^{k+1}_{\text{clim}} + \frac{t_l-t_k}{t_{k+1}-t_k} \delta^* F^l_{\text{clim}} \\
\delta^* F^l_{\text{clim}} & \leftarrow & 0
\end{array} \\
\quad \textsc{End} \qquad \textsc{End}
\end{array}
$$

Adjoint de la température en surface

La formule d'affectation de la température en surface à chaque point de grille est l'équation (1.74) (dans notre cas $H = S$). Nous notons T^l_s le vecteur contenant les températures à chaque point de grille au temps t_l et H^l le vecteur contenant les épaisseurs au même temps. Les variables actives de ce bloc sont F^l_{clim}, H^{l-1} et T^l_ss. Le tangent de cette routine vaut

$$\delta T^l_s = \delta F^l_{\text{clim}} + \gamma \delta H^l$$

Mis sous forme matricielle, nous avons :

$$
\begin{pmatrix} \delta T^l_s \\ \delta H^{l-1} \\ \delta F^l_{\text{clim}} \end{pmatrix}
=
\begin{pmatrix} \mathbf{0}_{n_x,n_x} & \gamma\, \mathbf{I}_{n_x} & \mathbf{1}_{n_x,1} \\ \mathbf{0}_{n_x,n_x} & \mathbf{I}_{n_x} & \mathbf{0}_{n_x,1} \\ \mathbf{0}_{1,n_x} & \mathbf{0}_{1,n_x} & 1 \end{pmatrix}
\begin{pmatrix} \delta T^l_s \\ \delta H^{l-1} \\ \delta F^l_{\text{clim}} \end{pmatrix}
$$

avec $\mathbf{0}_{r,s}$ et $\mathbf{1}_{r,s}$ matrices de taille $r \times s$ remplie respectivement de 0 et de 1. Pour rappel, δT^l_s et δH^{l-1} sont des vecteurs de taille n_x. Le code adjoint est obtenu en transposant la matrice précédente.

$$
\begin{pmatrix} \delta^* T^l_s \\ \delta^* H^{i-1} \\ \delta^* F^l_{\text{clim}} \end{pmatrix}
=
\begin{pmatrix} \mathbf{0}_{n_x,n_x} & \mathbf{0}_{n_x,n_x} & \mathbf{0}_{n_x,1} \\ \gamma\, \mathbf{I}_{n_x} & \mathbf{I}_{n_x} & \mathbf{0}_{n_x,1} \\ \mathbf{1}_{1,n_x} & \mathbf{0}_{1,n_x} & 1 \end{pmatrix}
\begin{pmatrix} \delta^* T^l_s \\ \delta^* H^{l-1} \\ \delta^* F^l_{\text{clim}} \end{pmatrix}
$$

Le code adjoint de ce bloc s'écrit donc :

$$
\begin{array}{l}
\textsc{Procedure}\ \mathbf{ad_Ts}\Big(\delta^* T^l_s, \delta^* H^{l-1}, \delta^* F^l_{\text{clim}}\Big) \\
\qquad \begin{array}{lcl}
\delta^* H^{l-1} & \leftarrow & \delta^* H^{i-1} + \gamma\ \delta^* T^l_s \\
\delta^* F^l_{\text{clim}} & \leftarrow & \delta^* F^l_{\text{clim}} + \displaystyle\sum_{j=1}^{n_x} \delta^* T^l_{s,j} \\
\delta^* T^i_s & \leftarrow & 0
\end{array} \\
\textsc{End}
\end{array}
$$

Adjoint de l'accumulation

L'accumulation Acc^l se calcule grâce à la relation (1.72). Les variables actives de ce bloc sont T^l_s et Acc^l. Le tangent de l'accumulation prend la forme suivante

$$\delta Acc^l = Acc_0\, c_1\, e^{c_1 T^l_s}\, \delta T^l_s$$

Nous passons directement à l'écriture du code adjoint :

PROCEDURE **ad_Acc**$\left(\delta^* T_s^l, \delta^* Acc^l\right)$
$\quad \delta^* T_s^l \leftarrow \delta^* T_s^l + Acc_0\, c_1\, e^{c_1 T_s^l}\, \delta^* Acc^i$
$\quad \delta^* Acc^i \leftarrow 0$
END

Adjoint de l'ablation

Dans Winnie, l'ablation Abl_j^l se calcule à chaque point de grille x_j par la relation (1.73). Les variables actives de ce bloc sont T_s^l et Abl^l. Pour chaque point de grille, nous avons

$$\delta Abl_j^l = \begin{cases} 2\, Abl_0\, \dfrac{T_{s,j}^l - T_{\text{no melt}}}{T_{\text{no melt}}^2}\, \delta T_s & \text{si } T_{s,j}^l > T_{\text{no melt}} \\ 0 & \text{sinon} \end{cases} \tag{3.25}$$

Nous passons directement au code adjoint de la routine d'ablation :

PROCEDURE **ad_Acc**$\left(\delta^* T_s^l, \delta^* Acc^l\right)$
$\quad$ IF $T_{s,j}^l > T_{\text{no melt}}$
$\qquad \delta^* T_s^l \leftarrow \delta^* T_s^l + 2\, Abl_0\, \dfrac{T_{s,j}^l - T_{\text{no melt}}}{T_{\text{no melt}}^2}\, \delta^* Abl_j^l$
$\quad$ END
$\quad \delta^* Abl^l \leftarrow 0$
END

Adjoint du bilan de masse

Le bilan de masse est la somme de l'accumulation et de l'ablation $\dot{b}_m^l = Acc^l + Abl^l$. Les variables de ce bloc sont toutes actives. Le tangent de ce bloc s'écrit tout simplement :

$$\delta \dot{b}_m^l = \delta Acc^l + \delta Abl^l$$

A partir de ce qui précède, on construit le code adjoint :

PROCEDURE **ad_Bilan**$\left(\delta^* \dot{b}_m^l, \delta^* Acc^l, \delta^* Abl^l\right)$
$\quad \delta^* Acc^l \leftarrow \delta^* Acc^l + \delta^* \dot{b}_m^l$
$\quad \delta^* Abl^l \leftarrow \delta^* Abl^l + \delta^* \dot{b}_m^l$
$\quad \delta^* \dot{b}_m^l \leftarrow 0$
END

Adjoint de la vitesse moyenne

Pour rappel, dans Winnie, la vitesse moyenne est calculée sur les nœuds mineurs. Considérons le vecteur $\overline{U}_x^l$ contenant les vitesses $\overline{U}_{x,j}^l$ aux nœuds mineurs $x_{j-\frac{1}{2}}$. En utilisant les formules de discrétisation (1.77) et de calcul de $\overline{U}_x$ (1.71), les $\overline{U}_{x,j}^l$ valent pour j variant

de $1, \ldots, n_x - 1$ (pour rappel, les nœuds majeurs sont numérotés de 0 à $n_x - 1$)

$$\overline{U}^l_{x,j} = -B_{AT3}\frac{(\rho_i\, g)^3}{5}\left(\frac{H^{l-1}_{j-1} + H^{l-1}_j}{2}\right)^4\left(\frac{H^{l-1}_j - H^{l-1}_{j-1}}{\Delta x}\right)^3 - B_{AT1}\frac{\rho_i\, g}{3}\left(\frac{H^{l-1}_{j-1} + H^{l-1}_j}{2}\right)^2\frac{H^{l-1}_j - H^{l-1}_{j-1}}{\Delta x} \quad (3.26)$$

Le tangent vaut ici

$$\delta\overline{U}^l_{x,j} = \alpha^l_j\, \delta H^{l-1}_{j-1} + \beta^l_j\, \delta H^{l-1}_j$$

avec

$$\alpha^l_j = -\frac{B_{AT3}}{80}\left(\frac{\rho_i\, g}{\Delta x}\right)^3 (H^{l-1}_{j-1} + H^{l-1}_j)^3 (H^{l-1}_j - H^{l-1}_{j-1})^2 (H^{l-1}_j - 7H^{l-1}_{j-1}) - \frac{B_{AT1}}{12}\frac{\rho_i\, g}{\Delta x}(H^{l-1}_{j-1} + H^{l-1}_j)(H^{l-1}_j - 3H^{l-1}_{j-1}) \quad (3.27)$$

$$\beta^l_j = -\frac{B_{AT3}}{80}\left(\frac{\rho_i\, g}{\Delta x}\right)^3 (H^{l-1}_{j-1} + H^{l-1}_j)^3 (H^{l-1}_j - H^{l-1}_{j-1})^2 (7H^{l-1}_j - H^{l-1}_{j-1}) - \frac{B_{AT1}}{12}\frac{\rho_i\, g}{\Delta x}(H^{l-1}_{j-1} + H^{l-1}_j)(3H^{l-1}_j - H^{l-1}_{j-1}) \quad (3.28)$$

A partir de ce qui précède, nous construisons le code adjoint :

```
PROCEDURE ad_Ux(δ*Ū^l_x, δ*H^{l-1})
  FOR j = 1, ..., n_x - 1
      δ*H^{l-1}_j   ←  δ*H^{l-1}_j + β^l_j δ*Ū^l_{x,j}
      δ*H^{l-1}_{j-1} ←  δ*H^{l-1}_{j-1} + α^l_j δ*Ū^l_{x,j}
  END
  δ*Ū^l_x ← 0
END
```

Adjoint de la hauteur de la calotte

Le bloc de calcul de l'évolution de la hauteur de la calotte est certainement le bloc où mettre en oeuvre l'adjoint est le plus difficile à faire. En effet, le calcul de H^l n'est pas un calcul direct. Il nécessite la résolution du système linéraire suivant

$$P(\overline{U}^l_x)\begin{pmatrix} H^l_1 \\ \vdots \\ H^l_{n_x-2} \end{pmatrix} = \begin{pmatrix} H^l_1 \\ \vdots \\ H^l_{n_x-2} \end{pmatrix} + \Delta t \begin{pmatrix} \dot{b}^l_{m,1} \\ \vdots \\ \dot{b}^l_{m,n_x-2} \end{pmatrix} \quad (3.29)$$

avec $P(\overline{U}^l_x)$ la matrice suivante

$$\begin{pmatrix} 1 + \frac{\Delta t}{\Delta x}(c^+_2 + c^-_1) & -\frac{\Delta t}{\Delta x}c^-_2 & 0 & \cdots & 0 \\ -\frac{\Delta t}{\Delta x}c^+_2 & 1 + \frac{\Delta t}{\Delta x}(c^+_3 + c^-_2) & -\frac{\Delta t}{\Delta x}c^-_3 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -\frac{\Delta t}{\Delta x}c^+_{n_x-3} & 1 + \frac{\Delta t}{\Delta x}(c^+_{n_x-2} + c^-_{n_x-3}) & -\frac{\Delta t}{\Delta x}c^-_{n_x-2} \\ 0 & \cdots & 0 & -\frac{\Delta t}{\Delta x}c^+_{n_x-2} & 1 + \frac{\Delta t}{\Delta x}(c^+_{n_x-1} + c^-_{n_x-2}) \end{pmatrix}$$

et $c_j^+ = \max(\overline{U}^l_{x,j}, 0)$ et $c_j^- = \max(-\overline{U}^l_{x,j}, 0)$.

L'algorithme de mise à jour de H^l est alors :

$$
\begin{array}{l}
H^l_{1:n_x-1} \leftarrow P^{-1}(\overline{U}^l_x)\left(H^{l-1}_{1:n_x-1} + \Delta t\ \dot{b}^l_{m,1:n_x-1}\right) \\
\textsc{For}\ j = 0, \cdots, n_x - 1 \\
\quad \textsc{If}\ H^l_j < 0 \\
\qquad H^l_j \leftarrow 0 \\
\quad \textsc{End} \\
\textsc{End}
\end{array}
$$

Les variables globales de cet algorithme sont H^l, H^l, $\overline{U}^l_x$ et $\dot{b}^l_m$, ces variables sont aussi actives. On passe maintenant au calcul du modèle tangent. On calcule ensuite le tangent de la résolution du système suivant l'équation (1.78), on a :

$$
\begin{aligned}
&\delta H^l_j + \frac{\Delta t}{\Delta x}\left(\left(c^+_{j+1}\delta H^l_j - c^-_{j+1}\delta H^l_{j+1}\right) - \left(c^+_j \delta H^l_{j-1} - c^-_j \delta H^l_j\right)\right) \\
&= \Delta t\ \dot{b}^l_{m,j} + \delta H^{l-1}_j + \frac{\Delta t}{\Delta x}\left(\left(r^+_{j-1} + r^-_j\right)\delta\overline{U}^l_{x,j} - \left(r^+_j + r^-_{j+1}\right)\delta\overline{U}^l_{x,j+1}\right)
\end{aligned}
$$

$$
\text{Avec}\quad r^+_j = \begin{cases} H^l_j & \text{si } \overline{U}^l_{x,j} > 0 \\ 0 & \text{sinon} \end{cases} \qquad \text{et}\quad r^-_j = \begin{cases} H^l_j & \text{si } \overline{U}^l_{x,j} < 0 \\ 0 & \text{sinon} \end{cases}
$$

On pose la matrice suivante :

$$
R(\overline{U}^l_x, H^l) = \frac{\Delta t}{\Delta x}\begin{pmatrix}
r^+_1 & -r^+_1 - r^-_2 & 0 & \cdots & 0 \\
0 & r^+_2 + r_3 & -r^+_2 - r^-_3 & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & r^+_{n_x-3} + r^-_{n_x-2} & -r^+_{n_x-2}
\end{pmatrix}
$$

D'où finalement :

$$
\delta H^l_{1:n_x-2} = P^{-1}(\overline{U}^l_x)\left(\delta H^{l-1}_{1:n_x-2} + \Delta t\ \delta\dot{b}^l_{m,1:n_x-2} + R(\overline{U}^l_x, H^l)\delta\overline{U}^l_{x,1:n_x-1}\right)
$$

Enfin, si $H^l_j < 0$, alors $\delta H^l_j = 0$. On vient donc d'écrire le modèle tangent du bloc de calcul de la hauteur de la calotte. Ecrivons maintenant le modèle adjoint, pour cela, écrivons d'abord l'adjoint de la résolution de système, on obtient :

$$
\begin{pmatrix}
\delta^* H^l_{1:n_x-2} \\
\delta^* H^{l-1}_{1:n_x-2} \\
\delta^* \dot{b}^l_{m,1:n_x-2} \\
\delta^* \overline{U}^l_{x,1:n_x-1}
\end{pmatrix}
=
\begin{pmatrix}
0 & 0 & 0 & 0 \\
(P(\overline{U})^T)^{-1} & \mathbf{I}_{n_x-2} & 0 & 0 \\
\Delta t(P(\overline{U})^T)^{-1} & 0 & \mathbf{I}_{n_x-2} & 0 \\
R(\overline{U}, H^i)^T(P(\overline{U})^T)^{-1} & 0 & 0 & \mathbf{I}_{n_x-1}
\end{pmatrix}
\begin{pmatrix}
\delta^* H^l_{1:n_x-2} \\
\delta^* H^{l-1}_{1:n_x-2} \\
\delta^* \dot{b}^l_{m,1:n_x-2} \\
\delta^* \overline{U}^l_{x,1:n_x-1}
\end{pmatrix}
$$

L'adjoint du restant du bloc est facile à écrire, on utilise les règles utilisées dans les blocs précédents. On obtient le code adjoint suivant :

$$
\begin{array}{l}
\textsc{Procedure}\ \mathbf{ad_H}\left(\delta^*\dot{b}^l_m, \delta^* H^l, \delta^* H^{l-1}, \delta^*\overline{U}^l_x\right) \\
\quad \textsc{For}\ j = 0, \cdots, n_x - 1
\end{array}
$$

```
        If H_j^l < 0
            δ*H_j^l ← 0
        End
    End
    δ*H^{l-1}_{1:n_x-2}      ←  (P(Ū)^T)^{-1} δ*H^l_{1:n_x-2} + δ*H^{l-1}_{1:n_x-2}
    δ*ḃ^l_{m,1:n_x-2}        ←  Δt(P(Ū)^T)^{-1} δ*H^l_{1:n_x-2} + δ*ḃ^l_{m,1:n_x-2}
    δ*Ū^l_{x,1:n_x-2}        ←  R(Ū, H^i)^T (P(Ū)^T)^{-1} δ*H^l_{1:n_x-2} + δ*Ū^l_{x,1:n_x-1}
    δ*H^l                    ←  0
End
```

Adjoint du bloc principal

D'après toutes les règles vues dans les paragraphes précédents, l'adjoint du bloc principal s'écrit de la manière suivante :

```
For l = n_t, n_t - 1, ..., 1
    ad_H (δ*ḃ^l_m, δ*H^l, δ*H^{l-1}, Ū^l_x)
    ad_Ux (δ*H^{l-1}, Ū^l_x)
    ad_Bilan (δ*ḃ^l_m, δ*Acc^l, δ*Abl^l)
    ad_Abl (δ*T^l_s, δ*Abl^l)
    ad_Acc (δ*T^l_s δ*Acc^l)
    ad_Ts (δ*T^l_s, δ*F^l_clim, δ*H^{l-1})
    ad_F clim (δ*F^l_clim, δ*T_clim)
End
```

3.2.3 Validation du code adjoint

Comment valider le code adjoint que nous venons de construire ? I nous faut valider chaque bloc construit mais en plus tester le gradient.

Validation bloc par bloc

La validation bloc par bloc consiste à comparer les valeurs d'entrées δX_e et de sortie δX_s du bloc tangent avec les valeurs d'entrées $\delta^* X_e$ et de sortie $\delta^* X_s$ du bloc adjoint. On pose F le bloc tangent, alors F^* correspond au bloc adjoint. On a l'égalité suivante :

$$(\delta X_s, \delta^* X_e) = (F(\delta X_e), \delta^* X_e) = (\delta X_e, F^*(\delta^* X_e)) = (\delta X_e, \delta^* X_s) \tag{3.30}$$

La validation bloc par bloc consiste simplement à vérifier que la deuxième égalité de l'équation précédente est valide pour des données de δX_e et $\delta^* X_e$ aléatoires. Cette technique nous permet de valider tous les blocs écrits plus haut.

Il nous reste maintenant plus qu'à tester le gradient.

Test du premier ordre

Le premier test de validation est élémentaire, il s'agit de vérifier l'approximation, au premier ordre en α proche de 0 :

$$\frac{\mathcal{J}(\mathbf{x}_0 + \alpha \mathbf{e}_i) - \mathcal{J}(\mathbf{x})}{\alpha} = (\nabla_{\mathbf{x}} \mathcal{J}(\mathbf{x}_0), \mathbf{e}_i) + o(1) \tag{3.31}$$

L'idée est donc de calculer pour diverses valeurs de la direction de perturbation $\mathbf{e}_i$, et pour diverses valeurs de α tendant vers 0 d'une part le quotient

$$\tau(\alpha, \mathbf{e}_i) = \frac{\mathcal{J}(\mathbf{x}_0 + \alpha \mathbf{e}_i) - \mathcal{J}(\mathbf{x}_0)}{\alpha} \tag{3.32}$$

calculé avec le code direct et d'autre part le produit scalaire

$$\delta(\mathbf{e}_i) = (\nabla_{\mathbf{x}} \mathcal{J}(\mathbf{x}_0), \mathbf{e}_i) \tag{3.33}$$

où $\nabla_{\mathbf{x}} \mathcal{J}(\mathbf{x}_0)$ est donné par le modèle adjoint. Ensuite, il suffit de vérifier que l'erreur relative $\varepsilon(\alpha, \mathbf{e}_i)$ définie ci-dessous tende bien vers 0 lorsque α tend vers 0.

$$\varepsilon(\alpha, \mathbf{e}_i) = \frac{|\tau(\alpha, \mathbf{e}_i) - \delta(\mathbf{e}_i)|}{|\delta(\mathbf{e}_i)|} \tag{3.34}$$

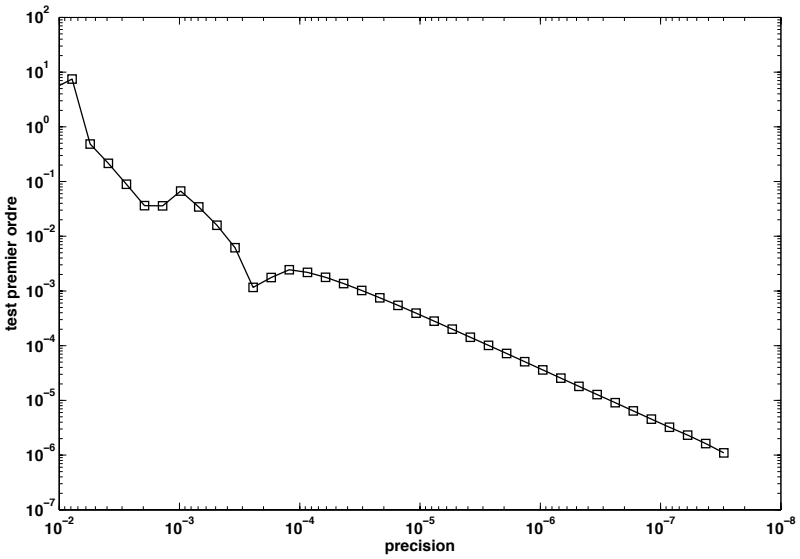

FIGURE 3.4 – Exemple de convergence pour le premier ordre du gradient. Sur cet exemple, la convergence semble assurée à partir de $\alpha = 10^{-4}$.

Test du deuxième ordre

Si la fonction coût $\mathcal{J}$ est quadratique, le développement de Taylor suivant est exact

$$\mathcal{J}(\mathbf{x}_0 + \alpha \mathbf{e}_i) = \mathcal{J}(\mathbf{x}_0) + \alpha \left(\nabla_{\mathbf{x}} \mathcal{J}(\mathbf{x}_0), \mathbf{e}_i\right) + \frac{\alpha^2}{2} \left(\mathbf{Hess}_{\mathbf{x}}(\mathbf{x}_0)\, \mathbf{e}_i, \mathbf{e}_i\right) \tag{3.35}$$

$\mathbf{Hess_x}$ est la hessienne de la fonction coût $\mathcal{J}$. Pour $\mathbf{e}_i$ fixé, le dernier terme est constant

$$\mathcal{J}(\mathbf{x}_0 + \alpha\mathbf{e}_i) = \mathcal{J}(\mathbf{x}_0) + \alpha\left(\nabla_{\mathbf{x}}\mathcal{J}(\mathbf{x}_0), \mathbf{e}_i\right) + \frac{\alpha^2}{2}C(\mathbf{e}_i) \tag{3.36}$$

La formule de Taylor donne alors

$$\frac{\tau(\alpha, \mathbf{e}_i) - \delta(\mathbf{e}_i)}{\alpha} = \frac{1}{2}C(\mathbf{e}_i) \tag{3.37}$$

Le deuxième test du gradient consiste donc à calculer la quantité $r(\alpha, \mathbf{e}_i)$

$$r(\alpha, \mathbf{e}_i) = \frac{\tau(\alpha, \mathbf{e}_i) - \delta(\mathbf{e}_i)}{\alpha} \tag{3.38}$$

pour diverses directions $\mathbf{e}_i$ et divers α et à vérifier que, lorsque α tend vers 0, cette quantité tend vers une constante dépendant de $\mathbf{e}_i$.

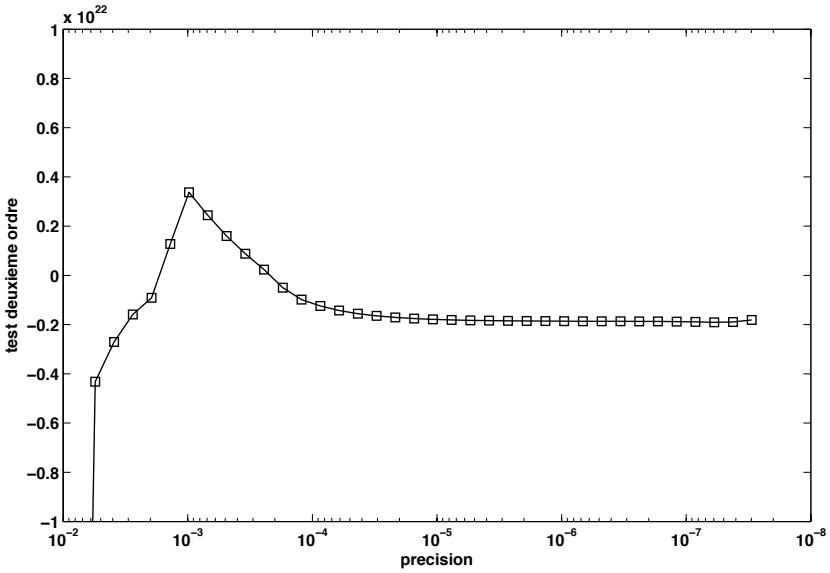

FIGURE 3.5 – Exemple de convergence pour le deuxieme ordre du gradient (même configuration que pour la figure 3.4). Sur cet exemple, la convergence semble assurée à partir de $\alpha = 10^{-4}$. Les résultats se dégradent pour des α très petits du fait des erreurs numériques incontournables à cette précision.

3.3 Expériences jumelles réalisées

Nous décrivons ici les différents types d'expériences jumelles menées durant nos travaux de thèse. Chacune d'entre elles souligne une des difficultés du problème étudié. Affirmons dès à présent que ceux-ci ne sont pas liés à la méthode d'assimilation utilisée.

3.3.1 Supposer un forçage climatique sinusoïdal

L'expérience jumelle proposée consiste à retrouver les paramètres (a_0, a_1, a_2). Nous observons l'épaisseur de la calotte à chaque pas de temps, nous avons donc beaucoup de données. Cependant, il a été impossible au 4D-Var de retrouver correctement ces 3 paramètres à chaque fois que l'ébauche s'éloignait des valeurs références prescrites. Il semblerait que pour des raisons physiques ces 3 paramètres ne soient pas indépendants entre eux en minimisant la fonction coût. Deux paramètres peuvent suffir à expliquer l'évolution de la calotte. Dans cette optique, nous avons utilié le 4D-Var afin de retrouver a_1 et a_2. Regardons tout d'abord l'allure de la fonction coût et l'écart aux observations.

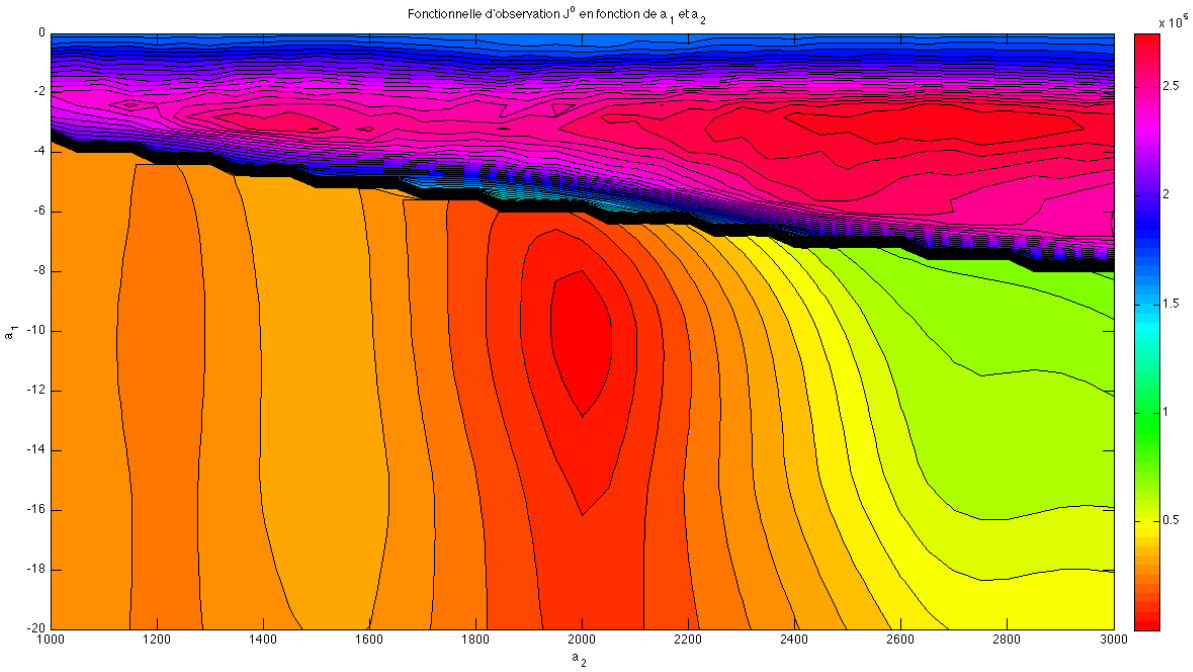

FIGURE 3.6 – Graphe de l'écart aux observations $\mathcal{J}^o$ en fonction de a_1 et a_2

En regardant le graphe, on remarque l'existence de minima locaux pour l'écart aux obsevations. L'algorithme de minimisation risque de donner des résultats érronés si nous ne régularisons pas correctement la fonction coût. Dans notre expérience, nous prenons pour ébauche $a_1^b = -8$ et $a_2^b = 2500$ alors que les valeurs de références sont de 10 e 2000. On voit bien la suppression de la plupart des minima locaux. Faisons maintenant fonctionner notre algorithme dans ce cadre. Le résultat obtenu est :

$$a_{1,res} = -9.871 \qquad \text{et} \qquad a_{2,res} = 1999.98$$

Celui-ci est plutôt satisfaisant. On veut aussi voir l'évolution du vecteur de contrôle (paramètres a_1 et a_2) pour chaque itération. On peut l'observer sur la figure 3.7. Remarquons que l'algorithme teste certaines directions qui s'avèrent infructueuses.

Dans cette expérience de 4D-Var, les observations dont nous disposions étaient très complètes, puisque nous observions la hauteur de glace H en tout point et en tout temps.

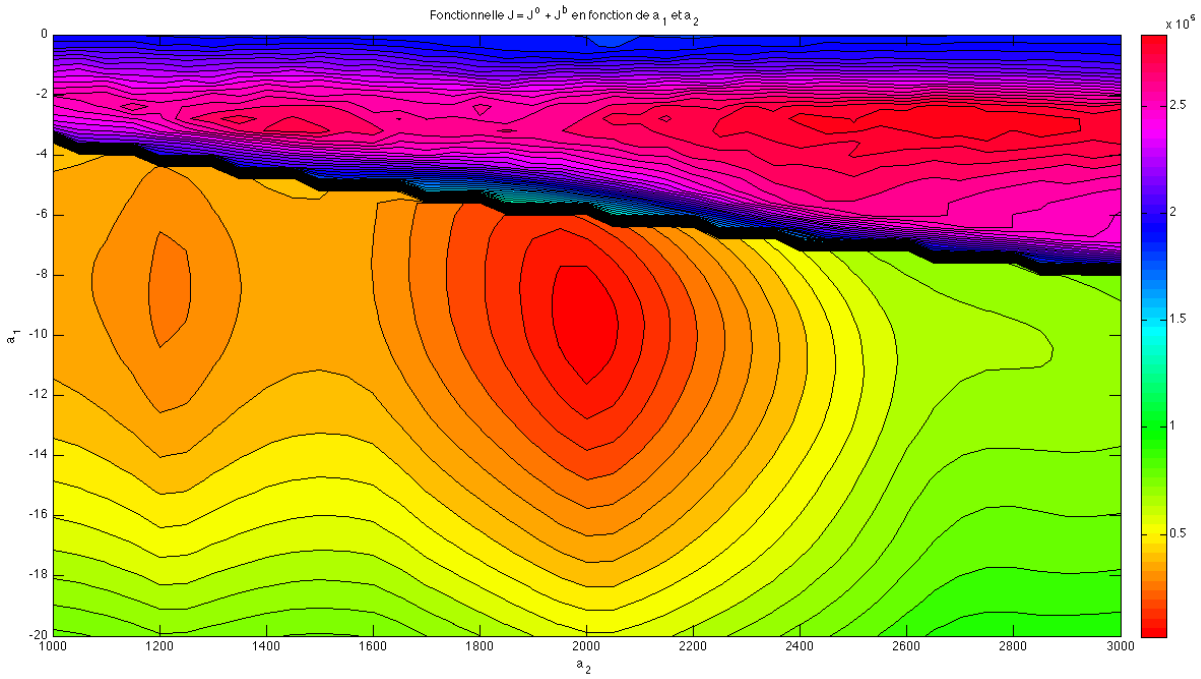

FIGURE 3.7 – Graphe de la fonction coût $\mathcal{J} = \mathcal{J}^o + \mathcal{J}^b$ avec $\mathcal{J}^b(a_1, a_2) = w_b \left(\frac{(a_1 - a_1^b)^2}{(a_1^b)^2} + \frac{(a_2 - a_2^b)^2}{(a_2^b)^2} \right)$ en fonction de a_1 et a_2. Ici $w_b = 20000$.

En glaciologie, particulièrement lorsque l'on considère des périodes de temps anciennes, on ne dispose pas d'informations aussi précises, mais uniquement du volume global de glace en tout temps.

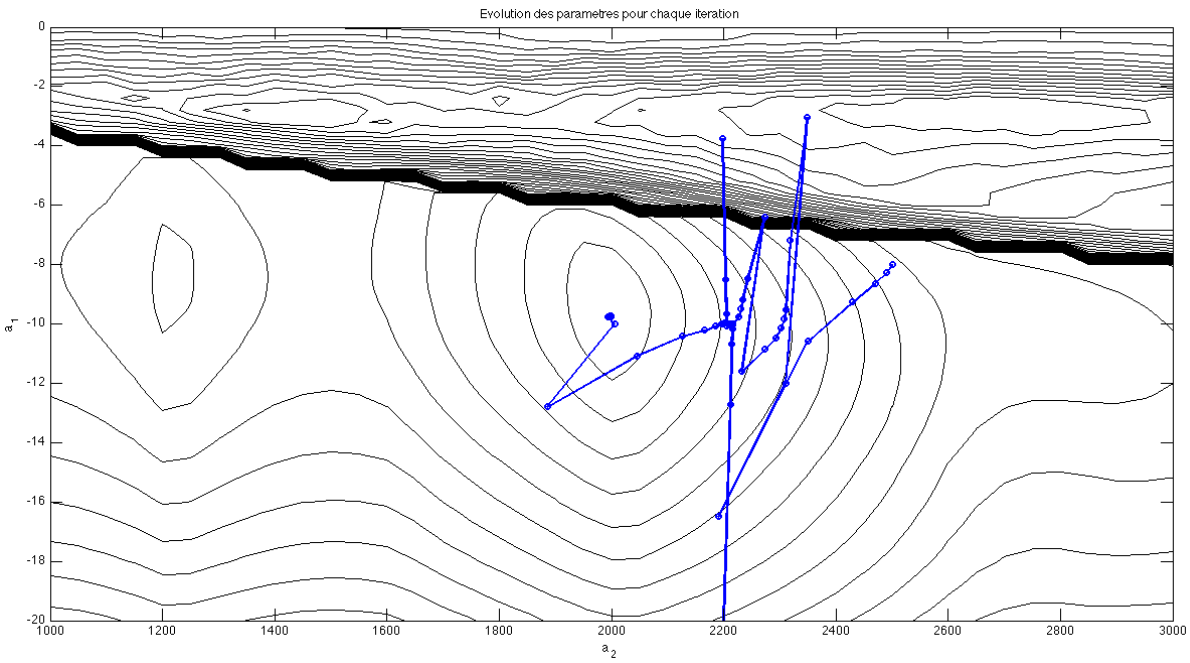

FIGURE 3.8 – Evolution des paramètres en fonction des itérations

3.3.2 Retrouver le forçage climatique avec une bonne ébauche

Cherchons mainteant à retrouver le forçage climatique (sous forme linéaire par morceau) sur une fenêtre d'assimilation de 20 000 ans à partir d'une ébauche proche de la vérité. Pour cela, nous assimilons les données de volume disponibles à chaque pas de temps du modèle ($\Delta t = 10$ ans).

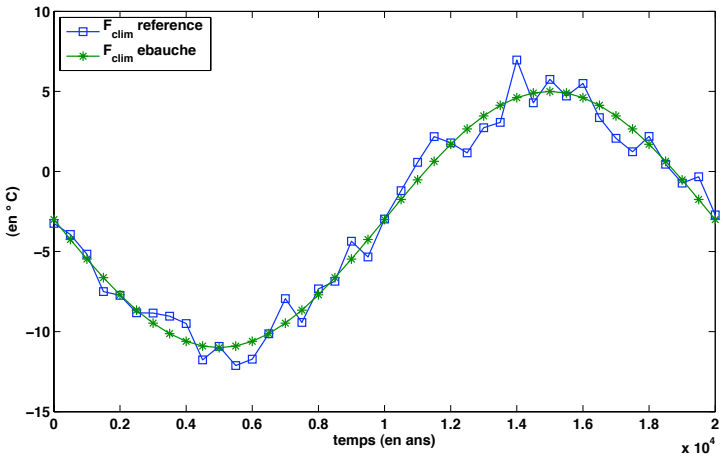

FIGURE 3.9 – Nous cherchons à reconstruire le forçage de référence (en bleu) à partir de l'ébauche sinusoïdale (en vert). Les termes de forçages sont à déterminer tous les 500 ans (entre ces deux termes, nous appliquons la régression linéaire entre ces deux points indiquée à l'équation (3.2)).

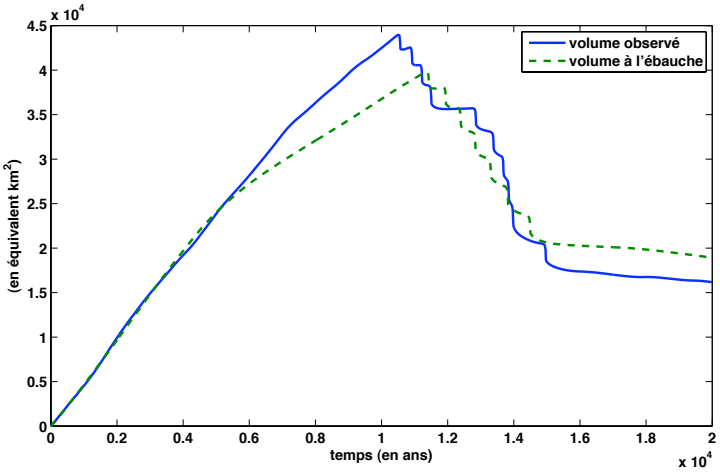

FIGURE 3.10 – Volume observé (obtenu par le forçage climatique de référence) et le volume à l'ébauche (obtenu par l'ébauche sinusoïdale du forçage climatique)

La matrice de covariance d'erreur d'observation $\mathbf{R}$ est prise diagonale et les écarts types valent $1\,km^2$. Les observations de volume sont donc nombreuses (2 000 observations sur 20 000 ans) et très précises.

Le forçage climatique de référence est bien retrouvé (voir figures 3.11 et 3.12) par la méthode d'assimilation utilisée. Cela valide notre approche.

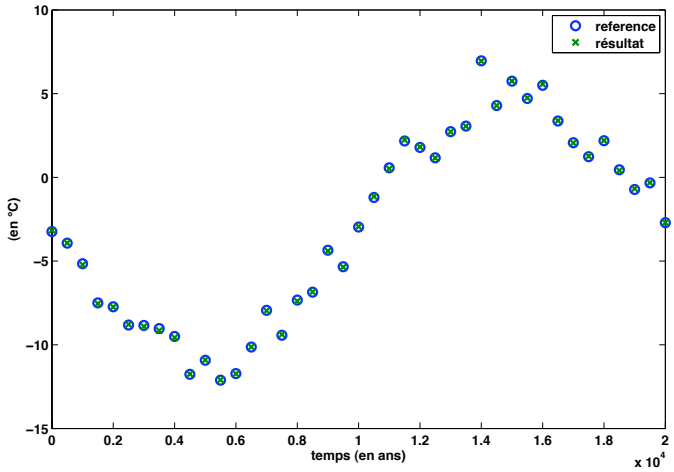

FIGURE 3.11 – Forçage climatique reconstruit (croix vertes) par méthode adjointe, comparaison avec la valeur de référence (ronds bleus)

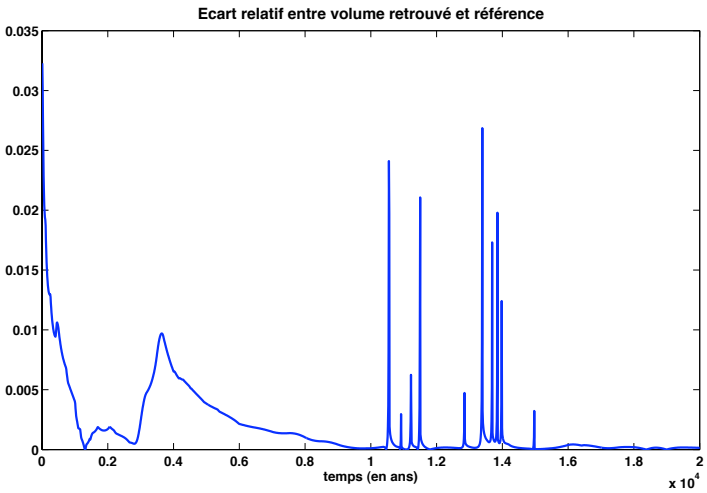

FIGURE 3.12 – Ecart relatif entre le volume reconstruit à l'aide de la méthode adjointe et le volume observé au cours du temps.

3.3.3 Que faire si nous n'avons pas une bonne ébauche

Dans l'expérience précédente, nous avions une ébauche du forçage climatique proche de sa vraie valeur. Ce n'est pas forcément le cas en pratique. Nous nous plaçons maintenant

dans le cas suivant où l'ébauche est prise nulle : $T^b_{clim} = \mathbf{0}$ et nous essayons de reconstruire cette fois-ci un forçage sinusoïdal (courbe verte de la figure 3.9). La fréquence et l'incertitude des observations de volume restent inchangées.

Nous avons commencé par reprendre l'approche directe de l'expérience précédente. Cependant, quelque soit la forme de la matrice de covariance d'ébauche $\mathbf{B}$, l'algorithme n'est jamais arrivé à converger vers un résultat un tant soit peu proche de la vérité. Ceci est principalement dû à la multiplicité des minima locaux de la fonction coût utilisée. Pour contourner ce problème d'ébauche, nous décidons d'utiliser une fenêtre grandissante, c'est à dire effectuer plusieurs expériences d'assimilation de données successives afin de reconstruire petit à petit le forçage climatique. Nous commençons par considérer la fenêtre temporelle $[0, 500]$, puis $[0, 1\,000]$, $[0, 1\,500]$ et ainsi de suite jusqu'à ce que toutes les observations sur $[0, 20\,000]$ soient prises en compte. L'ébauche est prise à zéro pour la première fenêtre temporelle puis est construite de la manière suivante pour les autres fenêtres. Pour la fenêtre d'assimilation $[0, t_i]$

— reprendre les valeurs trouvées précédemment pour l'ébauche $F^b_{\text{clim}}(0), \ldots, F^b_{\text{clim}}(t_{i-1})$
— construire une ébauche $F^b_{\text{clim}}(t_i)$ satisfaisante

Nous proposons deux manières d'effectuer le calcul de $F^b_{\text{clim}}(t_i)$, la première est simple, il suffit de prendre $F^b_{\text{clim}}(t_i) = F^b_{\text{clim}}(t_{i-1})$. Cependant, nous ne reconstruisons pas correctement le forçage climatique (voir 3.13 et 3.14). L'algorithme part dans un minimum local et n'arrive plus à en sortir par la suite.

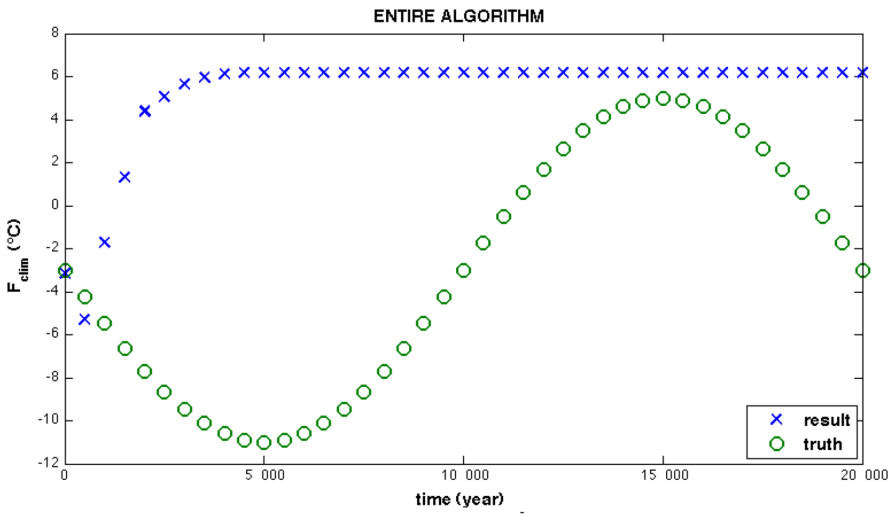

FIGURE 3.13 – Forçage climatique reconstitué (en croix bleu) en prenant pour ébauche $F^b_{\text{clim}}(t_i) = F^b_{\text{clim}}(t_{i-1})$.

Nous choisissons maintenant de calculer $F^b_{\text{clim}}(t_i)$ comme le prolongement de la tendance sur le forçage entre $F^b_{\text{clim}}(t_{i-2})$ et $F^b_{\text{clim}}(t_{i-1})$. Plus exactement, nous prenons

$$F^b_{\text{clim}}(t_i) = F^b_{\text{clim}}(t_{i-2}) + \frac{t_i - t_{i-2}}{t_{i-1} - t_{i-2}} \left(F^b_{\text{clim}}(t_{i-1}) - F^b_{\text{clim}}(t_{i-2}) \right) \tag{3.39}$$

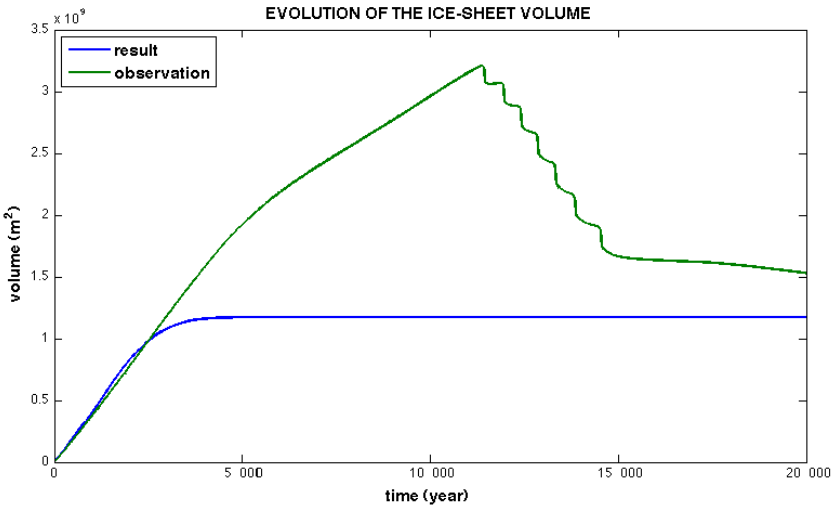

FIGURE 3.14 – Evolution du volume de la calotte reconstruit (en bleu) suite au forçage climatique trouvé dans la figure 3.13

Cette fois-ci le forçage climatique est correctement reconstitué (voir figure 3.15). Cependant, la convergence est facilitée au maximum par l'ébauche choisie qui est très proche de la vraie valeur recherchée (la fonction sinus très régulière est bien approchée localement par sa tangente).

Pour d'autres forçages climatiques choisis (non montrés ici), la première méthode de construction de l'ébauche est à préférer. Enfin dans certains cas, aucune des deux méthodes décrites n'est efficace. Pondérons néanmoins ces réflexions quelque peu pessimistes en soulignant que l'approche par fenêtre grandissante est préférable à l'approche directe de la première expérience lorsque le forçage climatique est peu ou pas connu (donc une mauvaise ébauche).

3.4 Conclusion

Ce chapitre détaille la méthode adjointe proposée afin de reconstruire l'évolution temporelle du paramètre de forçage climatique intervenant dans le calcul de la température en surface en utilisant des observations du volume de la calotte. Cette méthode n'avait jamais été utilisée même dans un cadre simplifié tel que proposé ici. Les résultats obtenus ont montré l'efficacité de cette méthode mais aussi que le problème de contrôle n'avait pas de solution évidente en soi. En effet, suite à nos diverses expériences, nous avons pu constaté l'importance du choix du terme d'ébauche sur la qualité des résultats (la plupart des méthodes d'assimilation y sont en général sensibles). D'autre part, il apparaît que ce problème n'a pas forcément de solution unique même avec régularisation par le terme d'ébauche. L'apport d'autres observations comme l'état actuel de la calotte serait pertinent dans cette optique.

La transposition d'une telle expérience à un cas plus opérationnel redouble de diffi-

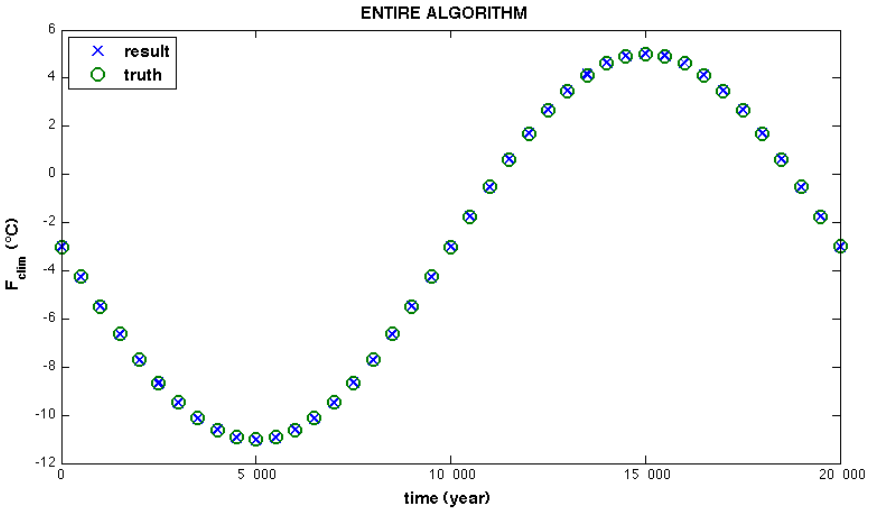

FIGURE 3.15 – Forçage climatique reconstitué (en croix bleu) en prenant pour ébauche $F^b_{\text{clim}}(t_i)$ calculé à l'aide de la relation (3.39).

cultés supplémentaires. En effet, l'évolution du volume de glace au cours du temps n'est pas dû qu'à une unique calotte mais à plusieurs. D'autre part, dans nos expériences, nous avons négligé des phénomènes importants dans l'évolution des calottes polaires à longue échelle de temps : l'isostasie et le frottement à la base. Ainsi, il est connu que le forçage climatique seul n'aurait pu faire disparaître feue la calotte des Laurentides il y a 7 000 ans environ s'il n'y avait pas eu de glissement à la base. Dans un cadre opérationnel, il faudrait donc rechercher des paramètres climatiques communs à toutes ces calottes sachant que certaines d'entre elles n'existent plus en plus de contrôler chaque topographie du socle rocheux et chaque coefficient de loi de frottement. Il est clair qu'utiliser le volume à lui seul n'est pas suffisant. Cependant il serait à nos yeux tout à fait pertinent d'envisager le problème de contrôle précédemment cité en utilisant comme observation outre l'évolution du volume de glace au cours du temps (obtenu à partir de carottes de sédiments marins), l'état actuel des calottes polaire, les isochrones des calottes encore existantes et des données géomorphologiques comme l'observation de l'étendue d'une calotte à certains instants. L'ébauche sur les paramètres climatiques pourrait elle être obtenue par les résultats d'assimilation de données obtenus à partir de carottes de glace.

Chapitre 4

Initialisation de modèle de calotte polaire par filtre de Kalman d'ensemble

Nous abordons ici l'un des problèmes phares de l'assimilation de données en glaciologie : le problème du *spin up* ou de l'initialisation de modèle. Nos travaux de thèse sur ce sujet font l'objet d'un article soumis au journal *Nonlinear Processes in Geophysics* dans un special issue consacré à l'utilisation des méthodes d'ensemble en géosciences [Bonan *et al.*, 2013]. Nous reproduisons dans ce chapitre l'article en anglais muni au préalable d'un résumé en français. Nous ajoutons aussi en fin de chapitre quelques résultats obtenus avec le filtre ETKF-N version duale (voir algorithme 2.11 pour les détails du filtre).

4.1 Résumé des travaux

Le problème étudié ici est la reconstruction de l'état actuel d'une calotte polaire soit les variables et certains paramètres influents mal connus de notre modèle d'évolution de calotte polaire simplifié Winnie. La fenêtre de temps des expériences est 20 ans. Le problème d'assimilation posé peut se résumer de la manière suivante :

— « *Que contrôlons nous ?* » Nous contrôlons à chaque point de grille :

1. L'épaisseur $H(x,t)$ de la calotte, seule variable prognostique de Winnie.
2. La topographie du socle rocheux $B_{soc}(x)$, paramètre de Winnie mal connu.
3. Le coefficient de frottement basal $\beta(x)$, paramètre de Winnie inconnu.

Rappelons que β est un paramètre forcément positif. Pour contourner cette difficulté, nous choisissons de chercher β sous la forme $\alpha = \log_{10}(\beta)$.

— « *Qu'observons nous ?* » Chaque année durant 20 ans, nous observons les quantités suivantes :

1. L'altitude de la surface $S(x,t)$ de la calotte à chaque point de grille.

2. La vitesse horizontale de la glace en surface $u_{s,x}(x,t)$ à chaque point de grille.

$$u_{s,x} = u_{b,x} - B_{AT3}\frac{(\rho_i\, g)^3}{4}H^4\left(\frac{\partial S}{\partial x}\right)^3 - B_{AT1}\frac{\rho_i\, g}{2}H^2\,\frac{\partial S}{\partial x} \tag{4.1}$$

$$\text{avec} \qquad u_{b,x} = -\frac{1}{\beta}\,\rho_i\, g\, H\,\frac{\partial S}{\partial x} \tag{4.2}$$

3. La topographie du socle rocheux $B_{soc}(x)$ en de rares points de grille.

Notons tout de suite que l'opérateur d'observation est fortement non-linéaire pour obtenir la vitesse en surface. D'autre part, la matrice de covariance d'erreur d'observation $\mathbf{R}$ est prise diagonale.

— « *Quel système d'assimilation utilisons nous ?* » : filtre de Kalman d'ensemble version ETKF (voir algorithme 2.9).

L'article se limite à une étude de faisabilité à l'aide d'expériences synthétiques plus communement appelées expériences jumelles en assimilation de données. Signalons que nous prenons en compte les temps d'observations. L'experience détaillée dans ce chapitre utilise volontairement pour ébauche des valeurs assez éloignées de l'état de référence.

L'utilisation d'un ETKF nécessite de créer un ensemble de depart. Pour cela, nous nous servons des observations à notre disposition (et nous foulons aux pieds une hypothèse importante du filtre de Kalman qui est l'indépendance de l'ébauche par rapport aux observations). Nous générons tout d'abord un ensemble de topographie de socle suivant une loi normale dont la moyenne est l'ébauche de socle que nous avons. La matrice de covariance est quant à elle construite de la maniere suivante :

— les variances dépendent de l'éloignement du point de grille à une observation de socle disponible (plus nous sommes loin d'une observation, plus la variance est grande)
— les corrélations entre deux points de grille dépendent uniquement de la distance entre ces deux points, nous choisissons ici de prendre en compte deux longueurs de corrélations différentes (une pour prendre en compte le comportement à grande échelle + une autre pour les échelles courtes). Cela nous évite d'obtenir des topographies de socle trop chahutés ou trop lisses.

L'epaisseur de la calotte est construite ensuite en soustrayant de l'altitude de surface observée à un temps donné une topographie de socle simulée. Il ne reste plus qu'à construire un ensemble de α. Pour cela, nous utilisons une loi normale ayant pour moyenne l'ébauche sur alpha et une matrice de covariance prescrite. Enfin, nous faisons tourner le modèle sur chaque membre de l'ensemble pendant 1 an afin d'obtenir des états plus physiques.

Nous testons notre systeme d'assimilation pour plusieurs tailles d'ensemble. En grande taille, l'ETKF fonctionne correctement excepté pour la reconstruction du coefficient de frottement. Cependant, ce dernier est bien reconstruit pour les zones où il a de l'influence (nous regardons pour cela les vitesses de glissement reconstruites par rapport à leurs vraies valeurs). Lorsque nous voudrons utiliser un modele complet 3D, utiliser une grande

taille d'ensemble (par rapport à la taille des quantités recherchées) nous sera impossible. Nous étudions donc le cas des ensemble de petite taille. Dans ces cas, il faut utiliser les procédures d'inflation et de localisation afin d'obtenir des résultats satisfaisants hormis aux zones de transitions où le glissement et la déformation de la glace jouent tous les 2 un rôle non-négligeable.

Enfin, en complément de cet article, nous choisissons d'employer un filtre ETKF-N afin de mieux investiguer le rôle de l'inflation. Pour un ensemble de taille modéré, l'ETKF-N fonctionne relativement bien, l'importance de l'inflation n'est donc pas à négliger (meme si elle paraissait peu claire par rapport à la localisation dans les expériences menées dans l'article). Pour des ensembles de plus petite taille, la localisation est incontournable. Cependant, la combinaison localisation - ETKF-N conduit à des calottes reconstruites non-physiques.

4.2 Une approche ETKF pour l'estimation jointe de l'état initial et de paramètres pour un modèle de calotte polaire

Submitted to Nonlinear Processes in Geophysics

Bertrand Bonan[1], Maëlle Nodet[1,2], Catherine Ritz[3], and Vincent Peyaud[3]

[1]INRIA, Laboratoire Jean Kuntzmann
[2]Université Joseph Fourier – Grenoble 1, Laboratoire Jean Kuntzmann
[3]Université Joseph Fourier – Grenoble 1 / CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement

4.2.1 Introduction

Antarctica and Greenland account for a significant fraction of today's sea-level rise. Good estimates of their future contribution are therefore crucial to produce sea-level change forecasts, as underlined by [Hanna *et al.*, 2013]. Producing pertinent estimates of polar ice sheets contribution to sea-level change rely on our ability to run a precisely calibrated ice sheet evolution model starting from a reliable initial state. By initial state we mean a representation of today's ice caps : bedrock topography, ice thickness, basal friction law.

In this work we will not address calibration of the full comprehensive model. Indeed, this problem covers a wide range of research fields : numerical analysis, mass-balance modelling, thermo-mechanical modelling. We will address the problem of initial state estimation. There exists a wide range of methods to solve this problem, all of them looking to produce a good fit between the estimated present state and the available observations.

There is a variety of models dealing with ice flow, all of them are based on the quasi-static equilibrium equation and consider ice as a viscous fluid with a non linear viscosity. Most models take advantage of the very small aspect ratio of ice sheets and use a thin layer

approximation but they differ on the order of the approximation. We use here the simplest approximation (shallow ice approximation). Although simple and fast, it preserves the non linearity of the system and is sensitive to the parameters we want to estimate, bedrock elevation and basal drag.

The Greenland and Antarctic ice sheets are huge ice bodies, more than 1000 km wide with harsh climatic conditions that limit field observations. We rely consequently on satellite observations that provide a global coverage. The two major characteristics that can be used for our study are surface elevation and ice velocities at the ice sheet surface. Although in this work we use synthetic experiments and generate our own observations, we stay as close as possible to availability, resolution and uncertainties of real observations. Ice sheet surface elevation is monitored by radar (ERS-1, ENVISAT), or laser (ICESat) satellite altimetry since approximately 20 years. The spatial resolution depends on the instrument but is typically kilometric (smaller along track). The measurements errors for these observations are usually small : for instance for the DEM (digital elevation model) of Antarctica, 42% of the map has a RMS smaller than 2 meters, but RMS can be up to 130 meters in mountainous regions, see [Griggs et Bamber, 2009]. Ice velocity is obtained by radar interferometry and maps are available over the major part of both ice sheets with a kilometric resolution ([Rignot *et al.*, 2011] for Antarctica, [Joughin *et al.*, 2010] for Greenland). The uncertainty ranges from 1 to 17 meters per year [Rignot *et al.*, 2011]. Beside satellite observations, there also are localized measurements of the bedrock topography generally obtained by radar measurements from planes and thus restricted to flight lines. A bedrock topography map has been recently published gathering all the measurements from any country [Fretwell *et al.*, 2013]. The coverage is however still heterogeneous, from closely spaced flight lines in some places to huge regions with no flight at all. The map comes thus with an associated error map, ranging from 20 meters (on measured points) up to 1000 meters (in unobserved areas), and it is worth noting that this error map is obtained thanks to kriging method, with no ice flow consideration.

Data assimilation (DA) covers the methods used to combine model and data in order to estimate initial states or parameters. There are two main classes of DA algorithms : variational, based on optimal control theory (the prototype being 4D-Var) and sequential, based on optimal statistical estimation (the prototype being Kalman filter). DA is widely known in weather and oceanography forecasting, but its introduction in glaciology is fairly recent, in particular for the initial state estimation problem for sea-level rise. [MacAyeal, 1992] and [MacAyeal, 1993] introduced control methods to infer basal drag in ice-streams models, using in particular the self-adjoint property of such models, leading to many application papers ([Rommelaere et MacAyeal, 1997], [Vieli et Payne, 2003]), and later for Full-Stokes models [Morlighem *et al.*, 2010], [Jay-Allemand *et al.*, 2011]). Later on, many DA and inverse methods were introduced in glaciology. The Best Linear Unbiased Estimation (BLUE) and Optimal Interpolation (OI) methods were introduced by [Arthern, 2003] and [Berliner *et al.*, 2008]. The Robin inverse method due to [Chaabane et Jaoua, 1999] has been introduced by [Arthern et Gudmundsson, 2010] for ice-sheet models. And finally, [Heimbach et Bugnion, 2009] presented the first adjoint ice-sheet model derived automatically.

As the ice surface elevation is pretty well (in terms of accuracy and data density) observed by satellite sensing, initial state estimation focuses on bedrock topography and basal sliding law estimation. The joint identification of bedrock topography and basal friction law is a difficult problem. Indeed, different configurations (e.g. a low bed and no sliding versus a higher bed with sliding) may lead to identical surface observations. This is why most of the previous works on this subject chose to investigate the basal sliding identification only, usually for local modelling (glaciers or ice-streams). Recently, a couple of papers addressed the ice-sheet initialisation problem : [Arthern et Hindmarsh, 2006] with a BLUE/OI method and [Gillet-Chaulet *et al.*, 2012] with control and Robin methods. And more recently still, researchers start to investigate coupled inversion of bedrock topography and basal drag simultaneously, see [Raymond-Pralong et Gudmundsson, 2011] (using Bayes theorem) and [van Pelt *et al.*, 2013] (with a simple Picard iteration to reduce the observation-model misfit).

Both variational and Kalman DA methods are equivalent and optimal under very restrictive hypotheses (in particular model linearity and errors gaussianity). Of course in the framework of realistic applications, models are seldom linear and errors are not Gaussian, so that both methods have different advantages and drawbacks, and may compare differently depending on the application. In this paper we made the choice to start our investigation with a Kalman based method. Kalman filter (KF) was introduced by [Kalman, 1960] for linear systems. Later on it was extended to non-linear systems. Because of the curse of dimensionality the Extended Kalman Filter is not practical on large realistic systems, as the covariance and gain matrices are too large to be computed or even stored on a computer. [Evensen, 1994] then introduced the Ensemble Kalman filter (EnKF), which is a Monte-Carlo approximation of the KF, that avoids the computation of the covariance matrices. Following the corrected version of the algorithm by [Burgers *et al.*, 1998], numerous authors used the EnKF and proposed improvements and variants.

To our best knowledge, neither KF nor EnKF have ever been used in glaciology, and we propose to implement and investigate its performance to tackle the sea-level rise problem. This paper is organized as follows : in Section 4.2.2 we give some details about the DA methods, in Section 4.2.3 we describe the ice-sheet model and in Section 4.2.4 we describe our numerical results. And finally we discuss the results and conclude in Section 4.2.5.

4.2.2 Methods

First introduced by [Evensen, 1994] and later corrected by [Burgers *et al.*, 1998], the Ensemble Kalman Filter (EnKF) is a common approach in data assimilation for state and parameter estimation. In the traditional Kalman filter, the state of a physical system at a time t_k is represented by 2 quantities : $\mathbf{x}_k$, the extended state vector of size $n_{\mathbf{x}}$ which contains an estimation of all variables and searched parameters, and $\mathbf{P}_k$, the error covariance matrix of this estimation. The EnKF as a Monte-Carlo method approximates both quantities by the use of an ensemble of N_e realisations $\mathbf{x}_k^{(i)}$, $i = 1, \ldots, N_e$. The

estimation of the extended state vector is given by the ensemble mean

$$\overline{\mathbf{x}}_k = \frac{1}{N_e}\sum_{i=1}^{N_e}\mathbf{x}_k^{(i)} \tag{4.3}$$

and $\mathbf{P}_k$ by the ensemble covariance matrix

$$\mathbf{P}_k = \frac{1}{N_e - 1}\mathbf{X}_k{\mathbf{X}_k}^T \tag{4.4}$$

with $\mathbf{X}_k = [\mathbf{x}_k^{(1)} - \overline{\mathbf{x}}_k, \ldots, \mathbf{x}_k^{(N_e)} - \overline{\mathbf{x}}_k]$ the ensemble perturbation matrix.

EnKF is a two step algorithm : forecast and analysis. Quantities produced during the forecast step (resp. analysis step) are denoted with a superscript ^{f} (resp. superscript ^{a}).

The *forecast* consists in propagating the ensemble using a numerical model :

$$\mathbf{x}_k^{(i)f} = \mathcal{M}_k\left({\mathbf{x}_{k-1}^{(i)}}^a\right) + \eta_k^{(i)} \tag{4.5}$$

with $\mathcal{M}_k$ a nonlinear model which propagates the state vector from time t_{k-1} to time t_k (of course, the model does not change the parameters) and $\eta_k^{(i)}$ a sample vector following the model error pdf. In this paper we assume a perfect model so that $\eta_k^{(i)} = 0$.

At time t_k observations are available. We denote them $\mathbf{y}_k^o$, which is a vector of size $n_{\mathbf{y}}$. The mapping between observations and model state is given by the following relation :

$$\mathbf{y}_k^o = \mathcal{H}_k\left(\mathbf{x}_k^t\right) + \epsilon_k \tag{4.6}$$

with $\mathcal{H}_k$ a possibly nonlinear observation operator, $\mathbf{x}_k^t$ the true extended model state vector and ϵ_k the observation error. ϵ_k is assumed unbiased (zero mean) and Gaussian with covariance matrix $\mathbf{R}_k$.

The aim of the *analysis* step is to correct the forecast ensemble thanks to the observations. Several versions of EnKF exist. Here we use a deterministic one called Ensemble Transform Kalman Filter (ETKF). In the following, we omit the time index k for readability.

ETKF Formulation

The ETKF was first introduced by [Bishop *et al.*, 2001]. The main idea of this formulation is the following : the analysis covariance matrix $\mathbf{P}^a$ can be written as a transformation of the forecast ensemble perturbations as

$$\mathbf{P}^a = \mathbf{X}^f\widetilde{\mathbf{P}}^a{\mathbf{X}^f}^T \tag{4.7}$$

with $\widetilde{\mathbf{P}}^a$ a N_e by N_e matrix. The same idea (in a slightly different formulation) was developed for the SEIK filter ([Pham, 1996], [Pham *et al.*, 1998], [Pham, 2001]). Here we follow the ETKF implementation given by [Hunt *et al.*, 2007]. In order to formulate the analysis step, we first define the forecast ensemble projected in observation space as follows :

$$\left\{{\mathbf{y}^{(1)}}^f = \mathcal{H}\left({\mathbf{x}^{(1)}}^f\right), \ldots, {\mathbf{y}^{(N_e)}}^f = \mathcal{H}\left({\mathbf{x}^{(N_e)}}^f\right)\right\} \tag{4.8}$$

where $\overline{\mathbf{y}}^f$ is the mean of the previous ensemble and $\mathbf{Y}^f$ is the associated perturbation matrix. The analysis step then consists in minimising the following cost function :

$$\mathcal{J}(\mathbf{w}) = \frac{N_e - 1}{2}\mathbf{w}^T\mathbf{w} + \frac{1}{2}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f - \mathbf{Y}^f\mathbf{w}\right) \tag{4.9}$$

where $\mathbf{w}$ is a vector of size N_e. The minimizer of $\mathcal{J}$ is given by :

$$\mathbf{w}^a = \widetilde{\mathbf{P}}^a \mathbf{Y}^{f^T}\mathbf{R}^{-1}\left(\mathbf{y}^o - \overline{\mathbf{y}}^f\right) \tag{4.10}$$

where

$$\widetilde{\mathbf{P}}^a = \left(\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + (N_e - 1)\,\mathbf{I}\right)^{-1} \tag{4.11}$$

is the inverse of the Hessian of $\mathcal{J}$ at the minimum. The ensemble covariances after the analysis are given by equation (4.7) and the ensemble mean by :

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^f + \mathbf{X}^f\mathbf{w}^a \tag{4.12}$$

The analysis ensemble perturbation matrix is obtained by :

$$\mathbf{X}^a = \mathbf{X}^f\left((N_e - 1)\widetilde{\mathbf{P}}^a\right)^{1/2}\mathbf{U} \tag{4.13}$$

where $\mathbf{U}$ is an arbitrary orthogonal matrix which preserves ensemble mean. For the sake of simplicity, we set $\mathbf{U}$ as the identity matrix in the sequel.

Inflation and ETKF

EnKF is known for suffering from under-sampling issues (the so-called *curse of dimensionality*). Two well-known procedures are developed to correct their symptoms : inflation and localisation. The first one scales up by a factor the ensemble perturbation matrix to avoid the underestimation of the variances ([Anderson et Anderson, 1999], as known as the forgetting factor in [Pham *et al.*, 1996]). In the ETKF formulation of [Hunt *et al.*, 2007], the inflation consists in replacing (4.11) by

$$\widetilde{\mathbf{P}}^a = \left(\mathbf{Y}^{f^T}\mathbf{R}^{-1}\mathbf{Y}^f + \frac{N_e - 1}{\rho}\,\mathbf{I}\right)^{-1} \tag{4.14}$$

with ρ the inflation factor. It is equivalent to rescale $\mathbf{Y}^f$ by $\sqrt{\rho}$. If the observation operator is linear, it is equivalent to rescale the ensemble perturbation matrix. This procedure is very popular due to its simplicity but can only resolve the chronic variance underestimation due to EnKF.

Localisation and ETKF

Another under-sampling issue of EnKF is the problem of long-range spurious correlations. These can be overcome by the use of localisation. The ETKF configuration in [Hunt *et al.*, 2007] does not allow the use of localisation on forecast error covariance matrix as in [Hamill *et al.*, 2001] or [Houtekamer et Mitchell, 2001]. But localisation on $\mathbf{R}$ is still possible. One way is to perform local analysis by assimilating a subset of observations on a given location [Ott *et al.*, 2004]. It has been implemented by [Hunt *et al.*, 2007]

and is known as LETKF. This method can lead to discontinuities in the analysis when an observation leaves the subset of used observation for a given location. To solve this, [Hunt *et al.*, 2007] select a subset of observations and increase error variance for observations close to the boundary of the subset. It is equivalent for a given location to replace $\mathbf{R}$ by $\rho_l^{-1} \circ \mathbf{R}$ with ρ_l some distance-based correlation matrix (such as the one suggested by [Gaspari et Cohn, 1999]), the origin being the given location and $\circ$ denoting the element-wise (Hadamard or Schur) product. Note that $\rho_l^{-1} \circ \mathbf{R}$ must be changed for each given location even if it uses the same observations as for another location. We use this corrected version of LETKF for our experiments.

Note that (L)ETKF can produce negative ice thicknesses. In those (rare) cases, we set back ice thickness to zero.

4.2.3 Ice sheet model

Basis of ice sheet model

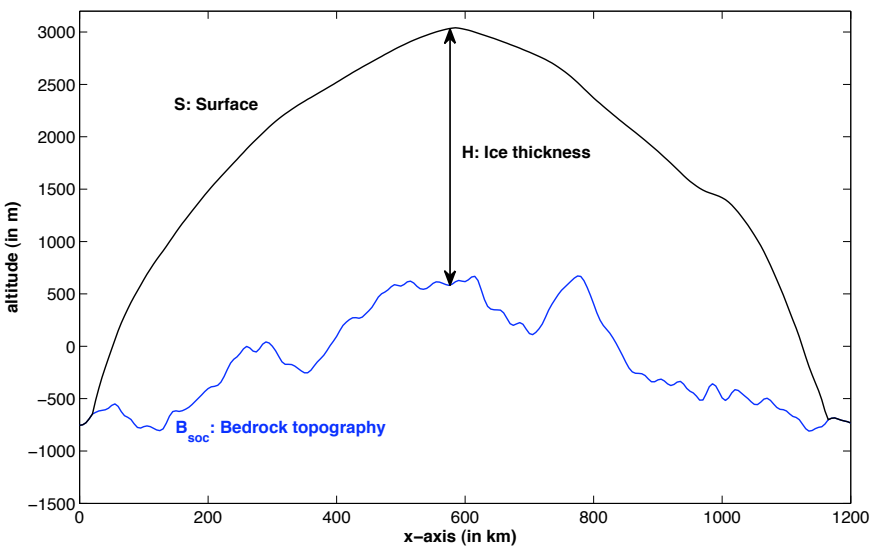

FIGURE 4.1 – Ice sheet geometry : an ice cap of thickness $H(x,t)$ lies on a bedrock (in blue) whose topography is $B_{\text{soc}}(x)$. $S(x,t)$ represents the surface altitude of the ice sheet.

The goal of an ice sheet model is to simulate the evolution of ice thickness $H(x,t)$ over a supposly fixed bedrock topography $B_{\text{soc}}(x)$ (Fig. 4.1). The total has a surface elevation denoted $S(x,t)$, according to :

$$S(x,t) = H(x,t) + B_{\text{soc}}(x) \tag{4.15}$$

where t represents the time, and the x axis represents the latitude.

Ice sheet thickness $H(x,t)$ is governed by the balance between surface mass budget, precipitation or surface melting, and ice flow that drains ice accumulated in the central parts toward the edges of the ice sheet. All these components can evolve with time in response to climatic changes and internal feedbacks. In the context of model initialisation, we will focus on ice flow, surface mass balance assumed to be known.

In this study, we use a parameterization of surface mass balance b_m as a function of atmospheric temperature T_s. More precisely b_m is the sum of accumulation $\mathrm{Acc} \geq 0$ (snow falls) and $\mathrm{Abl} \leq 0$ (melting). These two terms are parameterized as follows

$$\mathrm{Acc} = \mathrm{Acc}_0 e^{-c_1 T_s} \tag{4.16}$$

$$\mathrm{Abl} = \begin{cases} \mathrm{Abl}_0 \left(\dfrac{T_s - T_{\mathrm{no\,melt}}}{T_{\mathrm{no\,melt}}} \right)^2 & \text{if } T_s > T_{\mathrm{no\,melt}} \\ 0 & \text{else} \end{cases} \tag{4.17}$$

Ablation occurs when the temperature exceeds $T_{\mathrm{no\,melt}}$ and T_s depends on surface elevation of the ice sheet S, distance along the x axis (representing latitude) and climate F_{clim}, according to the formula :

$$T_s(x,t) = F_{\mathrm{clim}} + \lambda\, x + \gamma\, S(x,t) \tag{4.18}$$

In this work we have

$$\lambda = \frac{1}{111000}$$

and the vertical lapse rate is

$$\gamma = -0.0063°/m$$

This parameterization reproduces qualitatively the typical surface mass balance over an ice sheet and the feedbacks associated with its possible changes in geometry.

Ice flow is gravity driven, it follows the quasi static equilibrium equation, i.e.. the momentum conservation equation in which the only body force is due to gravity and acceleration is negligible.

Incompressibility is assumed for the whole domain. This is a classical approximation in ice sheet modelling. Although at the surface snow slowly densifies in firn and ice, this process is restricted to the top 100 m of the ice sheet (over about 3000 m). Thus, for large scale modelling, this layer is replaced by an ice equivalent one with density equal to ice density.

Ice behaves as a viscous fluid with a non linear viscosity. It is usual among glaciologists to use the empirical "Glen" flow law [Duval, 1979] with an exponent 3 (see below for more details). In the following however we will use a polynomial law associating the Glen flow law and a Newtonian one. This law was suggested by [Lliboutry, 1993], [Cuffey et Paterson, 2010], and can better represent deformation at very low stresses existing for instance in the upper part of the ice sheet. Moreover, it prevents infinite viscosity in case of no deformation. It is written as :

$$\dot{\epsilon}_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \left(A\tau^2 + \phi \right) \tau'_{ij} \tag{4.19}$$

where ϵ_{ij} is the train rate tensor component, u_i the velocity in the direction i, τ'_{ij} is the deviatoric stress component, A and ϕ are the coefficients of, respectively, the Glen flaw law ($n = 3$) and the linear (Newtonian) flow law and τ^2 is the effective stress, or formally the second invariant of the deviatoric stress tensor.

Boundary condition at the surface are atmospheric pressure and no tangential stress. At the ice bed interface, a dragging law (usually called sliding law by glaciologists) must be prescribed to derive the tangential stress τ_b (also called basal drag).

The dragging law depends on the substratum, and especially wether it is hard-bed or soft-bed (sediment). However, many of the sliding/dragging laws proposed in glaciology can be written as follows (see [Cuffey et Paterson, 2010]) :

$$\tau_b = KU^a_{\text{slid}} \tag{4.20}$$

where U_{slid} is the basal tangential velocity, K and a are respectively coefficient and exponent of the dragging law and their value depends on the characteristics of the rock-ice interface, basal temperature, sediment thickness, water pressure, roughness,... In this article, we will restrict our work to a linear law (a=1).

Note that when the ice sheet reaches the ocean, ice begins to float and forms structures called ice shelves (or floating tongues if they are narrow). Basal drag there is zero and equation (4.20) is not valid. In this work we will only consider an ice sheet resting on the bedrock (grounded ice sheet).

Associating quasi static equilibrium and viscous behavior leads to the Full Stokes equations. There is a whole hierarchy of approaches to solve this problem (see for instance the ISMIP inter-comparison project, [Pattyn *et al.*, 2008]). Recently, a few ice sheet models were designed to solve rigourously the Full Stokes problem but given the numerical cost they are usually restricted to regional scale or can only afford a few century simulations, see [Gillet-Chaulet *et al.*, 2012], [Larour *et al.*, 2012]. All the other ice sheet models use approximations based on the very small aspect ratio of ice sheets. Ice sheets are indeed a few kilometers thick and more than a thousand kilometers wide so that asymptotic development are possible. In this work, we will use an approximation called Shallow Ice approximation that preserves the main non linearity of the problem, linked to the deformation law, and the dependency on the dragging law coefficient that is one of the characteristics we want to assimilate.

Shallow ice approximation (SIA) for a flowline model

This approximation [Hutter, 1983] describes ice deformation in the vertical plane and allows to calculate the vertical profile of horizontal velocity. It is based on asymptotic approach and is vertically integrated. We use here a model, Winnie, which is a flow line and isothermal version of a 3D ice sheet model, GRISLI, (see [Ritz *et al.*, 2001], [Peyaud *et al.*, 2007], [Quiquet *et al.*, 2012]). According to the SIA, ice flow is in the direction of the steepest slope of the surface and can be thus relatively well represented by a flow line following that direction. The isothermal assumption allows uniform deformation coefficients, A and ϕ. It is taken as a first step in this study because taking thermal

processes into account would require to assimilate a third poorly known parameter that is the geothermal heat flow. Such a topic is beyond the scope of this study.

The vertically averaged horizontal velocity, $\overline{U}$, is given by :

$$\overline{U} = \overline{U_{\text{def}}} + U_{\text{slid}} \tag{4.21}$$

$\overline{U_{\text{def}}}$ comes from ice deformation in the vertical plane and according to the SIA equations and our constitutive equation (polynomial) reads as :

$$\overline{U_{\text{def}}} = \left[-\frac{A}{5}(\rho g)^3 H^4 \frac{\partial S}{\partial x}^2 - \frac{\phi}{3}\rho g H^2 \right] \frac{\partial S}{\partial x} \tag{4.22}$$

where x is the horizontal coordinate, ρ is ice density, g gravity, H and S are ice thickness and ice sheet surface elevation.

Note that the non linearity of the constitutive equation is implicitly taken into account by the exponents for surface slope and ice thickness.

The horizontal velocity at the surface, U_s is one of the observed characteristic and is given by

$$U_s = \left[-\frac{A}{4}(\rho g)^3 H^4 \frac{\partial S}{\partial x}^2 - \frac{\phi}{2}\rho g H^2 \right] \frac{\partial S}{\partial x} + U_{\text{slid}}. \tag{4.23}$$

To derive the basal velocity U_{slid}, we assume that basal drag and basal velocity are linked through a linear dragging law and that the basal shear stress τ_b is given by the SIA :

$$\tau_b = \rho g H \frac{\partial S}{\partial x} = -\beta\, U_{\text{slid}} \tag{4.24}$$

where β is positive and is one of the characteristics we assimilate. Because this parameter covers several order of magnitude, we often use $\alpha = \log_{10}(\beta)$.

Although simple, this method reproduces the various types of ice flow found in ice sheets : slow flow in the central parts where ice deformation is dominant, and fast flow in ice streams for low values of β. However, due to the thin layer approximations, one should be careful in the interpretation of sharp horizontal changes in the bedrock topography or basal drag coefficient.

Finally, the evolution of ice sheet geometry is derived from the mass conservation equation :

$$\frac{\partial H}{\partial t} = b_m - \frac{\partial \overline{U} H}{\partial x} \tag{4.25}$$

In our model, mass conservation is the only time dependent equation and H is the only prognostic variable. Ice velocities are diagnostic (completely determined by the ice sheet geometry) because acceleration terms are neglected.

Numerics

The velocity field is obtained analytically as a function of ice thickness and surface slope. The mass conservation equation is solved by a finite difference method with a staggered grid, velocities and slopes grid points are taken at midpoint between thickness grid points. Because velocity is strongly dependent on the ice sheet geometry, the time scheme is semi-implicit.

4.2.4 Numerical results

We now present our numerical results. This section is divided into four paragraphs : firstly we present the numerical setup of the various experiments. Secondly we explain how to build the initial ensemble members. Thirdly we present the preliminary results obtained with the ETKF (without inflation or localisation) and a large ensemble. Finally, we study the impact of smaller ensemble sizes, for ETKF with inflation and localisation.

We recall here that our problem is to estimate jointly the ice thickness H, the bedrock topography B_{soc}, and the basal sliding coefficient $\beta = 10^{\alpha}$. As this problem has never been studied before with ensemble filtering, this work is focused on twin experiments, because it will allow quantitative assessment of the method performance.

Numerical setup

Twin experiments consist in comparing three model states : the reference (or true) state, which is used to generate observations ; the background state, which is another model state, different from the reference ; the assimilated state, also called analysis, resulting from the DA procedure. We can then compare the misfit between the assimilation and the reference, and the misfit between reference and background (no assimilation).

We performed the twin experiments over a 20-year time window. This is realistic for this problem, as the available observation records roughly span 20 years. Note that this small window size makes our proposed DA experimental setup quite challenging.

We first explain how to choose the reference state. We set up the model Winnie with 241 grid-points and a space-step $\Delta x = 5$ km, so that the state vector $(H(x), B_{\text{soc}}(x), \alpha(x))$ length is 743. The reference run is initialised at $t = 0$ with the ice thickness H and bedrock topography B_{soc} presented in Fig. 4.2. This ice thickness is chosen as the stable state for a constant climate forcing. In practice, it is produced from a very long (50 000 years) model run, starting from $H = 0$. Also at $t = 0$ the sliding parameter coefficients $\beta(x) = 10^{\alpha(x)}$ are defined as in Fig. 4.3.

With this reference initial state, we now impose a linear climate change ($+0.2°$ for 20 years) and we perform a 20-year model run in order to get the reference state over the assimilation window. Note that after 20 years the ice thickness is very similar to the initial one : the dynamics, highly nonlinear, are also very slow.

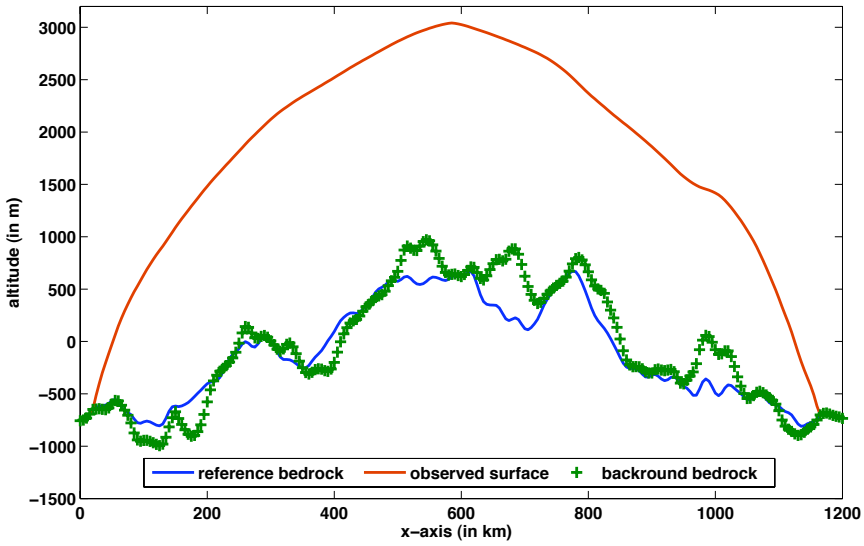

FIGURE 4.2 – Ice sheet geometry for the reference and background states. The x-axis represents the horizontal extent of the ice-sheet (in km). The y-axis is the elevation in meters from the sea-level. The reference bedrock topography is in blue, the background bedrock is in green crosses. Reference and background surface altitude are identical, in red.

We then use the reference state to generate synthetic observations. We choose $\Delta t = 1$ year as a realistic observation interval, therefore our assimilation experiments will have 20 observation times, and thus 20 analysis steps. Every year, we observe the surface elevation S and the surface velocity U_s at each grid point (we detailed the reference surface velocity profile with its sliding and deformation components in Fig. 4.4). We also observe the bedrock topography every year and every 30 grid points. To simulate observation errors, we add noise to these (perfect) observations. More precisely, we add a Gaussian white noise to each observation independently, with realistic standard deviations : 2 m for surface elevation, 3 m/yr for surface velocity and 20 m for bedrock topography.

Every DA experiment presented in this section uses the same background state (see Fig. 4.2 and 4.3) as initial ensemble mean.

Initial ensembles

As we said before, the initial ensemble mean is set to the background state. Before we describe the procedure, we have to acknowledge that some observations were used twice : once to build initial ensembles (i.e. as a priori information), and another time in the DA system (i.e. as observations). This violates a crucial hypothesis required by Kalman based filters : the independence between the a priori estimation and the observations. However this is a characteristic of realistic frameworks : indeed the state-of-the-art a priori bedrock topography Bedmap2 [Fretwell *et al.*, 2013] is produced using surface altitude observations (as well as all available bedrock measurements).

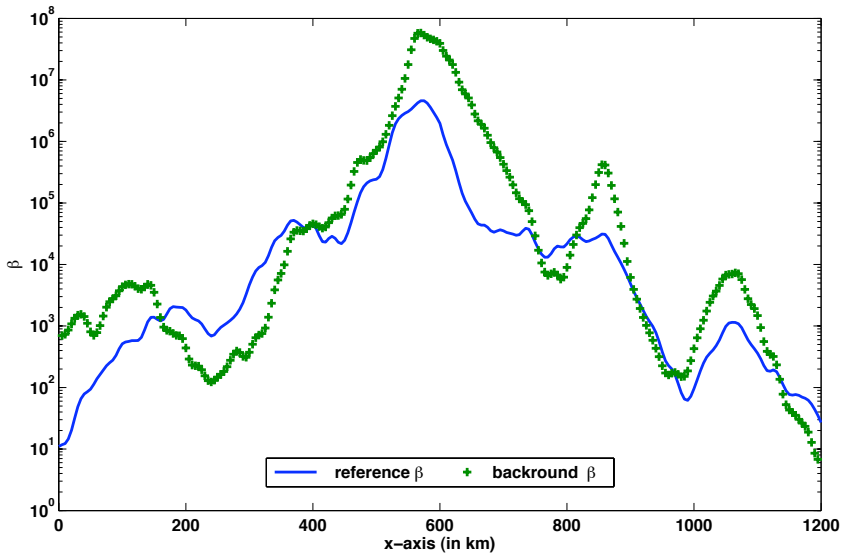

FIGURE 4.3 – Basal sliding parameter $\beta(x)$ as a function of the ice-sheet extent x, for the reference state (solid blue) and the background (green crosses), in log-scale.

Now let us describe how to build initial ensemble members. For $\beta = 10^{\alpha}$, we first generate an ensemble of $\alpha(x)$ parameters drawn from a Gaussian distribution with a given covariance matrix (a squared exponential covariance function). Then we transform this ensemble of α into an ensemble of β.

Then we generate an ensemble of bedrock topographies according to a Gaussian distribution, whose mean is the background and whose covariance matrix is consistent with observation statistics. The covariance matrix requires two ingredients : the variances and a correlation matrix, according to the formula $\mathbf{B} = \mathbf{\Sigma C \Sigma}$, where $\mathbf{\Sigma}$ is the square root of the diagonal matrix of variances, and $\mathbf{C}$ is a correlation matrix. Where bedrock topography observations are available, the variance is set to the observation variance (20 m in standard deviation). Where there is no ice, the variance is set to the surface altitude observation variance (2 m in standard deviation). Elsewhere the variance increases with the distance to the closest observation (bedrock or surface) available. Fig. 4.5 presents the standard deviations whose squares form the diagonal of $\mathbf{\Sigma}$. At worst the standard deviation is equal to 320 m, which is accurate compared to available real uncertainty estimation on B_{soc}.

For the correlation matrix, we assume the correlation of bedrock topography between two grid points only depends on distance (isotropy). Classically a Gaussian function is taken for the correlation function. However we choose a combination of two Gaussian functions (see Fig. 4.6), as using only one Gaussian function produced too smooth or too rough bedrock topographies. Indeed, this combination allows us to ensure the smoothness of simulated bedrock topographies at large scale and their slight roughness at shorter scales. This behaviour seems to produce realistic bedrock topographies.

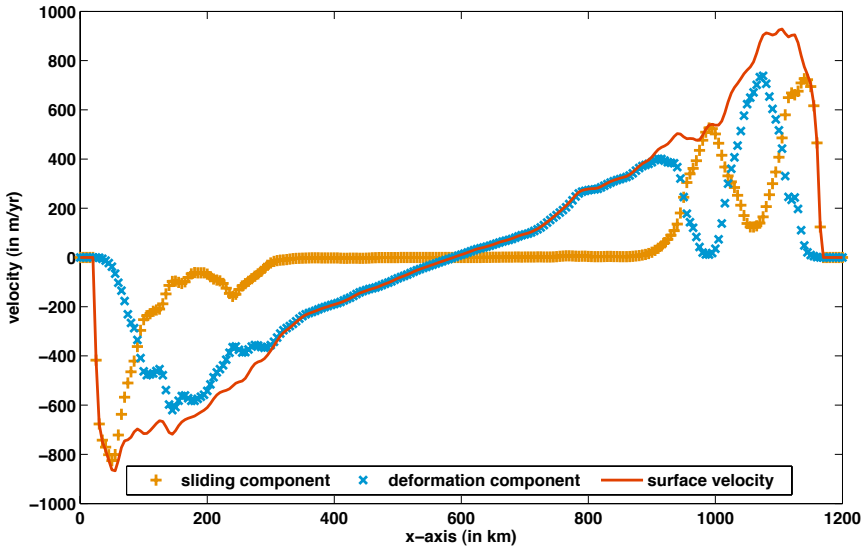

FIGURE 4.4 – Ice velocity at the surface of the ice-sheet for the reference state, in meters per year, as a function of the ice-sheet extent x. The total surface velocity U_s is represented in red, its sliding component U_{slid} is in orange crosses, its deformation component $U_{s,def}$ with cyan x (we recall that $U_s = U_{slid} + U_{s,def}$).

We now have an ensemble of bedrock topographies. We then generate an ensemble of surface elevations according to a Gaussian distribution. As before, its mean is the true surface elevation (from the reference run) and its covariance matrix is consistent with surface elevation observation statistics (white noise with standard deviation of 2 m). To obtain the initial ensemble for ice thickness, we subtract the ensemble of bedrock topography from the ensemble of surface elevation : $H = S - B_{soc}$. Then we correct this new ensemble to avoid negative ice thicknesses (we just set them back to zero). Finally we run the model for each ensemble member during 1 year in order to obtain more physically balanced ice sheets and we rescale the produced ensemble so that its mean remains equal to the background. An example of 50 ensemble members is presented in Fig. 4.7.

First results with a large size of ensemble

We first perform twin experiments with a large ensemble (size $N_e = 1000$) compared to the state vector dimension, in order to validate our ETKF approach, without having to deal immediately with under-sampling issue. We therefore run the ETKF without inflation nor localisation. The bedrock topography obtained after 20 years is presented in Fig. 4.8 (mean topography with or without assimilation) and 4.9 (standard deviation with or without assimilation). This clearly shows that the final results on ice thickness and bedrock topography are very accurate, as the average RMS error for bedrock topography is decreased from 207.5 m to 45.9 m and the maximum standard deviation from 671.6 m to 150.5 m. The average RMS error for ice thickness is also decreased from 207.5 m to 45.5 m and the maximum standard deviation from 671.6 m to 150.5 m.

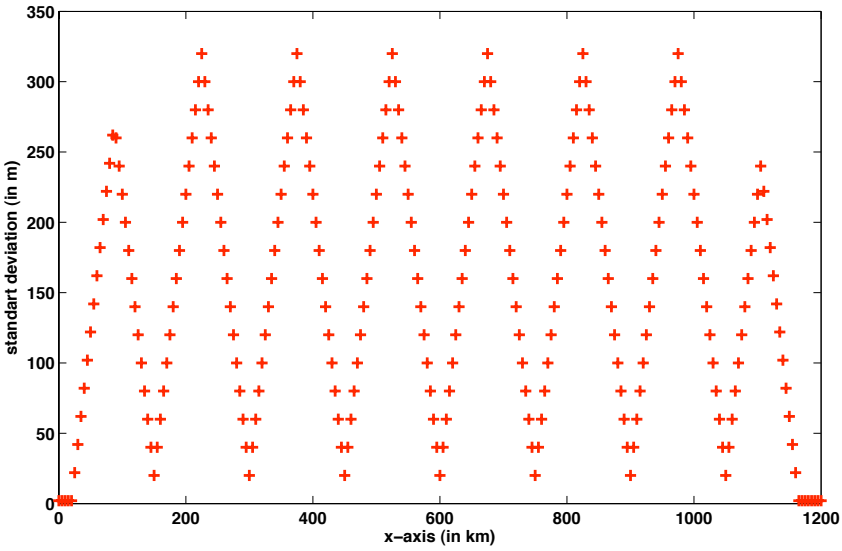

FIGURE 4.5 – Standard deviations (in meters) for the matrix $\mathbf{\Sigma}$ used to generate the initial bedrock topography ensembles, as a function of the extent x. Local minimums correspond to either "no ice" points (close to the boundaries of the domain) or observation points.

Fig. 4.10 presents the results for the β parameter. We can see that the accuracy after assimilation is quite good at the edges of the ice sheet and worse in the center, where β is large, and also where there is no ice. However, poor results on large β are not meaningful. Indeed, we are mostly interested in recovering the sliding component of the velocity, and it is well known among glaciologists that its sensitivity to variations of β for large β is very low : $\beta \geq 10^5$ leads to zero sliding velocity in any cases, so that $\beta = 10^6$ or 10^7 does not make any difference. Similarly, β in areas where there is no ice is meaningless. This is confirmed in Fig. 4.11, which shows the sliding counterpart of the surface velocity : we can see that the analysis is much closer to the reference than the background. More precisely, Fig. 4.12 shows the standard deviation of the error on surface sliding velocity with or without assimilation, and we can see that it is much reduced with assimilation. However, U_{slid} is poorly recovered at some points around $x = 1000$ km. This corresponds to a zone where $\frac{\partial S}{\partial x}$ is much smaller, which is the onset of the sliding. At this point the dynamics are close to being discontinuous, so that the Shallow approximation fails, and the ETKF as well. Therefore, this also explains that the average RMS error for U_{slid} is only decreased from 238.1 m/yr to 231.3 m/yr.

Dealing with small sizes of ensemble

The results obtained with a large ensemble are satisfying. Nevertheless we will not be able to perform such experiment with a full 3D large-scale ice sheet model. Indeed in that case the state vector dimension is larger than 100 000, and it would be impossible to use an ensemble that large. With this remark in mind we now perform ETKF experiments with smaller ensembles : 100, 50 and 30. Without localisation and/or the ETKF is known

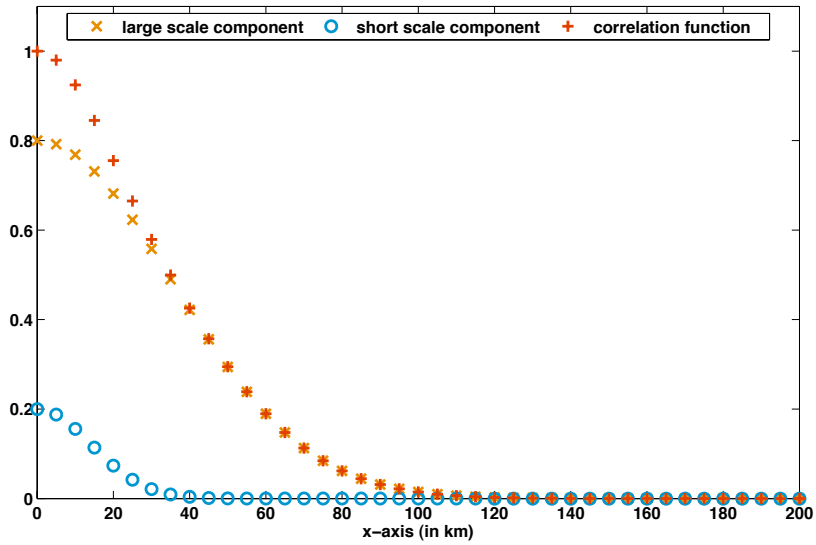

FIGURE 4.6 – Correlation function (in red crosses) used for the generation of initial bedrock topography ensembles. This is a combination of two Gaussian functions : one to capture large scale behaviour (orange x) and one for short scale behaviour (cyan circles).

for its divergence for small ensembles. To check this fact, we perform one experiment with 100 members, without localisation nor inflation. Fig. 4.13 presents the final analysed bedrock topography, which is clearly degraded with respect to the background. Actually, RMS error on bedrock topography raises from 207.5 m for the background state to 302.4 m for the final analysis step, which is a 50% increase.

Consequently, for the remaining part of this section, we use LETKF with manually tuned inflation and localisation. Let us recall briefly how localisation is performed. The observation error covariance matrix $\mathbf{R}$ is modified into $\rho_l^{-1} \circ \mathbf{R}$, with ρ_l some distance-based correlation matrix ([Gaspari et Cohn, 1999]). The distance l must be tuned manually in order to achive good results. It corresponds to the maximal distance between a given grid point and the observations used in analysis at this point.

We performed numerous experiments in order to manually determine the best l and the best inflation parameter for each ensemble. We chose the bedrock topography RMS error as a score to measure the performance of inflation and localisation. Regarding localisation, the best results were obtained with l between $16\Delta x$ (ensemble size 30) and $24\Delta x$ (ensembles size 50 and 100), so that this distance l is quite insensitive to ensemble sizes. On the contrary, the optimal value for the inflation parameter proved to be much more sensitive on ensemble sizes (ranging from 0.98 to 1.14).

Table 4.1 presents the RMS and maximal errors for B_{soc} and U_{slid} for the optimally tuned LETFK with 30, 50 and 100 members, as well as the ETKF for 100 and 1000 members. We can see that despite the small ensemble sizes the results are pretty good. In

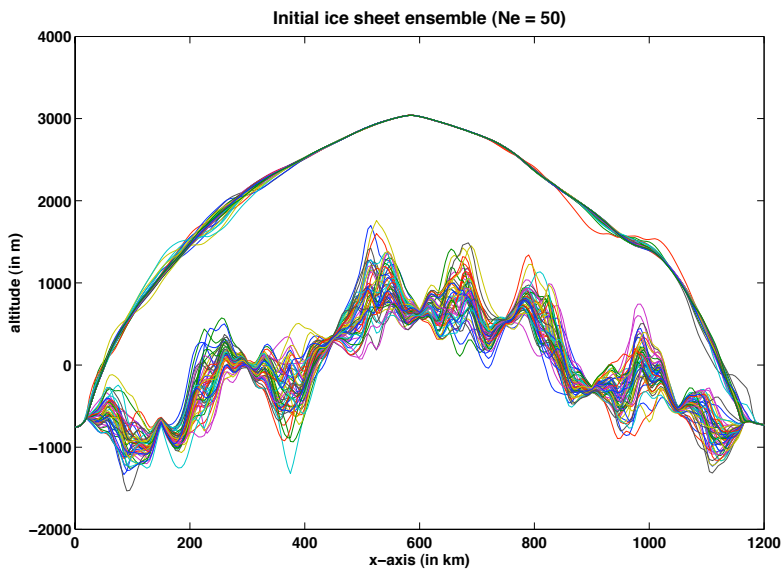

FIGURE 4.7 – Example of initial ice sheets ensemble (50 members) : bedrock topographies $B_{\text{soc}}(x)$ and surface elevation $S(x)$, as functions of the horizontal distance x.

particular, for the 100-member ensemble, the quality of the retrieved B_{soc} is even better than the ETKF results with large ensemble ($N_e = 1000$). Fig. 4.14 and Fig. 4.15 present the bedrock topography and its errors standard deviation with 30-, 50- and 100-member ensembles and confirm the good performance of the filters. As before, the performance on β itself is not significant, so we show only results on the sliding velocity. Fig. 4.16 shows the sliding velocity, and Fig. 4.17 presents the standard deviation for the sliding velocities. As previously these figures enlighten two different regimes. First, where the ice is either grounded or in full sliding the filters perform quite well. Second, where the ice just starts to slide (where the proportion between the sliding and deformation counterparts of the velocity changes) the filters fail and the RMS is large. In these areas we believe that the model itself is faulty, so that the filters are in difficulty.

We also performed LETFK (with inflation) experiments with smaller ensembles (size 10 and 20) but the results were very dissatisfying because of filter divergence (not shown here).

4.2.5 Conclusion and further directions

In this article, we developed an Ensemble Transform Kalman Filter [Hunt *et al.*, 2007] in order to estimate jointly the bedrock topography, the ice thickness and the basal sliding parameter of an ice sheet. The originality of this work is the application of an EnKF approach in ice sheet modelling. We performed twin experiments with a flowline simplified ice sheet model using the Shallow-ice approximation and a sliding law for basal velocities. Every experiment used surface elevation and surface velocity observations all over the ice sheet and a couple of observations of bedrock topography. First we successfully tested

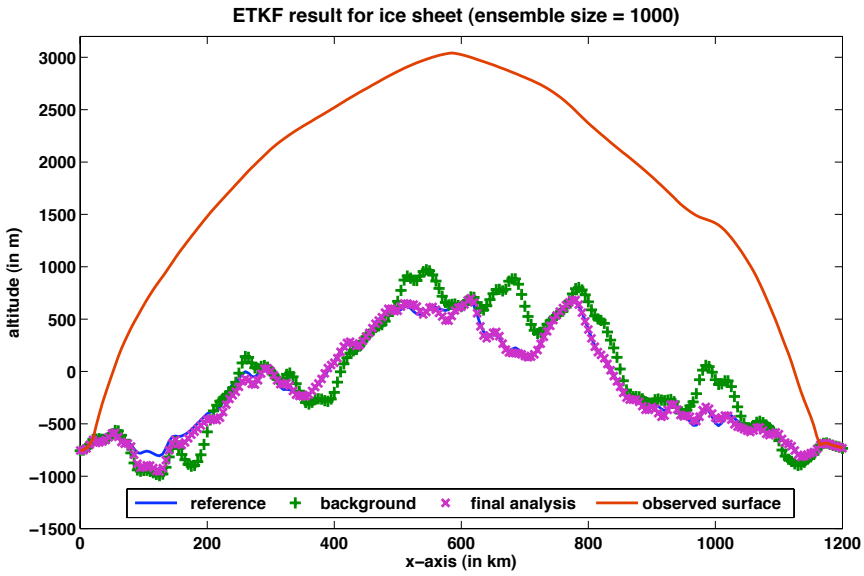

FIGURE 4.8 – Ice sheet geometries after 20 years, with the 1000-member ETKF. The analysis bedrock topography (final mean ensemble in purple x) is compared to the background (green crosse) and the reference (blue). The surface elevation is accurately observed (in red), so that background, reference and analysis are similar.

our DA approach with a large ensemble to validate the use of ETKF. Then we tried with smaller ensembles. In those cases, localisation and inflation are mandatory. Obtained performances were good even for ensembles with a size as small as 30.

However localisation and inflation were manually tuned. Localisation distance was not very sensitive to ensemble size but inflation was. In order to avoid the manual tuning of optimal inflation parameter, we could use online estimation. This is a growing interest in the DA community. Some works such as [Bocquet, 2011] or [Bocquet et Sakov, 2012] provided convincing theoretical arguments for the use of inflation and its automatic computation.

Finally, results shown here are preliminary as we used a flowline model. The logical choice to improve the model complexity would be to use a hybrid shallow ice-shallow shelf model as in GRISLI [Ritz *et al.*, 2001].

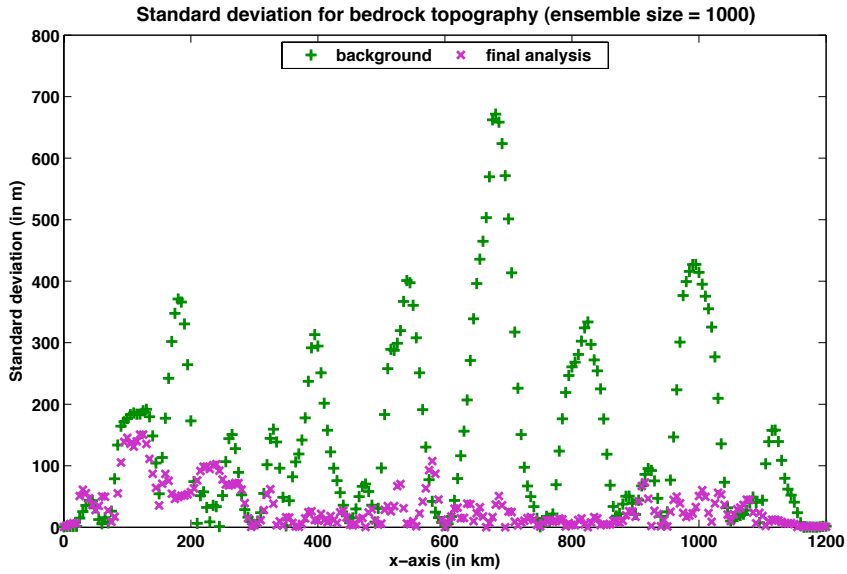

FIGURE 4.9 – Standard deviation for the bedrock topographies errors (compared to the reference) after 20 years of the 1000-member ETKF. The analysis standard deviation is in purple x, the background in green crosses. The average RMS error is decreased from 207.5 m to 45.9 m and the maximum standard deviation from 671.6 m to 150.5 m.

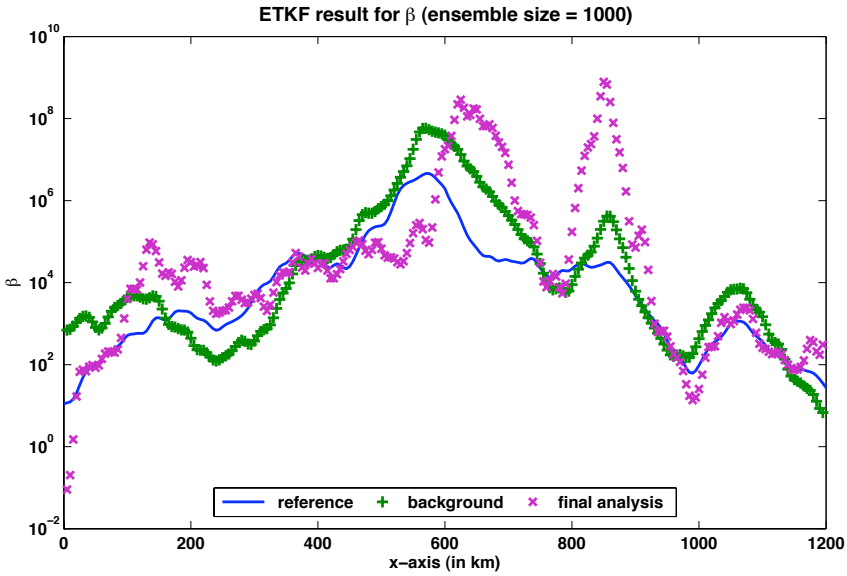

FIGURE 4.10 – β parameter after 20 years of the 1000-member ETKF. The analysis (purple x) is compared to the background (green crosse) and the reference (blue). Values of β above 10^5 are all equivalent in terms of sliding velocity. Similarly, β is meaningless where there is no ice (close to the boundaries).

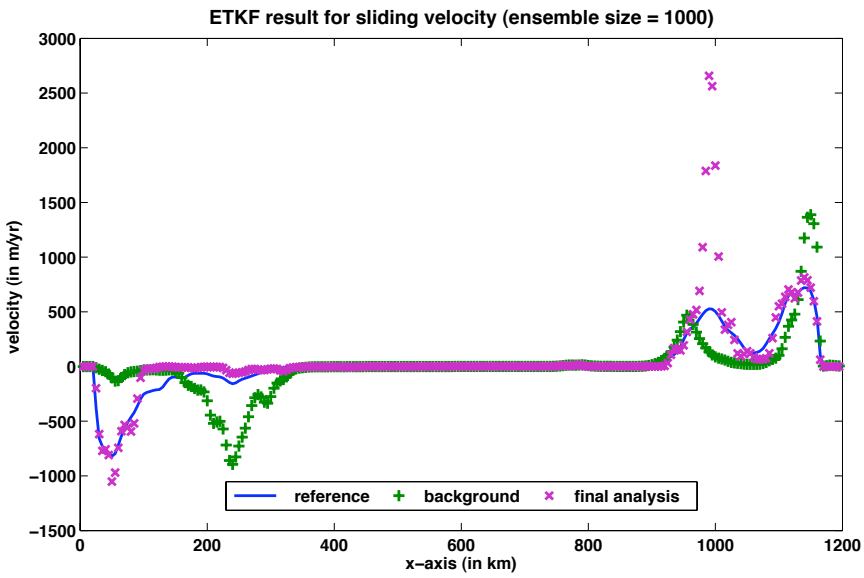

FIGURE 4.11 – Sliding component U_{slid} of the surface velocity after 20 years of the 1000-member ETKF. The analysis mean (purple x) is compared to the background (green crosses) and the reference (blue).

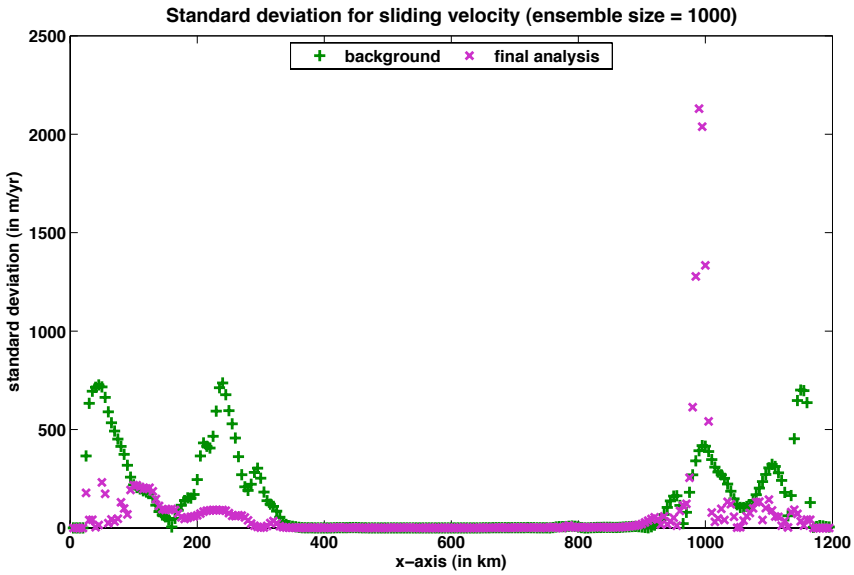

FIGURE 4.12 – Standard deviation between final analysis mean (purple x) or background (green crosses) and reference sliding component of surface velocity after 20 years of the 1000-member ETKF.

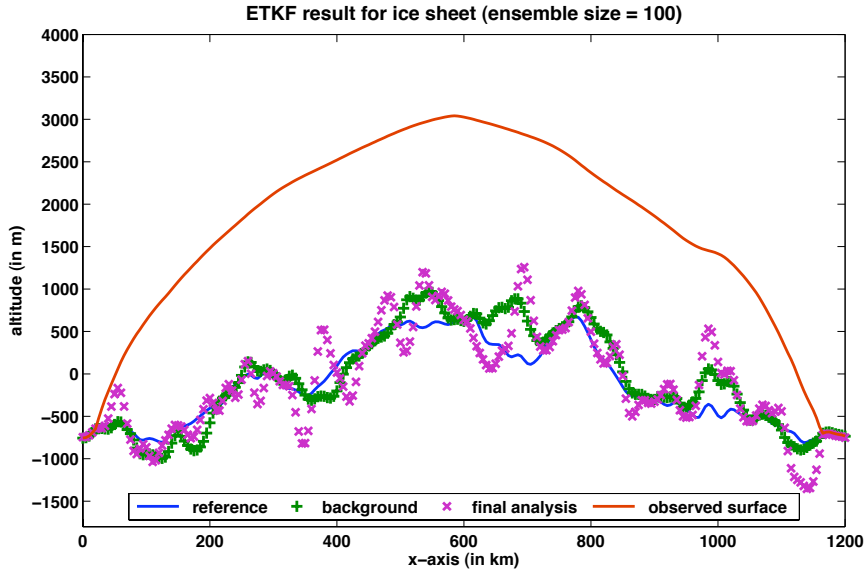

FIGURE 4.13 – Bedrock topography after 20 years, with the 100-member ETKF (no inflation, no localisation). The analysis (final mean ensemble in purple x) is compared to the background (green crosses) and the reference (blue). RMS error raises from 207.5 m for the background state to 302.4 m for the final analysis step and the maximal absolute error from 671.6 m to 1093.7 m.

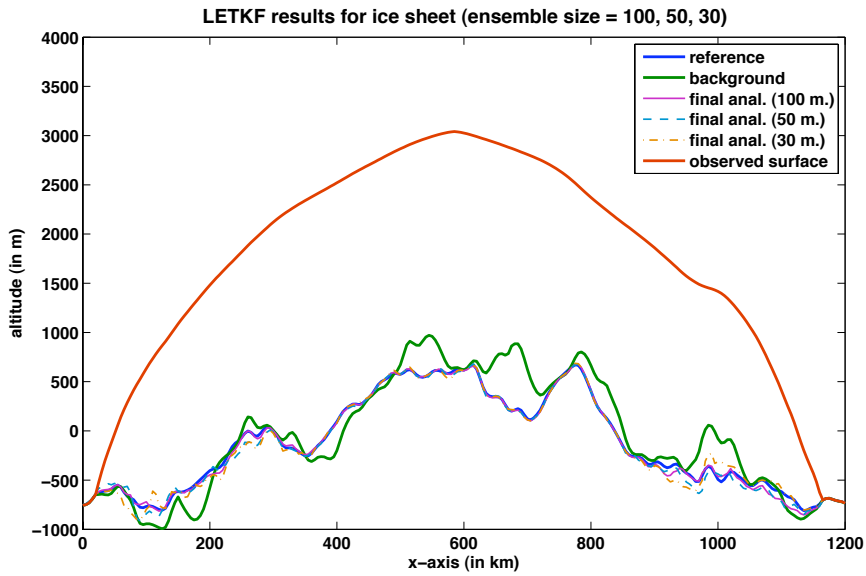

FIGURE 4.14 – Bedrock topography after 20 years of the LETKF with inflation. The background (green) is compared to reference (blue) and the analyses for various ensemble sizes : 100 members (purple), 50 members (cyan) and 30 members (orange). RMS errors evolution is synthesised in Table 4.1.

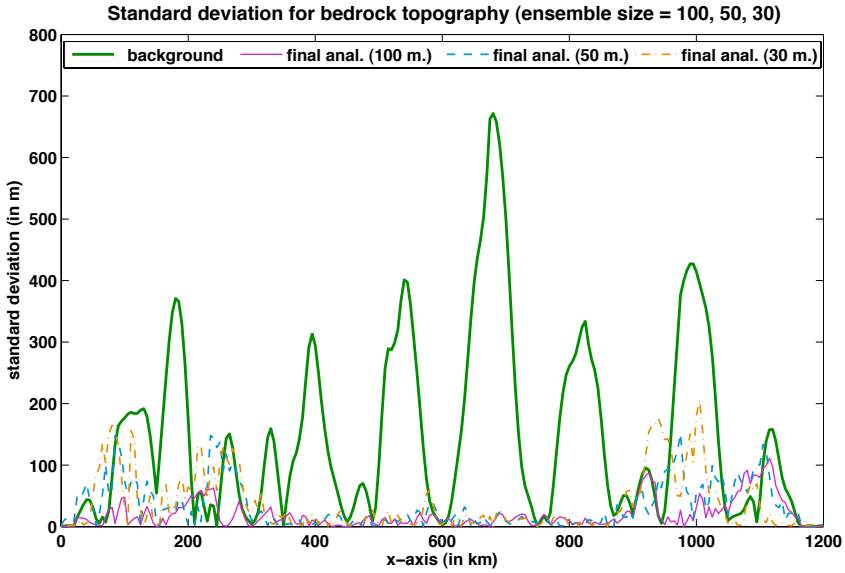

FIGURE 4.15 – Standard deviation of the errors (compared to the reference) for the bedrock topography after 20 years of the LETKF with inflation. The background (green) is compared to the analyses for various ensemble sizes : 100 members (purple), 50 members (cyan) and 30 members (orange). RMS errors evolution is synthesised in Table 4.1.

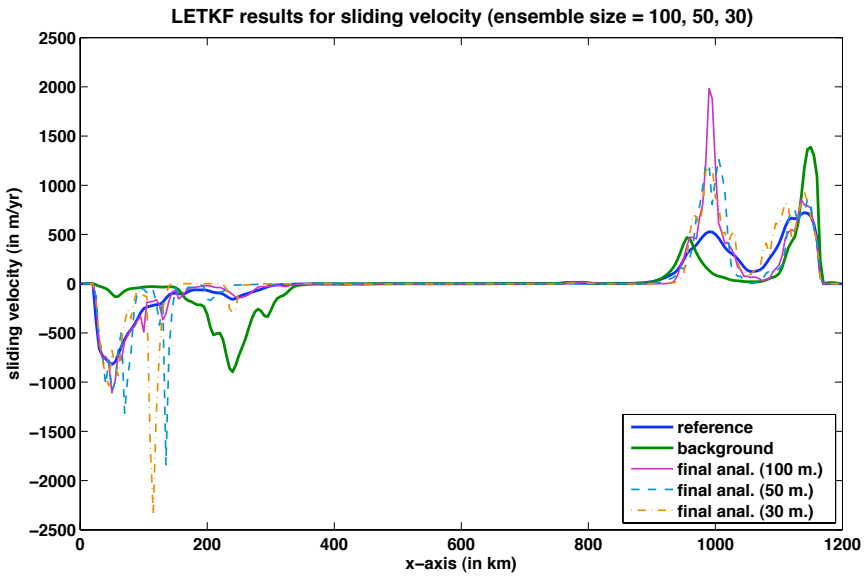

FIGURE 4.16 – Sliding component U_{slid} of the surface velocity after 20 years of the LETKF with inflation. The mean of the final analysis ensemble is compared to the background (green) and the reference (blue), for various ensemble sizes : 100 members (purple), 50 members (cyan) and 30 members (orange). Localisation and inflation parameters are described in Table 4.1.

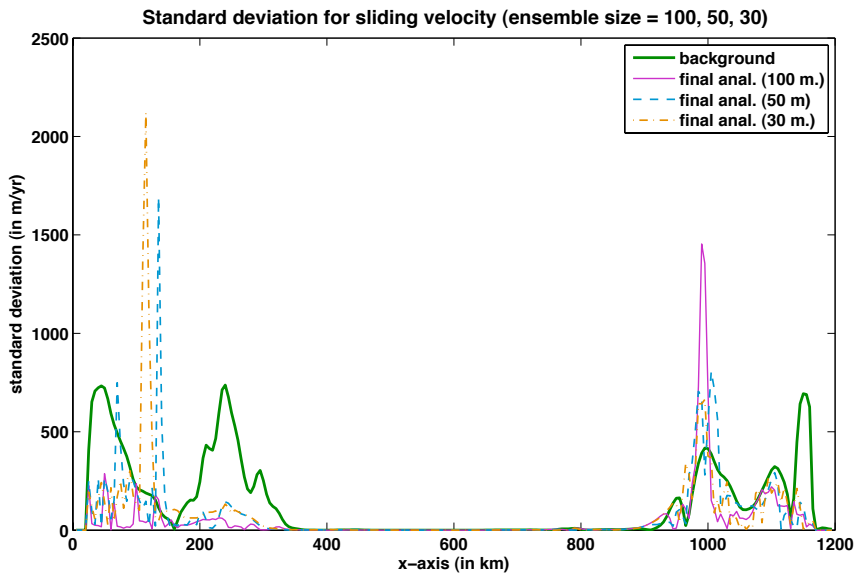

FIGURE 4.17 – Standard deviation of the errors (compared to the reference) for the sliding component U_{slid} of the surface velocity after 20 years of the LETKF with inflation. The background (green) is compared to the analyses for various ensemble sizes : 100 members (purple), 50 members (cyan) and 30 members (orange). RMS errors evolution is synthesised in Table 4.1.

TABLE 4.1 – Summary of best LETKF + inflation results in term of RMS and maximal deviation and comparison with standard ETKF results

Experiment	Size of ensemble	Loc. l	Inflation	RMSE B_{soc}	Max. error B_{soc}
Background	–	–	–	207.5 m	671.6 m
ETKF	1000	–	–	45.9 m	150.5 m
ETKF	100	–	–	302.4 m	1093.7 m
LETKF + inflation	100	$24\Delta x$	1.02	29.9 m	111.2 m
LETKF + inflation	50	$23\Delta x$	1.14	48.8 m	151.1 m
LETKF + inflation	30	$16\Delta x$	0.98	59.7 m	205.5 m

Experiment	Size of ensemble	Loc. l	Inflation	RMSE U_b	Max. error U_b
Background	–	–	–	238.1 m/yr	737.1 m/yr
ETKF	1000	–	–	237.9 m/yr	2130.0 m/yr
ETKF	100	–	–	347.5 m/yr	1895.6 m/yr
LETKF + inflation	100	$24\Delta x$	1.02	158.1 m/yr	1453.9 m/yr
LETKF + inflation	50	$23\Delta x$	1.14	188.0 m/yr	1691.2 m/yr
LETKF + inflation	30	$16\Delta x$	0.98	214.5 m/yr	2120.6 m/yr

4.3 Essais avec un ETKF-N version duale

Durant cette thèse, nous avons aussi mis en place un filtre de Kalman d'ensemble de taille finie. Nous avons choisi d'utiliser la version duale (voir algorithme 2.11). La minimisation de la fonction coût (2.116) se fait par simple dichotomie. Nous avons testé ce filtre dans les mêmes contextes que ceux évoqués durant l'article. Pour un ensemble de taille 1000, l'intérêt est très limité car la taille de l'ensemble est assez grande pour que le filtre soit efficace sans besoin d'inflation. Pour cette raison, nous n'affichons pas les résultats obtenus.

Étudions maintenant l'effet de l'ETKF-N sur l'ensemble de taille 100 utilisé précédemment dans l'article. Les résultats obtenus sont sensiblement meilleurs que ceux obtenus par simple ETKF que ce soit en terme de géométrie de la calotte (voir figure 4.18 à comparer avec la figure 4.13) et surtout en terme de vitesse de glissement (voir figure 4.19).

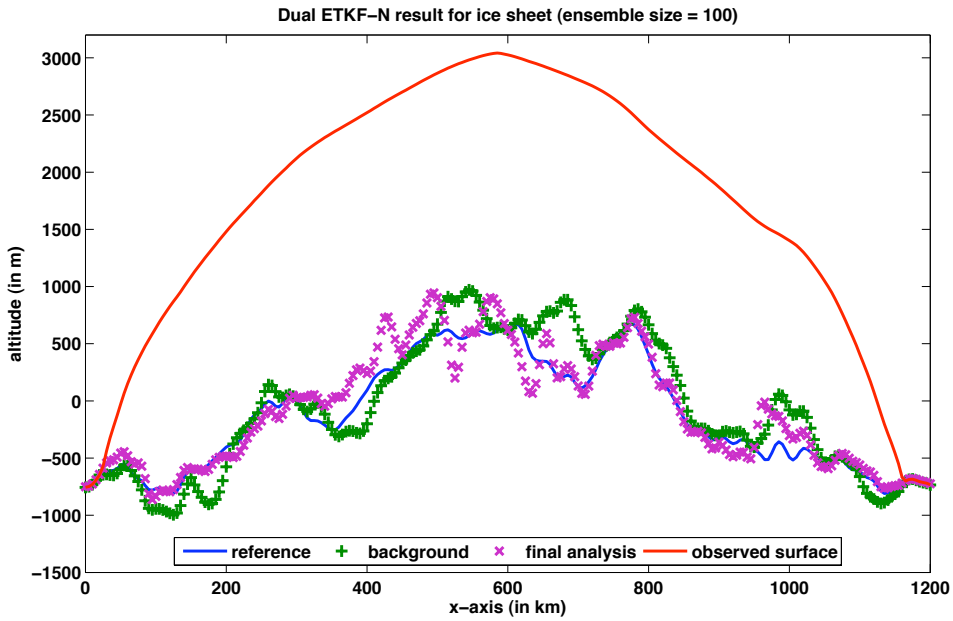

FIGURE 4.18 – Calotte reconstruite par ETKF-N version duale pour un ensemble de taille 100. L'erreur RMS obtenue pour le socle rocheux vaut 155.8 m et l'écart maximal 502.5 m. Ces valeurs sont à comparer avec les valeurs obtenues pour l'ETKF simple (les résultats sont meilleurs) et pour le LETKF avec localisation et inflation réglé manuellement (les résultats sont moins bons), voir tableau 4.1.

Enfin, notons sur cette expérience les forts paramètres d'inflation trouvés comme l'indique la figure 4.20.

Nous avons essayé d'appliquer l'ETKF-N version duale aux ensembles de tailles 30 et 50 mais leur trop faible taille font que l'algorithme diverge. Bien que l'expérience précédente ait mis en lumière l'importance de l'inflation, utiliser la localisation semble être le seul moyen d'obtenir de bons résultats pour des ensembles de si petites tailles.

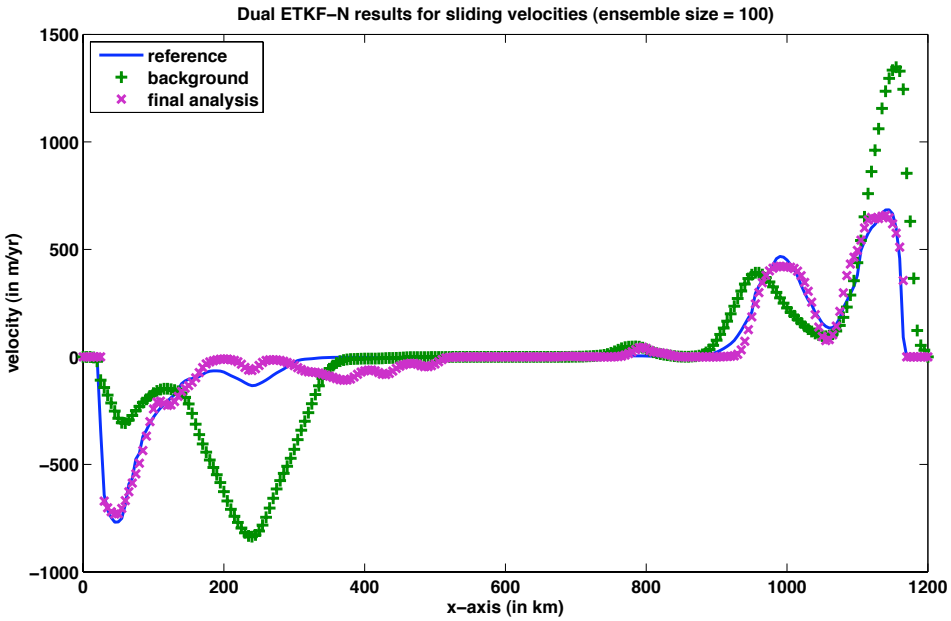

FIGURE 4.19 – Vitesse de glissement reconstruite par ETKF-N version duale pour un ensemble de taille 100. L'erreur RMS vaut 53.7 $m.an^{-1}$ et l'écart maximal 341.7 $m.an^{-1}$. Ces valeurs sont à comparer avec les valeurs obtenues pour l'ETKF simple (les résultats sont meilleurs) et pour le LETKF avec localisation et inflation réglé manuellement (les résultats sont meilleurs), voir tableau 4.1.

Avant de conclure, nous avons essayé de tester l'ETKF-N avec localisation. Cependant, la plupart des expériences ont menées à des discontinuités numériques artificielles sur la reconstitution du socle rocheux et du paramètre de glissement conduisant potentiellement à l'arrêt du programme d'assimilation. Les explosions numériques apparaissent sur les zones proches des bords là où le glissement est prépondérant. Il semble difficile de distinguer quelles sont leur source : remise à zéro des épaisseurs négatives par exemple. Il est aussi possible que le filtre ne soit pas adapté à notre cas en cas d'utilisation de la localisation. En effet, la localisation a tendance à produire des discontinuités artificielles dans les états reconstruits (puisqu'elle tronque les covariances à longue distance). La question suivante peut alors se poser. L'ETKF-N exarcerbe t-il cette tendance dans notre cas ?

4.4 Conclusion

Dans ce chapitre, nous avons mis en place un filtre de Kalman d'ensemble ETKF dans le but de résoudre le problème du *spin up* soit la reconstruction conjointe du paramètre de glissement β, de la topographie du socle rocheux et de l'épaisseur actuelle de la calotte. Les tests réalisés ont montré la validité de l'approche sur un cas synthétiques. Notons que, grâce aux procédures d'inflation et de localisation, l'ensemble utilisé peut être pris de petite taille. Ceci est de bonne augure afin d'envisager l'utilisation de l'ETKF dans des cadres plus opérationnels.

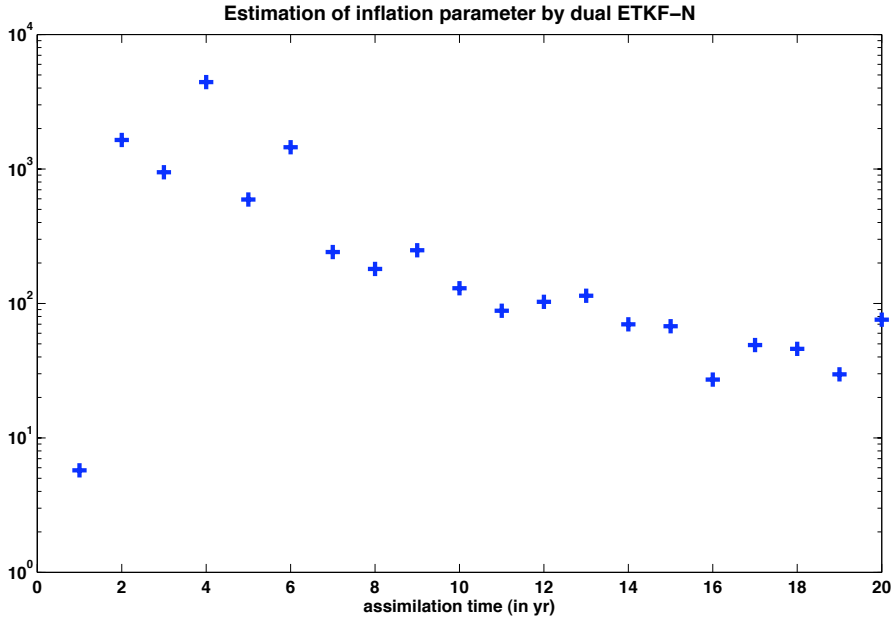

FIGURE 4.20 – Évolution de l'inflation estimée à chaque itération par l'ETKF-N version duale pour un ensemble de taille 100. Nous parlons du paramètre $\rho^2 = \frac{N_e - 1}{\zeta^a}$ avec ζ^a paramètre estimé par ETKF-N version duale. Remarquons les valeurs conséquentes prescrites en inflation.

Cependant, quelques éléments sont à améliorer. Tout d'abord, dans ce chapitre, nous observons la surface de la calotte à chaque instant d'observation à tous les points de grille. Or ce n'est pas le cas des observations réelles à notre disposition. Il faudrait sans doute tester le cas où nous avons à disposition des observations partielles de la calotte. D'autre part, nous avons noté une mauvaise reconstruction du paramètre de frottement basal lorsque celui-ci n'était pas prépondérant. Un critère thermique sur le glissement permettrait d'éviter de rechercher ce paramètre aux endroits où il n'intervient que très marginalement. Il nous faut aussi faire attention aux zones de transition où la vitesse de la glace dépend et du glissement et de la déformation de la glace. D'autres perspectives en terme d'amélioration du systeme d'assimilation peuvent être envisagées, toujours en gardant l'idée d'une méthode d'ensemble comme base.

Conclusions et perspectives

Conclusions ...

Durant ces dernières années, l'utilisation de l'assimilation de données pour les calottes polaires a connu un développement croissant. Les travaux menés durant cette thèse s'inscrivent dans cette tendance et se focalisent plus particulièrement sur les questions méthodologiques.

Au cours de cette thèse, nous avons travaillé sur deux problèmes inverses chacun représentatif d'une échelle de temps caractéristique de l'étude des calottes polaires. Ainsi, la première partie de ce travail (chapitre 3) est consacrée à la reconstruction de l'évolution à long terme (20 000 ans) d'un paramètre climatique, grâce à l'observation de l'évolution du volume d'une calotte polaire. Ce problème inverse de paléoclimatologie est résolu par une méthode variationnelle (4D-Var). Il a nécessité le développement du code adjoint du modèle d'évolution simplifié utilisé. Les différentes expériences d'assimilation conduites ont souligné l'importance de prendre une bonne ébauche pour ce paramètre climatique et les difficultés rencontrées si ce n'était pas le cas.

La seconde partie de cette thèse (chapitre 4) a été consacrée au problème du *spin up*. La reconstruction conjointe de la topographie du socle rocheux, de l'épaisseur de la calotte et des paramètres de glissement à la base a été effectuée à l'aide d'un filtre de Kalman d'ensemble ETKF. L'initialisation de l'ETKF nous a conduit notamment à développer un algorithme de simulation de topographies de socle rocheux aléatoires réalistes. Nous avons aussi mis en place un filtre de Kalman d'ensemble de taille finie ETKF-N permettant d'estimer automatiquement le paramètre d'inflation. Les expériences d'assimilation menées ont montré que l'approche ETKF utilisée donnait de bons résultats, même pour des ensembles de faible taille, grâce à l'utilisation de la localisation. Ceci est très encourageant pour l'utilisation de cette technique avec des données réelles, même si certaines améliorations sont nécessaires comme nous le verrons plus bas.

... et perspectives

Tournons nous maintenant du côté des perspectives suite aux travaux menés durant cette thèse. Celles-ci sont nombreuses. En voici quelques unes.

Vers un problème de paléoclimatologie appliquée aux calottes plus global

Dans le cas du problème inverse en paléoclimatologie, l'extension directe à un cas concret semble hors de portée. En effet, dans un tel cadre, les paramètres climatiques s'appliqueraient à plusieurs calottes polaires en même temps (typiquement les calottes du Groenland, de la Fennoscandie, des Laurentides et l'Antarctique pour un problème inverse s'étalant du dernier maximum glaciaire à aujourd'hui). D'autre part, pour être réaliste à cette échelle, les modèles d'évolution doivent prendre en compte l'isostasie et le glissement basal. Outre les paramètres climatiques, il faudrait aussi contrôler l'épaisseur de la calotte, la topographie du socle rocheux (ou l'enfoncement isostatique pour le cas des calottes disparues) et le glissement basal pour chaque calotte. Observer uniquement l'évolution du volume de glace n'est clairement pas suffisant dans ce cadre. Par contre, il nous semble envisageable de contrôler toutes ces variables si nous ajoutons les observations suivantes : observations de l'état actuel des calottes, isochrones, données issues de l'étude du relèvement isostatique (voir [Tarasov et Peltier, 2004]) et données géomorphologiques (moraines, eskers, drumlins, ... voir [Peyaud, 2006] pour un descriptif des différents objets de la géomorphologie glaciaire). Ces dernières données permettent notamment de reconstituer l'évolution de l'étendue d'une calotte au cours du temps. A un instant donné, nous savons alors où il y avait de la glace et où il n'y en avait pas. Cette observation n'est pas continue mais binaire. Nous pourrions contourner le problème en choisissant d'assimiler la distance au bord de la calotte pour chacun de ses points ou dire que nous avons des observation d'épaisseur nulle aux endroits où nous n'observons pas de glace. Mais ce serait perdre de l'information et quid de l'incertitude sur les observations prescrites. La question qui se pose ici est comment intégrer complètement ces observations apportant de l'information dans nos systèmes d'assimilation. Enfin, l'ébauche sur les paramètres climatiques pourrait être fournie par les résultats obtenus par assimilation de données sur les carottes de glace et les carottes de sédiments marins.

Vers une utilisation opérationnelle du filtre de Kalman d'ensemble pour le problème du spin up

Dans le cas du problème du *spin up*, de nombreuses perspectives nous semble intéressantes. Nous listons ici les principales :

— Améliorer le modèle utilisé : ajout d'un couplage thermomécanique dans Winnie ainsi qu'un critère de température pour le glissement à la base. Cette contrainte physique ajoutée devrait rendre le problème plus contrôlable. A terme, il nous semble tout à fait envisageable d'employer un filtre de Kalman d'ensemble avec un modèle plus complet tel que le modèle GRISLI (3D thermomécanique) sur des applications réalistes.

— Comparer pour une même expérience les résultats obtenus par filtrage avec ceux obtenus par méthode adjointe.

— Mieux prendre en compte les non-linéarités du modèle et des opérateurs d'observation : les filtres de Kalman d'ensemble itératifs récemment développés (cf. [Sakov *et al.*, 2012] et [Bocquet et Sakov, 2012]) semblent donner d'excellents ré-

sultats pour des modèles particulièrement non-linéaires, leur utilisation dans le contexte du *spin up* serait sans doute bénéfique.

— Prendre en compte l'erreur modèle : une première idée serait de générer des observations synthétiques avec un modèle Full Stokes, puis de comparer les résultats de l'assimilation obtenus par le modèle Full Stokes en question ou par un modèle simplifié. Une première étape pourrait être de se placer dans le cadre d'une ligne d'écoulement et d'utiliser Winnie et une version 2D d'Elmer/Ice.

Vers une meilleure prise en compte des « non-gaussianités » en assimilation opérationnelle

L'assimilation de données est souvent confrontée aux très fortes non-linéarités du modèle et/ou des opérateurs d'observation. Le cadre de l'évolution des calottes polaires ne fait pas exception à cette règle. Nous avons vu dans le chapitre 2 que le BLUE, base du filtre de Kalman d'ensemble notamment, était optimal dans le cas linéaire si les variables recherchées et observées étaient gaussiennes. En pratique, celles ci ne le sont que très rarement (parce que certaines variables sont contraintes par exemple). Il faudrait donc utiliser des méthodes prenant mieux en compte les non-linéarités et le caractère non-gaussien des variables employées. De telles méthodes existent, rappelons notamment le filtre particulaire ou les méthodes hybrides. Une autre idée prometteuse est l'anamorphose gaussienne (consulter notamment [Simon et Bertino, 2009], [Simon et Bertino, 2012] et [Béal *et al.*, 2010]). Elle consiste à transformer l'ensemble des vecteurs d'états et d'observation en vecteurs gaussiens avant la phase d'analyse d'un filtre de Kalman d'ensemble. Classiquement dans ce cas, la transformation est construite en reliant les quantiles calculés à partir des ensembles aux quantiles de la loi normale. Cependant, d'autres méthodes pour transformer des variables non-gaussiennes en variables gaussiennes existent. Toute une littérature a été développée à ce sujet en statistiques (citons par exemple la méthode de Box-Cox [Box et Cox, 1964] qui est sans doute la plus connue dans ce domaine). Elle mériterait d'être exploitée en assimilation et comparée avec ce qui est fait à l'heure actuelle. Il faudrait aussi vérifier que la transformation donne entière satisfaction c'est à dire qu'après transformation l'ensemble obtenu suit bien une loi normale. Pour cela, l'utilisation de tests d'adéquation à une loi statistique tel le test de Shapiro-Wilk pour le cas gaussien serait d'une grande aide.

Comme nous pouvons le constater, il reste beaucoup de travail à faire en terme d'assimilation de données entre autre appliquée aux les calottes polaires. Pour terminer, l'auteur de ces lignes voudrait souligner que l'apport de l'assimilation de données à la glaciologie est de son point de vue évident mais que l'apport de la glaciologie à l'assimilation de données est tout aussi important !

Annexe A

Quelques rappels sur les probabilités

Le but de cette annexe est de rapidement redonner les notions de bases en probabilités nécessaires à la compréhension du chapitre sur l'assimilation de données. Nous nous restreignons au cas des variables aléatoires continues.

A.1 Variable aléatoire et densité de probabilité

Soit $\mathbf{X}$ une variable aléatoire de $\mathbb{R}^{n_\mathbf{x}}$, on note $p_\mathbf{X}(\mathbf{x})$ sa densité de probabilité. $p_\mathbf{X}$ est une fonction de $\mathbb{R}^{n_\mathbf{x}}$ ayant les propriétés suivantes

$$p_\mathbf{X}(\mathbf{x}) \geq 0 \qquad \text{pour tout } \mathbf{x} \in \mathbb{R}^{n_\mathbf{x}} \tag{A.1}$$

$$\int_{\mathbb{R}^{N_\mathbf{x}}} p_\mathbf{X}(\mathbf{x})\, d\mathbf{x} = 1 \tag{A.2}$$

Par abus de langage, nous avons confondu la variable aléatoire $\mathbf{X}$ et la variable muette de la densité de probabilité $\mathbf{x}$ dans tout le chapitre sur l'assimilation de données. Dans ce cas, l'indice sur la densité p sert uniquement à distinguer dans quel espace vit la variable aléatoire traitée. S'il s'agit d'une variable aléatoire de taille $\mathbb{R}^{N_\mathbf{x}}$ dans l'espace d'état, nous notons la densité $p_\mathbf{X}$. S'il s'agit d'une variable aléatoire de taille $\mathbb{R}^{N_\mathbf{y}}$ dans l'espace des observations, nous notons la densité $p_\mathbf{Y}$.

A.2 Espérance mathématique

L'opérateur $\mathbb{E}$ désigne l'espérance mathématique. Si $\mathbf{X}$ est une variable aléatoire de densité $p_\mathbf{X}(\mathbf{x})$, alors son espérance mathématique (appelée aussi moyenne) notée $\mathbb{E}[\mathbf{X}]$ vaut

$$\mathbb{E}[\mathbf{X}] = \int_{\mathbb{R}^{n_\mathbf{x}}} \mathbf{x}\, p_\mathbf{X}(\mathbf{x})\, d\mathbf{x} \tag{A.3}$$

L'espérance d'une fonction arbitraire de $\mathbf{X}$, $g(\mathbf{X})$ vaut de son côté

$$\mathbb{E}[g(\mathbf{X})] = \int_{\mathbb{R}^{n_\mathbf{x}}} g(\mathbf{x})\, p_\mathbf{X}(\mathbf{x})\, d\mathbf{x} \tag{A.4}$$

Enfin, notons que l'espérance est un opérateur linéaire. Ainsi si $\mathbf{X}$ et $\mathbf{Y}$ sont deux variables aléatoires alors

$$\mathbb{E}[\mathbf{X}+\mathbf{Y}] = \mathbb{E}[\mathbf{X}] + \mathbb{E}[\mathbf{Y}] \tag{A.5}$$

A.3 Matrice de covariance

Par définition, la matrice de covariance d'une variable aléatoire $\mathbf{X}$ notée $\mathbf{Cov}(\mathbf{X})$ vaut

$$\mathbf{Cov}(\mathbf{X}) = \mathbb{E}\left[(\mathbf{X}-\mathbb{E}[\mathbf{X}])(\mathbf{X}-\mathbb{E}[\mathbf{X}])^T\right] \tag{A.6}$$

En utilisant la définition de l'espérance, nous avons aussi

$$\mathbf{Cov}(\mathbf{X}) = \int_{\mathbb{R}^{n_\mathbf{x}}} (\mathbf{x}-\mathbb{E}[\mathbf{X}])(\mathbf{x}-\mathbb{E}[\mathbf{X}])^T \, p_\mathbf{X}(\mathbf{x}) \, d\mathbf{x} \tag{A.7}$$

A.4 Probabilités conditionnelles

A.4.1 Variables à densité jointe

Considérons le couple $(\mathbf{X}, \mathbf{Y})$ de variables aléatoires de densité jointe $p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y})$. Les variables aléatoires $\mathbf{X}$ et $\mathbf{Y}$ possèdent toutes les deux des densités notées $p_\mathbf{X}$ et $p_\mathbf{Y}$ données par les relations suivantes

$$p_\mathbf{X}(\mathbf{x}) = \int_{\mathbb{R}^{n_\mathbf{y}}} p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} \tag{A.8}$$

$$p_\mathbf{Y}(\mathbf{y}) = \int_{\mathbb{R}^{n_\mathbf{x}}} p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} \tag{A.9}$$

A.4.2 Probabilité conditionnelle

Revenons un bref instant aux probabilités discrètes et notons A et B deux événements. La loi de probabilité de A sachant B notée $P(A|B)$ est classiquement définie par la relation

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{A.10}$$

avec $P(A \cap B)$ la loi de probabilité des deux événements joints.

Par analogie au cas discret, on peut aussi définir la loi de probabilité de la variable $\mathbf{X}$ connaissant la variable $\mathbf{Y}$ notée $p_\mathbf{X}(\mathbf{x}|\mathbf{y})$ par la relation suivante

$$p_\mathbf{X}(\mathbf{x}|\mathbf{y}) = \frac{p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y})}{p_\mathbf{Y}(\mathbf{y})} = \frac{p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y})}{\displaystyle\int_{\mathbb{R}^{n_\mathbf{x}}} p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y}) \, d\mathbf{x}} \tag{A.11}$$

La densité de probabilité jointe peut alors se réécrire

$$p_{(\mathbf{X},\mathbf{Y})}(\mathbf{x}, \mathbf{y}) = p_\mathbf{X}(\mathbf{x}|\mathbf{y}) \, p_\mathbf{Y}(\mathbf{y}) = p_\mathbf{Y}(\mathbf{y}|\mathbf{x}) \, p_\mathbf{X}(\mathbf{x}) \tag{A.12}$$

En combinant les relations (A.8) et (A.12) ou (A.9) et (A.12), nous obtenons les égalités suivantes

$$p_{\mathbf{X}}(\mathbf{x}) = \int_{\mathbb{R}^{n_{\mathbf{y}}}} p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})\, p_{\mathbf{Y}}(\mathbf{y})\, d\mathbf{y} \tag{A.13}$$

$$p_{\mathbf{Y}}(\mathbf{y}) = \int_{\mathbb{R}^{n_{\mathbf{x}}}} p_{\mathbf{Y}}(\mathbf{y}|\mathbf{x})\, p_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \tag{A.14}$$

Les égalités précédentes sont généralisables. Notons maintenant la variable aléatoire $\mathbf{Z}$ de densité $p_{\mathbf{Z}}(\mathbf{z})$. Nous avons

$$p_{\mathbf{X}}(\mathbf{x}|\mathbf{z}) = \int_{\mathbb{R}^{n_{\mathbf{y}}}} p_{\mathbf{X}}(\mathbf{x}|\, \mathbf{y}, \mathbf{z})\, p_{\mathbf{Y}}(\mathbf{y}|\mathbf{z})\, d\mathbf{y} \tag{A.15}$$

A.4.3 Espérance conditionnelle

L'espérance conditionnelle de $\mathbf{X}$ sachant $\mathbf{Y}$ est définie par la relation suivante

$$\mathbb{E}[\mathbf{X}|\mathbf{Y} = \mathbf{y}] = \int_{\mathbb{R}^{n_{\mathbf{x}}}} \mathbf{x}\, p_{\mathbf{X}}(\mathbf{x}|\mathbf{y})\, d\mathbf{x} \tag{A.16}$$

Annexe B

Bibliographie

[Amodei, 1995] L. Amodei (1995). Solution approchée pour un problème d'assimilation de données météorologiques avec prise en compte de l'erreur de modèle. *Comptes rendus de l'Académie des sciences. Série 2. Sciences de la terre et des planètes*, 321(12):1087–1094.

[Anderson, 2001] J. L. Anderson (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129(12):2884–2903.

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

[Anderson et Anderson, 1999] J. L. Anderson et S. L. Anderson (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilation and forcasts. *Monthly Weather Review*, 127:2741–2758.

[Andersson *et al.*, 1998] E. Andersson, J. Haseler, P. Undén, P. Courtier, G. Kelly, D. Vasiljevic, C. Brankovic, C. Gaffard, A. Hollingsworth, C. Jakob, P. Janssen, E. Klinker, A. Lanzinger, M. Miller, F. Rabier, A. Simmons, B. Strauss, P. Viterbo, C. Cardinali et J.-N. Thépaut (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). III : Experimental results. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1831–1860.

[Andrews, 1968] A. Andrews (1968). A square root formulation of the Kalman covariance equations. *AIAA Journal*, 6(6):1165–1166.

[Arnaud, 2004] E. Arnaud (2004). *Méthodes de filtrage pour du suivi dans des séquences d'images – Application au suivi de points caractéristiques.* Thèse de doctorat, Université de Rennes I.

[Arthern, 2003] R. J. Arthern (2003). Optimal estimation of changes in the mass of ice sheets. *Journal of Geophysical Research*, 108(F1).

[Arthern et Gudmundsson, 2010] R. J. Arthern et G. H. Gudmundsson (2010). Initialization of ice-sheet forecasts viewed as an inverse Robin problem. *Journal of Glaciology*, 56(197):527–533.

[Arthern et Hindmarsh, 2006] R. J. Arthern et R. C. A. Hindmarsh (2006). Determining the contribution of Antarctica to sea-level rise using data assimilation methods. *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences*, 364(1844):1841–1865.

[Avenel *et al.*, 2013] C. Avenel, E. Mémin et P. Pérez (2013). Stochastic level set dynamics to track closed curves through image data. *Journal of Mathematical Imaging and Vision*, accepté pour publication, http ://hal.inria.fr/hal-00854420.

[Bamber *et al.*, 2009] J. L. Bamber, J. L. Gomez-Dans et J. A. Griggs (2009). A new 1 km digital elevation model of the Antarctic derived from combined satellite radar and laser data–Part 1 : Data and methods. *The Cryosphere*, 3(11):101–111.

[Bamber *et al.*, 2013] J. L. Bamber, J. A. Griggs, R. T. W. L. Hurkmans, J. A. Dowdeswell, S. P. Gogineni, I. Howat, J. Mouginot, P. J., S. Palmer, E. Rignot et D. Steinhage (2013). A new bed elevation dataset for Greenland. *The Cryosphere*, 7(2):499–510.

[Béal *et al.*, 2010] D. Béal, P. Brasseur, J.-M. Brankart, Y. Ourmières et J. Verron (2010). Characterization of mixing errors in a coupled physical biogeochemical model of the North Atlantic : implications for nonlinear estimation using Gaussian anamorphosis. *Ocean Science*, 6(1):247–262.

[Bengtsson *et al.*, 2008] T. Bengtsson, P. Bickel et B. Li (2008). Curse-of-dimensionality revisited : Collapse of the particle filter in very large sclae systems. *Probability and statistics : Essays in honor of David A. Freedman*, 2:316–334.

[Bennett et Thorburn, 1992] A. F. Bennett et M. A. Thorburn (1992). The generalized inverse of a nonlinear quasigeostrophic ocean circulation model. *Journal of Physical Oceanography*, 3:213–230.

[Berliner *et al.*, 2008] L. Berliner, K. Jezek, N. Cressie, Y. Kim, C. Q. Lam et C. J. van der Veen (2008). Modeling dynamic controls on ice streams : a Bayesian statistical approach. *Journal of Glaciology*, 54(1):705–714.

[Beyou *et al.*, 2013] S. Beyou, A. Cuzol, S. S. Gorthi et E. Mémin (2013). Weighted ensemble transform Kalman filter for image assimilation. *Tellus A*, 65(18803).

[Bickel *et al.*, 2008] P. Bickel, B. Li et T. Bengtsson (2008). Sharp failure rates for the bootstrap particle filter in high dimensions. *IMS Collections : Pushing the Limits of Contemporary Statistics : Contributions in Honor of Jayanta K. Ghosh*, 3:318–329.

[Bintanja *et al.*, 2005a] R. Bintanja, R. S. van de Wal et J. Oerlemans (2005a). Modelled atmospheric temperatures and global sea levels over the past million years. *Nature*, 437(7055):125–128.

[Bintanja *et al.*, 2005b] R. Bintanja, R. S. van de Wal et J. Oerlemans (2005b). A new method to estimate ice age temperatures. *Climate Dynamics*, 24(2-3):197–211.

[Bishop *et al.*, 2001] C. H. Bishop, B. J. Etherton et S. J. Majumdar (2001). Adaptive sampling with the Ensemble Transform Kalman Filter. Part I : Theoretical aspects. *Monthly Weather Review*, 129:420–436.

[Bishop et Hodyss, 2009] C. H. Bishop et D. Hodyss (2009). Ensemble covariances adaptatively localized with ECO-RAP. Part 1 : Tests on simple error models. *Tellus A*, 61(1):84–96.

[Blatter, 1995] H. Blatter (1995). Velocity and stress fields in grounded glaciers : a simple algorithm for including deviatoric stress gradients. *Journal of Glaciology*, 41(138):333–344.

[Blayo *et al.*, 1998] E. Blayo, J. Blum et J. Verron (1998). Assimilation variationnelle de données en océanographie et réduction de la dimension de l'espace de contrôle. Équations aux dérivées partielles et applications. Articles dédiés à Jacques-Louis Lions. Gauthier-Villars : Paris. 199-219.

[Blayo et Nodet, 2012] E. Blayo et M. Nodet (2012). Introduction à l'assimilation de données variationnelle. Notes de cours, M2 Mathématiques, Informatique, Université Joseph Fourier – Grenoble 1 et 3eme année Ensimag.

[Bocquet, 2011] M. Bocquet (2011). Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlinear Processes in Geophysics*, 18(5):735–750.

[Bocquet, 2013] M. Bocquet (2013). Introduction aux principes et méthodes de l'assimilation de données en géophysique. Notes de cours du Master M2 OACOS et de l'École Nationale Supérieure des Techniques Avancées ParisTech et de l'École des Ponts ParisTech, Révision 1.29.

[Bocquet et Sakov, 2012] M. Bocquet et P. Sakov (2012). Combining inflation-free and iterative Ensemble Kalman Filters for strongly nonlinear systems. *Nonlinear Processes in Geophysics*, 19(3):383–399.

[Bonan *et al.*, 2012] B. Bonan, M. Nodet, O. Ozenda et C. Ritz (2012). Data assimilation in glaciology. *In Advanced Data Assimilation for Geosciences*, Les Houches, France.

[Bonan *et al.*, 2013] B. Bonan, M. Nodet, C. Ritz et V. Peyaud (2013). An ETKF approach for initial state and parameter estimation in ice sheet modelling. *Nonlinear Processes in Geophysics (article soumis).*

[Bouttier et Courtier, 1999] F. Bouttier et P. Courtier (1999). Data assimilation, concepts and methods. Meteorological training course lecture series, ECMWF.

[Böðvarsson, 1955] G. Böðvarsson (1955). On the flow of ice-sheets and glaciers. *Jökull*, 5:1–8.

[Box et Cox, 1964] G. E. P. Box et D. R. Cox (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 211–252.

[Brankart *et al.*, 2010] J.-M. Brankart, E. Cosme, C.-E. Testut, P. Brasseur et J. Verron (2010). Efficient adaptive error parameterizations for square root or ensemble Kalman filters : application to the control of ocean mesoscale signals. *Monthly Weather Review*, 138(3):932–950.

[Broquet, 2007] G. Broquet (2007). *Caractérisation des erreurs de modélisation pour l'assimilation de données dans un modèle océanique régional du Golfe de Gascogne.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Bueler et Brown, 2009] E. Bueler et J. Brown (2009). Shallow shelf approximation as a "sliding law" in a thermomechanically coupled ice sheet model. *Journal of Geophysical Research : Earth Surface*, 114(F3).

[Burgers *et al.*, 1998] G. Burgers, P. J. van Leeuwen et G. Evensen (1998). Analysis scheme in the ensemble Kalman filter. *Monthly Weather Review*, 126:1719–1724.

[Chaabane et Jaoua, 1999] S. Chaabane et M. Jaoua (1999). Identification of Robin coefficients by the means of boundary measurements. *Inverse Problems*, 15:1425.

[Chorin *et al.*, 2010] A. J. Chorin, M. Morzfeld et X. Tu (2010). Implicit particle filters for data assimilation. *arXiv preprint arXiv :1005.4002*.

[Cohn et Todling, 1996] S. E. Cohn et R. Todling (1996). Approximate data assimilation schemes for stable and unstable dynamics. *Journal of the Meteorological Society of Japan*, 74:63–75.

[Corazza *et al.*, 2003] M. Corazza, E. Kalnay, D. J. Patil, S.-C. Yang, R. Morss, M. Cai, I. Szunyogh, B. R. Hunt et J. A. Yorke (2003). Use of the breeding technique to estimate the structure of the analysis "errors of the day". *Nonlinear Processes in Geophysics*, 10(3):233–243.

[Cornford *et al.*, 2013] S. L. Cornford, D. F. Martin, D. T. Graves, D. F. Ranken, A. M. Le Brocq, R. M. Gladstone, A. J. Payne, E. G. Ng et W. H. Lipscomb (2013). Adaptive mesh, finite volume modeling of marine ice sheets. *Journal of Computational Physics*, 232:529–549.

[Courtier, 1997] P. Courtier (1997). Dual formulation of four-dimensional variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 123(544):2449–2461.

[Courtier *et al.*, 1998] P. Courtier, E. Andersson, W. Heckley, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, M. Fisher et J. Pailleux (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I : Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550):1783–1807.

[Courtier *et al.*, 1994] P. Courtier, J.-N. Thépaut et A. Hollingsworth (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519):1367–1387.

[Cuffey et Paterson, 2010] K. M. Cuffey et W. S. B. Paterson (2010). *The physics of glaciers*. Butterworth-Heinemann.

[Da Silva *et al.*, 1995] A. Da Silva, J. Pfaendtner, J. Guo, M. Sienkiewicz et S. Cohn (1995). Assessing the effects of data selection in DAO's Physical-space Statistical Analysis System. *In Proceedings of the second international WMO symposium on assimilation of observations in meteorology and oceanography, Tokyo, 13 – 17 March 1995*, volume WMO.TD 651, pages 273–278.

[Daget, 2008] N. Daget (2008). *Estimation d'ensemble des paramètres des covariances d'erreur d'ébauche dans un système d'assimilation variationnelle de données océanographique*. Thèse de doctorat, Université Toulouse III – Paul Sabatier.

[de Fleurian, 2010] B. de Fleurian (2010). *Développement d'un modèle d'hydrologie sous-glaciaire dédié à la simulation du glissement basal des glaciers*. Thèse de doctorat, Université de Grenoble.

[Dee, 1995] D. P. Dee (1995). On-line estimation of error covariance parameters for atmospheric data assimilation. *Monthly Weather Review*, 123(4):1128–1145.

[Desroziers, 2007] G. Desroziers (2007). *Mise en œuvre, diagnostic et optimisation des schémas d'assimilation de données*. Habilitation à diriger des recherches, Université Toulouse III – Paul Sabatier.

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

[Di Lorenzo *et al.*, 2007] E. Di Lorenzo, A. M. Moore, H. G. Arango, B. D. Cornuelle, A. J. Miller, B. Powell, B. S. Chua et A. F. Bennett (2007). Weak and strong constraint data assimilation in the inverse Regional Ocean Modeling System (ROMS) : Development and application for a baroclinic coastal upwelling system. *Ocean Modelling*, 16(3):160–187.

[Doucet *et al.*, 2000] A. Doucet, S. Godsill et C. Andrieu (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10(3): 197–208.

[Drouet, 2012] A.-S. Drouet (2012). *Dynamique du glacier émissaire : des processus à l'application sur un glacier école, l'Astrolabe, Antarctique de l'Est.* Thèse de doctorat, Université de Grenoble.

[Dubinkina et Goosse, 2013] S. Dubinkina et H. Goosse (2013). An assessment of particle filtering methods and nudging for climate state reconstructions. *Climate of the Past*, 9:1141–1152.

[Dubinkina *et al.*, 2011] S. Dubinkina, H. Goosse, Y. Sallaz-Damaz, E. Crespin et M. Crucifix (2011). Testing a particle filter to reconstruct climate changes over the past centuries. *International Journal of Bifurcation and Chaos*, 21(12):3611–3618.

[Dumas, 2002] C. Dumas (2002). *Modélisation de l'évolution de l'Antarctique depuis le dernier cycle glaciaire-interglaciaire jusqu'au futur : importance relative des différents processus physiques et rôle des données d'entrée.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Durbiano, 2001] S. Durbiano (2001). *Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Duval, 1979] P. Duval (1979). Creep and recrystallization of polycrystalline ice. *Bulletin de Minéralogie*, 102(2-3):80–85.

[EPICA community members, 2004] EPICA community members (2004). Eight glacial cycles from an Antarctic ice core. *Nature*, 429:623–628.

[Ettema *et al.*, 2009] J. Ettema, M. R. van den Broeke, E. van Meijgaard, W. J. van de Berg, J. L. Bamber, J. E. Box et R. C. Bales (2009). Higher surface mass balance of the Greenland ice sheet revealed by high-resolution climate modeling. *Geophysical Research Letters*, 36(12).

[Evensen, 1992] G. Evensen (1992). Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research : Oceans*, 97(C11): 17905–17924.

[Evensen, 1994] G. Evensen (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5):10143 – 10162.

[Evensen, 2003] G. Evensen (2003). The Ensemble Kalman filter : theoretical formulation and practical implementation. *Ocean Dynamics*, 53:343–367.

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

[Evensen, 2007] G. Evensen (2007). *Data assimilation : the Ensemble Kalman Filter*. Springer.

[Fettweis, 2007] X. Fettweis (2007). Reconstruction of the 1979–2006 Greenland ice sheet surface mass balance using the regional climate model mar. *The Cryosphere*, 1:21–40.

[Fisher, 2010] H. Fisher (2010). A short primer on ice cores. Lecture notes, Karthaus summer school 2010.

[Flament *et al.*, 2013] T. Flament, E. Berthier et F. Rémy (2013). Cascading water underneath Wilkes Land, East Antarctic Ice Sheet, observed using altimetry and digital elevation models. *The Cryosphere Discussions*, 7(2):841–871.

[Flament et Rémy, 2012] T. Flament et F. Rémy (2012). Dynamic thinning of Antarctic glaciers from along-track repeat radar altimetry. *Journal of Glaciology*, 58(211):830–840.

[Fretwell *et al.*, 2013] P. Fretwell, H. D. Pritchard, D. J. Vaughan, J. L. Bamber, N. E. Barrand, R. Bell, C. Bianchi, R. G. Bingham, D. D. Blankenship, G. Casassa, G. Catania, D. Callens, H. Conway, A. Cook, H. F. J. Corr, D. Damaske, V. Damm, F. Ferraccioli, R. Forsberg, S. Fujita, Y. Gim, P. Gogineni, J. Griggs, R. C. A. Hindmarsh, P. Holmlund, J. W. Holt, R. W. Jacobel, A. Jenkins, W. Jokat, T. Jordan, E. C. King, J. Kohler, W. Krabill, M. Riger-Kust, K. A. Langley, G. Leitchenkov, C. Leuschen, B. P. Luyendyk, K. Matsuoka, J. Mouginot, N. O. Nitsche, Y. Nogi, O. A. Nost, S. V. Popov, E. Rignot, D. M. Rippin, A. Rivera, J. Roberts, N. Ross, M. J. Siegert, A. M. Smith, D. Steinhage, M. Studinger, B. Sun, B. K. Tinto, B. C. Welch, D. Wilson, D. A. Young, C. Xiangbin et A. Zirizzotti (2013). Bedmap2 : improved ice bed, surface and thickness datasets for Antarctica. *The Cryosphere*, 7:375–393.

[Fujita *et al.*, 1999] S. Fujita, H. Maeno, S. Uratsuka, T. Furukawa, S. Mae, Y. Fujii et O. Watanabe (1999). Nature of radio echo layering in the Antarctic Ice Sheet detected by a two-frequency experiment. *Journal of Geophysical Research : Solid Earth*, 104(B6):13013–13024.

[Furrer et Bengtsson, 2007] R. Furrer et T. Bengtsson (2007). Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. *Journal of Multivariate Analysis*, 98:227–255.

[Gagliardini *et al.*, 2013] O. Gagliardini, T. Zwinger, F. Gillet-Chaulet, G. Durand, L. Favier, B. de Fleurian, R. Greve, M. Malinen, C. Martín, P. Råback, J. Ruokolainen, M. Sacchettini, M. Schäfer, H. Seddik et J. Thies (2013). Capabilities and performance of elmer/ice, a new generation ice-sheet model. *Geoscientific Model Development*, 6:1299–1318.

[Gaspari et Cohn, 1999] G. Gaspari et S. E. Cohn (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125:723–757.

[Gauthier *et al.*, 1993] P. Gauthier, P. Courtier et P. Moll (1993). Assimilation of simulated wind lidar data with a Kalman filter. *Monthly Weather Review*, 121(6):1803–1820.

[Giering et Kaminski, 1998] R. Giering et T. Kaminski (1998). Recipes for adjoint code construction. *ACM Transactions on Mathematical Software (TOMS)*, 24(4):437–474.

[Gillet-Chaulet, 2006] F. Gillet-Chaulet (2006). *Modélisation de l'écoulement de la glace polaire anisotrope et premières applications au forage de Dôme C.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Gillet-Chaulet *et al.*, 2012] F. Gillet-Chaulet, O. Gagliardini, H. Seddik, M. Nodet, G. Durand, C. Ritz, T. Zwinger, R. Greve et D. G. Vaughan (2012). Greenland Ice Sheet contribution to sea-level rise from a new-generation ice-sheet model. *The Cryosphere*, 6:1561–1576.

[Goldberg, 2011] D. N. Goldberg (2011). A variationally derived, depth-integrated approximation to a higher-order glaciological flow model. *Journal of Glaciology*, 57(201):157–170.

[Goldberg et Heimbach, 2013] D. N. Goldberg et P. Heimbach (2013). Parameter and state estimation with a time-dependant adjoint marine ice sheet model. *The Cryosphere Discussion*, 7:2845–2890.

[Goldberg et Sergienko, 2011] D. N. Goldberg et O. V. Sergienko (2011). Data assimilation using a hybrid ice flow model. *The Cryosphere*, 5(2):315–327.

[Goosse *et al.*, 2010] H. Goosse, E. Crespin, A. de Montety, M. Mann, H. Renssen et A. Timmermann (2010). Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation. *Journal of Geophysical Research*, 115(D9):D09108.

[Gordon *et al.*, 1993] N. J. Gordon, D. J. Salmond et A. F. Smith (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *In IEE Proceedings F (Radar and Signal Processing)*, volume 140(2), pages 107–113. IET.

[Gottwald et Majda, 2013] G. A. Gottwald et A. J. Majda (2013). A mechanism for catastrophic filter divergence in data assimilation for sparse observation networks. *Nonlinear Processes in Geophysics*, 20:705–712.

[Greve, 1997] R. Greve (1997). A continuum–mechanical formulation for shallow polythermal ice sheets. *Philosophical Transactions of the Royal Society of London. Series A : Methematical, Physical and Engineering Sciences*, 355(1726):921–974.

[Greve et Blatter, 2009] R. Greve et H. Blatter (2009). *Dynamics of ice sheets and glaciers.* Springer.

[Griewank et Walther, 2008] A. Griewank et A. Walther (2008). *Evaluating derivatives : principles and techniques of algorithmic differentiation.* Society for Industrial and Applied Mathematics.

[Griggs et Bamber, 2009] J. A. Griggs et J. L. Bamber (2009). A new 1 km digital elevation model of Antarctica derived from combined radar and laser data–Part 2 : Validation and error estimates. *The Cryosphere*, 3:113–123.

[Grinsted et Dahl-Jensen, 2002] A. Grinsted et D. Dahl-Jensen (2002). A Monte Carlo-tuned model of the flow in the NorthGRIP area. *Annals of Glaciology*, 35(1):527–530.

[Hamill *et al.*, 2001] T. M. Hamill, J. S. Whitaker et C. Snyder (2001). Distance-dependant filtering of background error covariance in an ensemble Kalman filter. *Monthly Weather Review*, 129:2776–2790.

[Hanna *et al.*, 2013] E. Hanna, F. J. Navarro, F. Pattyn, C. M. Domingues, X. Fettweis, E. R. Ivins, R. J. Nicholls, C. Ritz, B. Smith, S. Tulaczyk, P. L. Whitehouse et H. J. Zwally (2013). Ice-sheet mass balance and climate change. *Nature*.

[Harlim et Hunt, 2007] J. Harlim et B. R. Hunt (2007). A non-Gaussian Ensemble Filter for assimilating infrequent noisy observations. *Tellus A*, 59(2):225–237.

[Harlim et Majda, 2010] J. Harlim et A. J. Majda (2010). Catastrophic filter divergence in filtering nonlinear dissipative systems. *Communication in Mathematical Sciences*, 8(1):27–43.

[Hascoet et Pascual, 2013] L. Hascoet et V. Pascual (2013). The tapenade automatic differentiation tool : principles, model, and specification. *ACM Transactions on Mathematical Software (TOMS)*, 39(3):20.

[Heimbach et Bugnion, 2009] P. Heimbach et V. Bugnion (2009). Greenland ice-sheet volume sensitivity to basal, surface and initial conditions derived from an adjoint model. *Annals of Glaciology*, 50:67–80.

[Hindmarsh, 2001] R. C. A. Hindmarsh (2001). Notes on basic glaciological computational methods and algorithms. *In Continuum Mechanics and Applications in Geophysics and the Environment*, pages 222–249. Springer.

[Hindmarsh, 2004] R. C. A. Hindmarsh (2004). A numerical comparison of approximations to the Stokes equations used in ice sheet and glacier modeling. *Journal of Geophysical Research*, 109(F1):F01012.

[Hoteit, 2001] I. Hoteit (2001). *Filtres de Kalman réduits et efficaces pour l'assimilation de données en océanographie*. Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Houtekamer et Mitchell, 1998] P. L. Houtekamer et H. L. Mitchell (1998). Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126:796–811.

[Houtekamer et Mitchell, 2001] P. L. Houtekamer et H. L. Mitchell (2001). A sequential ensemble Kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129:123–137.

[Hunt *et al.*, 2007] B. R. Hunt, E. J. Kostelich et I. Szunyogh (2007). Efficient data assimilation for spatiotemporal chaos : A local ensemble transform Kalman filter. *Physica D*, 230:112 – 126.

[Hutter, 1983] K. Hutter (1983). *Theoretical Glaciology : Mathematical Approaches to Geophysics*. D. Reidel.

[Ide *et al.*, 1997] K. Ide, P. Courtier, M. Ghil et A. Lorenc (1997). Unified notation for data assimilation : Operational, Sequential and Variational. *Journal of the Meteorological Society of Japan*, 75(1B):181–189.

[Jay-Allemand *et al.*, 2011] M. Jay-Allemand, F. Gillet-Chaulet, O. Gagliardini et M. Nodet (2011). Investigating changes in basal conditions of Variegated Glacier prior to and during its 1982–1983 surge. *The Cryosphere*, 5:659–672.

[Jeffreys, 1961] H. Jeffreys (1961). *The theory of probability, 3rd edition*. Oxford University Press.

[Joughin *et al.*, 2010] I. Joughin, B. E. Smith, I. M. Howat, T. Scambos et T. Moon (2010). Greenland flow variability from ice-sheet-wide velocity mapping. *Journal of Glaciology*, 56(197):415–430.

[Julier et Uhlmann, 1997] S. J. Julier et J. K. Uhlmann (1997). New extension of the Kalman filter to nonlinear systems. *In AeroSense'97*, pages 182–193. International Society for Optics and Photonics.

[Kalman, 1960] R. E. Kalman (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45.

[Kleman *et al.*, 1997] J. Kleman, C. Hättestrand, I. Borgström et A. Stroeven (1997). Fennoscandian palaeoglaciology reconstructed using a glacial geological inversion rrlodel. *Journal of Glaciology*, 43(144):283–299.

[Kushner, 1967] H. Kushner (1967). Approximations to optimal nonlinear filters. *IEEE Transactions on Automatic Control*, 12(5):546–556.

[Larour *et al.*, 2012] E. Larour, H. Seroussi, M. Morlighem et E. Rignot (2012). Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM). *Journal of Geophysical Research*, 117:F01022.

[Lawson et Hansen, 2004] W. G. Lawson et J. A. Hansen (2004). Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Monthly Weather Review*, 132(8):1966–1981.

[Le Dimet, 1980] F.-X. Le Dimet (1980). Une étude générale d"analyse objective variationnelle des champs météorologiques. Rapport scientifique LAMP 28, Université Clermont II.

[Le Dimet et Blum, 2002] F.-X. Le Dimet et J. Blum (2002). Assimilation de données pour les fluides géophysiques. *Matapli*, 67:37–68.

[Le Dimet *et al.*, 2002] F.-X. Le Dimet, I. M. Navon et D. N. Daescu (2002). Second-order information in data assimilation. *Monthly Weather Review*, 130(3):629–648.

[Le Dimet et Talagrand, 1986] F.-X. Le Dimet et O. Talagrand (1986). Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects. *Tellus A*, 38(2):97–110.

[Le Gland *et al.*, 2011] F. Le Gland, V. Monbet et V.-D. Tran (2011). Large sample asymptotics for the ensemble Kalman filter. *In* D. Crisan et B. Rozovskii, éditeurs : *The Oxford handbook on Nonlinear Filtering*, pages 598–631. Oxford University Press.

[Le Meur, 1996] E. Le Meur (1996). *Spécificité de l'isostasie en contexte glaciaire. Présentation et application d'un modèle de réponse terrestre.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Leeuwenburgh *et al.*, 2005] O. Leeuwenburgh, G. Evensen et L. Bertino (2005). The impact of ensemble filter definition on the assimilation of temperature profiles in the tropical pacific. *Quarterly Journal of the Royal Meteorological Society*, 131(613):3291–3300.

[Lemieux-Dudon *et al.*, 2010] B. Lemieux-Dudon, E. Blayo, J.-R. Petit, C. Waelbroeck, A. Svensson, C. Ritz, J.-M. Barnola, B. M. Narcisi et P. F. (2010). Consistent dating for Antarctic and Greenland ice cores. *Quaternary Science Reviews*, 29.

[Lermusiaux et Robinson, 1999] P. F. J. Lermusiaux et A. R. Robinson (1999). Data assimilation via error subspace statistical estimation. Part I : Theory and schemes. *Monthly Weather Review*, 127(7):1385–1407.

[Li *et al.*, 2009] H. Li, E. Kalnay et T. Miyoshi (2009). Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Quarterly Journal of the Royal Meteorological Society*, 135(639):523–533.

[Li et Navon, 2001] Z. Li et I. M. Navon (2001). Optimality of variational data assimilation and its relationship with the Kalman filter and smoother. *Quarterly Journal of the Royal Meteorological Society*, 127(572):661–683.

[Lions, 1968] J.-L. Lions (1968). *Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles*. Dunod.

[Livings *et al.*, 2008] D. M. Livings, S. L. Dance et N. K. Nichols (2008). Unbiased Ensemble Square Root Filters. *Physica D*, 237(8):1021–1028.

[Lliboutry, 1993] L. Lliboutry (1993). Anisotropic, transversely isotropic nonlinear viscosity of rock ice and rheological parameters inferred from homogenization. *International journal of plasticity*, 9(5):619–632.

[MacAyeal, 1989] D. R. MacAyeal (1989). Large-scale ice flow over a viscous basal sediment : Theory and application to ice stream B, Antarctica. *Journal of Geophysical Research : Solid Earth*, 94(B4):4071–4087.

[MacAyeal, 1992] D. R. MacAyeal (1992). The basal stress distribution of Ice Stream E, Antarctica, inferred by control methods. *Journal of Geophysical Research*, 97:595–603.

[MacAyeal, 1993] D. R. MacAyeal (1993). A tutorial on the use of control methods in ice-sheet modeling. *J. Glaciol*, 39(131):91–98.

[Mallat *et al.*, 1998] S. Mallat, G. Papanicolaou et Z. Zhang (1998). Adaptative covariance estimation of locally stationary processes. *Annals of Statistics*, 26(1):1–47.

[Mandel *et al.*, 2011] J. Mandel, L. Cobb et J. D. Beezley (2011). On the convergence of the ensemble Kalman filter. *Applications of Mathematics*, 56(6):533–541.

[Mitchell et Houtekamer, 2000] H. L. Mitchell et P. L. Houtekamer (2000). An adaptative ensemble Kalman filter. *Monthly Weather Review*, 128:416–433.

[Mitchell et Houtekamer, 2009] H. L. Mitchell et P. L. Houtekamer (2009). Ensemble Kalman filter configurations and their performance with the Logistic Map. *Monthly Weather Review*, 137:4325–4343.

[Miyoshi, 2011] T. Miyoshi (2011). The Gaussian approach to adaptive covariance inflation and its implementation with the local ensemble transform Kalman filter. *Monthly Weather Review*, 139(5):1519–1535.

[Mölg, 2010] T. Mölg (2010). Tropical glaciers and climate dynamics. Lecture notes, Karthaus summer school 2010.

[Morland, 1984] L. W. Morland (1984). Thermomechanical balances of ice sheet flows. *Geophysical & Astrophysical Fluid Dynamics*, 29(1-4):237–266.

[Morland, 1987] L. W. Morland (1987). Unconfined ice-shelf flow. *In Dynamics of the West Antarctic Ice Sheet*, pages 99–116. Springer.

[Morlighem, 2011] M. Morlighem (2011). *Détermination de propriétés des glaciers polaires par modélisation numérique et télédétection.* Thèse de doctorat, École Centrale Paris.

[Morlighem *et al.*, 2010] M. Morlighem, E. Rignot, H. Seroussi, E. Larour, H. Ben Dhia et D. Aubry (2010). Spatial patterns of basal drag inferred using control methods from a full-Stokes and simpler models for Pine Island Glacier, West Antarctica. *Geophysical Research Letters*, 37(14).

[Morzfeld et Chorin, 2012] M. Morzfeld et A. J. Chorin (2012). Implicit particle filtering for models with partial noise, and an application to geomagnetic data assimilation. *Nonlinear Processes in Geophysics*, 19:365–382.

[Nakano et Ueno, 2012] S. Nakano et G. Ueno (2012). Hybrid approach of ensemble transform and importance sampling for non-linear data assimilation. *In International Conference on Ensemble Methods in Geophysical Sciences (Toulouse).*

[NEEM community members, 2013] NEEM community members (2013). Eemian interglacial reconstructed form a Greenland folded ice core. *Nature*, 493:489–494.

[Nerger *et al.*, 2012] L. Nerger, T. Janjić, J. Schröter et W. Hiller (2012). A unification of Ensemble Square Root Kalman Filters. *Monthly Weather Review*, 140(7): 2335–2345.

[Neveu, 2011] E. Neveu (2011). *Application des méthodes multigrilles à l'assimilation variationnelle de données en géophysique.* Thèse de doctorat, Université de Grenoble.

[Nodet, 2005] M. Nodet (2005). *Modélisation mathématique et assimilation de données lagrangienne pour l'océanographie.* Thèse de doctorat, Université de Nice Sophia Antipolis.

[Ott *et al.*, 2004] E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. J. Kostelich, M. Corazza, E. Kalnay, D. Patil et J. A. Yorke (2004). A local Ensemble Kalman filter for atmospheric data assimilation. *Tellus A*, 56:415–428.

[Paillard et Parrenin, 2004] D. Paillard et F. Parrenin (2004). The Antarctic ice sheet and the triggering of deglaciations. *Earth and Planetary Science Letters*, 227(3):263–271.

[Papadakis, 2007] N. Papadakis (2007). *Assimilation de données image : application au suivi de courbes et de champs de vecteurs.* Thèse de doctorat, Université de Rennes I.

[Papadakis *et al.*, 2010] N. Papadakis, E. Mémin, A. Cuzol et N. Gengembre (2010). Data assimilation with the weighted ensemble Kalman filter. *Tellus A*, 62(5):673–697.

[Parrenin et Hindmarsh, 2007] F. Parrenin et R. C. A. Hindmarsh (2007). Influence of a non-uniform velocity field on isochrone geometry along a steady flowline of an ice sheet. *Journal of Glaciology*, 53(183):612–622.

[Parrenin *et al.*, 2006] F. Parrenin, R. C. A. Hindmarsh et F. Rémy (2006). Analytical solutions for the effect of topography, accumulation rate and lateral flow divergence on isochrone layer geometry. *Journal of Glaciology*, 52(177):191–202.

[Parrenin *et al.*, 2001] F. Parrenin, J. Jouzel, C. Waelbroeck, C. Ritz et J.-M. Barnola (2001). Dating the Vostok ice core by an inverse method. *Journal of Geophysical Research. D. Atmospheres*, 106:31.

[Parrenin *et al.*, 2004] F. Parrenin, F. Remy, C. Ritz, M. J. Siegert et J. Jouzel (2004). New modeling of the Vostok ice flow line and implication for the glaciological chronology of the Vostok ice core. *J. Geophys. Res*, 109.

[Parrish et Derber, 1992] D. F. Parrish et J. C. Derber (1992). The National Meteorological Center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8):1747–1763.

[Pattyn, 2003] F. Pattyn (2003). A new three-dimensional higher-order thermomechanical ice-sheet model : Basic sensitivity, ice stream development, and ice flow across subglacial lakes. *Journal of Geophysical Research : Solid Earth*, 108(B8).

[Pattyn *et al.*, 2008] F. Pattyn, L. Perichon, A. Aschwanden, B. Breuer, B. de Smedt, O. Gagliardini, G. H. Gudmundsson, R. C. A. Hindmarsh, A. Hubbard, J. V. Johnson, T. Kleiner, Y. Konovalov, C. Martin, A. J. Payne, D. Pollard, S. Price, M. Rückamp, F. Saito, O. Souček, S. Sugiyama et T. Zwinger (2008). Benchmark experiments for higher-order and full-Stokes ice sheet models (ISMIP–HOM). *The Cryosphere*, 2:95–108.

[Peyaud, 2006] V. Peyaud (2006). *Rôle de la dynamique des calottes glaciaires dans les grands changements climatiques des périodes glaciaires-interglaciaires.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Peyaud *et al.*, 2007] V. Peyaud, C. Ritz et G. Krinner (2007). Modelling the Early Weichselian Eurasian Ice Sheets : role of ice shelves and influence of ice-dammed lakes. *Climate of the Past*, 3:375–386.

[Pham, 1996] D.-T. Pham (1996). A singular evolutive interpolated Kalman filter for data assimilation in oceanography. Rapport technique 163, IMAG-LMC.

[Pham, 2001] D.-T. Pham (2001). Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review*, 129(5):1194–1207.

[Pham *et al.*, 1998] D.-T. Pham, J. Verron et L. Gourdeau (1998). Filtres de Kalman singuliers évolutif pour l'assimilation de données en océnographie. *C. R. Acad. Sci., Paris, Sci. terre planètes*, 326(4):255–260.

[Pham *et al.*, 1996] D.-T. Pham, J. Verron et M.-C. Roubaud (1996). A singular evolutive extended Kalman filter for data assimilation in oceanography. Rapport technique 162, IMAG-LMC.

[Philippon *et al.*, 2006] G. Philippon, G. Ramstein, S. Charbit, M. Kageyama, C. Ritz et C. Dumas (2006). Evolution of the Antarctic ce sheet throughout the last deglaciation : a study with a new coupled climate-north and south hemisphere ice sheet model. *Earth and Planetary Science Letters*, 248(3):750–758.

[Pollard et DeConto, 2012] D. Pollard et R. M. DeConto (2012). A simple inverse method for the distribution of basal sliding coefficients under ice sheets, applied to Antarctica. *The Cryosphere*, 6:953–971.

[Pritchard *et al.*, 2009] H. D. Pritchard, R. J. Arthern, D. G. Vaughan et L. A. Edwards (2009). Extensive dynamic thinning on the margins of the Greenland and Antarctic ice sheets. *Nature*, 461(7266):971–975.

[Quiquet, 2012] A. Quiquet (2012). *Reconstruction de la calotte polaire du Groenland au cours du dernier cycle glaciaire-interglaciaire à partir de l'association de la modéli-*

sation numérique 3D et des enregistrements des carottages glaciaires profonds. Thèse de doctorat, Université de Grenoble.

[Quiquet *et al.*, 2012] A. Quiquet, H. J. Punge, C. Ritz, X. Fettweis, H. Gallée, M. Kageyama, G. Krinner, D. Salas y Mélia et J. Sjolte (2012). Sensitivity of a Greenland ice sheet model to atmospheric forcing fields. *The Cryosphere*, 6(5):999–1018.

[Raymond-Pralong et Gudmundsson, 2011] M. Raymond-Pralong et G. H. Gudmundsson (2011). Bayesian estimation of basal conditions on Rutford Ice Stream, West Antarctica, from surface data. *Journal of Glaciology*, 57(202).

[Raynaud *et al.*, 2009] L. Raynaud, L. Berre et G. Desroziers (2009). Objective filtering of ensemble-based background-error variances. *Quarterly Journal of the Royal Meteorological Society*, 135:1177–1199.

[Ricci, 2004] S. Ricci (2004). *Assimilation variationnelle océanique : modélisation multivariée de la matrice de covariance d'erreur d'ébauche.* Thèse de doctorat, Université Toulouse III – Paul Sabatier.

[Rignot *et al.*, 2011] E. Rignot, J. Mouginot et B. Scheuchl (2011). Ice flow of the Antarctic ice sheet. *Science*, 333(6048):1427–1430.

[Ritz, 1992] C. Ritz (1992). *Un modèle thermo-mécanique d'évolution pour le bassin glaciaire Antarctique Vostok-Glacier Bird : Sensibilité aux valeurs des paramètres mal connus.* Thèse d'État, Université Joseph Fourier – Grenoble 1.

[Ritz *et al.*, 2001] C. Ritz, V. Rommelaere et C. Dumas (2001). Modeling the evolution of Antarctic ice sheet over the last 420,000 years : Implications for altitude changes in the Vostok region. *Journal of Geophysical Research : Atmospheres*, 106(D23):31943–31964.

[Robert, 2004] C. Robert (2004). *Développement et comparaison de méthodes d'assimilation de données de rang réduit dans un modèle de circulation océanique : application à l'océan Pacifique Tropical.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Rommelaere, 1997] V. Rommelaere (1997). *Trois problèmes inverses en glaciologie.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Rommelaere et MacAyeal, 1997] V. Rommelaere et D. R. MacAyeal (1997). Large-scale rheology of the Ross Ice Shelf, Antarctica, computed by a control method. *Annals of Glaciology*, 24:43–48.

[Sacher et Bartello, 2008] W. Sacher et P. Bartello (2008). Sampling errors in Ensemble Kalman Filtering. Part I : Theory. *Monthly Weather Review*, 136:3035–3049.

[Sakov et Bertino, 2011] P. Sakov et L. Bertino (2011). Relation between two common localisation methods for the EnKF. *Computational Geosciences*, 15(2):225–237.

[Sakov et Oke, 2008a] P. Sakov et P. R. Oke (2008a). A deterministic formulation of the Ensemble Kalman Filter : an alternative to ensemble square root filters. *Tellus A*, 60(2):361–371.

[Sakov et Oke, 2008b] P. Sakov et P. R. Oke (2008b). Implications of the form of the ensemble transformation in the ensemble square root filters. *Monthly Weather Review*, 136(3):1042–1053.

[Sakov *et al.*, 2012] P. Sakov, D. S. Oliver et L. Bertino (2012). An iterative EnKF for strongly nonlinear systems. *Monthly Weather Review*, 140(6):1988–2004.

[Sasaki, 1955] Y. Sasaki (1955). A fundamental study of the numerical prediction based on the variational principle. *Journal of the Meteorological Society of Japan*, 33(6):262–275.

[Sasaki, 1958] Y. Sasaki (1958). An objective analysis based on the variational method. *Journal of the Meteorological Society of Japan*, 36(3):77–88.

[Sasaki, 1970] Y. Sasaki (1970). Some basic formalisms in numerical variational analysis. *Monthly Weather Review*, 98(12):875–883.

[Scambos et Haran, 2002] T. A. Scambos et T. Haran (2002). An image-enhanced DEM of the Greenland Ice Sheet. *Annals of Glaciology*, 34(1):291–298.

[Schoof et Hindmarsh, 2010] C. Schoof et R. C. A. Hindmarsh (2010). Thin-film flows with wall slip : an asymptotic analysis of higher order glacier flow models. *The Quarterly Journal of Mechanics and Applied Mathematics*, 63(1):73–114.

[Simon et Bertino, 2009] E. Simon et L. Bertino (2009). Application of the Gaussian anamorphosis to assimilation in a 3-D coupled physical-ecosystem model of the North Atlantic with the EnKF : a twin experiment. *Ocean Science*, 5(4):495–510.

[Simon et Bertino, 2012] E. Simon et L. Bertino (2012). Gaussian anamorphosis extension of the DEnKF for combined state parameter estimation : Application to a 1D ocean ecosystem model. *Journal of Marine Systems*, 89(1):1–18.

[Snyder, 2012] C. Snyder (2012). Introduction to the Kalman Filter. *In Advanced Data Assimilation for Geosciences*, Les Houches, France.

[Snyder *et al.*, 2008] C. Snyder, T. Bengtsson, P. Bickel et J. Anderson (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, 136(12): 4629–4640.

[Souopgui, 2010] I. Souopgui (2010). *Assimilation d'images pour les fluides géophysiques*. Thèse de doctorat, Université de Grenoble.

[Tarasov et Peltier, 2004] L. Tarasov et W. R. Peltier (2004). A geophysically constrained large ensemble analysis of the deglacial history of the North American ice-sheet complex. *Quaternary Science Reviews*, 23(3):359–388.

[Tippett *et al.*, 2003] M. K. Tippett, J. L. Anderson, C. H. Bishop, T. M. Hamill et J. S. Whitaker (2003). Ensemble Square Root Filters. *Monthly Weather Review*, 131(7):1485–1490.

[van de Berg *et al.*, 2006] W. J. van de Berg, M. R. van den Broeke, C. H. Reijmer et E. van Meijgaard (2006). Reassessment of the antactic surface mass balance using calibrated output of a regional atmospheric climate model. *Journal of Geophysical Research : Atmospheres*, 111(D11).

[van Leeuwen, 1999] P. J. van Leeuwen (1999). Comments on "Data assimilation using an ensemble Kalman filter technique". *Monthly Weather Review*, 127:1374–1377.

[van Leeuwen, 2003] P. J. van Leeuwen (2003). A variance-minimizing filter for large-scale applications. *Monthly Weather Review*, 131(9):2071–2084.

[van Leeuwen, 2009] P. J. van Leeuwen (2009). Particle filtering in geophysical systems. *Monthly Weather Review*, 137(12):4089–4114.

[van Pelt *et al.*, 2013] W. J. J. van Pelt, J. Oerlemans et C. H. Reijmer (2013). An iterative inverse method to estimate basal topography and initialize ice flow models. *The Cryosphere*, pages 873–920.

[Verlaan et Heemink, 1995] M. Verlaan et A. W. Heemink (1995). Tidal flow forecasting using reduced rank square root filters. Report 95–75, Technische Universiteit Delft.

[Vidard, 2001] A. Vidard (2001). *Vers une prise en compte des erreurs modèle en assimilation de données 4D variationnelle. Application à un modèle d'océan.* Thèse de doctorat, Université Joseph Fourier – Grenoble 1.

[Vieli et Payne, 2003] A. Vieli et A. J. Payne (2003). Application of control methods for modelling the flow of Pine Island Glacier, West Antarctica. *Annals of Glaciology*, 36(1):197–204.

[Wang et Bishop, 2003] X. Wang et C. H. Bishop (2003). A comparison of breeding and Ensemble Transform Kalman Filter ensemble forecast schemes. *Journal of the atmospheric sciences*, 60(9):1140–1158.

[Wang *et al.*, 2004] X. Wang, C. H. Bishop et S. J. Julier (2004). Which is better, an ensemble of positive-negative pairs or a centered spherical simplex ensemble ? *Monthly Weather Review*, 132(7):1590–1605.

[Wang *et al.*, 1992] Z. Wang, I. M. Navon, F.-X. Le Dimet et X. Zou (1992). The second order adjoint analysis : theory and applications. *Meteorology and atmospheric physics*, 50(1-3):3–20.

[Whitaker et Hamill, 2002] J. S. Whitaker et T. M. Hamill (2002). Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130(7):1913–1924.

[Zhou *et al.*, 2006] Y. Zhou, D. McLaughlin et D. Entekhabi (2006). Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Monthly Weather Review*, 134(8):2128–2142.

Résumé

L'évolution des calottes polaires est régie à la fois par une dynamique d'écoulement complexe et par des mécanismes tel le glissement à la base, la température de la glace ou le bilan de masse en surface. De plus, de nombreuses boucles de rétroactions sont constatées entre les différents phénomènes impliqués. Tout ceci rend la modélisation de cette évolution complexe. Malgré tout, un certain nombre de modèles ont été développés dans cette optique. Ceux-ci font tous intervenir des paramètres influents qui dans certains cas sont peu ou pas connus. Ils nécessitent donc d'être correctement spécifiés. L'assimilation de données peut permettre une meilleure estimation de ces paramètres grâce à l'utilisation d'observations qui sont peu nombreuses en glaciologie. Dans cette thèse, nous nous intéressons à la mise en place de systèmes d'assimilation performants pour deux problèmes inverses concernant l'évolution des calottes polaires. Pour mieux nous concentrer sur ce point, nous avons travaillé avec un modèle d'évolution de calotte simplifié (appelé Winnie) qui, cependant, représente bien la plupart des processus complexes de la dynamique de la glace, et permet de travailler à différentes échelles de temps. Dans un premier temps, nous mettons en place une approche 4D-Var pour la reconstruction de l'évolution d'un paramètre climatique influant sur l'évolution d'une calotte sur une échelle de temps typique de 20 000 ans. Elle nécessite notamment l'écriture du code adjoint du modèle. Dans un second temps, nous nous intéressons au problème du spin-up. Ce problème de calibration du modèle pour des simulations à échelle de temps courtes (pas plus de 100 ans) consiste plus particulièrement en la reconstruction conjointe de l'état initial, de la topographie du socle rocheux et des paramètres de glissement basal. Nous développons ici une approche filtre de Kalman d'ensemble pour résoudre ce problème.

Mots clés : Assimilation de données, Méthodes inverses, Estimation de paramètres, Filtre de Kalman d'ensemble, Glaciologie, Calottes polaires.

Abstract

Ice sheet evolution is both driven by complex flow dynamics and by physical mechanisms such as basal sliding, ice temperature or surface mass balance. In addition to those, many feedback loops are observed between the different implicated phenomena. This explains how complex is to model this evolution. However several models have been developed. These models depend on influent parameters, which often are unfortunately poorly known. So they need to be correctly specified. Data assimilation can give a better estimation of these parameters thanks to observations which are quite rare in glaciology. During this thesis, we work with a simplified ice sheet evolution model called Winnie in order to focus on the setup. Nevertheless Winnie depicted the main complex processes of ice dynamics and can be used for studies with different time scales. The first part of the thesis focuses on on developping a 4D-Var approach in order to retrieve the evolution of a climatic parameter for a typical time scale of 20 000 years. This approach requires the implementation of the adjoint code of the evolution model. In a second part, we focus on the spin-up problem. This calibration problem for short term (maximum 100 years) simulations involve retrieving jointly the initial state, the bedrock topography and basal sliding parameters. In order to solve this problem we develop an Ensemble Kalman Filter approach.

Keywords : Data assimilation, Inverse methods, Parameter estimation, Ensemble Kalman Filter, Glaciology, Ice sheets.