


# THÈSE

**En vue de l'obtention du**

# DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

**Délivré par :**

Institut National Polytechnique de Toulouse (INP Toulouse)

**Discipline ou spécialité :**

Océan, Atmosphère et surfaces continentales

---

## Présentée et soutenue par :

M. NIELS OGER

le jeudi 2 juillet 2015

## Titre :

OBSERVATION ADAPTATIVE: LIMITES DE LA PREVISION ET DU CONTROLE DES INCERTITUDES

---

**Ecole doctorale :**

Sciences de l'Univers de l'Environnement et de l'Espace (SDUEE)

**Unité de recherche :**

Centre National de Recherches Météorologiques (CNRM)

**Directeur(s) de Thèse :**

M. OLIVIER PANNEKOUCKE

M. ALEXIS DOERENBECHER

**Rapporteurs :**

M. OLIVIER TALAGRAND, ENS PARIS

M. YANNICK TREMOLET, ECMWF READING

**Membre(s) du jury :**

M. OLIVIER THUAL, INP TOULOUSE, Président

M. ALEXIS DOERENBECHER, ECOLE NATIONALE METEOROLOGIE TOULOUSE, Membre

M. ARTHUR VIDARD, INRIA GRENOBLE - RHONE ALPES, Membre

M. ETIENNE MEMIN, INRIA RENNES, Membre

M. OLIVIER PANNEKOUCKE, CENTRE NATIONAL RECHERCHE METEOROLOGIQUE, Membre

# Remerciements

Les premiers remerciements vont naturellement à mes directeurs de thèse Olivier Pannekoucke et Alexis Doerenbecher, ainsi qu'à Philippe Arbogast pour leur support et leurs conseils durant toutes ces années. Vous avez toujours cru qu'on arriverait quelque part.

Je remercie aussi toute l'équipe RECYF pour leur accueil et leur bonne humeur à chacune de mes missions à Toulouse.
Et à la *colloc'* (Nathan, Florent, Thomas, Khem, Pierrot et Vincent) pour m'avoir gardé un coin où dormir.

Merci à Etienne Arbogast pour son aide pour l'installation et la compréhension de OOPS.

Merci à la direction de la DIRAG et du CNRM pour m'avoir permis de réaliser ce travail de thèse malgré le contexte peu favorable.

Un grand merci à toute l'équipe TTI de la DIRAG qui m'a permis de garder la motivation pour aller jusqu'au bout dans tout ce que nous avons entrepris. Même si cela ne contribue pas directement à mon travail de thèse, cela permet de garder le cap. Merci aussi pour votre patience quand je n'étais pas aussi disponible que j'aurais dû l'être.

Et une mention spéciale pour Anaïs pour sa générosité et sa très grande patience. Je n'aurais pas pu espérer mieux comme soutien moral et logistique.

Je remercie les rapporteurs et examinateurs du jury de thèse pour leurs questions et remarques pertinentes sur la première version du manuscrit.

# Résumé

L'observation adaptative (OA) est une pratique de prévision numérique du temps (PNT) qui cherche à prévoir quel jeu (ou réseau) d'observations supplémentaires à déployer et à assimiler dans le futur améliorera les prévisions. L'objectif est d'accroître la qualité des prévisions météorologiques en ajoutant des observations là où elles auront le meilleur impact (optimal). Des méthodes numériques d'OA apportent des réponses objectives mais partielles. Elles prennent en compte à la fois les aspects dynamiques de l'atmosphère à travers le modèle adjoint, et aussi le système d'assimilation de données. Le système d'assimilation de données le plus couramment utilisé pour l'OA est le 4D-Var. Ces méthodes linéaires (technologie de l'adjoint) reposent cependant sur une réalisation déterministe (ou trajectoire) unique. Cette trajectoire est entachée d'une incertitude qui affecte l'efficacité de l'OA. Le point de départ de ce travail est d'évaluer l'impact de l'incertitude associée au choix de cette trajectoire sur une technique : la KFS. Un ensemble de prévisions est utilisé pour étudier cette sensibilité.

Les expériences réalisées dans un cadre simplifié montrent que les solutions de déploiement peuvent changer en fonction de la trajectoire choisie. Il est d'autant plus nécessaire de prendre cette incertitude en considération que le système d'assimilation utilisé n'est pas vraiment optimal du fait de simplifications liées à sa mise en œuvre.

Une nouvelle méthode d'observation adaptative, appelée *Variance Reduction Field* (VRF), a été développée dans le cadre de cette thèse. Cette méthode permet de déterminer la réduction de variance de la fonction score attendue en assimilant une pseudo-observation supplémentaire pour chaque point de grille. Deux approches de la VRF sont proposées, la première est basée sur une simulation déterministe. Et la seconde utilise un ensemble d'assimilations et de prévisions. Les deux approches de la VRF ont été implémentées et étudiées dans le modèle de Lorenz 96. Le calcul de la VRF à partir d'un ensemble est direct si l'on dispose déjà des membres de l'ensemble. Le modèle adjoint n'est pas nécessaire pour le calcul.

L'implémentation de la VRF dans un système de prévision du temps de grande taille, tel qu'un système opérationnel, n'a pas pu être réalisée dans le cadre de cette thèse. Cependant, l'étude de faisabilité de la construction de la VRF dans l'environnement OOPS a été menée. Une description de OOPS (version 2013) est d'abord présentée dans le manuscrit, car cet environnement est une nouveauté en soi. Elle est suivie de la réflexion sur les développements à introduire pour l'implémentation de la VRF.

# Abstract

The purpose of adaptive observation (AO) strategies is to design optimal observation networks in a prognostic way to provide guidance on how to deploy future observations. The overarching objective is to improve forecast skill. Most techniques focus on adding observations. Some AO techniques account for the dynamical aspects of the atmosphere using the adjoint model and for the data assimilation system (DAS), which is usually either a 3D or 4D-Var (ie. solved by the minimization of a cost function). But these techniques rely on a single (linearisation) trajectory. One issue is to estimate how the uncertainty related to the trajectory affects the efficiency of one technique in particular : the KFS. An ensemble-based approach is used to assess the sensitivity to the trajectory within this deterministic approach (ie. with the adjoint model).

Experiments in a toy model show that the trajectory uncertainties can lead to significantly differing deployments of observations when using a deterministic AO method (with adjoint model and VDAS). This is especially true when we lack knowledge on the VDAS component.

During this work a new tool for observation targeting called Variance Reduction Field (VRF) has been developed. This technique computes the expected variance reduction of a forecast Score function that quantifies forecast quality. The increase of forecast quality that is a reduction of variance of that function is linked to the location of an assimilated test probe. Each model grid point is tested as a potential location. The VRF has been implemented in a Lorenz 96 model using two approaches. The first one is based on a deterministic simulation. The second approach consists of using an ensemble data assimilation and prediction system. The ensemble approach can be easily implemented when we already have an assimilation ensemble and a forecast ensemble. It does not need the use of the adjoint model.

The implementation in real NWP system of the VRF has not been conducted during this work. However a preliminary study has been done to implement the VRF within OOPS (2013 version). After a description of the different components of OOPS, the elements required for the implementation of the VRF are described.

# Table des matières


**1 Introduction** **13**
- 1.1 Contexte de la Prévision Numérique du Temps . . . 13
- 1.2 Sensibilité aux observations . . . 14
- 1.3 Prévision d'ensemble . . . 17
- 1.4 Problématique de la thèse . . . 17
- 1.5 Organisation du manuscrit . . . 19

**2 Assimilation de données** **21**
- 2.1 Formalisme du filtrage . . . 21
  - 2.1.1 Filtrage non linéaire . . . 21
  - 2.1.2 Filtre de Kalman . . . 22
  - 2.1.3 Exemple d'une assimilation à une observation . . . 23
  - 2.1.4 Rôle de la matrice de covariance d'erreur d'ébauche . . . 24
  - 2.1.5 Limitation de l'utilisation du filtre de Kalman . . . 26
- 2.2 Méthodes variationnelles . . . 26
  - 2.2.1 3D-Var . . . 26
  - 2.2.2 Extension temporelle : 4D-Inc . . . 26
- 2.3 Méthodes ensemblistes . . . 28
  - 2.3.1 Filtre de Kalman d'ensemble . . . 28
  - 2.3.2 Ensemble reduced rank square root Kalman filter . . . 28
- 2.4 Méthodes hybrides . . . 30
- 2.5 Les limites de l'utilisation des ensembles . . . 30
  - 2.5.1 Localisation . . . 30
- 2.6 Diagnostics sur le système d'assimilation de données opérationnel . . . 31
  - 2.6.1 Screening et sélection des observations . . . 31
  - 2.6.2 Sélection des observations pour l'assimilation de données . . . 32
  - 2.6.3 Diagnostic de la variance et des longueurs de portée . . . 34
  - 2.6.4 Illustration des DFS MF . . . 35
- 2.7 Conclusion chapitre 2 . . . 37

**3 Réduction de la variance d'erreur d'analyse due à une observation** **41**
- 3.1 Développement analytique . . . 41
  - 3.1.1 Assimilation d'une observation unique . . . 41
  - 3.1.2 Evaluation de l'assimilation d'une observation supplémentaire . . . 43
- 3.2 Validation expérimentale . . . 43
- 3.3 Conclusion . . . 45

**4 Sensibilité aux conditions initiales** **47**
- 4.1 Généralité et objectifs . . . 47
  - 4.1.1 Formalisme général . . . 47
- 4.2 Méthodes diagnostiques : impact des observations . . . 48

4.2.1 Formulation par développement de Taylor . . . 49
4.2.2 Méthode Géométrique : Daescu and Todling . . . 51
4.2.3 Ensemble Forecast Sensitivity to Observations . . . 51
4.2.4 Limite des méthodes d'impact d'observation pour l'observation adaptative . . . 52
4.3 Méthodes pronostiques . . . 52
4.3.1 Principe de l'observation adaptative . . . 53
4.3.2 Méthode de la Kalman Filter Sensitivity . . . 54
4.3.3 Lien entre la KFS et la HRR . . . 55
4.3.4 Application de l'ETKF à l'observation adaptative . . . 56
4.3.5 Variance Reduction Field . . . 57
4.3.6 Développement à l'ordre 2 . . . 57
4.4 Utilisation des sensibilités . . . 58
4.4.1 Déploiement d'observations supplémentaires . . . 58
4.4.2 Efficacité des déploiements d'observations supplémentaires . . . 60
4.4.3 Utilisation de la sensibilité aux observations pour améliorer le screening . . . 61
4.5 Conclusion du chapitre 4 . . . 62

**5 Modèles numériques utilisés 63**
5.1 Modèle Quasi-Géostrophique à deux couches . . . 63
5.1.1 Description du modèle Quasi-Géostrophique à deux couches . . . 63
5.1.2 Implémentation du modèle . . . 65
5.1.3 Validation du modèle non linéaire . . . 67
5.1.4 Comportement du modèle . . . 69
5.2 Modèle de Lorenz . . . 71
5.2.1 Equation du modèle de Lorenz . . . 71
5.2.2 Implémentation du modèle tangent linéaire et de l'adjoint dans Lorenz . . . 73
5.3 Bilan . . . 74

**6 Etude de la sensibilité à la trajectoire 75**
6.1 Protocole et cadre expérimental . . . 76
6.1.1 Choix de la condition initiale . . . 76
6.1.2 Système d'assimilation de données . . . 77
6.1.3 Description des expériences réalisées . . . 77
6.1.4 Exemple de résultats préliminaires . . . 78
6.2 Résultats et discussion . . . 80
6.2.1 Exp I : Variabilité du champ de KFS à la trajectoire . . . 80
6.2.2 Exp II : Valorisation de l'ensemble pour optimiser le déploiement . . . 85
6.2.3 Exp III : Validation du formalisme . . . 86
6.3 Conclusion des expériences . . . 91

**7 Implémentation de la VRF dans le modèle de Lorenz 93**
7.1 Implémentation de la VRF . . . 93
7.2 Cadre expérimental des équations . . . 94
7.3 Résultats . . . 94
7.4 Discussion sur l'implémentation de la VRF . . . 94
7.5 Forecast variance reduction field, a new tool for adaptive observation : ensemble versus adjoint computation . . . 95

**8 Implémentation de la VRF dans OOPS** **105**
8.1 Contexte et objectifs de OOPS . . . 105
8.2 Architecture générale du code . . . 106
8.2.1 Classes génériques : 1er niveau . . . 106
8.2.2 Classes spécifiques : 2e niveau . . . 107
8.2.3 Lien entre le C++ et le FORTRAN : 3e niveau . . . 109
8.3 Description des objets utilisés dans OOPS . . . 113
8.3.1 Organisation des classes dans OOPS . . . 113
8.3.2 Configuration d'une simulation . . . 114
8.4 Algorithme de la VRF dans OOPS . . . 120
8.4.1 Cadre des simulations . . . 120
8.4.2 Avec un adjoint et une assimilation variationnelle . . . 122
8.4.3 Avec un ensemble . . . 125
8.5 Commentaires sur l'implémentation dans OOPS . . . 128

**9 Conclusion et perspectives** **129**
9.1 Retour sur la problématique de la thèse . . . 129
9.2 Analyse de la sensibilité de la KFS à trajectoire de référence . . . 130
9.3 Sensibilité aux observations à l'aide d'un ensemble . . . 130
9.4 Perspectives de travail à la suite de cette thèse . . . 131

**A Modélisation de B** **139**

**B Algorithme combiné du gradient conjugué et de Lanczos** **141**

**C Algorithme de l'ETKF** **143**

**D Création d'une distribution d'erreur à partir d'un opérateur** **145**

**E Calcul du gradient de la fonction score par régression linéaire** **147**

**F Développement de la VRF au second ordre à partir d'un ensemble** **149**
F.1 Résolution de la régression linéaire polynomiale . . . 149
F.2 Calcul de la variance de la fonction score au second ordre . . . 150
F.3 Réduction de variance de la fonction au second ordre . . . 151

**G Calcul de la trace d'une matrice par une méthode de Monte-Carlo** **153**

**H Assessing the trajectory influence in adaptive observation Kalman filter sensitivity method.** **155**

**I Code C++ pour le calcul de la VRF avec un ensemble** **169**

# Chapitre 1

# Introduction

**Préambule**

Ce travail de thèse est la continuité de mon stage de fin d'études réalisé dans l'équipe GMAP/RECYF du CNRM-GAME. La thèse s'est déroulée en parallèle de mon activité professionnelle à temps plein en tant que responsable de la division Télécom et Traitement de l'Information (TTI) de la Direction Antilles-Guyane (DIRAG) de Météo-France, située en Martinique. Le contexte d'éloignement géographique a imposé des méthodes de travail permettant le plus d'autonomie possible. Ainsi, les développements réalisés durant cette thèse sont plutôt théoriques avec des mises en œuvre numériques dans des environnements légers, comme par exemple le logiciel Scilab (http ://www.scilab.org/fr). Les simulations réalisées pendant le stage au CNRM ont été réalisées sur le cluster de calcul ANETO, alors que les simulations propres au travail de thèse ont été réalisées sur mon ordinateur personnel. Le contexte opérationnel est cependant étudié dans le dernier chapitre du manuscrit, à travers une implémentation du principal résultat de cette thèse (la méthode VRF) dans un nouvel environnement de développement : OOPS (*Object-Oriented Prediction System*), qui préfigure les codes informatiques pour l'assimilation opérationnelle de demain.

## 1.1 Contexte de la Prévision Numérique du Temps

Parmi les missions institutionnelles de Météo-France figure la sécurité météorologique des personnes et des biens. Cela passe par la surveillance et la prévision du comportement de l'atmosphère, du manteau neigeux, de l'océan superficiel, à la fois grâce aux réseaux observations et aux modèles issus de la Prévision Numérique du Temps (PNT). Plusieurs modèles sont utilisés à Météo-France pour répondre aux besoins des prévisionnistes, ARPEGE (Courtier et al., 1991) et le modèle déterministe du Centre Européen de Prévision Météorologie à Moyen Terme (CEPMMT) fournissent des prévisions déterministes sur l'ensemble du globe sur la semaine à venir. Le modèle à aire limité AROME (Seity et al., 2011) fournit une prévision à haute résolution spatiale pour les deux jours à venir. Deux ensembles de prévision ARPEGE (PEARP, Descamps et al. (2014)) et bientôt AROME (PEARO, Bouttier et al. (2012)) permettent d'évaluer l'incertitude sur la prévision déterministe et d'envisager différents scénarios possibles sur des échéances courtes, à très courtes, respectivement. Et finalement la prévision d'ensemble du CEPMMT offre une prévision de la tendance jusqu'à 15 jours (moyenne échéance). Tous ces modèles numériques, effectuent des prévisions quotidiennes, voire plusieurs fois par jour (réseaux). Météo-France, tout comme le CEPMMT, dispose de supercalculateurs pour cela. Ces prévisions successives permettent aux prévisionnistes d'avoir à disposition des informations récentes et à jour, notamment grâce aux dernières observations disponibles. Historiquement, ces prévisions sont réparties dans la journée en fonction des heures des principales collectes d'observation (ou réseaux synoptiques principaux à 0h, 6h, 12h et 18h UTC).

Un système de PNT opérationnel doit trouver un compromis entre,

- le coût de calcul, qui dépend de la résolution du modèle et de la durée de la prévision,
- le délai après l'heure du réseau d'observation pour réceptionner un maximum des observations disponibles, aussi appelé *cut-off*,

– le délai de mise à disposition des données aux prévisionnistes.

Le rafraîchissement des prévisions est réalisé grâce aux procédures d'assimilation des données d'observation. L'assimilation de données consiste à produire l'analyse objective, qui correspond à la condition initiale, point de départ nécessaire pour réaliser une nouvelle prévision. L'analyse est construite comme la correction d'une ébauche (en pratique, la prévision la plus récente) par l'information issue des observations. Les avancées scientifiques mises en œuvre pour le calcul de cette analyse sont une importante source d'amélioration des prévisions de ces dernières années. L'assimilation de données représente environ 40% [1] du temps de calcul pour réaliser une prévision. C'est pourquoi le *cut-off* et le choix des observations sont cruciaux pour ajuster l'analyse et la prévision en fonction des besoins du prévisionnistes.

Les observations utilisées en PNT sont de différents types et proviennent de nombreuses sources. Les observations dites conventionnelles sont fournies par des stations de mesure au sol ou en mer (SYNOP), par les bouées (BUOY), par les radiosondages d'altitude (TEMP) ou les mesures issues des avions de ligne (données de type AMDAR). Mais les observations majoritaires sont fournies par les satellites géostationnaires ou défilants. Ces mesures sont effectuées par différents instruments passifs (de type sondeur ou imageur) qui explorent différentes longueurs d'onde (dans le visible, l'infrarouge ou la bande d'absorption de la vapeur d'eau) ou actifs (type radar). Depuis peu de temps, les modèles non hydrostatiques sont capables d'assimiler des données de vent et de réflectivité collectées de radars de précipitation.
La répartition spatiale et temporelle des observations est hétérogène, et varie suivant les types d'observation. La figure 1.1 montre la localisation des mesures de stations de surface et de type SYNOP (au sens large) assimilées dans le modèle ARPEGE pour la journée du 15/04/2015 au réseau de 0h UTC. Cet exemple montre que la répartition des observations (ici entre la terre et la mer) peut-être inégale à la surface du globe. Ce que l'on voit ici sur l'horizontal est aussi vrai pour la verticale : certaines couches de l'atmosphère sont très peu observées par des systèmes d'observation in-situ. Les observations satellite assurent désormais une couverture régulière et relativement uniforme du globe, mais ces techniques ont d'autres limitations : par exemple, la mesure de vent est relativement pauvre en comparaison des informations relatives à la température et à l'humidité. En outre, l'assimilation des mesures infrarouge en présence de nuages reste délicate. Les possibles hiatus ou discontinuités spatio-temporels des réseaux d'observations sont dommageables à la qualité des prévisions car des objets météorologiques (tels des précurseurs de dépressions par exemple) pourraient échapper à l'assimilation et ainsi priver les prévisions des ingrédients nécessaires à la prévision d'un phénomène à fort impact.

## 1.2 Sensibilité aux observations

Les méthodes de sensibilité aux observations ont été développées à la fin des années 90 à la suite de la campagne de mesure FASTEX [2] (Joly et al., 1999). Cette campagne a été la première mise en application de l'observation adaptative en zone tempérée (i.e. hors cyclones tropicaux) et avec l'appui de méthodes numériques. Les premières méthodes cherchaient à déterminer les zones dynamiquement sensibles à travers de techniques numériques utilisant le modèle adjoint (e.g. les vecteurs singuliers, voir Palmer et al. (1998)) ou ensemblistes (*Ensemble Transform*, voir Bishop and Toth (1999)). L'objectif de ces techniques était de déterminer quelles étaient les zones géographiques où de possibles défauts sur les conditions initiales auraient eu un effet important quant à la dégradation des prévisions. En y déployant des observations supplémentaires, on essayait d'améliorer les conditions initiales et donc la prévision. Des méthodes plus avancées ont ensuite été développées pour prendre en compte le système d'assimilation de données en plus de la dynamique. En effet, il a été montré dans un système opérationnel, que l'impact des observations sur la prévision dépendent du système d'assimilation (Bergot, 1999). Parmi ces méthodes de seconde génération on retrouve l'*Ensemble Transform Kalman Filter* (ETKF)(Bishop et al., 2001), la *Hessian Reduced Rank estimate* (HRR) (Leutbecher, 2003) ou la *Kalman Filter Sensitivity* (KFS)(Doerenbecher and Bergot, 2001) développée à Météo-France. Appliquées de manière pronostique ces techniques permettent

1. Cas du modèle global ARPEGE à Météo-France (2008)
2. Fronts and Atlantic Storm-Track EXperiment

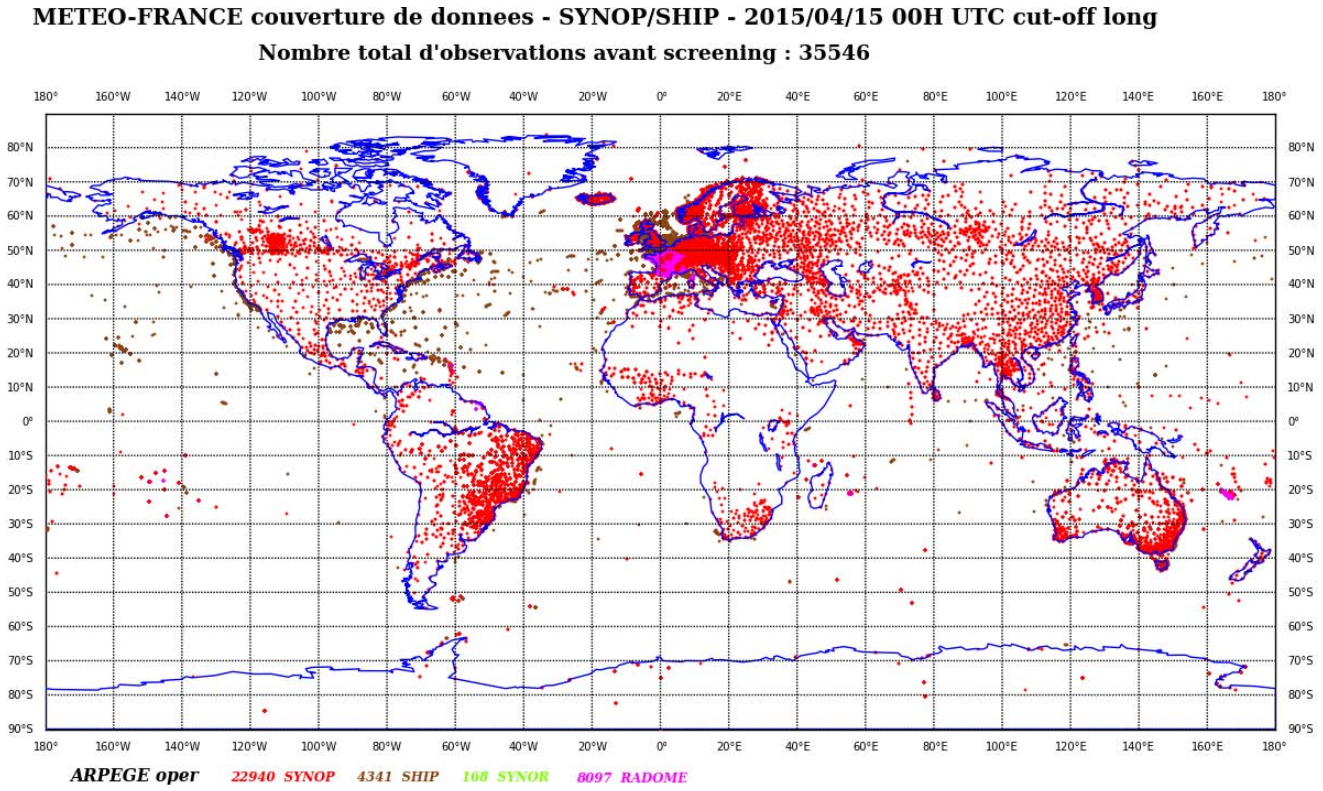



FIGURE 1.1 – Positionnement des observations sol entrant dans le système d'assimilation de ARPEGE. Les SYNOP et SYNOR correspondent aux mesures sol transmises via le Système Mondial de Télécommunication (SMT), RADOME est le réseau propriétaire de Météo-France et les SHIP sont les mesures effectuées depuis des navires volontaires (observation d'opportunité). Cette carte ne représente pas les données collectées par les bouées (BUOY) ancrées ou dérivantes.

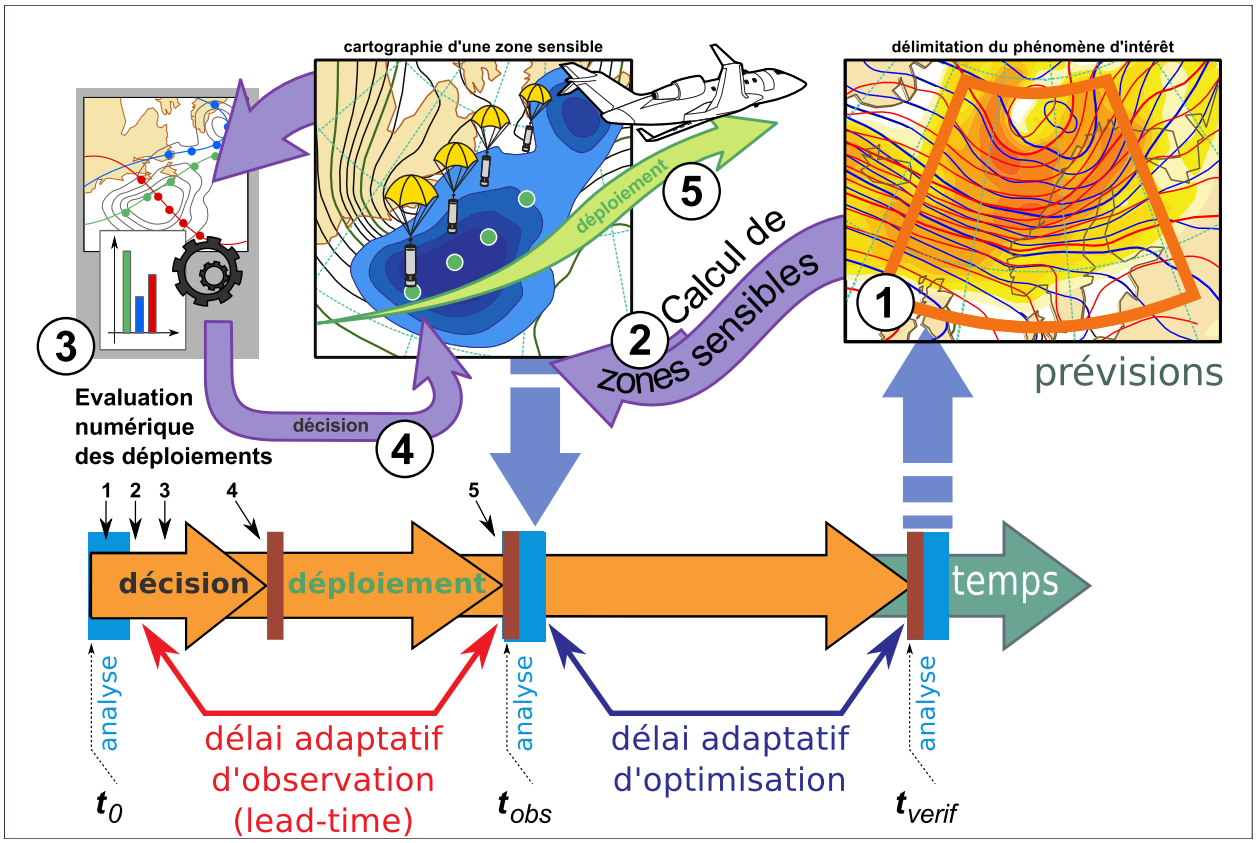



FIGURE 1.2 – Schéma récapitulatif des étapes relatives à un système d'observation adaptative. La première étape est le calcul d'une prévision jusqu'à $t_{verif}$. La deuxième consiste à calculer le champ de sensibilité. A partir de ce champ de sensibilité des stratégies de déploiement sont proposées et évaluées (3). Si une décision de déploiement supplémentaire est prise (4), une demande est transmise au gestionnaire du moyen de mesure qui réalise le déploiement d'observations à $t_{obs}$ (5).

de fournir des indications pour le déploiement d'observations supplémentaires dans le futur. La philosophie de l'observation adaptative est de déterminer aujourd'hui (à $t_0$), l'endroit où il faudra ajouter des observations demain (à $t_{obs}$), pour améliorer la prévisibilité du modèle pour la prévision d'après-demain (à $t_{verif}$). Développer un système d'observation adaptative complet est complexe car en plus de calcul les zones de sensibilité, il est nécessaire de pouvoir déployer ces observations supplémentaires à $t_{obs}$, voir figure 1.2. Il faut donc coordonner de façon originale producteurs et utilisateurs des observations. Le DTS (*Data Targeting System*) est un système quasi-opérationnel et coopératif développé par le CEPMMT qui fournit une interface web aux différents acteurs pour suivre toutes les étapes du processus et même participer aux étapes de décision qui aboutissent au déploiement d'observation supplémentaires (Prates et al., 2009).

La plate-forme d'observation (ou vecteur) est un élément important. Elle est soit dédiée à l'observation adaptative, tel un avion de recherche larguant des sondes (dropsondes), soit opportuniste en utilisant des systèmes déjà en place, comme des sites de radiosondages. Le choix dépend essentiellement des ressources disponibles. Les déploiements réalisés lors de la campagne FASTEX étaient ambitieux, avec des plans de vol pour plusieurs avions de mesure. Ce genre de déploiement est toujours utilisé aux USA (NCEP et NOAA) dans des campagnes hivernales au-dessus de l'Océan Pacifique Nord (`http://www.aoc.noaa.gov/article_winterstorm.htm`). Pour la campagne CONCORDIASI (`http://www.cnrm.` en Antarctique, se sont des ballons dérivant larguant des sondes à la demande qui ont été déployés. Pour la dernière campagne HyMeX (Ducrocq et al. (2014)), les observations supplémentaires étaient fournies par le réseau de radiosondages conventionnels. Des lâchers supplémentaires (en dehors des horaires de sondages routiniers) étaient réalisés à la demande, en fonction de stratégies de déploiements.

Devant la complexité de mise en œuvre des systèmes complets d'observation adaptative, et les questions d'optimisation des réseaux d'observation qu'elles ont soulevées, des approches diagnostiques ont été développées sur la base de la sensibilité aux observations. Ces méthodes d'étude d'impact des observations ont connu un essor ces dernières années. En effet, elles permettent à peu de frais d'estimer l'impact effectif d'une observation ou de tout groupe d'observations sur la prévision. Les méthodes les plus avancées, comme l'*Ensemble Forecast Sensitivity to Observations* (EFSO, Hotta et al. (2014)), permettent aussi de mettre en œuvre une optimisation des paramètres du système d'assimilation (ajustement des variances d'erreur d'observation). L'objet de ce travail de thèse est de revisiter les méthodes d'observation adaptative, à partir des développements effectués dans le cadre des études d'impact d'observations et de tirer profit des méthodes ensemblistes.

## 1.3 Prévision d'ensemble

L'atmosphère est connue pour avoir une évolution chaotique, c'est-à-dire que deux états initiaux proches peuvent donner des prévisions très éloignées. Dans les conditions initiales, cette incertitude est en partie prise en compte par le système d'assimilation de données, mais ensuite, le système de prévision déterministe n'est pas conçu pour prévoir l'évolution de ces incertitudes dans le temps (échéances). Pour représenter cette évolution de l'incertitude dans le temps, des méthodes de Monte-Carlo sont utilisées. Le principe est d'échantillonner la distribution de probabilité des erreurs sur les conditions initiales. On applique le modèle de prévision à chaque réalisation dans l'échantillon, ou membre de l'ensemble, voir l'illustration figure 1.3.

L'intégration temporelle de chaque membre est réalisée avec le modèle non linéaire complet. Cela permet d'avoir une estimation complète de l'incertitude et pas uniquement une évaluation dans un cadre linéaire comme cela est le cas pour le filtre de Kalman.

Pour réaliser ce travail de thèse, plusieurs ensembles sont utilisés. La manière de générer un ensemble dépend de l'utilisation que l'on souhaite en faire. C'est là un thème de recherche en soi. Cependant pour un usage dans les modèles simplifiés, la création des perturbations est relativement simple.

## 1.4 Problématique de la thèse

La méthode d'observation adaptative utilisée à Météo-France, la KFS, nécessite un certain nombre d'hypothèses pour être mise en œuvre. Une trajectoire de référence est utilisée pour transporter l'information de $t_{verif}$ à $t_{obs}$ en utilisant le modèle adjoint. De plus, comme $t_{obs}$ se trouve dans le futur par rapport à l'instant où sont réalisés les calculs, il est nécessaire de prévoir ou simuler un système d'assimilation complet, c'est-à-dire un réseau d'observation et une matrice de covariance d'erreur d'ébauche. Le réseau d'observation ainsi simulé est appelé réseau de routine. Par commodité, un tel réseau peut être simulé à partir des observations d'un réseau passé.

**Sensibilité de la KFS à la trajectoire non linéaire** Le premier objectif de la thèse est d'évaluer la sensibilité de la KFS au choix de la trajectoire de référence, au choix du réseau de routine simulé et à la matrice d'erreur d'ébauche utilisée dans l'assimilation. Pour répondre à cette question, une maquette numérique a été développée permettant pour construire différentes expériences. Cette approche permet de maîtriser l'ensemble des éléments de la simulation, la sensibilité de la KFS est étudiée hors des contraintes de temps réel. Un ensemble de prévisions est utilisé pour construire différents scénarios possibles et donc obtenir des variations dans les évaluations de déploiements produites par la KFS. Les résultats présentés dans cette thèse, sur ce sujet, ont fait l'objet d'un article publié au *Quarterly Journal of the Royal Meteorological Society* (Oger et al., 2012).

**Une méthode pour mieux tirer partie de l'information dans un ensemble** Le second objectif est de proposer une méthode permettant de tirer profit de l'information de l'ensemble pour réduire l'incertitude

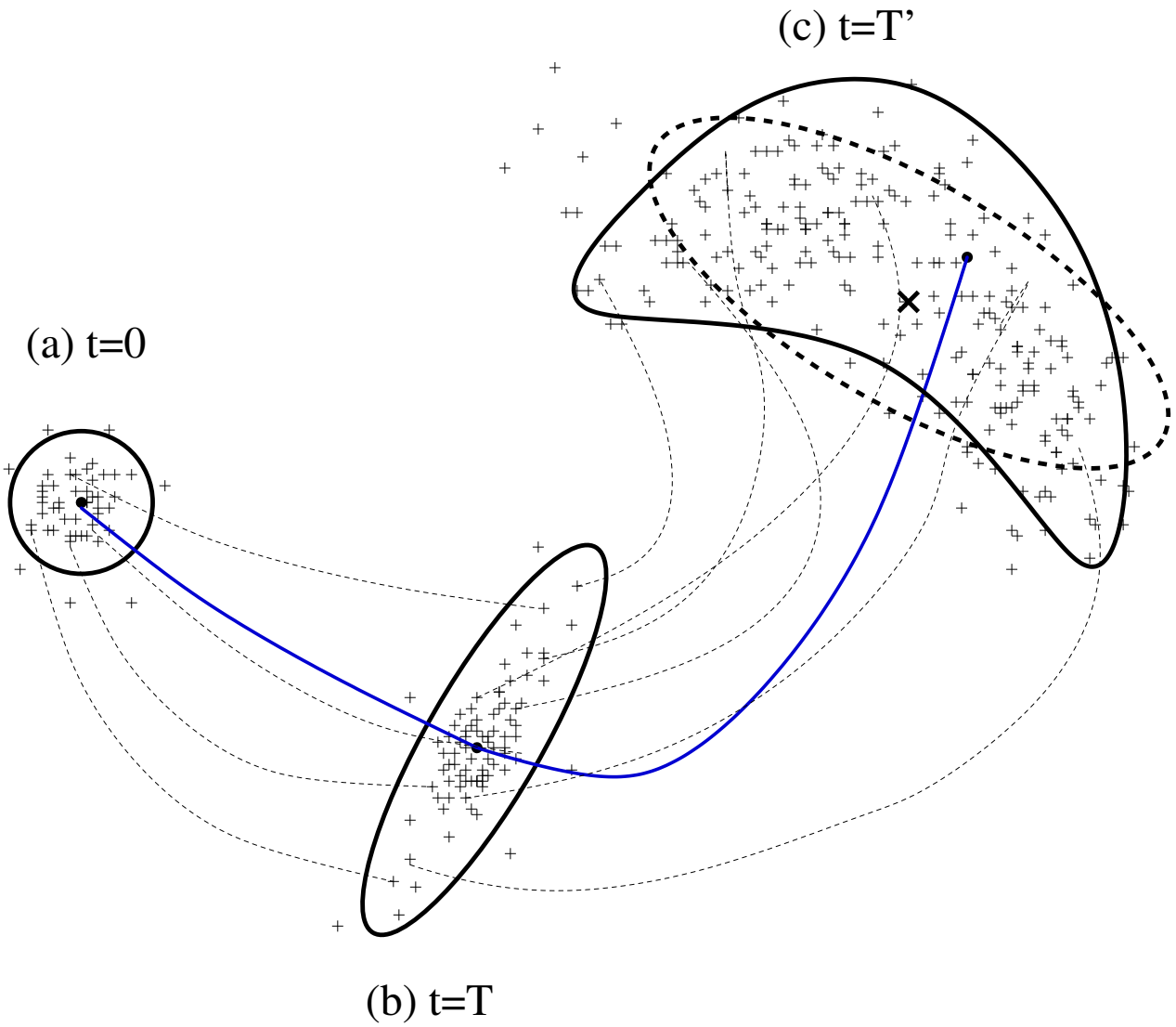



FIGURE 1.3 – Schéma d'évolution d'un ensemble de prévisions. Les + représentent le vecteur d'état de chaque membre de l'ensemble. Les traits pointillés représentent certaines des trajectoires non linéaires entre deux instants. Le point et la ligne bleue indiquent l'état de référence autour duquel est construit l'ensemble. Le cercle représente l'incertitude autour de cet état de référence. Entre $t = 0$ et $t = T$, l'incertitude subit une croissance linéaire . L'axe privilégié de croissance de l'incertitude génère une ellipse. Entre $t = T$ et $t = T'$ la croissance est non linéaire, la répartition de l'incertitude dans l'ensemble est beaucoup plus complexe et ne peut plus être représentée par une gaussienne.

liée à la trajectoire de référence et à la modélisation de la matrice de covariance d'erreur d'ébauche. Les aspects théoriques de la KFS sont repris pour développer la méthode *Variance Reduction Field* (VRF) qui fournit une estimation de la réduction de variance associée à une observation supplémentaire en chaque point de grille. Cette méthode a la particularité de pouvoir être implémentée dans un système d'assimilation variationnelle et utilisant un adjoint pour le transport du gradient de la fonction score, ou d'être implémentée avec un ensemble de prévisions et d'assimilations. Le développement théorique autour de la VRF, ainsi que les premiers résultats dans un cadre académique, ont été compilés sous la forme d'une note soumise au *Monthly Weather Review* (Oger et al., soumis en 2014).

La validation de la VRF dans une application grandeur nature n'a pas pu être réalisé dans le cadre de la thèse. Cependant, l'étude de faisabilité a été menée dans l'environnement OOPS. L'objectif poursuivi est de s'appuyer sur OOPS pour ensuite pouvoir documenter le comportement de la VRF dans les modèle de la PNT (ARPEGE/IFS, AROME).

**Résultat connexe non attendu initialement** Si l'on applique le formalisme de la VRF en se restreignant à l'étape d'analyse, il est alors possible d'écrire l'impact d'une observation sur la réduction de variance. En particulier, sous des hypothèses de stationnarité locales il est possible d'exprimer la réduction de variance en faisant intervenir le diagnostique des longueurs de portée, une longueur caractéristique de la corrélation des erreurs de prévision (voir chapitre 3). Ce résultat n'était pas attendu dans l'organisation initial du travail de thèse. L'intérêt principal de ce diagnostic est de pouvoir contribuer à la sélection des observations dans le screening (étape correspondant à la sélection des jeux d'observations entrant dans la production de l'analyse).

## 1.5 Organisation du manuscrit

Le manuscrit est composé de 6 chapitres, hors introduction et conclusion. Les chapitres 2, 4 et 5 sont des chapitres théoriques qui présentent respectivement l'assimilation de données, la sensibilité aux observations et les deux modèles utilisés pour les expériences (le modèle quasi-géostrophique à 2 couches et le modèle de Lorenz 1996). Le chapitre 3 propose une nouvelle méthode pour estimer la réduction de variance à l'analyse associée l'assimilation d'une observation, sans directement prendre en compte la prévision. Les chapitres 6 et 7 présentent les simulations numériques réalisées et les résultats associés. Le chapitre 8 ne contient pas de simulation ou de résultat. Il décrit l'environnement de développement orienté objet, OOPS, et la manière d'implémenter la VRF dans ce contexte qui doit permettre une implémentation à l'échelle d'un système de PNT quasi-opérationnel.

Le chapitre 2 introduit l'assimilation de données à travers le formalisme du filtrage non linéaire et le filtre de Kalman. Un exemple d'assimilation à une observation est utilisé pour illustrer le rôle de la matrice de covariance d'erreur d'ébauche $\mathbf{B}$. Une présentation des méthodes variationnelles et d'assimilation par transformée d'ensemble est ensuite réalisée. La fin du chapitre est consacrée aux diagnostics relatifs aux observations (contrôle, *screening* et DFS) dans les systèmes d'assimilation opérationnels à Météo-France.

Le chapitre 3 est organisé sous la forme d'une note. Une méthode est proposée pour calculer la réduction de variance d'erreur d'analyse associée à une observation supplémentaire, en utilisant en particulier une modélisation des fonctions de covariance. Cette méthode est illustrée dans un cadre idéalisé.

Le chapitre 4 traite des méthodes de sensibilité aux observations, en commençant par le calcul d'impact d'observations et ensuite l'observation adaptative. Les méthodes d'impact d'observations proposent plusieurs manières d'effectuer des développements au second ordre dans un contexte diagnostic (c'est-à-dire que les calculs sont réalisés à posteriori de la validité de l'assimilation). Trois méthodes d'observation adaptative sont présentées, l'ETKF, la HRR et la KFS, utilisée à Météo-France, avec un focus particulier sur cette dernière. Une évolution de la KFS est proposée sous la forme d'un calcul de champ de réduction de variance nommée VRF, pour *Variance Reduction Field*. Le chapitre se termine sur une présentation d'un usage des sensibilités par rapport aux besoins et aux possibilités opérationnelles.

Le chapitre 5 est le dernier chapitre principalement théorique. Il décrit les équations des deux modèles dynamiques utilisés pour réaliser les expériences, le modèle quasi-géostrophique à 2 couches (QG) et le

modèle de Lorenz 96. La dynamique du modèle QG est décrite en particulier, en lien avec la validation des modèles linéaire et tangent associés qui ont été développés. Le cadre des simulations avec le modèle QG est aussi introduit, avec un écoulement de base et le développement de perturbations baroclines.

Le chapitre 6 répond à la première problématique de la thèse, à savoir la sensibilité de la KFS à la trajectoire de référence, et le rôle respectif de la matrice de covariance d'erreur d'ébauche et du réseau d'observation de routine. Plusieurs expériences sont réalisées avec le modèle QG et présentées pour illustrer ces différents points. Ces premiers résultats montrent qu'il peut être intéressant de tirer profit d'un ensemble de prévisions pour réduire l'incertitude liée à la trajectoire de référence. Quelques pistes sont proposées à la fin du chapitre. Le travail rapporté dans ce chapitre a fait l'objet d'un article publié au QJRMS.

Le chapitre 7 reprend les expériences d'observation adaptative à partir d'un ensemble mais cette fois-ci en illustrant l'utilisation de la VRF dans un modèle de Lorenz 96. Deux manières d'implémenter la VRF sont proposées, la première en utilisant un système d'assimilation variationnel et le modèle adjoint, la seconde en utilisant un ensemble d'assimilations et de prévisions. Le cadre simplifié du modèle de Lorenz 96 permet de comparer ces deux méthodes au calcul analytique de la réduction de variance. Le travail rapporté dans ce chapitre a fait l'objet d'une notre soumise au MWR et actuellement en révision.

Le chapitre 8 reprend les éléments d'implémentation de la VRF en suivant les deux approches présentées au chapitre 7 mais cette fois ci dans l'environnement quasi-opérationnel OOPS. Une introduction à l'environnement de développement orienté objet est réalisée ainsi qu'une présentation des classes principales et de la manière dont fonctionne le code. Les deux implémentations sont ensuite proposées en indiquant les modifications nécessaires du code et les configurations associées (fichiers XML équivalents à des *namelists*). Une implémentation opérationnelle de la VRF dans l'environnement OOPS n'a pas être réalisée dans le cadre de la thèse par manque de temps.

Le manuscrit se termine par une conclusion au chapitre 9 qui présente les principaux résultats de la thèse et leurs implications dans l'utilisation des méthodes d'observation adaptative. Des propositions pour un nouvel usage de ces méthodes, adaptées aux nouvelles observations disponibles, sont présentées.

# Chapitre 2

# Assimilation de données

Ce chapitre présente les principes de l'assimilation de données en s'appuyant sur les notations couramment utilisées (Ide et al., 1997). Lorsque l'on souhaite suivre l'évolution d'un système complexe comme l'atmosphère, il faut avoir à l'esprit qu'il n'est jamais possible de connaître exactement l'état du système à un instant donné. Ceci est dû aux simplifications introduites par la modélisation et au caractère chaotique du système qui se traduit par une forte sensibilité de la prévision à ses conditions initiales et des incertitudes sur la prévision. De même, les observations du système sont aussi entachées d'erreurs (erreur de mesure, de représentativité). L'objectif de l'assimilation de données en météorologie opérationnelle est d'offrir la meilleure représentation numérique de l'atmosphère à partir des informations disponibles : une prévision et des observations.

Ce chapitre introduit le formalisme général de l'assimilation de données dans la section 2.1. Ensuite les deux principales méthodes d'assimilation utilisées en observation adaptative pour la PNT sont présentées avec le formalisme variationnel (2.2) et de la Transformée d'Ensemble (ETKF), 2.3). La fin du chapitre (2.6) est consacrée aux différents diagnostics sur le système d'assimilation de données utilisés à Météo-France.

## 2.1 Formalisme du filtrage

### 2.1.1 Filtrage non linéaire

Le formalisme présenté ici est très général et s'inscrit dans le cadre de l'approche Bayésienne conduisant au processus de filtrage non linéaire (Del Moral, 2004). Dans ce cadre probabiliste, la problématique de l'assimilation de donnée, consiste à décrire l'évolution de l'information décrivant la connaissance de l'état d'un système. Formellement, l'état au temps $i$ est modélisé par une variable aléatoire de dimension $n$ notée $\boldsymbol{X}_i$. L'objectif est de caractériser $\boldsymbol{X}_i$ en s'appuyant sur la dynamique et les observations au cours du temps. Ces observations sont également modélisées par une variable aléatoire, notée $\boldsymbol{y}_i$ au temps $i$.

L'espace des observations ne correspondant pas nécessairement à l'espace du modèle (ensemble des états possibles pour $\boldsymbol{X}_i$), il est nécessaire d'introduire un opérateur d'observation $\mathcal{H}$ permettant de traduire l'état $\boldsymbol{X}_i$ dans les quantités observées contenues dans $\boldsymbol{Y}_i$. En particulier, cet opérateur permet de passer de l'espace du modèle, de dimension $n$, à l'espace des observations, de dimension $p$. Ainsi, le lien entre état du modèle et observation s'écrit $\boldsymbol{Y}_i = \mathcal{H}_i(\boldsymbol{X}_i) + \varepsilon_i^o$, où $\varepsilon_i^o$ désigne les erreurs d'observation, modélisées sous la forme d'un vecteur aléatoire de moyenne nulle. Ces erreurs peuvent être liées à l'interpolation de l'état du système vers l'espace des observations (erreur de représentativité) ou tout simplement à des erreurs de mesure liées à l'instrument. L'information contenue dans les observations, reliée à l'état du système, est représentée par une distribution de probabilité qu'on notera ici $\Pr(\boldsymbol{Y}_i|\boldsymbol{X}_i)$. On suppose de plus que cette probabilité est à densité (par rapport à la mesure de Lebesgue), notée ici $\mathrm{p}(\boldsymbol{y}_i|\boldsymbol{x}_i)$.

Dans le cas où $p < n$, le système est sous-déterminé et ne peut être résolu de manière directe. Un état à priori du système est alors introduit pour fermer le problème. Le modèle dynamique $\mathcal{M}$ transporte l'état

de l'instant $i-1$ à l'instant $i$ par la relation

$$\boldsymbol{X}_i = \mathcal{M}_i(\boldsymbol{X}_{i-1}) + \varepsilon_i^m,$$

où $\varepsilon_i^m$ désigne l'erreur de modèle, modélisée comme étant un vecteur aléatoire de moyenne nulle. L'erreur modèle représente les défauts de la modélisation numérique du problème physique. La loi de transition est notée $\Pr(\boldsymbol{X}_i|\boldsymbol{X}_{i-1})$ et la fonction de densité associée est $\mathrm{p}(\boldsymbol{x}_i|\boldsymbol{x}_{i-1})$.

En utilisant la formule de Bayes [1] il est possible d'estimer la densité à l'instant $i$ en fonction des états précédents et des observations

$$\mathrm{p}(\boldsymbol{x}_i|\boldsymbol{y}_{1:i}) = \frac{\mathrm{p}(\boldsymbol{y}_i|\boldsymbol{x}_i)\,\mathrm{p}(\boldsymbol{x}_i|\boldsymbol{y}_{1:i-1})}{Z}, \tag{2.1}$$

où $Z$ est un facteur de normalisation permettant d'obtenir une densité de probabilité. $\mathrm{p}(\boldsymbol{x}_i|\boldsymbol{y}_{1:i-1})$ peut s'écrire sous la forme $\int \mathrm{p}(\boldsymbol{x}_i|\boldsymbol{x}_{i-1})\,\mathrm{p}(\boldsymbol{x}_{i-1}|\boldsymbol{y}_{1:i-1})d\boldsymbol{x}_{i-1}$, il apparaît ainsi une récursivité pour l'estimation de l'état à l'instant $i$ à partir de l'estimation à l'instant $i-1$. Cette récursivité est traduite par un cycle d'assimilations et de prévisions, l'information des observations aux instants antérieurs étant contenue dans la distribution de l'estimateur.

Il faut alors distinguer deux états du système à un instant $i$, l'estimateur à priori noté $\boldsymbol{X}_i^b$ aussi appelé ébauche (*background*), dont l'information provient des instants précédents, et l'estimateur à posteriori noté $\boldsymbol{X}_i^a$ qui contient en plus l'information issue des dernières observations, aussi appelé analyse. Les estimateurs sont définis suivant les relations $\boldsymbol{X}_i^b = \mathbb{E}(\boldsymbol{X}_i|\boldsymbol{Y}_{1:i-1})$ et $\boldsymbol{X}_i^a = \mathbb{E}(\boldsymbol{X}_i|\boldsymbol{Y}_{1:i})$, où $\mathbb{E}$ désigne l'espérance mathématique. Ainsi définis, $\boldsymbol{X}_i^b$ et $\boldsymbol{X}_i^a$ sont les estimateurs optimaux (estimateurs qui réduisent la variance d'erreur, van Leeuwen (2003)).

Il est possible d'appliquer ce formalisme du filtrage non linéaire en discrétisant les différentes distributions de probabilité (méthode de Monte-Carlo). Il s'agit alors de l'approche particulaire aussi appelé filtre particulaire (Gordon et al., 1993; Del Moral, 2004). Ce formalisme se trouve être simplifié dans le cas de distribution gaussienne et d'évolution linéaire : C'est le filtre de Kalman (Kalman, 1960).

### 2.1.2 Filtre de Kalman

Dans la suite, les erreurs sont supposées suivre des lois gaussiennes centrées telles que $\forall i\, \varepsilon_i^o \sim \mathcal{N}(0, \boldsymbol{R}_i)$, $\varepsilon_i^m \sim \mathcal{N}(0, \boldsymbol{Q}_i)$, et $\varepsilon_i^b \sim \mathcal{N}(0, \boldsymbol{B}_i)$ où la notation $\mathcal{N}(\boldsymbol{m}, \boldsymbol{C})$ correspond à une loi gaussienne d'espérance $\boldsymbol{m}$ et de matrice de covariance $\boldsymbol{C}$. Pour simplifier le problème d'estimation, le modèle ainsi que l'opérateur d'observation sont supposés linéaires, $\mathcal{M}_i = \boldsymbol{M}_i$ et $\mathcal{H}_i = \boldsymbol{H}_i$.

Ces hypothèses ont alors deux rôles : la linéarité permet de conserver le caractère Gaussien, et la Gaussiannité permet de produire des expressions analytiques pour la solution. A partir de l'équation (2.1), il est possible de définir une fonction de log-vraisemblance (ou fonction coût) $J(\boldsymbol{x}_i) = -log(p(\boldsymbol{x}_i|\boldsymbol{y}_{1:i}))$. Dans le cas gaussien, cette fonction de log-vraisemblance s'écrit (à une constante près)

$$J(\boldsymbol{x}_i) = J^b(\boldsymbol{x}) + J^o(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}^b\|_{\boldsymbol{B}_i^{-1}} + \frac{1}{2}\|\boldsymbol{y}^o - \boldsymbol{H_i x}\|_{\boldsymbol{R}_i^{-1}} \tag{2.2}$$

où $J_b$ est lié à l'ébauche ($p(\boldsymbol{x}_i|\boldsymbol{y}_{1:i-1})$) et $J_o$ est lié aux observations ($p(\boldsymbol{y}_i|\boldsymbol{x}_i)$). Cette fonctionnelle est quadratique en $\boldsymbol{x}$, avec $\boldsymbol{B}$ et $\boldsymbol{R}$ définies positives, ce qui implique qu'elle est convexe et admet donc un unique minimum. En se plaçant au point $\boldsymbol{x}^*$ qui annule le gradient avec $\nabla J = \boldsymbol{B}^{-1}(\boldsymbol{x}^* - \boldsymbol{x}^b) - \boldsymbol{H}^T\boldsymbol{R}^{-1}(\boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x}^*)$, il vient que l'optimum vérifie $\boldsymbol{x}^* = \boldsymbol{x}^b + (\boldsymbol{B}^{-1} + \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{R}^{-1}(\boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x}^b) = \boldsymbol{x}^b + \boldsymbol{K}\boldsymbol{d}$. L'analyse optimale est alors le minimum de la fonction coût $J$ (2.2) et correspond à une combinaison linéaire de l'ébauche et des observations $\boldsymbol{x}_i^a = \boldsymbol{x}_i^b + \boldsymbol{K}_i\boldsymbol{d}_i$, où $\boldsymbol{d}_i = \boldsymbol{y}_i - \boldsymbol{H}\boldsymbol{x}_i^b$ est appelé vecteur d'innovation et où $\boldsymbol{K}_i = \boldsymbol{B}_i\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{B}_i\boldsymbol{H}^T + \boldsymbol{R})^{-1}$ est la matrice de gain. Le formalisme du meilleur estimateur linéaire non biaisé (*BLUE*) montre également que $\boldsymbol{K}_i$ minimise la variance d'erreur d'analyse, conduisant à une solution optimale. Ce qui donne pour la matrice de covariance d'erreur d'analyse,

$$\boldsymbol{A}_i = (\boldsymbol{I} - \boldsymbol{K}_i\boldsymbol{H})\,\boldsymbol{B}_i. \tag{2.3}$$

1. Formule de Bayes : $\Pr(A/B) = \frac{\Pr(B/A)\Pr(A)}{\Pr(B)}$

Dans les conditions décrites précédemment la dynamique linéaire conserve le caractère gaussien des distributions, il suffit donc de connaître l'espérance et la matrice de covariance d'erreur d'analyse pour estimer la distribution dans le futur. L'évolution temporelle de l'ébauche s'écrit alors $\boldsymbol{x_i^b} = \boldsymbol{M}_i \boldsymbol{x}_{i-1}^a$, quant à l'erreur elle s'exprime alors $\varepsilon_i^b = \boldsymbol{M}_i \varepsilon_{i-1}^a + \varepsilon_i^m$. Ainsi, la matrice de covariance d'ébauche $\boldsymbol{B}_i = \mathbb{E}(\varepsilon_i^b \varepsilon_i^{bT})$ s'écrit $\boldsymbol{B}_i = \boldsymbol{M}_i \boldsymbol{A}_{i-1} \boldsymbol{M}_i^T + \boldsymbol{Q}_i$, sous l'hypothèse $\mathbb{E}(\varepsilon^a \varepsilon^{mT}) = 0$. Le premier terme $\boldsymbol{M}_i \boldsymbol{A}_{i-1} \boldsymbol{M}_i^T$ correspond au transport linéaire de l'incertitude, et le terme $\boldsymbol{Q}_i$ est associé à l'erreur modèle. Les calculs et le cyclage sont résumés ci-dessous.

**Équation du filtre de Kalman**

Finalement les équations du filtre de Kalman se résument à

1. Correction

$$\begin{aligned} \boldsymbol{x}_i^a &= \boldsymbol{x}_i^b + \boldsymbol{K}_i(\boldsymbol{y}_i^o - \boldsymbol{H}_i \boldsymbol{x}_i^b) \\ \boldsymbol{K}_i &= \boldsymbol{B}_i \boldsymbol{H}_i^T (\boldsymbol{H}_i \boldsymbol{B}_i \boldsymbol{H}_i^T + \boldsymbol{R}_i)^{-1} \\ \boldsymbol{A}_i &= (\boldsymbol{I}_i - \boldsymbol{K}_i \boldsymbol{H}_i) \boldsymbol{B}_i \end{aligned}$$

2. Prévision

$$\begin{aligned} \boldsymbol{x}_{i+1}^b &= \boldsymbol{M}_{i+1} \boldsymbol{x}_i^a \\ \boldsymbol{B}_{i+1} &= \boldsymbol{M}_{i+1} \boldsymbol{A}_i \boldsymbol{M}_{i+1}^T + \boldsymbol{Q}_{i+1} \end{aligned}$$

Une illustration du processus d'assimilation est présentée ci-dessous dans un cas simple. Cela permet de comprendre le rôle des différents composants d'un système d'assimilation.

### 2.1.3 Exemple d'une assimilation à une observation

L'exemple de l'impact d'une observation unique permet de mieux appréhender le rôle joué par la matrice de covariance $\boldsymbol{B}$. Dans un domaine 2D, on considère une observation en position centrale. L'ébauche est choisie nulle, l'observation a une valeur 1 et la matrice $\boldsymbol{B}$ est supposée homogène (corrélations identiques sur tout le domaine) de fonction de covariance $\rho(\boldsymbol{x}) = \exp(-\frac{x^2}{2L_{px}^2} - \frac{y^2}{2L_{py}^2})$. Dans ce cas, l'analyse correspond à l'incrément d'analyse $\delta \boldsymbol{x}^a = \boldsymbol{x}^a - \boldsymbol{x}^b = \boldsymbol{K}\boldsymbol{d}$ dont l'amplitude s'exprime comme le rapport des variances d'erreur d'observation et d'ébauche $\delta = \sigma^{b2}/(\sigma^{b2} + \sigma^{o2})$. La forme de l'incrément est donnée par les longueurs de portée selon $x$ et $y$, notées $L_{px}$ et $L_{py}$, qui caractérise la zone d'influence spatiale d'une observation.

La figure 2.1 représente l'incrément en 2D sur le domaine et son profil selon $x$. Entre les différents cas, les longueurs de portée ou la variance sont modifiées. Dans les cas 1 et 2, pour des variances égales $\sigma^{b2} = \sigma^{o2}$, l'amplitude théorique de l'incrément est égale à $\delta = 1/2$. Ce qui se retrouve sur les figures. Pour le cas 3, $\sigma^{b2} = 9\sigma^{o2}$ ce qui donne comme amplitude $\delta = 9/10$. Augmenter la variance de l'erreur d'ébauche revient à accorder plus d'importance à l'observation et l'analyse s'en rapproche plus nettement. En augmentant les longueurs de portée, on élargit la structure de l'incrément d'analyse. Dans le cas où un nombre plus important d'observations est utilisé, les longueurs de portée permettent de lisser le champ analysé.

Cette illustration permet de mieux comprendre le rôle de la matrice $\boldsymbol{B}$. La matrice de covariance d'erreur d'ébauche $\boldsymbol{B}$ caractérise la dispersion des erreurs d'ébauche $\varepsilon^b$. Les valeurs sur sa diagonale sont les variances d'erreur d'ébauche pour chaque variable du vecteur d'état. Des forts écart-types indiquent une zone de forte incertitude dans la prévision, des faibles écart-types indiquent une zone de confiance en l'ébauche. Ce raisonnement s'applique aussi pour $\boldsymbol{A}$. La relation $\boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{B}$ identifie explicitement le lien entre les matrices $\boldsymbol{A}$ et $\boldsymbol{B}$ et le rôle des observations. Les observations réduisent la variance dans les régions où les mesures sont effectuées, ce qui revient à réduire l'incertitude ($-\boldsymbol{K}\boldsymbol{H}\boldsymbol{B}$).

Un incrément d'analyse s'écrit $\delta \boldsymbol{x}^a = \boldsymbol{x}^a - \boldsymbol{x}^b = \boldsymbol{K}\boldsymbol{d}$ avec $\boldsymbol{K} = \boldsymbol{B}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1}$. Il est donc directement proportionnel aux colonnes de $\boldsymbol{B}$, c'est-à-dire aux fonctions de corrélation. La matrice de covariance $\boldsymbol{B}$ ne renseigne donc pas seulement sur les zones d'incertitude de modèle mais elle conditionne aussi la forme des incréments d'analyse.

### 2.1.4 Rôle de la matrice de covariance d'erreur d'ébauche

Comme illustré dans l'exemple d'assimilation avec une seule observation, la matrice **B** joue un rôle essentiel dans la construction de l'incrément d'analyse puisqu'elle contribue au filtrage de l'erreur d'observation et que c'est elle qui propage dans l'ébauche l'information issue de l'innovation (distance initiale entre ébauche et observations).

#### Caractérisation et propriétés des matrices de covariance

La matrice de covariance d'erreur de prévision est caractérisée par différentes propriétés. Tout d'abord, il s'agit d'une matrice de covariance multivariée qui contient l'information statistique reliant les différentes variables thermodynamiques entre elles. En particulier, la matrice de covariance rend compte des équilibres tels que le géostrophisme de grande échelle qui relie le vent moyen et le géopotentiel.

Ensuite, pour une variable thermodynamique donnée, la matrice de covariance (mono-variée) est caractérisée par le champ de variance (diagonale de la matrice) et par l'extension spatiale des fonctions de corrélation $\rho(\mathbf{x}, \mathbf{y})$ souvent diagnostiquées par la longueur de portée moyenne $L_p$ définie, en 2D, d'après Daley (1991) comme

$$L_p^2 = -\frac{2\rho(0)}{\nabla^2 \rho(0)} \tag{2.4}$$

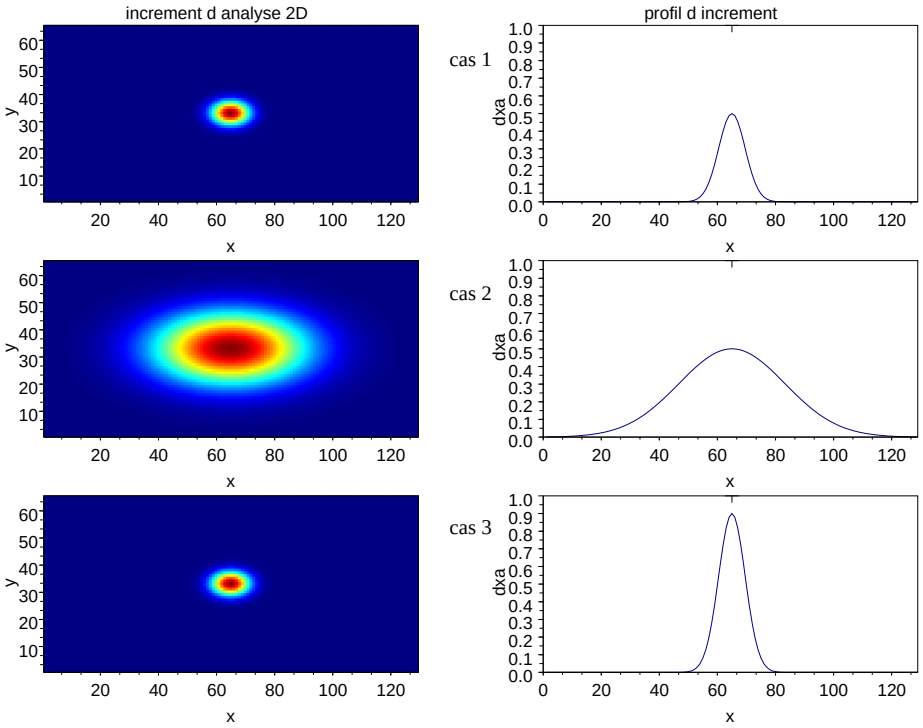



FIGURE 2.1 – Incrément d'analyse 2D, à gauche et profil selon $x$, à droite. La croix représente l'observation de valeur 1. Chaque ligne correspond à un cas d'assimilation. Cas 1 : $L_{px} = 550km$ , $L_{py} = 400km$, $\sigma^{b2} = \sigma^{o2}$, $\delta = 0.5$. Cas 2 : $L_{px} = 2200km$, $L_{py} = 1100km$, $\sigma^{b2} = \sigma^{o2}$, $\delta = 0.5$. Cas 3 : $L_{px} = 550km$, $L_{py} = 400km$, $\sigma^{b2} = 9\sigma^{o2}$, $\delta = 0.9$.

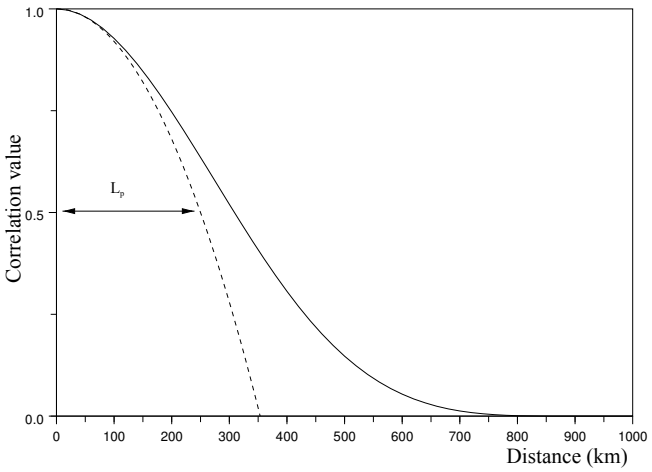



FIGURE 2.2 – Illustration 1D de la longueur de portée (extrait de Pannekoucke et al. (2008) avec accord des auteurs). La fonction de corrélation $\rho$ est représentée en trait continu, la parabole de longueur de portée $L_p$ est en traits pointillés.

Une manière de comprendre ce que représente la longueur de portée est d'introduire le développement de Taylor à l'ordre 2 de la fonction de corrélation $\rho(x, x+\delta x) = \rho(\delta x)$, dans la limite des petits $\delta x$ (Pannekoucke et al., 2008),

$$\rho(\delta x) \approx \rho(0) + \frac{\delta x^2}{2}\frac{d^2\rho}{dx^2}(0) = 1 - \frac{\delta x^2}{L_p^2}. \tag{2.5}$$

Dans un espace à 2 dimensions, il existe donc une longueur de portée selon $x$, $L_{p_x}$, et une longueur de portée selon $y$, $L_{p_y}$. La longueur de portée peut être interprétée comme la distance à l'origine où la fonction de corrélation vaut $\rho(L_p) = 0.5$ si $\rho$ est une parabole, comme cela est illustré sur la figure 2.2.

Dans le cas multidimensionnel, le développement de Taylor

$$\rho(\mathbf{x}, \mathbf{x}+\delta\mathbf{x}) = 1 - \frac{1}{2}|\delta\mathbf{x}|^2_{\boldsymbol{g}(\mathbf{x})}, \tag{2.6}$$

fait intervenir un tenseur de métrique locale, $\boldsymbol{g}(\mathbf{x})$, défini à partir de la fonction de covariance $\rho(\boldsymbol{x})$ comme $(g_{\boldsymbol{x}})_{ij} = -\partial^2_{ij}\rho$, où $i, j$ sont les indices de la matrice en 2D selon les axes $x, y$.

Enfin, un autre caractère important pour la matrice de covariance d'erreur de prévision en météorologie est que les fonctions de corrélation sont non-séparables, (Courtier et al., 1998) au sens où la propagation des corrections suivant la verticale dépend de l'extension spatiale de la correction horizontale (comportement *shallow water* pour les corrections horizontales de grande échelle, et *deep water* pour les corrections horizontales de petite échelle).

### Homogénéité et hétérogénéité des fonctions de corrélation de B

Les fonctions de corrélation sont appelées homogènes quand le champ de fonction de corrélation est invariant en translation et hétérogène dans le cas contraire. De plus, les fonctions de corrélation sont dites isotropes si la corrélation entre deux points ne dépend que de la distance séparant ces deux positions. Pour le plan, on peut rencontrer des fonctions homogènes et anisotropes. Pour la sphère, il y a équivalence entre homogène et isotrope (Gaspari and Cohn, 1999). À nouveau, le diagnostic des longueurs de portée, ou des tenseurs locaux, permet de proposer une description synthétique de la forme des fonctions de corrélation et renseigne donc sur le caractère homogène ou hétérogène, ainsi que sur l'anisotropie de ces fonctions.

### Éléments de modélisation de la matrice B

Dans les formulations variationnelles de l'assimilation de données, la modélisation des fonctions de covariance se présente souvent sous la forme d'un produit d'opérateurs linéaires permettant de construire

une matrice multivariée (représentation des interactions statistiques entre les erreurs affectant des variables modèles de nature différente) et non-séparable (Courtier et al., 1998).

Pour permettre de représenter des fonctions de corrélation non triviales, différentes stratégies ont été développées, en voici les principales variantes que l'on trouve dans la littérature :

- l'hypothèse diagonale qui consiste à supposer connue les directions principales de la matrice de corrélation, telle que la formulation diagonale spectrale (Courtier et al., 1998) (homogène) ou diagonale ondelette (Fisher, 2003) (hétérogène) ;
- l'opérateur de diffusion qui permet de construire des fonctions de corrélation hétérogènes et anisotropes (Weaver and Courtier, 2001; Pannekoucke and Massart, 2008; Weaver and Mirouze, 2013) ;
- l'utilisation de filtres récursifs (Purser et al., 2003a,b) dont la mise en œuvre comporte des similitudes avec l'opérateur de diffusion (Mirouze and Weaver, 2010)

### 2.1.5 Limitation de l'utilisation du filtre de Kalman

L'approche linéaire du filtre de Kalman permet une écriture simple du processus d'assimilation. Cependant sa mise en place nécessite de connaître et d'inverser des matrices de la dimension du vecteur d'état, ce qui n'est pas toujours possible dès que sa dimension devient très grande. De plus le formalisme du filtre de Kalman ne peut prendre en compte que des observations contemporaines du même instant que l'ébauche, ce qui nécessite de réaliser plusieurs assimilations si l'on considère une séries d'observations réparties dans le temps (fenêtre temporelle) pour l'analyse. Une autre difficulté rédhibitoire pour les système de grande taille, est le coût de la propagation temporelle de la matrice de covariance d'erreur d'analyse. Deux types de méthodes permettent de prendre en compte ces défauts, les méthodes variationnelles et les méthodes ensemblistes. Les principales méthodes utilisées dans les centres de prévision numérique sont présentées dans les sections suivantes.

## 2.2 Méthodes variationnelles

En météorologie opérationnelle la dimension du vecteur d'état est de l'ordre de $\mathcal{O}(10^8)$, ce qui implique des matrices de taille $\mathcal{O}(10^{16})$, impossibles à représenter entièrement en mémoire. Il est donc impossible d'inverser directement ces matrices pour résoudre l'équation du *BLUE*. Une stratégie est de construire la solution de manière itérative sous la forme d'un problème de minimisation de la fonction coût (2.2), en utilisant un minimiseur (comme le gradient conjugué ou l'algorithme de Newton).

### 2.2.1 3D-Var

Cette fonction coût mesure les distances de l'estimateur à l'ébauche et aux observations. Ainsi $\boldsymbol{x}^a$ est construit comme le minimum de la fonction coût 2.2. Cette méthode est appelée 3D variationnel (3D-Var). Une limitation du 3D-Var est que la dynamique n'est pas prise en compte, en particulier si $\boldsymbol{B}$ est modélisée. Toutes les observations sont supposées être valables en un même instant. Pour éliminer ce défaut une extension temporelle dite 4D-Var peut être introduite, comme expliqué ci-dessous.

### 2.2.2 Extension temporelle : 4D-Inc

Le formalisme du filtrage peut s'étendre au cas d'une trajectoire pour former un lisseur. Dans ce cadre, la fonctionnelle 4D-Var s'écrit alors $J(\boldsymbol{x}_0) = J^b(\boldsymbol{x}_0) + \sum J^o(\boldsymbol{x}_i)$, où $\boldsymbol{x}_i$ correspond à l'état, de trajectoire d'état initiale $\boldsymbol{x}_0$, à l'instant $i$. Dans l'esprit, chaque observation est comparée à l'état du modèle à l'instant où elle est mesurée. Les observations utilisées sont comprises dans une fenêtre temporelle d'assimilation. En pratique, une assimilation par lot est réalisée. On regroupe les observations contemporaines de petites périodes (redécoupage de la fenêtre en tranches ou *time-slots*), autour d'un seul instant par période. L'analyse est alors choisie comme l'état à l'instant du début de la fenêtre d'assimilation dont la trajectoire est optimale par rapport à la trajectoire de l'ébauche initiale et aux observations. Une illustration du schéma de principe du 4D-Var est donnée dans la figure 2.3

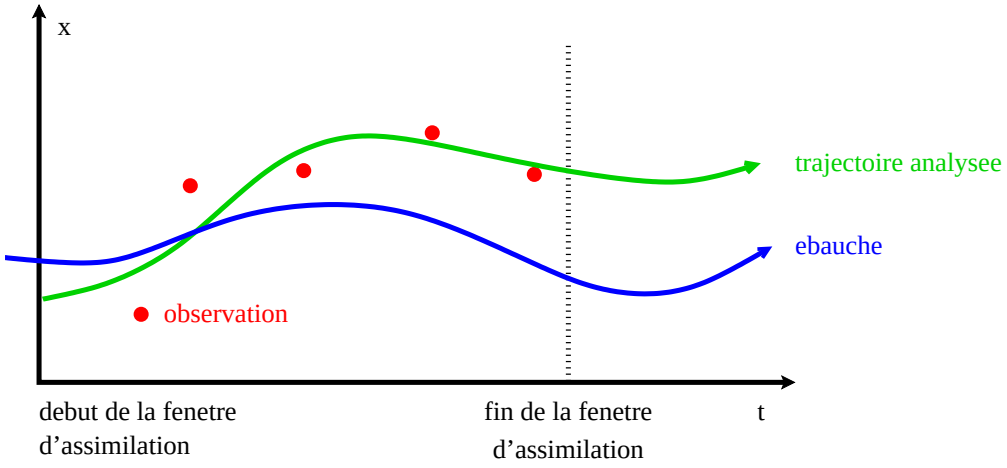



FIGURE 2.3 – Schéma de lissage du 4D-Var. La trajectoire initiale est en bleu (ébauche), les observations sont représentées par les points rouges et la trajectoire analysée est en vert.

La méthode présentée dans cette section est le 4D-Var incrémental (Courtier et al., 1994), c'est sous cette forme que l'assimilation de données est implantée dans le modèle ARPÈGE de Météo-France et au Centre Européen de Prévision Météorologique à Moyen Terme (CEPMMT ou ECMWF [2]). En supposant que l'ébauche est peu éloignée de l'analyse, on considère des petites perturbations $\delta\boldsymbol{x}$ autour de $\boldsymbol{x}^b$, avec $\delta\boldsymbol{x} \ll \boldsymbol{x}^b$. A l'aide d'un développement de Taylor à l'ordre 1, on peut alors écrire, $\mathcal{M}(\boldsymbol{x}^b + \delta\boldsymbol{x}) = \mathcal{M}(\boldsymbol{x}^b) + \boldsymbol{M}_{\boldsymbol{x}^b}\delta\boldsymbol{x} + \mathcal{O}(\|\delta\boldsymbol{x}\|)$. Le modèle linéaire tangent, noté $\boldsymbol{M}$, est dérivé du modèle non linéaire $\mathcal{M}$. Il décrit l'évolution de petites perturbations le long de la trajectoire non linéaire de $\boldsymbol{x}^b$. Le modèle adjoint, noté $\boldsymbol{M}^T$, est dérivé à partir des équations discrétisées du modèle tangent (Errico, 1997). Il permet d'exprimer le gradient à l'instant initial en fonction du gradient à l'instant $i$, $\nabla_{\boldsymbol{x}_0} J = \boldsymbol{M}^T_{0\to i}\nabla_{\boldsymbol{x}_i} J$. Le modèle adjoint rétro-propage l'information. Ces deux modèles sont définis autour de la trajectoire associée à l'ébauche. Dans le cas où les perturbations sont de trop grande amplitude, le domaine d'application n'est plus valable puisque les non linéarités ne peuvent plus être négligées. L'utilisation du modèle tangent et de son adjoint permet de transporter les incréments et de les comparer aux innovations. Une intégration par ces modèles n'a de sens que sur quelques jours à quelques heures (en fonction de la non linéarité des processus représentés par le modèle), ce qui limite la taille des fenêtres d'assimilation. A titre d'exemple le modèle IFS du CEPMMT utilise des fenêtres de 12h, ARPEGE 6h et AROME 1h.

La formulation incrémentale permet de prendre en compte les non-linéarités du modèle par une mise à jour de la trajectoire utilisée lors de la minimisation. Lorsqu'un premier minimum $\mathbf{x}^1$ est obtenu, la nouvelle trajectoire de référence issue de ce minimum est calculée. Une nouvelle minimisation en utilisant cette trajectoire peut être effectuée, ce qui permet de construire une série d'états $\mathbf{x}^n$ à partir de l'ébauche $\mathbf{x}^b$. Le calcul d'une nouvelle trajectoire est appelé *boucle externe*, alors que les boucles de minimisation sont dites *boucles internes*. Le vecteur d'innovation $\boldsymbol{d}_i^n$ est calculé avec la forme non linéaires des opérateurs. L'expression de la fonction coût en formulation incrémentale est donnée par

$$J(\delta\boldsymbol{x}^n) = \frac{1}{2}(\delta\boldsymbol{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b)^T\boldsymbol{B}^{-1}(\delta\boldsymbol{x}^n + \mathbf{x}^{n-1} - \mathbf{x}^b) + \frac{1}{2}\sum_{i=0}^{T_f}(\boldsymbol{d}_i^{n-1} - \boldsymbol{H}_i\boldsymbol{M}_{0\to i}\delta\boldsymbol{x})^T\boldsymbol{R}^{-1}(\boldsymbol{d}_i^{n-1} - \boldsymbol{H}_i\boldsymbol{M}_{0\to i}\delta\boldsymbol{x}), \tag{2.7}$$

avec $\mathbf{x}^n = \mathbf{x}^{n-1} + \delta\mathbf{x}^n$ ($\mathbf{x}^0 = \mathbf{x}^b$ et $\mathbf{x}^{-1} = \mathbf{x}^b$) et $\boldsymbol{d}_i^n = \mathcal{H}(\mathcal{M}_{0\to i}(\mathbf{x}^n))$.

Lors de la minimisation, en utilisant un algorithme du gradient conjugué, il est possible d'obtenir une estimation de la matrice $\boldsymbol{A}$ à partir de l'algorithme de Lanczos (plus de détails sont données en annexe B). Cette approximation de la matrice de covariance d'analyse est utilisée lorsque l'on souhaite mettre en place les méthodes de sensibilité aux observations (voir 4).

2. European Centre for Medium-Range Weather Forecasts

Un des inconvénients de la méthode variationnelle présentée ici est l'absence de calcul explicite de la matrice de covariance $\boldsymbol{B}$. Dans la formulation classique du 4D-Var, les corrélations sont modélisées et les valeurs de variance ont un caractère climatologique. Or, comme nous l'avons vu au paragraphe 2.1.3, la matrice $\boldsymbol{B}$ détermine la forme des incréments d'analyse. Pour pallier à ce défaut, des méthodes d'assimilation basées sur des ensembles de prévisions issues d'ensemble d'analyses perturbées ont été développées (Berre et al., 2007). L'inspiration de cette approche provient de l'essor des méthodes d'ensemble présentées ci-dessous.

## 2.3 Méthodes ensemblistes

Les méthodes ensemblistes en prévision numérique sont basées sur la méthode de Monte Carlo (Epstein, 1969). Cette approche a connu une forte progression dans la communauté de l'assimilation de donnée depuis les travaux de Evensen (1994).

### 2.3.1 Filtre de Kalman d'ensemble

Le filtre de Kalman d'ensemble (EnKF, Evensen (2003)) est une adaptation du filtre de Kalman permettant de contourner deux principales limitations dans son application à des systèmes de grande dimension, la propagation des matrices de covariance et la prise en compte des non linéarités.

L'esprit de cette méthode est de remplacer la propagation des incertitudes décrite dans la phase de prévision du filtre de Kalman (très coûteuse), par la propagation temporelle d'un échantillon de $N_e$ membres, à partir duquel la matrice de covariance d'erreur de prévision est estimée.

Formellement, un ensemble de conditions initiales est construit pour produire les ébauches $\boldsymbol{x}_k^b(t)$. L'ensemble est supposé non biaisé, c'est à dire que $\mathbb{E}(\boldsymbol{x}^b) = \boldsymbol{x}^t$. Les erreurs d'ébauche sont donc calculées à partir de l'ensemble $\boldsymbol{\varepsilon}_k^b = \boldsymbol{x}_k^b - \overline{\boldsymbol{x}^b}$, où $\overline{\boldsymbol{x}^b}$ est la moyenne de l'ensemble. On obtient alors,

$$\begin{aligned} \boldsymbol{B} &= \mathbb{E}(\boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^{b^T}) \\ &= \frac{1}{N_e - 1} \sum_{k=1}^{N_e} \boldsymbol{\varepsilon}_k^b \boldsymbol{\varepsilon}_k^{b^T} \end{aligned} \tag{2.8}$$

Les équations de l'étape de correction du filtre de Kalman restent valables pour les modèles linéaires gaussiens. L'EnKF permet donc de contourner le calcul de $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{A}\boldsymbol{M}^T$, et évite donc les très nombreux appels aux modèles tangent-linéaire et adjoint, mais il ne résout pas le problème de l'inversion de la matrice de grande dimension ($\boldsymbol{K} = \boldsymbol{B}\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1}$).

### 2.3.2 Ensemble reduced rank square root Kalman filter

Pour réduire la dimension de la matrice à inverser, un changement de variable est réalisé, $\boldsymbol{x}^b = \overline{\boldsymbol{x}^b} + \boldsymbol{X}^b\boldsymbol{w}$, où $\boldsymbol{X}^b$ est tel que $\boldsymbol{B} = \boldsymbol{X}^b\boldsymbol{X}^{b^T}$. $\boldsymbol{X}^b$ est la matrice des perturbations de l'ensemble, de dimension $n \times N_e$. Les colonnes de $\boldsymbol{X}^b$ sont les vecteurs $\boldsymbol{x}_i^b$ normés par $1/\sqrt{N_e - 1}$.

$$\boldsymbol{X}^b = \frac{1}{\sqrt{N_e - 1}} \left( \boldsymbol{x}_1^b - \overline{\boldsymbol{x}^b} \quad \boldsymbol{x}_2^b - \overline{\boldsymbol{x}^b} \quad \cdots \quad \boldsymbol{x}_{N_e}^b - \overline{\boldsymbol{x}^b} \right) \tag{2.9}$$

En prenant $\overline{\boldsymbol{x}^b}$ comme ébauche et en remplaçant dans l'équation (2.7) on obtient,

$$\begin{aligned} \widetilde{J} &= \frac{1}{2}(\overline{\boldsymbol{x}^b} - \boldsymbol{x})^T \boldsymbol{B}^{-1}(\overline{\boldsymbol{x}^b} - \boldsymbol{x}) + \frac{1}{2}(\boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x})^T \boldsymbol{R}^{-1}(\boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x}) \\ &= \frac{1}{2}\boldsymbol{w}^T \boldsymbol{X}^{b^T} \boldsymbol{B}^{-1} \boldsymbol{X}^b \boldsymbol{w} + \frac{1}{2}(\boldsymbol{d} - \boldsymbol{H}\boldsymbol{X}^b\boldsymbol{w})^T \boldsymbol{R}^{-1}(\boldsymbol{d} - \boldsymbol{H}\boldsymbol{X}^b\boldsymbol{w}) \\ &= \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \frac{1}{2}(\boldsymbol{d} - \boldsymbol{Y}^b\boldsymbol{w})^T \boldsymbol{R}^{-1}(\boldsymbol{d} - \boldsymbol{Y}^b\boldsymbol{w}) \end{aligned} \tag{2.10}$$

avec $\boldsymbol{d} = \boldsymbol{y}^o - \boldsymbol{H}\overline{\boldsymbol{x}^b}$ et $\boldsymbol{Y}^b = \boldsymbol{H}\boldsymbol{X}^b$. La Héssienne devient $\widetilde{J}'' = \boldsymbol{I} + {\boldsymbol{Y}^b}^T\boldsymbol{R}^{-1}\boldsymbol{Y}^b$, de dimension $N_e \times N_e$, et peut alors être diagonalisée et inversée facilement.

$$\begin{aligned} \widetilde{\boldsymbol{A}}^{-1} &= \boldsymbol{I} + {\boldsymbol{Y}^b}^T\boldsymbol{R}^{-1}\boldsymbol{Y}^b \\ &= \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T \\ \text{d'où } \widetilde{\boldsymbol{A}} &= \boldsymbol{Q}\boldsymbol{D}^{-1}\boldsymbol{Q}^T \end{aligned}$$

A partir de $\widetilde{\boldsymbol{A}}$, l'inverse de la Héssienne dans l'espace réduit, la matrice de gain est obtenue avec $\widetilde{\boldsymbol{K}} = \widetilde{\boldsymbol{A}}{\boldsymbol{Y}^b}^T\boldsymbol{R}^{-1}$. Le retour dans l'espace physique se fait en multipliant par $\boldsymbol{X}^b$. Ceci permet de calculer l'espérance de l'analyse, qui correspond par construction à la moyenne de l'ensemble des membres analysés,

$$\overline{\boldsymbol{x}^a} = \overline{\boldsymbol{x}^b} + \boldsymbol{X}^b\widetilde{K}\boldsymbol{d} \tag{2.11}$$

Pour construire l'ensemble d'analyses, on reprend la formule du BLUE avec $\boldsymbol{x}_i^b = \overline{\boldsymbol{x}^b} + \boldsymbol{\varepsilon}_i^b$,

$$\begin{aligned} \boldsymbol{x}_i^a &= \boldsymbol{x}_i^b + \boldsymbol{X}^b\widetilde{K}(\boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x}_i^b) \\ &= \overline{\boldsymbol{x}^b} + \varepsilon_i^b + \boldsymbol{X}^b\widetilde{\boldsymbol{K}}(\boldsymbol{d} - \boldsymbol{H}\varepsilon_i^b) \\ &= \overline{\boldsymbol{x}^a} + (\boldsymbol{I} - \boldsymbol{X}^b\widetilde{\boldsymbol{K}}\boldsymbol{H})\varepsilon_i^b \end{aligned} \tag{2.12}$$

Cet algorithme correspond à l'*Ensemble Transform Kalman Filter* (ETKF) tel que proposé par Bishop et al. (2001) et repris dans Hunt et al. (2006).

**Prise en compte de la distribution temporelle des observations**

Dans la formulation de l'équation (2.10), toutes les observations sont supposées avoir lieu à l'instant d'assimilation comme pour un 3D-Var. En augmentant la dimension du vecteur de contrôle, il est possible de prendre en compte les observations sur une fenêtre d'assimilation. On définit $\underline{\boldsymbol{y}^{oT}} = \begin{pmatrix} {\boldsymbol{y}_0^o}^T & {\boldsymbol{y}_1^o}^T & \cdots & {\boldsymbol{y}_K^o}^T \end{pmatrix}$ où $\boldsymbol{y}_t^o$ contient les observations de la tranche $t$ de la fenêtre d'assimilation, pour une fenêtre d'assimilation découpée en $K+1$ tranches temporelles. Cette extension temporelle est réalisée pour l'ensemble des composants de la fonction score,

$$\underline{\boldsymbol{H}} = \begin{pmatrix} \boldsymbol{H}_0 & & & \\ & \boldsymbol{H}_1 & & \\ & & \ddots & \\ & & & \boldsymbol{H}_K \end{pmatrix}, \underline{\boldsymbol{R}} = \begin{pmatrix} \boldsymbol{R}_0 & & & \\ & \boldsymbol{R}_1 & & \\ & & \ddots & \\ & & & \boldsymbol{R}_K \end{pmatrix},$$

$$\underline{\boldsymbol{X}^b} = \begin{pmatrix} \boldsymbol{X}_0^b & & & \\ & \boldsymbol{X}_1^b & & \\ & & \ddots & \\ & & & \boldsymbol{X}_K^b \end{pmatrix} \text{ avec } \boldsymbol{X}_i^b = \begin{pmatrix} \mathcal{M}_i(\boldsymbol{x}_1^b) & \mathcal{M}_i(\boldsymbol{x}_2^b) & \cdots & \mathcal{M}_i(\boldsymbol{x}_K^b) \end{pmatrix}$$

$$\text{et } \underline{\boldsymbol{d}} = \begin{pmatrix} \boldsymbol{d}_0 \\ \boldsymbol{d}_1 \\ \vdots \\ \boldsymbol{d}_K \end{pmatrix} \text{ avec } \boldsymbol{d}_i = \boldsymbol{H}_i\mathcal{M}_i(\boldsymbol{x}^b).$$

La fonction coût s'écrit alors

$$\widetilde{J}(\boldsymbol{w}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \frac{1}{2}\left(\underline{\boldsymbol{d}} - \underline{\boldsymbol{Y}^b}\boldsymbol{w}\right)^T\underline{\boldsymbol{R}}\left(\underline{\boldsymbol{d}} - \underline{\boldsymbol{Y}^b}\boldsymbol{w}\right). \tag{2.13}$$

avec $\underline{\boldsymbol{Y}^b} = \underline{\boldsymbol{H}\boldsymbol{X}^b}$. Dans cette formulation les observations sont directement comparées aux prévisions à l'instant où elles ont lieu. La propagation temporelle est traitée implicitement, sans nécessité d'un modèle linaire tangent ou d'un modèle adjoint.

## 2.4 Méthodes hybrides

L'écriture de la fonction coût avec l'extension temporelle (2.13) peut aussi être utilisée dans le cadre d'une minimisation variationnelle (Desroziers et al., 2014), en utilisant le même changement de variable $\underline{\boldsymbol{x}} = \overline{\underline{\boldsymbol{x}}^b} + \underline{\boldsymbol{X}}^b\boldsymbol{w}$ que pour l'ETKF. On retrouve alors les équations du 4D-En-Var, qui reprend la fonction coût de l'ETKF mais en calculant l'analyse via une minimisation incrémentale. Différents algorithmes de minimisation pour le 4D-En-Var peuvent être trouvés dans Desroziers et al. (2014).

Le 4D-En-Var retrouve les avantages et les inconvénients du 4D-Var et de la transformée d'ensemble, en particulier la matrice $\underline{\boldsymbol{B}}$ issue de l'ensemble est déficiente en rang. Le 4D-En-Var est donc d'avantage utilisée sous une forme dite hybride, dont la matrice de covariance d'erreur d'ébauche devient

$$\underline{\boldsymbol{B}}^h = \alpha^{c2}\underline{\boldsymbol{B}}^c + \alpha^{e2}\underline{\boldsymbol{B}}^e, \tag{2.14}$$

avec $\alpha^{c2} + \alpha^{e2} = 1$, où $\underline{\boldsymbol{B}}^c$ correspond à la matrice d'erreur d'ébauche issue de la climatologie (modélisée et de plein rang) et $\underline{\boldsymbol{B}}^e$ est la matrice de covariance issue de l'ensemble.

## 2.5 Les limites de l'utilisation des ensembles

Les paragraphes précédents ont montrés quelles informations complémentaires l'utilisation d'un ensemble d'assimilation peut apporter. Cependant la construction d'un ensemble d'assimilation ou de prévisions cohérent avec la réalité de l'atmosphère n'est pas trivial. La qualité d'un ensemble d'assimilation est directement liée à la fiabilité de l'ensemble lui même, il faut donc répondre à certains critères.

Comme chaque membre nécessite le calcul d'une trajectoire non linéaire complète, le nombre de membres par ensemble est limité par le coût de calcul. L'ordre de grandeur de la taille d'un ensemble en prévision numérique est entre 5 et 200 membres, ce qui est très faible devant le nombre de degrés de liberté du système ($\mathcal{O}(10^8)$). De fait, le rang des matrices estimées est largement déficient. Pour tirer le meilleur profit de l'ensemble plusieurs approches sont utilisées lors de la génération des conditions initiales, l'objectif est d'intégrer dans l'ensemble toutes les sources d'incertitude liées à la prévision.

L'erreur modèle est une première source d'incertitude sur la représentation par le modèle de la réalité. Pour simuler cette erreur, plusieurs paramétrisations du modèle (pour les schémas de convection, de couche limite, de micro-physique) sont utilisés selon les membres. L'échantillonnage de l'incertitude sur la prévision repose (a) sur l'assimilation d'ensemble qui produit un échantillonnage cohérent des incertitudes d'erreur d'analyse en construisant un ensemble d'états analysés, puis (b) l'intégration non linéaire de chacun des états analysés.

Pour obtenir un système d'assimilation cohérent, il est aussi nécessaire de perturber les observations en accord avec la matrice d'erreur d'observation (Burger et al., 1998). La matrice de covariance d'erreur d'analyse calculée à partir de l'ensemble est alors cohérente avec celle utilisée dans le processus d'assimilation de données.

Une des principales limitations de l'emploi des méthodes d'ensemble est que la matrice de covariance estimée, à l'aide d'un ensemble de petite taille, présente des corrélations non pertinente à longue distance. Une méthode couramment employées pour réduire ce bruit d'échantillonnage est de limiter géographiquement les corrélations, c'est à dire de les filtrer. On parle alors de localisation de l'information.

### 2.5.1 Localisation

La localisation a été introduite pour supprimer les corrélations à grande distance (Houtekamer and Mitchell, 2001). En effet, l'estimation directe des corrélations à partir d'un ensemble de petite taille est entachée d'une erreur d'estimation. On peut vérifier que pour des points géographiquements éloignés, la corrélation théorique est faible. Ainsi le filtrage de Schur en supprimant artificiellement la corrélation entre les points géographiquement distant, permet de réduire l'erreur d'échantillonnage.

D'autre part, l'estimation de la matrice de corrélation à partir d'un ensemble produit une information statistique dans l'espace vectoriel engendré par l'ensemble. Ainsi, le rang de la matrice est strictement

inférieure à la taille de l'ensemble. Un corollaire du filtrage de Schur est d'augmenter le rang de la matrice de corrélation. On écrit alors

$$\underline{\boldsymbol{B}}^e = \underline{\tilde{\boldsymbol{B}}}^e \circ \underline{\boldsymbol{C}}, \tag{2.15}$$

avec $\circ$ indiquant le produit élément par élément entre 2 matrices appelé aussi le produit d'Hadamard-Schur. Dans l'équation (2.15) $\underline{\tilde{\boldsymbol{B}}}^e$ est la matrice brute issue de l'ensemble, $\underline{\boldsymbol{C}}$ est une matrice de corrélation de même dimension que $\underline{\tilde{\boldsymbol{B}}}^e$ et $\underline{\boldsymbol{B}}^e$ est la matrice filtrée utilisée dans le système d'assimilation. Il est à noter que cette localisation, généralement réalisée en point de grille, peut également être réalisée dans d'autres représentations, comme en spectral Buehner and Charron (2007), conduisant à des propriétés de filtrage particulières, similaires à celles obtenues par la modélisation de la matrice $\mathbf{B}$ en ondelettes (Pannekoucke et al., 2007; Varella et al., 2011).

Le choix de la méthode de localisation ou de filtrage des covariances pour la matrice $\boldsymbol{B}$ font partie des très nombreuses possibilités d'ajuster un système d'assimilation de données. Il est aussi possible de mentionner des questions équivalentes concernant la prise en compte des incertitudes sur les observations par la modélisation des matrices de covariance d'erreur d'observation $\boldsymbol{R}$. Souvent traitée comme une matrice diagonale, les corrélations d'erreurs existent bel et bien et ne doivent pas être ignorées. Avec une telle complexité, il est très difficile de maintenir un système d'assimilation parfaitement optimal, au sens où chaque observation supplémentaire doit rajouter de l'information utile au système et réduire la variance d'erreur d'analyse. Plusieurs outils diagnostics ont été développés pour contrôler le bon fonctionnement du système d'assimilation.

## 2.6 Diagnostics sur le système d'assimilation de données opérationnel

Dans ce chapitre nous ne nous intéressons qu'aux diagnostics directs à partir du système d'assimilation de données, sans prise en compte de la dynamique comme le font les méthodes d'étude d'impact d'observation ou d'observation adaptative. Les critères de sélection des observations dans le système d'assimilation opérationnel sont présentés d'abord, suivis des diagnostics sur la matrice de covariance d'erreur d'ébauche $\mathbf{B}$. La dernière partie est consacrée au calcul de la réduction des degrés de liberté (Degree of Freedom Signal) qui permet de mesure l'efficacité d'un type d'observation sur la réduction de l'incertitude de l'analyse.

### 2.6.1 Screening et sélection des observations

Le premier élément de contrôle du système d'assimilation se trouve en amont de celui-ci. Il s'agit du mécanisme d'acception et de rejet des observations, aussi appelé *screening*. Le Système Mondial de Télécommunication (SMT) permet à chaque centre de prévision numérique de récupérer un nombre considérable d'observations effectuées partout dans le monde. Il y a actuellement environ $10^8$ observations traitées dans le processus d'assimilation. La nature de ces observations peut être très variable, par exemple entre une mesure in-situ de radiosondage ou une mesure de radiance satellite qui reflète les propriétés de larges épaisseurs d'atmosphère, de même que la qualité (mesure automatiques ou mesures humaines). La distribution inégales de certaines observations les rend aussi trop denses en certains endroits en particulier par rapport à la résolution spatiale du modèle.

L'efficacité du processus d'assimilation dépend aussi de la composition de l'ensemble des observations qu'il traite. L'augmentation de la quantité d'observation augmente le coût numérique de l'assimilation qui représente environ autant qu'une prévision à plusieurs jours d'échéances et n'améliore pas toujours la qualité des analyses. Il convient donc de sélectionner les observations utilisées pour l'assimilation.

Un tri est effectué pour supprimer les observations dont la qualité est douteuse, ou qui pourraient avoir un impact négatif sur l'analyse. Des observations trop éloignées de l'ébauche peuvent générer des incréments d'analyse de grande amplitude qui vont créer des discontinuités dans l'analyse. Ces perturbations sont dissipées par le modèle dynamique lors des premiers pas de temps pour que celui-ci retrouve les équilibres physiques (équilibre géostrophique, équilibre thermique, etc...). Nous obtenons alors des ondes qui se propagent dans le modèle et perturbent les premières échéances du modèle, on parle de *spin-up*.

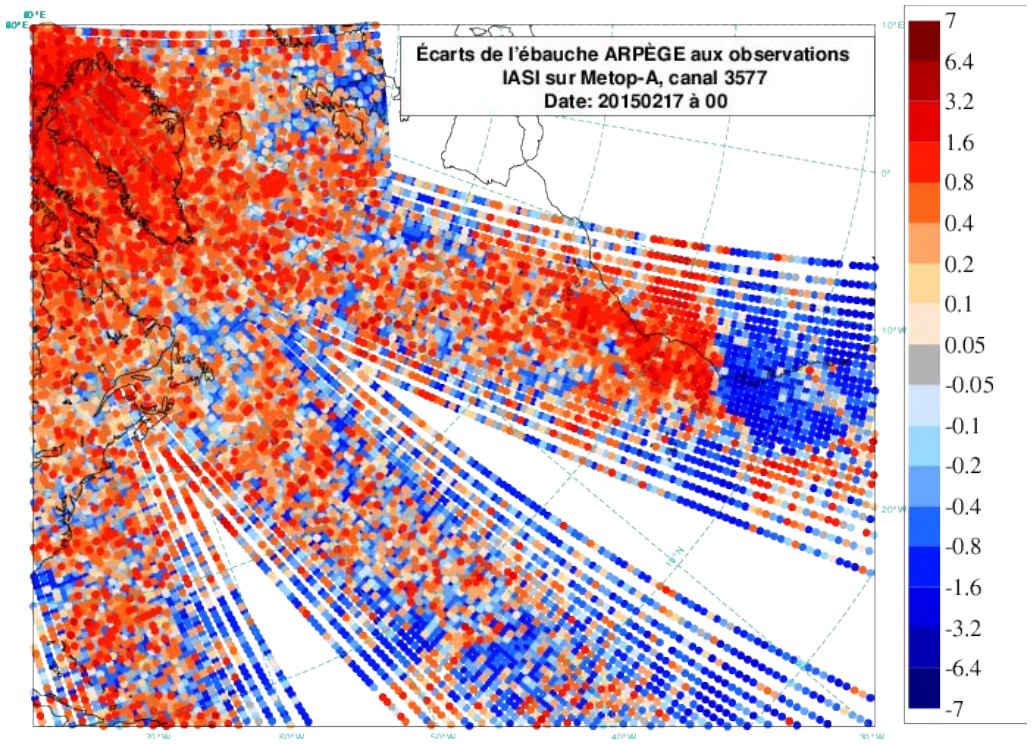



FIGURE 2.4 – Ecarts entre l'ébauche ARPEGE pour l'assimilation ARPEGE du 17/02/2015 pour 0UTC et les observations IASI (canal 3577, canal vapeur d'eau pointant dans stratosphère) du satellite METOP-A. Les forts écarts sont liés à la présence de systèmes nuageux sur l'ouest de l'Afrique. Les données satellites IR sont très affectées par la présence de nuages et demandent un traitement particulier pour permettre leur utilisation.

Même si les observations sont de bonne qualité, il est aussi parfois nécessaire de n'utiliser qu'une partie des observations disponibles quand celles-ci sont trop concentrées et que la matrice $\boldsymbol{R}$ représente mal les corrélations. C'est le cas pour les satellites défilants où seule une faible partie des observations sont conservées et aussi des données d'avion de lignes qui sont très concentrées en certaines régions du monde.

### 2.6.2 Sélection des observations pour l'assimilation de données

La sélection est réalisée en deux étapes, la surveillance et le contrôle, puis le *screening*.

#### Surveillance et contrôle des observations

La surveillance et le contrôle des observations sont réalisés en dehors du système d'assimilation de données. Des statistiques journalières et mensuelles sont réalisées entre la valeur des observations et l'ébauche du modèle $\boldsymbol{y}^o - \mathcal{H}(\boldsymbol{x}^b)$. En comparant ces statistiques entre observations de différents types sur la même zone (géographique et temporelle) et sur plusieurs périodes de temps, il est possible de détecter les observations douteuses. Les erreurs d'observations peuvent être de plusieurs types,

- erreur de mesure de l'instrument (faible le plus souvent)
- erreur dans le codage/décodage des observations
- mauvaise représentation ou résolution du modèle trop faible
- erreur lors de l'application de $\boldsymbol{H}$

Ces statistiques sont utilisées pour placer certaines observations en liste noire tant qu'elles sont considérées douteuses. Ces listes sont régulièrement mises à jour pour ne pas rejeter des observations correctes. Une illustration de carte de comparaison entre les observations et l'ébauche est montrée en figure 2.4.

| Type d'observation | Avant screening | % | Après screening | % | Assimilé |
|---|---|---|---|---|---|
| SYNOP (sol+GPS) | 251420 | 0.21% | 44233 | 0.47% | 17.59% |
| AIREP et AMDAR | 547288 | 0.47% | 371822 | 3.98% | 67.94% |
| SATOB et SAT WIND | 1773537 | 1.52% | 62214 | 0.67% | 3.51% |
| BUOY | 14815 | 0.01% | 5578 | 0.06% | 37.65% |
| TEMP (radiosondage) | 227230 | 0.19% | 122799 | 1.31% | 54.04% |
| PILOT et PROFILEURS | 94686 | 0.08% | 21496 | 0.23% | 22.7% |
| SATEM (radiances) | 113168976 | 96.77% | 8573917 | 91.73% | 7.58% |
| SCATT (diffusiomètres) | 715114 | 0.61% | 38684 | 0.41% | 5.41% |
| LIMB (GPS radio-occultation) | 158657 | 0.14% | 105760 | 1.13% | 66.66% |
| TOUTES OBS | 116951723 | 100% | 9346503 | 100% | 7.99% |

TABLE 2.1 – Nombre d'observations pour l'assimilation ARPEGE du réseau de 0h du 17/02/2015 par type d'observation. Les pourcentages avant et après screening indiquent la proportion par type d'observation sur l'ensemble des observations (avant et après). La dernière colonne indique le pourcentage des données effectivement assimilées par type d'observation.

Des seuils d'erreur standard sont définis par l'Organisation Mondial de la Météorologie pour certains types d'observations. Par exemple pour la station de surface, $15hPa$ d'écart de $Pmer$ ou $25m/s$ pour le vent à $10m$. Des seuils équivalents sont définis pour les mesures par radiosondage.

Cette première étape permet de filtrer grossièrement les instruments ayant des erreurs systématiques mais ne prend pas en compte l'impact effectif des observations sur l'ébauche lors de l'assimilation. C'est pourquoi une deuxième étape est nécessaire.

### Screening

Le screening est réalisé directement dans le processus d'assimilation de données. L'objectif est de conserver les observations qui seront les plus bénéfiques à l'analyse. Le screening est encore décomposé en plusieurs validations.

On vérifie d'abord la cohérence des données. On s'assure aussi qu'il n'y a pas deux fois la même observation qui a été prise en compte (par exemple des données AIREP avec une localisation et un horaire identique). Pour une mesure de radiance, il faut que les températures soient comprises entre $150K$ et $350K$. Des profils de radiance type sont aussi pré-définis, si une mesure ne correspond à aucun de ces profils, elle est rejetée.

Un second calcul d'écart à l'ébauche est réalisé mais cette fois ci en utilisant directement l'ébauche qui servira à l'assimilation. Dans le cadre d'un 4D-Var, l'ébauche est intégrée dans le temps.

Le dernier filtrage est le *thining*, il consiste à réduire la densité des observations à haute résolution. En effet, aucune corrélation horizontale n'est prise en compte dans la matrice $\boldsymbol{R}$. Donc si l'on conserve deux observations trop proches, le système les considérera comme indépendantes et leur associera plus de valeur que normalement prévu. Par exemple pour les mesures d'avions de ligne (type AIREP/AMDAR), on ne conserve qu'une mesure sur un rayon de $170km$, pour les mesures de vent SATOB (vents dérivés du déplacement de nuages par traitement d'images satellite successives), ce rayon est de $250km$. Le choix de la mesure à l'intérieur de chaque boîte est réalisé à partir de l'observation la plus complète, et de celle la plus proche du centre de la tranche temporelle. On préférera aussi une mesure satellite au dessus de la mer plutôt que de la terre, et au maximum avec un ciel clair. Un *thining* horizontale est réalisé pour les mesures AIREP, SATOB et TOVS. Un *thining* vertical est réalisé pour les AIREP et SATOB.

Le tableau 2.1 décrit le nombre d'observation de chaque type utilisées pour l'assimilation ARPEGE du réseau de 0h du 17/02/2015. On note que les données de radiances satellite composent une écrasante majorité des observations disponibles, alors que seulement 7.58% des observations passent le screening.

### 2.6.3 Diagnostic de la variance et des longueurs de portée

Dans les expériences d'observations adaptatives qui sont menée au chapitre 6 il est nécessaire de spécifier la matrice de covariance d'erreur de prévision $\mathbf{B}$. La grande dimension interdit la manipulation directe de ces matrices de covariance d'erreur, et pour les besoins de la modélisation ou des diagnostics, il est nécessaire de caractériser les matrices de covariances.

Ainsi cette section présente quelques méthodes classiques utilisées pour obtenir la variance et les longueurs de portée de $\mathbf{B}$ ou $\mathbf{A}$ dans un système de PNT opérationnel. L'estimation de ces deux quantités, variance et longueur de portée, repose principalement sur l'estimation ensembliste.

#### Variance d'erreur d'ébauche

Dans le cas d'un ensemble de prévision ou d'assimilation de $N_e$ membres, la variance d'erreur d'ébauche ou d'analyse est calculée à partir de la variance de chaque paramètre en chaque point de grille $i$ du vecteur d'état.

$$\sigma_i^{b2} = \frac{1}{\sqrt{1-N_e}} \sum_{k=1}^{N_e} ((\mathbf{x}_k^b)_i - \overline{\mathbf{x}^b}_i)^2$$

$$\sigma_i^{a2} = \frac{1}{\sqrt{1-N_e}} \sum_{k=1}^{N_e} ((\mathbf{x}_k^a)_i - \overline{\mathbf{x}^a}_i)^2$$

où $\overline{\mathbf{x}}$ est la moyenne de $\mathbf{x}$.

Lorsque l'on ne dispose pas d'un ensemble, la matrice $\mathbf{B}$ doit être modélisée. Le champ de variance est obtenu à partir d'une climatologie du modèle, il est donc directement connu. Les variances d'erreur d'analyse sont plus compliquées à calculer. Dans un système d'assimilation variationnel la matrice $\mathbf{A}$ est un sous-produit de la minimisation de la fonction coût. A partir d'un algorithme de Lanczos, on déduit une estimations des vecteurs propres et valeurs propres de $\mathbf{A}^{-1}$, voir annexe B. On peut alors obtenir une estimation de $\mathbf{A}^{1/2}$ sous forme d'opérateurs et en déduire une estimation de la variance en passant par la génération d'un ensemble de perturbations $\boldsymbol{\varepsilon}^a = \boldsymbol{A}^{1/2}\eta$ avec $\eta \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$ suivant une loi normale (voir annexe D). On retrouve alors le diagnostic de la variance à partir d'un ensemble.

#### Diagnostic des longueurs de portée

Les longueurs de portée (1D) et le tenseur de métrique local (2D) ont été présentés au chapitre précédent 2.1.4. Ils modélisent tous les 2 la forme de la fonction de corrélation en un point. Plusieurs méthodes existent pour ce diagnostic.

**Estimation de la métrique à partir du calcul de $\rho$** La première méthode vient directement de la définition de la longueur de portée à partir du développement de Taylor à l'ordre 2. En utilisant les perturbations (écarts à la moyenne) d'un ensemble,, il est possible de construire la valeur des corrélations $\rho_{\mathbf{x},\mathbf{x}+\delta\mathbf{x}}$ dans le voisinage d'un point $\mathbf{x}$ et d'en déduire la valeur de la dérivée seconde $\frac{d^2\rho_{\mathbf{x}}}{dx^2}$ par différence finie. Une approche équivalente a été proposée par Pannekoucke et al. (2008), et généralisée par Pannekoucke and Massart (2008); Pannekoucke (2009) pour l'estimation du tenseur de métrique locale.

**Formule de Belo Pereira et Berre (2006)** La deuxième méthode, introduite par (Belo Pereira and Berre, 2006), met en œuvre l'utilisation d'un ensemble de prévisions ou d'assimilations. On note $\varepsilon(\mathbf{x})$ l'erreur de prévision ou d'analyse, et $\sigma(\mathbf{x}) = \sqrt{\mathbb{E}[\varepsilon(\mathbf{x})^2]}$ le champ d'écart-type associé. On définit l'erreur normalisée comme $\underline{\varepsilon}(\mathbf{x}) = \frac{\varepsilon(\mathbf{x})}{\sigma(\mathbf{x})}$. On peut montrer que le tenseur de métrique locale $\boldsymbol{g}_{\mathbf{x}}$ (Eq. 2.5) s'écrit comme,

$$\boldsymbol{g}_{\mathbf{x}} = \mathbb{E}\left(\nabla\underline{\varepsilon}(\mathbf{x})\nabla\underline{\varepsilon}(\mathbf{x})^T\right). \tag{2.16}$$

L'avantage de cette méthode est de permettre une caractérisation mathématique des erreurs d'estimations de portée Raynaud and Pannekoucke (2013); Michel (2013b).

**Utilisation des portées pour la calibration objective de la matrice de covariance d'erreur de prévision**

Si la longueur de portée est d'abord un diagnostic, elle peut aussi être utilisée comme un outil permettant la paramétrisation des matrices de covariance en grande dimension.

Par exemple, Pannekoucke and Massart (2008) ont proposé d'utiliser le diagnostic des portées pour estimer de manière objective le tenseur de diffusion $\boldsymbol{\nu}$ intervenant dans la modélisation des covariances d'erreur de prévision à l'aide de l'opérateur de diffusion (Weaver and Courtier, 2001)

$$\partial_t x = \nabla \cdot (\boldsymbol{\nu} \nabla x). \tag{2.17}$$

Dans ce cadre, sous l'hypothèse d'homogénéité locale (l'échelle de variation des longueurs de portées est beaucoup plus grande que la valeur des portée) alors le tenseur de diffusion local peut être estimé à partir de l'inverse de la métrique locale (Eq. 2.5), $\boldsymbol{g}^{-1}$, suivant

$$\boldsymbol{\nu} = \frac{\boldsymbol{g}^{-1}}{2}. \tag{2.18}$$

En particulier, dans le cas de $\rho(\boldsymbol{x}_j) = \exp(-\frac{x_j^2}{2{L_{px}}_j^2} - \frac{y_j^2}{2{L_{py}}_j^2})$, les composantes diagonales du tenseur de diffusion valent alors $\frac{1}{2}{L_{px}}^2$ et $\frac{1}{2}{L_{py}}_j^2$. La figure 2.5 illustre des fonctions de corrélation diagnostiquées dans un ensemble (a) et des fonctions modélisées par l'équation de diffusion paramétrée par le calcul des tenseurs locaux (b). À partir de l'ensemble, panel (a), le tenseur de diffusion local (en rouge) est déduit du calcul du tenseur de métrique local. Il est ensuite introduit dans le modèle de corrélation basé sur la diffusion suivant un produit d'opérateur comme présenté en annexe A. Le résultat de la modélisation est reproduit panel (b) : la correspondance entre l'estimation et la modélisation des corrélations montre qu'il est possible de représenter les anisotropies et d'inclure par l'ensemble une dépendance à l'écoulement.

Une autre modélisation intéressante pour construire des fonctions de corrélation anisotrope est celle proposée par Desroziers (1997) qui repose sur un changement de coordonnées. Cette approche est séduisante car elle permet de produire des anisotropies à l'aide de simples interpolations. La principale difficulté est de trouver un changement de coordonnées adapté au problème. Michel (2013a,b) a proposé une manière objective de construire une telle transformation en s'inspirant d'une technique d'image par analyse en ondelette Clerc and Mallat (2003). En s'appuyant sur le formalisme de la géométrie Riemannienne, Pannekoucke et al. (2014) ont proposé la construction d'un changement de coordonnées objectif à l'aide de la métrique locale (simple à diagnostiquer comme nous l'avons énoncé précédemment).

Les deux parties précédentes 2.6.1 et 2.6.3 permette de contrôler en amont de l'assimilation les éléments que l'on va fournir au système. Il est aussi possible de diagnostiquer en aval de l'assimilation l'utilisation effective des observations à travers le calcul des DFS.

### 2.6.4 Illustration des DFS MF

La méthode dite des DFS (*Degree of Freedom Signal*) est utilisée à Météo-France pour quantifier l'impact d'un type d'observation sur le système d'assimilation. La méthode consiste à estimer le nombre de degrés de liberté qui sont contraints par les observations (Fischer 2003).

On effectue le changement de variable $\boldsymbol{z} = \boldsymbol{L}\boldsymbol{x}$ avec $\boldsymbol{L}$ telle que $\boldsymbol{L}\boldsymbol{B}\boldsymbol{L}^T = \boldsymbol{I}$, dans le nouvel espace la matrice de covariance d'erreur d'ébauche est donc l'identité. On définit alors la matrice $\boldsymbol{A}_z = \boldsymbol{L}\boldsymbol{A}\boldsymbol{L}^T$, qui correspond à la matrice d'erreur d'analyse dans le nouvel espace. Soit $\boldsymbol{V}$ la matrice des vecteurs propres de $\boldsymbol{A}_z = \boldsymbol{V}\boldsymbol{D}_z\boldsymbol{V}^T$. Comme $\boldsymbol{V}$ est orthogonal, $\boldsymbol{L}$ peut être remplacé par $\boldsymbol{V}^T\boldsymbol{L}$ pour définir le nouvel espace. Les degrés de liberté sont alors calculés comme la trace de la différence des deux matrices de covariance

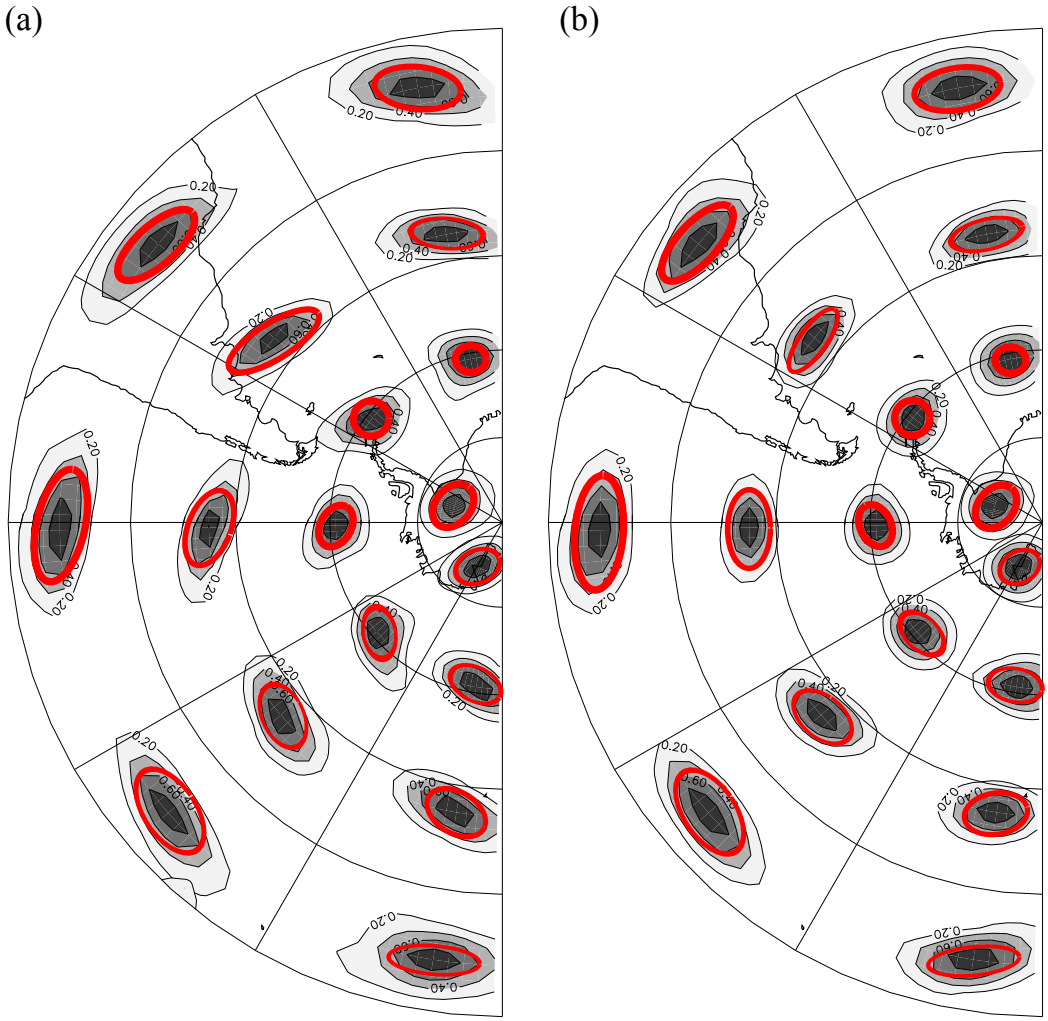



FIGURE 2.5 – Illustration des fonctions de corrélation 2D sur la sphère, vue du pôle sud. (a) fonctions de corrélation calculées à partir d'un ensemble de prévision. (b) fonctions de corrélation modélisées à partir d'une méthode de perturbations aléatoires gaussiennes. (d'après Pannekoucke and Massart (2008) avec l'autorisation des auteurs pour l'utilisation et la modification de la figure)

dans ce nouvel espace,

$$\begin{aligned} d &= \text{Trace}(\boldsymbol{V}^T\boldsymbol{L}\boldsymbol{B}\boldsymbol{L}^T\boldsymbol{V} - \boldsymbol{V}^T\boldsymbol{L}\boldsymbol{A}\boldsymbol{L}^T\boldsymbol{V}) \\ &= \text{Trace}(\boldsymbol{I} - \boldsymbol{D}_z) \\ &= N - \sum_{i=1}^{N} \lambda_i \end{aligned} \tag{2.19}$$

avec $\lambda_i$ les valeurs propres de $\boldsymbol{A}_z$.

Le calcul de la trace d'une matrice peut se faire en utilisant des produits matrice-vecteur, plusieurs méthodes sont décrites en annexe G. Pour une assimilation incrémentale, la Héssienne de la fonction coût pré-conditionnée $\boldsymbol{A}_z^{-1}$ est obtenue lors de la minimisation en utilisant l'algorithme de Lanczos. Cette estimation de la Héssienne peut alors être diagonalisée et inversée. Cette expression permet de calculer les produits matrice-vecteur pour obtenir la trace.

Les DFS sont calculés à chaque nouveau cycle (ou version) des modèles pour suivre le nombre d'observation assimilés en opérationnel et l'impact de chaque type d'observation. Ce diagnostic est appliqué sur une journée considérée représentative. La figure 2.6 montre le nombre d'observations assimilées pour la journée du 27/12/2013 dans le modèle ARPEGE. Celle-ci est à comparer à la figure 2.7 qui représente la part de réduction des degrés de liberté par type d'observation. Dans cette illustration de 2013, les données de radiance satellite représentent déjà plus de 90% des observations assimilées. Cependant elles ne jouent que pour 70% dans la réduction des degrés de liberté. A l'inverse les données de radiosondage (TEMP) et les mesures avions (AIRCRAFT) sont proportionnellement très exploités (TEMP 0.98% des données utilisées pour 7.87% de réduction, AIRCRAFT 4.76% des données utilisées pour 16.48% de réduction).

On constate donc que malgré un nombre très important d'observations disponibles, toutes les observations ne se valent pas par rapport à un modèle dynamique donné. De plus ces approches uniquement basées sur le diagnostic du processus d'assimilation ne prennent pas en compte l'état dynamique futur de l'atmosphère.

## 2.7 Conclusion chapitre 2

Les objectifs et les concepts de l'assimilation de données ont été rappelés dans ce chapitre. En particulier, le filtrage non linéaire a été introduit montrant le cadre général de l'assimilation de données dans le cas non-Gaussien et non linéaire. Un cas particulier important est celui où la dynamique et l'opérateur d'observation sont linéaires, et où les distributions de probabilité sont supposée Gaussienne. Dans ce cadre, le caractère Gaussien est conservé et le filtrage non linéaire peut être résolut analytiquement : il s'agit des équations du filtre de Kalman. D'un point de vue conceptuel le filtre de Kalman joue un rôle essentiel en assimilation de donnée et dans l'essor de la discipline. Cependant, dans la pratique, le filtre de Kalman n'est pas implémenté directement et deux familles d'algorithmes ont été développées : le variationnel et le filtre de Kalman d'ensemble.

Hérité du contrôle optimal, le formalisme variationnel correspond à la résolution des équations du BLUE par une approche de type minimisation d'une fonction coût, le 3D-Var. Cette méthode est équivalente à l'étape d'analyse du filtre de Kalman dans le cas où la matrice de covariance d'erreur d'ébauche est égale à la matrice de covariance d'erreur de prévision. L'extension temporelle du 3D-Var permettant de bénéficier de la dynamique est formulée par le 4D-Var, apparenté à un lisseur et non plus à un filtre.

L'approche 4D-Var a d'abord été présentée comme une solution pour prendre en compte la dynamique du modèle dans l'assimilation. L'avantage de cette approche est de relaxer l'hypothèse de Gaussiannité à l'intérieur de la fenêtre d'assimilation. La méthode efficace pour minimiser la fonction coût du 4D-Var est de procéder par méthode adjointe en développant le code tangent linéaire du modèle non linéaire ainsi que le code adjoint associé au code tangent-linéaire. Malgré cela, l'approche 4D-Var reste très coûteuse et la solution opérationnelle pour le mettre en œuvre repose sur une reformulation incrémentale, le 4D-Inc, où la minimisation s'appuie sur une boucle interne (minimisation du problème linéarisée autour d'une

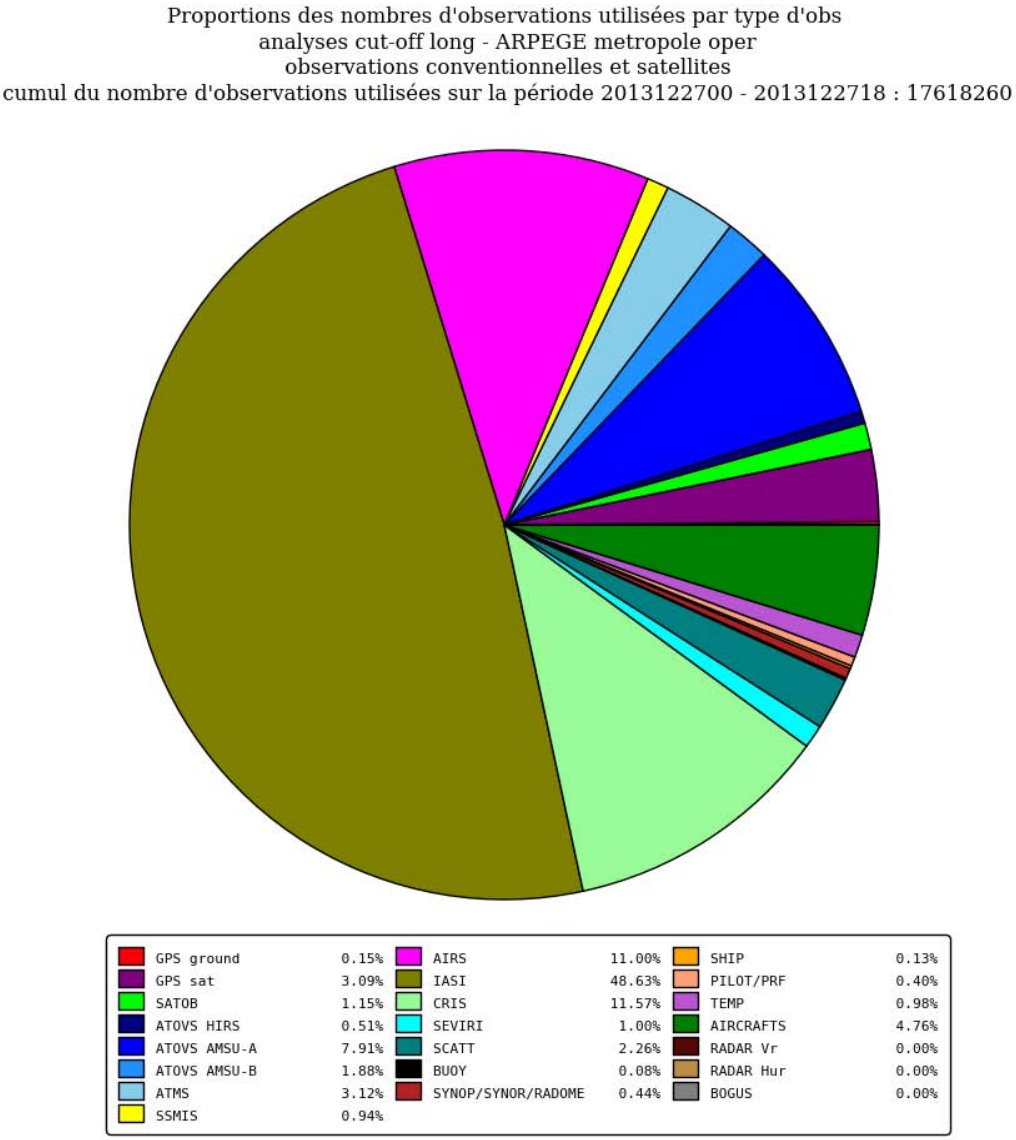



FIGURE 2.6 – Les mesures satellites représentent plus de 90% des observations assimilées, en particulier l'instrument IASI produit un très grand nombre de données à partir de ces nombreux canaux. Les mesures des stations au sol représentent moins d'1% des données et les radiosondages à peine 1%. Les profils issus des messages AMDAR des avions de ligne sont eux nombreux (4.76%) mais se concentrent autour des grands couloirs aériens.

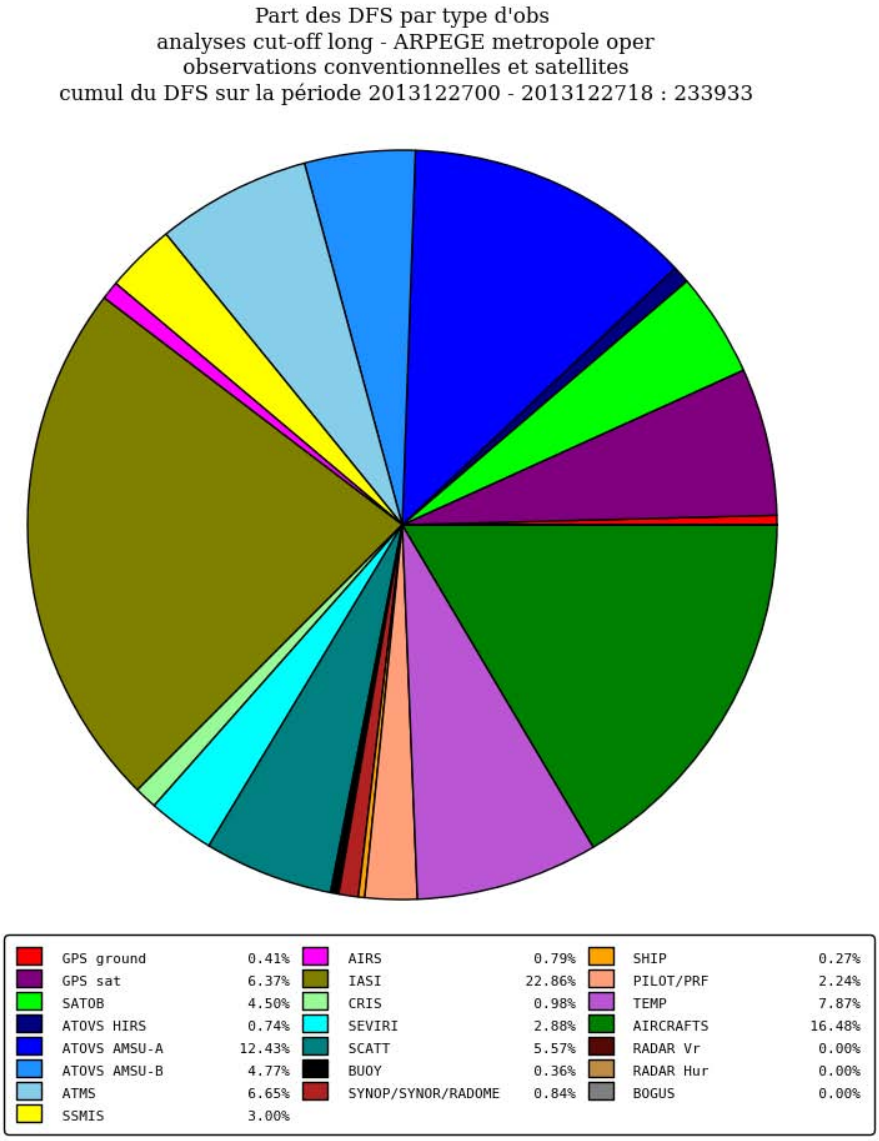



FIGURE 2.7 – Les données IASI représentent un peu de moins de 50% des données assimilées mais que 22.86% de la réduction des degrés de liberté. A contrario l'assimilation tire un grand profit des AMDAR (AIRCRAFTS) avec 16.48%. Les observations les plus efficaces restent cependant les radiosondages, avec 7% de réduction de DFS pour moins d'1% des données assimilées. Ceci explique en partie pourquoi les radiosondages sont encore utilisés comme observations de référence pour le calcul des scores des modèles de PNT.

trajectoire) et une boucle externe (mise à jour d'une trajectoire non linéaire). Cette formulation a été introduite par Courtier et al. (1994).

Il est intéressant de noter que la même année, 1994, est apparu une autre solution pratique pour résoudre un écueil : l'équation de propagation des incertitudes dans le filtre de Kalman. En effet, Evensen (1994) à introduit la formulation ensembliste du filtre de Kalman qui s'appuie sur l'estimation de la matrice de covariance à partir d'un ensemble de prévision (intégration par le modèle non linéaire). Parmi les variantes de cette formulation, la méthode par transformée d'ensemble a été détaillée dans ce chapitre. Cette formulation montre que l'analyse est construite dans le sous-espace engendré par l'ensemble de prévision.

Les méthodes ensemblistes sont cependant contraintes par le coût numérique associée à l'intégration temporelle d'ensembles de grandes tailles. En pratique, l'utilisation de cette approche se limite à un ensemble de petite taille (par rapport à la dimension du problème) et une résolution dégradée. L'estimation de la matrice de covariance est alors polluée par le bruit d'échantillonnage. Pour filtrer ce bruit d'échantillonnage, des techniques de filtrage ont été introduites, en particulier l'utilisation de la localisation. Cette localisation permet de supprimer les corrélations à grande distance et correspond à une augmentation artificielle de la taille de l'échantillon, ce qui est une réponse partielle à la déficience en rang de la matrice de covariance estimée.

L'approche ensembliste a engendré un bouillonnement d'idées dans la communauté scientifique et a permis de lever certaines difficultés rencontrées dans les algorithmes de l'assimilation de données. Pour exemple, dans la formulation 4D-Var, on peut remplacer la matrice de covariance d'erreur d'ébauche par une estimation de la matrice d'erreur de prévision déduite d'une assimilation d'ensemble : mise en œuvre de 4D-Var en parallèle. Une autre des difficultés techniques du 4D-Var ainsi levée par l'approche ensembliste est la nécessité de disposer des modèles tangent linéaire et adjoint pour le transport des perturbations au sein de la fenêtre d'assimilation. Assurer la mise en place, le maintien en conditions opérationnelles et l'évolution de ces deux modèles est très lourd et complexe. Les méthodes d'ensemble s'appuyent sur les trajectoires calculées, conduisant aux algorithmes 4D-En-Var.

Le 4D-En-Var permet de combiner les avantages des approches variationnelle et ensembliste. Un ensemble permet de disposer d'une matrice $\mathbf{B}$ *du jour* souvent complétée d'une matrice issue de la climatologie pour s'assurer que le résultat final soit de plein rang. La minimisation par une méthode variationnelle permet de prendre en compte une partie de la non linéarité de la trajectoire et de réduire les coûts de calcul.

Dans ce chapitre, nous avons également introduit les différents diagnostics du système d'assimilation de données qui permettent de mettre en évidence l'hétérogénéité des observations. Les observations satellites représentent une écrasante majorité ($> 90\%$) des observations assimilées et sont les plus difficiles à intégrer car il s'agit de mesures indirectes et nécessitent un débiasage. Un compromis doit être fait entre le nombre d'observations assimilées, leur qualité et le coût de calcul associé. Les observations in-situ dites conventionnelles (radiosondage, station au sol) sont toujours vitales pour contrôler le système d'assimilation.

Le chapitre suivant propose un calcul de réduction de variance d'erreur d'analyse pour une observation supplémentaire sans nécessité d'effectuer une assimilation complète. Ce développement est ensuite utilisé dans le chapitre 4, après la présentation du contexte de la sensibilité aux observations, pour introduire une nouvelle méthode d'observation adaptative appelée VRF.

# Chapitre 3

# Réduction de la variance d'erreur d'analyse due à une observation

Le chapitre 2 a présenté le formalisme et les méthodes d'un système d'assimilation de données. L'impact d'une observation sur l'analyse est illustré dans la section 2.1.3 et permet d'appréhender le rôle de la matrice de covariance d'erreur d'ébauche $\mathbf{B}$ lors de l'assimilation. Ce chapitre explore plus en profondeur l'impact de l'assimilation d'une observation cette fois ci sur les matrices de covariance d'erreur, en particulier $\mathbf{B}-\mathbf{A}$.

Comme il sera présenté en section 4.3 du chapitre 4, l'observation adaptative ne cherche pas à évaluer l'impact d'une observation future (dont la valeur est a priori inconnue) directement sur l'analyse mais plutôt la réduction de variance d'erreur d'analyse associée à l'assimilation de cette observation. L'objet du développement proposé ici est d'évaluer analytiquement cette réduction de variance sans avoir à réaliser une assimilation de données.

Le développement analytique et les hypothèses liées au calcul sont présentées dans la section 3.1. Ensuite la section 2.6.3 décrit les méthodes permettant d'obtenir les éléments nécessaires au calcul dans le contexte de la PNT. Une illustration de la méthode est réalisée dans un cadre simplifié dans la section 3.2.

## 3.1 Développement analytique

### 3.1.1 Assimilation d'une observation unique

#### Formalisme général de la réduction de variance pour une observation

Par définition, $\mathbf{B}-\mathbf{A}=\mathbf{B}-[(\mathbf{I}-\mathbf{KH})\mathbf{B}]$ tel que $\mathbf{B}-\mathbf{A}=\mathbf{KHB}=\mathbf{BH}^T(\mathbf{HBH}^T+\mathbf{R})^{-1}\mathbf{HB}$. On souhaite quantifier l'impact d'une observation d'un paramètre du modèle et positionnée exactement en un point $j$ de la grille de ce modèle. Il vient alors que l'opérateur d'observation est la forme linéaire $\mathbf{H}=[\delta_{ji}]$.

Par le calcul, il vient que $\mathbf{HB}=\mathbf{B}(j,:)=\mathbf{B}(:,j)^T$ correspond à la $j^{ime}$ ligne de la matrice $\mathbf{B}$. De plus, $(\mathbf{HBH}^T+\mathbf{R})^{-1}=\frac{1}{\sigma_j^{b^2}+\sigma_j^{o^2}}$ est un scalaire, d'où on déduit que

$$\mathbf{B}-\mathbf{A}=\frac{1}{\sigma_j^{b^2}+\sigma_j^{o^2}}\mathbf{B}(:,j)\mathbf{B}(:,j)^T. \tag{3.1}$$

Dans le cadre de l'impact d'une observation, on s'intéresse à la réduction de variance apportée par cette observation, ce qui se traduit par le calcul de la trace de cette différence. Or comme $Tr(\boldsymbol{u}\boldsymbol{u}^T)=|\boldsymbol{u}|^2$, il vient que

$$Tr(\mathbf{B}-\mathbf{A})=\frac{1}{\sigma_j^{b^2}+\sigma_j^{o^2}}||\mathbf{B}(j,:)||_2^2, \tag{3.2}$$

où $||\mathbf{B}(:,j)||_2^2$ désigne la norme $\mathcal{L}^2$ de la fonction (du vecteur) de covariance relative au point $j$ $\mathbf{B}(:,j)$.

Sans hypothèse supplémentaire sur la forme de la fonction de structure au point observé, il n'est pas possible de déduire une expression analytique de la réduction de variance.

### Hypothèse simplificatrice pour construire une expression analytique

On souhaite ici aller plus loin en supposant une forme particulière pour la fonction de structure.

**Hypothèse de travail :** On suppose, pour simplifier, que la fonction $\mathbf{B}(:,j)$ est localement Gaussienne, caractérisée par une variance $\sigma_j^{b^2}$ et un tenseur de diffusion locale $\boldsymbol{\nu}_j = \frac{1}{2}\boldsymbol{g}_j^{-1}$ tel que $\rho(\mathbf{x}, \mathbf{x}+\delta\mathbf{x}) \approx 1 - \frac{1}{2}|\delta\mathbf{x}|^2_{\boldsymbol{g}_j}$ (voir Eq. 2.5). Ainsi, la fonction de covariance prend localement la forme suivante :

$$\rho(\mathbf{x}_j, \mathbf{x}_j + \delta\mathbf{x}) = \sigma_j^{b^2} e^{-\frac{1}{2}|\delta\mathbf{x}|^2_{\boldsymbol{g}_j}} \tag{3.3}$$

avec la normalisation

$$\int_{\mathbb{R}} \rho(\mathbf{x}_j, \mathbf{x}_j + \delta\mathbf{x}) d\delta\mathbf{x} = \sqrt{|2\pi \boldsymbol{g}_j^{-1}|} = \sigma_j^{b^2}\sqrt{|4\pi\boldsymbol{\nu}_j|}. \tag{3.4}$$

**Expression analytique pour la réduction de variance** Une fois l'hypothèse sur la forme locale prise en compte, l'intégrale équivalente à la norme $||\mathbf{B}(:,j)||_2^2$ s'écrit sous une forme analytique obtenue comme suit :

$$\begin{aligned}\int_{\mathbb{R}} \left(\sigma_j^{b^2} e^{-\frac{1}{2}|\delta\mathbf{x}|^2_{\boldsymbol{g}_j}}\right)^2 d\delta\mathbf{x} &= \sigma_j^{b^4} \int_{\mathbb{R}} e^{-|\delta\mathbf{x}|^2_{\boldsymbol{g}_j}} d\delta\mathbf{x} \\ &= \sigma_j^{b^4} \int_{\mathbb{R}} e^{-|\delta\mathbf{x}|^2_{\frac{1}{2}\boldsymbol{\nu}_j^{-1}}} d\delta\mathbf{x} = \sigma_j^{b^4} \int_{\mathbb{R}} e^{-\frac{1}{2}|\delta\mathbf{x}|^2_{\boldsymbol{\nu}_j^{-1}}} d\delta\mathbf{x} = \sigma_j^{b^4}\sqrt{|2\pi\boldsymbol{\nu}|}. \end{aligned} \tag{3.5}$$

Ainsi, il vient que

$$Tr(\mathbf{B}-\mathbf{A}) = \frac{\sigma_j^{b^4}}{\sigma_j^{b^2} + \sigma_j^{o^2}}\sqrt{|2\pi\boldsymbol{\nu}_j|}. \tag{3.6}$$

**Expression analytique pour la réduction de variance dans le cas discret** Pour un problème discrétisé sur une grille de résolution $dx$, il vient que la réduction de variance est donnée par

$$Tr(\mathbf{B}-\mathbf{A}) = \frac{\sigma_j^{b^4}}{\sigma_j^{b^2} + \sigma_j^{o^2}}\sqrt{|2\pi\boldsymbol{\nu}_j|}\frac{1}{dx}, \tag{3.7}$$

pour un problème $1D$, et par

$$Tr(\mathbf{B}-\mathbf{A}) = \frac{\sigma_j^{b^4}}{\sigma_j^{b^2} + \sigma_j^{o^2}}\sqrt{|2\pi\boldsymbol{\nu}_j|}\frac{1}{dx^d},$$

pour un problème en dimension $d$. $dx$ correspond à la métrique locale qui intervient dans le calcul d'intégrale discrète, tel que décrit dans l'annexe C de Pannekoucke and Massart (2008).

### Synthèse sur la réduction de variance et interprétation

Le résultat de cette section peut se retenir sous la forme de la propriété suivante.
Dans le cas localement Gaussien et stationnaire où la fonction de structure en un point $j$ s'écrit

$$\rho(\mathbf{x}_j, \mathbf{x}_j + \delta\mathbf{x}) = \sigma_j^{b^2} e^{-\frac{1}{2}|\delta\mathbf{x}|^2_{\boldsymbol{g}_j}},$$

la réduction de variance associée à l'assimilation d'une observation au point $j$ s'écrit

$$Tr(\mathbf{B}-\mathbf{A}) = \frac{{\sigma_j^b}^4}{{\sigma_j^b}^2+{\sigma_j^o}^2}\sqrt{|2\pi\boldsymbol{\nu}_j|}, \tag{3.8}$$

dans le cas continu et

$$Tr(\mathbf{B}-\mathbf{A}) = \frac{{\sigma_j^b}^4}{{\sigma_j^b}^2+{\sigma_j^o}^2}\sqrt{|2\pi\boldsymbol{\nu}_j|}\frac{1}{dx^d}, \tag{3.9}$$

dans le cas discret en dimension $d$.

L'expression (3.9) permet de comparer l'impact potentiel de plusieurs observations sur la réduction d'incertitude en prenant en compte toute la dimension de la matrice de covariance d'erreur d'ébauche. Par exemple, on peut comparer la réduction de variance liée à une observation placée à un endroit de faible variance d'erreur d'ébauche mais avec de très grandes longueurs de portée, à celle produite par une observation placée à un endroit de très forte variance mais avec les longueurs de portée faibles. Ces calculs ne nécessitent pas de connaître la valeur de l'observation et ou même de réaliser une assimilation, seules les matrices $\mathbf{B}$ et $\mathbf{R}$ doivent être connues.

### 3.1.2 Evaluation de l'assimilation d'une observation supplémentaire

Le développement présenté ici peut être repris dans un contexte où l'on souhaite rajouter des observations supplémentaires à un réseau déjà existant. On distingue la matrice de covariance d'erreur d'analyse $\mathbf{A}$ qui correspond à la matrice issue de l'assimilation avec les observations standards, et la matrice $\widetilde{\mathbf{A}}$ qui correspond à l'assimilation des observations standard plus une observation supplémentaire. Cette observation supplémentaire est supposée indépendante du réseau initial. Assimiler cette observation en plus revient alors à effectuer une seconde assimilation après l'assimilation initiale dont le résultat (l'analyse) est utilisé comme ébauche pour la seconde. Dans les équations, on remplace alors $Tr(\mathbf{B}-\mathbf{A})$ par $Tr(\mathbf{A}-\widetilde{\mathbf{A}})$. L'équation (3.2) s'écrit alors,

$$Tr(\mathbf{A}-\widetilde{\mathbf{A}}) = \frac{1}{{\sigma_j^a}^2+{\sigma_j^o}^2}||\mathbf{A}(j,:)||_2^2. \tag{3.10}$$

On en déduit de la même manière l'équation (3.8) pour une observation supplémentaire,

$$Tr(\mathbf{A}-\widetilde{\mathbf{A}}) = \frac{{\sigma_j^a}^4}{{\sigma_j^a}^2+{\sigma_j^o}^2}\sqrt{|2\pi\boldsymbol{\nu}_j^a|}, \tag{3.11}$$

avec $\boldsymbol{\nu}_j^a$ le tenseur de diffusion local en $\mathbf{A}$.

Ce calcul permet d'estimer l'impact d'observations supplémentaires indépendantes des observations standards et de fournir un critère quantitatif pour prendre la décision d'effectuer le déploiement supplémentaire ou non. Cette méthode d'observation adaptative ne prend en compte que le système d'assimilation de données, les techniques plus avancées seront présentées au chapitre 4. Cette méthode peut aussi permettre d'identifier les observations ayant peu d'impact à priori sur la réduction de variance de l'analyse, pour prioriser l'assimilation des observations à fort impact.

## 3.2 Validation expérimentale

Pour illustrer la méthode d'estimation de la réduction de variance proposée ici, on se place sur un domaine $1D$ périodique (un cercle) décomposé sur une grille régulière. Une seule variable est considérée. La variance d'erreur d'ébauche est fixée à $\sigma^b = 1$, et la fonction de corrélation est construite homogène et isotrope. Les corrélations sont prises en compte selon l'expression $\rho_L^{\mathbf{x}}(\delta\mathbf{x})$ où $\mathbf{x}$ est le point du cercle, $\delta\mathbf{x}$ la distance à ce point et $L$ la longueur de portée. Pour cette expérience, 3 fonctions sont utilisées, une fonction gaussienne, une fonction de soar et une fonction suivant Gaspari and Cohn (1999) (Eq. 4.10). Les corrélations obtenues avec chacune des 3 fonctions sont représentées sur la figure 3.1.

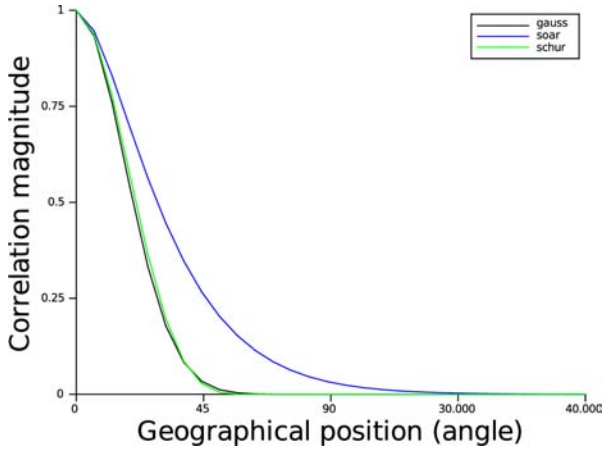



FIGURE 3.1 – Comparaison des fonctions de corrélation obtenues avec les 3 fonctions utilisées, gaussiènne, soar et G&C4.10.

Pour obtenir des corrélations hétérogènes, on utilise la méthode présentée dans Pannekoucke et al. (2007). Une transformée de Schmidt est appliquée pour déformer la grille du domaine (Courtier and Geleyn, 1988) à partir de l'équation

$$h(x) = a\left[\pi - 2\tan^{-1}\left\{\frac{1}{c}\tan\left(\frac{\pi}{2} - \frac{x}{2a}\right)\right\}\right]. \tag{3.12}$$

avec un coefficient d'étirement de $c = 2.4$. La fonction d'étirement est inversible, la fonction de corrélation devient alors,

$$\rho^{\mathbf{x}}(\delta\mathbf{x}) = \rho_L^{h^{-1}(\mathbf{x})}\left\{h^{-1}(\mathbf{x}+\delta\mathbf{x}) - h^{-1}(\mathbf{x})\right\}. \tag{3.13}$$

La figure 3.2 montre la répartition des longueurs de portée pour la fonction de corrélation gaussienne. Cette méthode de construction de corrélations hétérogènes à partir de corrélation homogènes permet de s'éloigner d'un cas idéalisé trop simple, tout en conservant la propriété définie et positive de la matrice de covariance d'erreur $\mathbf{B}$. Les longueurs de portée utilisées pour les 3 fonctions sont identiques, c'est-à-dire que le coefficient du développement de Taylor à l'ordre 2 en $\mathbf{x}$ de la fonction de corrélation est identique pour chaque fonction. Cependant, comme le montre la figure 3.1 la fonction soar a une décroissance plus faible due à la présence d'un terme d'ordre 3 positif. Définir une longueur de portée ne suffit donc pas pour décrire complètement la forme de la fonction de corrélation. Ceci explique le choix des fonctions utilisées pour l'expérience.

La figure 3.3 montre les réductions de variance obtenues par l'assimilation d'une observation en chaque point de grille. Les graphes sont des composites où chaque point représente la réduction de variance induite par l'assimilation d'une observation à ce point de grille. Les courbes noires sont obtenues par un calcul explicite de l'équation (3.2) en calculant $||\mathbf{B}(j,:)||_2^2$ à partir des expressions analytiques de $\rho^{\mathbf{x}}$. Les courbes bleues sont obtenues à partir de l'équation (3.9) et du calcul du tenseur local de diffusion .

Pour les courbes analytiques (bleues) le calcul suppose donc que les fonctions de corrélation sont gaussiennes. On remarque un très bon accord entre les 2 courbes pour les fonctions de corrélation gaussienne et G&C4.10. Un écart est cependant visible lorsque les longueurs de portée sont très faibles, autour de l'angle 180 degrés, à cause de la discrétisation numérique. L'écart est cependant beaucoup plus flagrant pour la fonction soar. Comme indiqué précédemment, la décroissance de la fonction soar est plus faible que la gaussienne. Lorsque l'on effectue le calcul de l'intégrale, on obtient donc une valeur plus forte de réduction de variance. Dans ce cas, l'équation (3.9) sous-estime l'impact de l'observation à cause de son hypothèse de gaussiannité.

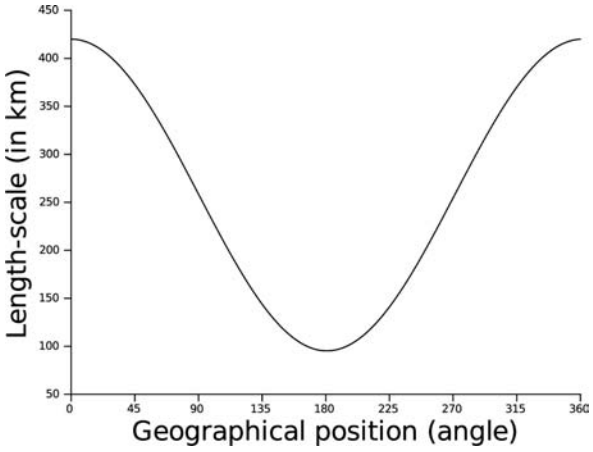



FIGURE 3.2 – Longueurs de portée sur le cercle après l'application de la fonction d'étirement. La fonction d'étirement permet d'obtenir des fonctions de corrélation non homogènes.

## 3.3 Conclusion

Une expression analytique a été présentée dans ce chapitre pour exprimer la réduction de variance d'erreur d'analyse associée à l'assimilation d'une observation. Ce calcul ne nécessite pas de réaliser une assimilation supplémentaire et ne nécessite que des quantités connues sur les matrices de covariance d'erreur (d'ébauche ou d'analyse selon le contexte). Cette méthode très simple permet de réaliser une première évaluation de la réduction d'incertitude associée à une observation supplémentaire. Des applications possibles sont par exemple la comparaison de stratégies de déploiement d'observations supplémentaires, ou une optimisation du choix des observations à travers le screening (identifier les zones où les observations auront un faible impact pour adapter les critères de sélection).

Pour une application dans le cadre de l'observation adaptative, une méthode basique pour prendre en compte la prévision de la réduction de variance pourrait être réalisée en associant le gradient d'une fonction score à la matrice de covariance d'erreur (voir l'annexe E). Cependant, une telle méthode reposerait sur des hypothèses fortes supplémentaires portant sur la modélisation de $\mathbf{A}$ et nécessiterait un ensemble de grande dimension pour représenter le transport des incertitudes. Des méthodes de calcul plus avancées pour traiter ces aspects existent. Elles sont présentées dans le chapitre suivant.

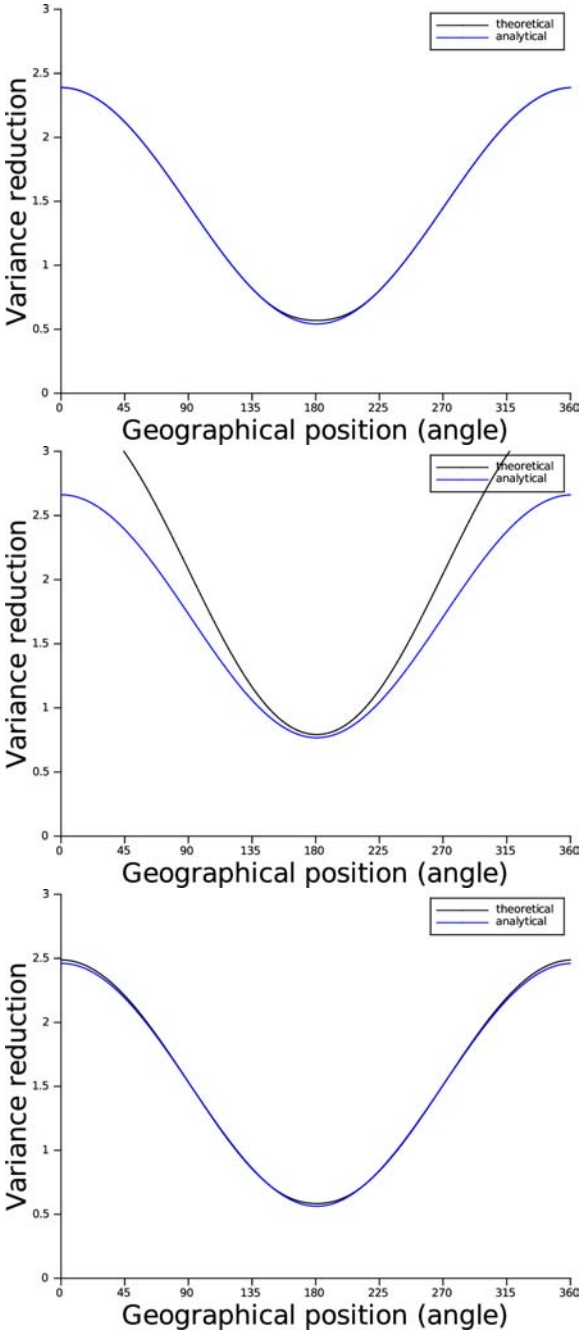



FIGURE 3.3 – Réduction de variance calculée à partir de l'équation (3.9) associée à l'assimilation d'une observation pour chaque point de grille (en bleu), comparée à la valeur théorique (en noir). Les fonctions de corrélation utilisées sont : gaussienne (haut), soar (milieu) et G&C4.10 (bas).

# Chapitre 4

# Sensibilité aux conditions initiales

Ce chapitre brosse un panorama des approches numériques utilisant des méthodes de sensibilité de la prévision aux conditions initiales, en particulier aux paramètres qui influent sur la prévision via le système d'assimilation de données. On se focalise ainsi sur des méthodes de sensibilité qui prennent en compte à la fois la dynamique atmosphérique et les observations qui influent sur la représentation de cette dynamique. Ces approches ont joué un rôle important dans l'avènement de l'observation adaptative mais aussi dans l'estimation de l'impact des observations sur l'assimilation de données. Ces approches s'appuient sur un formalisme linéaire qui permet d'accéder à l'information de sensibilité par un calcul unique, contrairement aux approches empiriques. Les sensibilités décrites dans ce chapitre utilisent soit la technologie de l'adjoint, soit des approches ensemblistes. Presque toutes sont décrites dans la littérature, sauf la VRF (*Variance Reduction Field*) qui est présentée en seconde partie et constitue un des résultats de cette thèse.

L'étude d'impact d'observation et l'observation adaptative décrivent tout un ensemble de techniques mêlant les aspects dynamiques et le système d'assimilation de données. La première partie de ce chapitre 4.1 est consacrée à la présentation du contexte et du formalisme commun entre ces deux familles de méthodes. Ensuite les méthodes d'estimation d'impact d'observation 4.2 et d'observation adaptative sont présentées 4.3. Pour finir une présentation des différentes utilisations pratiques de ces méthodes en faite en section 4.4.

## 4.1 Généralité et objectifs

### 4.1.1 Formalisme général

Le formalisme de la sensibilité est commun aux techniques d'étude d'impact d'observation et d'observation adaptative. Une présentation générale des différentes techniques est disponible dans Ancell and Hakim (2007). L'idée principale des méthodes de sensibilité est d'exprimer les variations d'une fonction score, notée $S$ et définie à un instant dit de vérification et noté $t_{verif}$, en fonction des grandeurs à un instant d'observation, $t_{obs}$. $S$ est définie de manière à mesurer l'erreur de prévision à une échéance donnée. La valeur de $S$ décroît lorsque l'état prévu $\boldsymbol{x}$ se rapproche de l'état vrai $\boldsymbol{x}^t$. Un exemple de fonction score est donné par la fonction quadratique

$$S(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x^t})^T \boldsymbol{E}(\boldsymbol{x} - \boldsymbol{x^t}),$$

où $\langle ., . \rangle_{\boldsymbol{E}}$ est le produit scalaire de métrique $\boldsymbol{E}$. Cette norme détermine la distance entre la prévision et l'état vrai à $t_{verif}$. L'état vrai $\boldsymbol{x^t}$ n'est pas connu précisément. La valeur de cette fonction ne peut donc être que approchée. Cependant, il est possible de s'intéresser à ses variations. L'objectif est donc de relier un $\delta S$ à une variation de l'état prévu.

**Sensibilité de la fonction score aux variations de l'état analysé**

La relation entre ce $\delta S$ et une variation d'analyse $\delta(\boldsymbol{x}^a)$ est obtenue par un développement de Taylor au premier ordre

$$\begin{aligned}\delta S &= S(\boldsymbol{x^b}(t_{verif}) + \delta(\boldsymbol{x}(t_{verif}))) - S(\boldsymbol{x^b}(t_{verif})),\\ &= [\nabla_{\boldsymbol{x^b}} S]^T \delta(\boldsymbol{x}(t_{verif})),\\ &= [\nabla_{\boldsymbol{x^b}} S]^T \boldsymbol{M}_{t_{obs}\rightarrow t_{verif}} \delta(\boldsymbol{x}(t_{obs})),\\ &= [\nabla_{\boldsymbol{x^a}} S]^T \delta(\boldsymbol{x^a}),\end{aligned}$$

avec $[\nabla_{\boldsymbol{x^a}} S] = \boldsymbol{M}^T_{t_{obs}\rightarrow t_{verif}} [\nabla_{\boldsymbol{x^b}} S]$, le gradient de $S$ par rapport à l'analyse. Cette forme du gradient est celle introduite par Rabier et al. (1996) dans leurs travaux fondateurs sur l'identification des erreurs dans les conditions initiales des prévisions.

On attire l'attention du lecteur sur le fait que dans la suite de ce chapitre, la notation $\delta(\boldsymbol{x}^a)$ est utilisée pour désigner une variation de l'état analysé. Il ne faut pas confondre cette variation avec l'incrément d'analyse dont la notation usuelle en assimilation de données est $\delta\boldsymbol{x}^a$

En supposant que le modèle adjoint peut s'appliquer entre $t_{obs}$ et $t_{verif}$, une variation de l'état analysé à $t_{obs}$ est donc directement reliée à une variation du score à $t_{verif}$ par la relation

$$\delta S = [\nabla_{\boldsymbol{x^a}} S]^T \delta(\boldsymbol{x^a}). \tag{4.1}$$

C'est à partir de ce point (la relation (4.1)), que les méthodes d'impact des observations et de calcul de prévisibilité divergent du fait de leurs cadres d'utilisation.

**Les deux contextes d'utilisation du calcul de sensibilité**

Deux cas sont à distinguer dans l'utilisation du calcul de sensibilité : le cas diagnostique et le cas pronostique.

Dans le cas diagnostique, on s'intéresse à la valeur de la fonction score que l'on cherche à expliquer par l'influence des observations dans la fabrication de l'état initial de la prévision. La fonction est calculée grâce à un état de référence, qui sert d'état vrai $\boldsymbol{x^t}$. En pratique on peut utiliser une analyse dite de vérification (c'est-à-dire la dernière analyse disponible à $t_{verif}$).

Dans le cadre pronostique, c'est-à-dire le cadre de l'observation adaptative, on se projette dans le futur. Les valeurs des observations ne sont logiquement pas connues, mais des hypothèses sur les statistiques sont possibles. Ainsi, la valeur de la fonction score ne peut être calculée, contrairement à sa variance. Et celle-ci est suffisante pour décider comment rajouter des observations choisies pour leur capacité à optimiser la réduction de cette variance.

Dans la suite, les deux contextes d'utilisation sont abordés successivement en commençant par le plus simple : toutes les données sont connues à posteriori et pour lequel l'approche en sensibilité permet d'accéder à une information complexe et non-triviale reliant qualité de la prévision et observations.

## 4.2 Méthodes diagnostiques : impact des observations

Dans cette section, on détaille des techniques déjà décrites dans la littérature et qui permettent d'exprimer l'impact des observations sur la prévision. Toutes peuvent se référer à la figure 4.1 qui place sur un diagramme les différents paramètres considérés. On cherche à exprimer en fonction des ingrédients à $t_{obs}$ des variations de la fonction $S$ à $t_{verif}$.

En reprenant l'équation (2.4), il est possible d'exprimer la relation (4.1) en fonction des différents paramètres d'entrée comme une dérivée composée

$$\nabla_{\boldsymbol{u}} S = \nabla_{\boldsymbol{u}} \boldsymbol{x}^a \nabla_{\boldsymbol{x}^a} S,$$

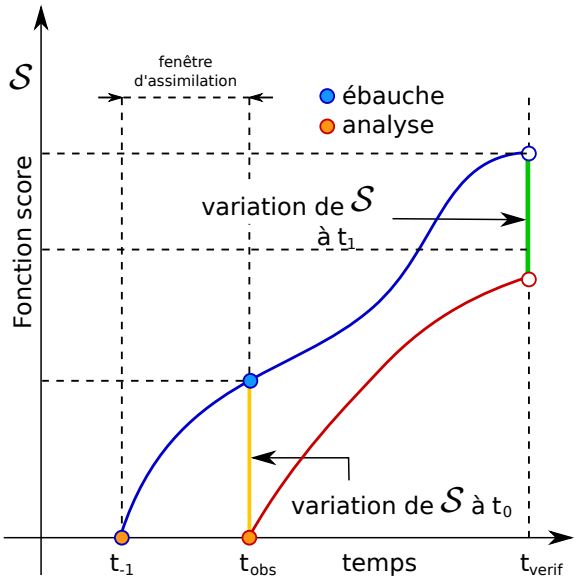



FIGURE 4.1 – Illustration de l'impact de l'analyse sur la fonction score $S$. L'ébauche (en bleu) est issue d'une assimilation précédente. L'assimilation des observations à $t_{obs}$ (ou $t_0$) permet de réduire la valeur de $S$ à $t_{verif}$ (ou $t_1$). La valeur de référence pour calculer $S_{t_{verif}}$ sera l'analyse $\mathbf{x}^a(t_{verif})$.

où $\boldsymbol{u}$ peut représenter $\boldsymbol{y^o}$, $\boldsymbol{x^b}$, $\boldsymbol{B}$ ou $\boldsymbol{R}$. Il est donc possible de calculer le gradient de $S$ par rapport à un type d'observation particulier, à l'ébauche ou par rapport aux variances théoriques, le détail des calculs est donné dans Daescu (2008). Pour les observations, ce gradient s'exprime par

$$\begin{aligned} \nabla_{\boldsymbol{y^o}} S &= \boldsymbol{K}^T \nabla_{\boldsymbol{x^a}} S, \\ &= \boldsymbol{R}^{-1} \boldsymbol{H} \boldsymbol{A} \nabla_{\boldsymbol{x^a}} S, \end{aligned}$$

où $\boldsymbol{K}^T$ est l'adjoint de la matrice de gain du système d'assimilation. Cette sensibilité est définie dans l'espace des observations assimilées : on connaît la sensibilité de $S$ pour chaque observation (ou sous groupe d'observation). L'expression de $\boldsymbol{K}^T$ est possible dans le cadre d'un 4D-var (Trémolet, 2008) mais nécessite la prise en compte des boucles externes et boucles internes de la minimisation. La principale difficulté réside dans le calcul de $\boldsymbol{A}\nabla_{\boldsymbol{x^a}} S$ .

Il est alors possible de déterminer l'impact d'une observation sur la prévision en effectuant le produit par le vecteur d'innovation, $\delta S = [\nabla_{\boldsymbol{y^o}} S]^T \delta(\boldsymbol{d})$, où $\delta(\boldsymbol{d})$ est le vecteur d'innovation associé aux observations considérées (Gelaro and Zhu (2009), Cardinali (2009)). Cette forme est une estimation strictement linéaire et correspond à un développement de Taylor au premier ordre d'une variation de la valeur de $S$ liée à l'assimilation des observations (par rapport au cas où il n'y a pas eu d'assimilation).

Cependant, certaines études ont montré qu'un développement au deuxième ordre (voire plus) est nécessaire pour obtenir des résultats plus robustes, notamment pour une meilleure prise en compte des effets liés à la dynamique non linéaire du modèle et que l'on cherche à estimer avec une méthode linéaire. Nous présentons ci-dessous deux méthodes de calcul à l'ordre 2 de l'impact des observations , suivies d'une méthode basée sur un ensemble.

### 4.2.1 Formulation par développement de Taylor

Cette méthode est présentée dans Errico (2007) et repose sur un développement de Taylor à l'ordre 2. Les équations utilisent une forme indiciaire pour simplifier les écritures. On cherche à exprimer la variance de la fonction score $S$ en fonction d'une perturbation de la condition initiale, $\delta S(\delta(\boldsymbol{x}^0)) = S(\mathcal{M}(\boldsymbol{x}^0 +$

$\delta(\boldsymbol{x}^0))) - S(\mathcal{M}(\boldsymbol{x}^0))$. Avec les notations suivantes,

$$\begin{aligned} \forall i, \boldsymbol{M}\delta(\boldsymbol{x}^0) &= \sum_j m_{ij}\delta x_j^0 \\ \forall k, \frac{\partial S}{\partial x_k^0} &= \sum_{i,j} m_{ik}e_{ij}(x_j^f - x_j^t), \end{aligned}$$

avec $m_{ij}$ les composantes du modèle tangent linéaire associé à la trajectoire $\mathcal{M}(\boldsymbol{x}^b(t_{obs}))$ et $e_{ij}$ les composantes de $\boldsymbol{E}$. On note $\mathcal{M}(\boldsymbol{x}^b(t_{obs})) = \boldsymbol{x}^f$ et $\boldsymbol{x}^t(t_{verif}) = \boldsymbol{x}^t$ pour ne pas alourdir les notations. On retrouve alors l'équation (4.1) au premier ordre,

$$\delta S_1 = \sum_k \delta x_k^0 \frac{\partial S}{\partial x_k^0} = (\delta \boldsymbol{x}^0)^T \boldsymbol{M}^T \boldsymbol{E}(\boldsymbol{x}^f - \boldsymbol{x}^t)$$

Une deuxième dérivation est ensuite réalisée,

$$\forall n, k, \frac{\partial^2 S}{\partial x_n^0 \partial x_k^0} = \sum_{i,j} \frac{\partial m_{ik}}{\partial x_n^0} e_{ij}(x_j^f - x_j^t) + m_{ik}e_{ij}m_{jn}. \tag{4.2}$$

La variation à l'ordre 2 s'écrit alors

$$\begin{aligned} \delta S_2 &= \sum_n \delta x_n^0 \left( \frac{\partial S}{\partial x_n^0} + \frac{1}{2}\sum_k \frac{\partial^2 S}{\partial x_n^0 \partial x_k^0}\delta x_k^0 \right), \\ &= \frac{1}{2}\sum_n \delta x_n^0 \left( \sum_{i,j}(m_{in} + \sum_k \frac{\partial m_{ik}}{\partial x_n^0}\delta x_k^0)e_{ij}(x_j^f - x_j^t) + \sum_{i,j} m_{in}e_{ij}(\sum_k m_{jk}\delta x_k^0 + x_j^f - x_j^t) \right) \end{aligned} \tag{4.3}$$

Une nouvelle étape est maintenant de considérer $\delta\boldsymbol{x}^0$ comme le vecteur d'incrément d'analyse de manière à exprimer $\delta S_2$ est fonction de l'ébauche $\boldsymbol{x}^b(t_{obs})$ et de l'analyse $\boldsymbol{x}^a(t_{obs})$. Tant que l'hypothèse linéaire reste valide le long de la trajectoire le dernier terme de l'équation (4.3) peut s'écrire,

$$\sum_k m_{jk}\delta x_k^0 + x_j^f - x_j^t = \delta x_j^f + x_j^f - x_j^t = x_j^{af} - x_j^t,$$

avec $\boldsymbol{x}^{af}$ la prévision issue de l'analyse $\boldsymbol{x}^a(t_{obs})$. Le premier terme de la parenthèse (4.3) peut aussi être simplifié en écrivant,

$$\begin{aligned} \sum_n \sum_{i,j}(m_{in} + \sum_k \frac{\partial m_{ik}}{\partial x_n^0}\delta x_k^0)\delta x_n^0 &= (\boldsymbol{M} + \delta\boldsymbol{M})\delta(\boldsymbol{x}^0) \\ &= \boldsymbol{M}_a\delta(\boldsymbol{x}^0), \end{aligned}$$

en notant $\boldsymbol{M}_a$ le modèle tangent linéaire issu de $\boldsymbol{x}^a(t_{obs})$, et $\boldsymbol{M}_b$ le modèle tangent linéaire issu de $\boldsymbol{x}^b(t_{obs})$. L'écriture ci-dessus permet de s'affranchir de la dérivée à l'ordre deux qui correspond formellement à un adjoint du second ordre (Le Dimet et al., 2002). Un tel outil n'existe pas pour les grands systèmes de PNT. L'équation (4.3) peut alors s'écrire sous la forme,

$$\delta S_2 = \frac{1}{2}(\boldsymbol{x}^a - \boldsymbol{x}^b)^T(\boldsymbol{M}_b^T E(\boldsymbol{x}^{af} - \boldsymbol{x}^t) + \boldsymbol{M}_a^T E(\boldsymbol{x}^f - \boldsymbol{x}^t)). \tag{4.4}$$

On obtient alors la réduction d'erreur de prévision due à l'assimilation de données. Pour effectuer ce calcul, il est nécessaire de connaître $\boldsymbol{x}^b$, la prévision $\boldsymbol{x}^f$ et le modèle adjoint associé à cette trajectoire. Il faut aussi connaître $\boldsymbol{x}^a$, la prévision qui en découle.

On attire l'attention du lecteur sur le fait que la première parenthèse de l'équation (4.4), correspondant à l'écart entre l'analyse et l'ébauche, peut être directement reliée aux observations : il peut s'écrire sous la forme $\boldsymbol{Kd}$. Ainsi, on fait apparaître le lien entre variation de la valeur observée et variation de la fonction score. C'est cette forme de l'équation qui est utilisée dans de nombreuses études sur l'impact des observations, telle que celle récente de Boullot et al. (2015). Cette approche est connue sous le nom de FSOI (*Forecast Sensitivity to Observation Impact*).

### 4.2.2 Méthode Géométrique : Daescu and Todling

La deuxième méthode présentée est une méthode dite géométrique (Daescu and Todling, 2009) qui s'appuie sur une approche paramétrique. Une telle approche est un atout pour accéder à des formulations correspondant à des développements d'ordre supérieur à 2 en s'affranchissant complètement d'un adjoint d'ordre supérieur à 1. La valeur $\delta S = S(\mathcal{M}(\boldsymbol{x}^a)) - S(\mathcal{M}(\boldsymbol{x}^b))$ est définie comme une fonction à paramètre sur le segment $[\boldsymbol{x}^b, \boldsymbol{x}^a]$, en notant $\boldsymbol{x}^0 = \boldsymbol{x}^b + p(\boldsymbol{x}^a - \boldsymbol{x}^b)$, avec $p \in [0,1]$. On peut alors définir une nouvelle fonction $\hat{S}(p) = S(\mathcal{M}(\boldsymbol{x}^0))$. $\delta S$ s'écrit alors,

$$\delta S = \hat{S}(1) - \hat{S}(0) = \int_0^1 \hat{S}'(p)dp. \tag{4.5}$$

A partir de différentes méthodes d'approximation d'une intégrale, il est possible de calculer une estimation de $\delta S$. On pose la fonction $\boldsymbol{g}(\boldsymbol{x}^0)^T = (\mathcal{M}(\boldsymbol{x}^0) - \boldsymbol{x}^t)^T \boldsymbol{E}\boldsymbol{M}_{\boldsymbol{x}^0}$, avec $\boldsymbol{M}_{\boldsymbol{x}^0}$ le modèle tangent linéaire issu de la trajectoire $\boldsymbol{x}^0$et la fonction à intégrer est $\hat{S}'(p) = \boldsymbol{g}(\boldsymbol{x}^0)^T(\boldsymbol{x}^a - \boldsymbol{x}^b)$. Avec la méthode des trapèzes, l'intégrale devient,

$$\delta S_{trap} = \frac{1}{2}(\boldsymbol{g}(\boldsymbol{x}^a) + \boldsymbol{g}(\boldsymbol{x}^b))^T(\boldsymbol{x}^a - \boldsymbol{x}^b). \tag{4.6}$$

Comme pour le calcul à partir du développement de Taylor, il est nécessaire de connaître l'état initial, l'état prévu et le modèle adjoint associé aux deux trajectoires : celle issue de l'ébauche et celle issue de l'analyse. Il faut aussi choisir un état de référence pris comme la vérité. On note cependant que les deux formules ne sont pas équivalentes. Effectivement, le développement de Taylor mélange des dérivées faisant références à des trajectoires distinctes (termes croisés), alors que l'approche géométrique utilise des dérivées basées sur une seule trajectoire à chaque fois.

D'autres méthodes comme celle de Simpson ou du point milieu peuvent être utilisées. Celles ci sont présentées dans Langland and Baker (2004), Daescu and Todling (2009). Elles permettent d'accéder à des développements d'ordre supérieur en tirant partie de la formulation paramétrique. Ainsi, le point milieu nécessite de calculer une $3^{\text{ème}}$ trajectoire issue d'un état à mi-chemin entre ébauche et analyse, correspondant au point défini par le paramètre $p = \frac{1}{2}$, mais cela donne accès à une développement d'une précision équivalente à l'ordre 3. Le détail des calculs et des commentaires sur les différentes approches sont décrits dans Fischer (2014).

### 4.2.3 Ensemble Forecast Sensitivity to Observations

Les deux méthodes précédentes ont été développées pour une implémentation dans un système déterministe et doté d'un modèle adjoint. Celle présentée maintenant reprend les développements de Langland and Baker (2004) pour réaliser les calculs à partir d'un ensemble de prévisions et d'assimilations (Kalnay et al., 2012).

On repart de la définition de $\delta S$ en notant $\boldsymbol{e}_a = \overline{\mathcal{M}(\boldsymbol{x}^a)} - \boldsymbol{x}^t$ et $\boldsymbol{e}_b = \overline{\mathcal{M}(\boldsymbol{x}^b)} - \boldsymbol{x}^t$, où $\overline{\mathcal{M}(\boldsymbol{x})}$ est la moyenne des trajectoires non linéaires de l'ensembe,

$$\begin{aligned}
\delta S &= \frac{1}{2}(\boldsymbol{e}_a^T \boldsymbol{E}\boldsymbol{e}_a - \boldsymbol{e}_b^T \boldsymbol{E}\boldsymbol{e}_b) \\
&= \frac{1}{2}(\boldsymbol{e}_a - \boldsymbol{e}_b)^T \boldsymbol{E}(\boldsymbol{e}_a + \boldsymbol{e}_b) \\
&= \frac{1}{2}\left[\mathcal{M}(\boldsymbol{x}^a) - \mathcal{M}(\boldsymbol{x}^b)\right]^T \boldsymbol{E}(\boldsymbol{e}_a + \boldsymbol{e}_b) \\
&= \frac{1}{2}\left[\boldsymbol{M}(\boldsymbol{x}^a - \boldsymbol{x}^b)\right]^T \boldsymbol{E}(\boldsymbol{e}_a + \boldsymbol{e}_b) \\
&= \frac{1}{2}\left[\boldsymbol{M}\boldsymbol{K}\delta(\boldsymbol{y}^o)\right]^T \boldsymbol{E}(\boldsymbol{e}_a + \boldsymbol{e}_b)
\end{aligned} \tag{4.7}$$

avec $\boldsymbol{K}^T$ l'adjoint du système d'assimilation de données et $\delta(\boldsymbol{y}^o) = \boldsymbol{y}^o - \mathcal{H}(\boldsymbol{x}^b)$.

En reprenant les calculs introduits pour l'ETKF (2.3.2), on peut réécrire la matrice de gain sous la forme, $\boldsymbol{K} = \frac{1}{N_e-1}\boldsymbol{X}^a\boldsymbol{X}^{aT}\boldsymbol{H}^T\boldsymbol{R}^{-1}$. Il est alors possible de remplacer la multiplication par l'adjoint de manière implicite à travers le transport des perturbations par l'ensemble,

$$\boldsymbol{M}\boldsymbol{K} = \frac{1}{N_e-1}\boldsymbol{M}\boldsymbol{X}^a\boldsymbol{X}^{aT}\boldsymbol{H}^T\boldsymbol{R}^{-1} \approx \frac{1}{N_e-1}\boldsymbol{X}^f\boldsymbol{Y}^{aT}\boldsymbol{R}^{-1}$$

avec $\boldsymbol{Y}^a = \boldsymbol{H}\boldsymbol{X}^a$.

L'équation (4.7) devient alors,

$$\delta S = \frac{1}{N_e-1}\delta(\boldsymbol{y}^o)^T\boldsymbol{R}^{-1}\boldsymbol{Y}^a\boldsymbol{X}^{fT}\boldsymbol{E}(\boldsymbol{e}_a + \boldsymbol{e}_b) \tag{4.8}$$

Comme pour les méthodes précédentes, $\boldsymbol{e}_a$ et $\boldsymbol{e}_b$ utilisent l'analyse issue de l'ensemble du réseau ultérieur pour être calculés. En choisissant les observations à inclure dans $\delta(\boldsymbol{y}^o)$ (restriction de $\boldsymbol{y}^o$ et $\mathcal{H}$), il est possible de déterminer l'impact des différents types d'observations. Il est intéressant de noter que l'équation (4.8) repose sur une hypothèse de transport linéaire des perturbations par l'ensemble.

### 4.2.4 Limite des méthodes d'impact d'observation pour l'observation adaptative

Le travail présenté dans ce manuscrit ne relève pas de l'impact d'observation au sens diagnostic comme présenté dans la section précédente. Cependant les nombreuses études réalisées dans ce domaine présentent des analogies avec certaines approches numériques pour l'observation adaptative. Il est donc intéressant de présenter ces développements et la manière dont ils peuvent ou non être repris pour l'observation adaptative.

Chacune des trois méthodes présentées ci-dessus utilise l'analyse $\boldsymbol{x}^a$ pour réaliser les calculs. Ce qui implique que les observations soient connues, à la fois de par leur position (influence sur $\mathcal{H}$), en terme de valeur (influence sur l'innovation) et en terme d'incertitude (influence sur $\boldsymbol{R}$), ce qui est évidemment le cas pour un travail diagnostic mais ne l'est pas pour un calcul pronostique. De même, le vecteur $\boldsymbol{x}^t$ n'étant jamais connu, c'est l'analyse $\boldsymbol{x}^a(t_{verif})$ qui est utilisée comme référence. Le calcul de l'impact des observations du réseau d'assimilation $j$ n'est donc possible que si l'assimilation $j+1$ a été réalisée. Ceci limite donc les applications opérationnelles pour adapter le système d'assimilation en temps réel. La figure 4.2 illustre le schéma d'application des deux types de méthodes avec en particulier l'instant *temps-réel* auquel ont lieu les calculs de sensibilité.

Cette limitation peut cependant être levée en utilisant des cycles d'assimilation très courts, comme dans le cas de l'assimilation de certains modèles de méso-échelle qui peut être reprise aussi fréquemment que toutes les heures. Dans ce cas, l'hypothèse que le réseau d'observation reste stationnaire est acceptable (une observation mauvaise dans l'heure précédente a toutes les chances d'être encore mauvaise une heure après). On dispose alors d'une solution en quasi temps réel d'adaptation du réseau d'observation. Pour les modèles globaux, une assimilation horaire n'est cependant pas possible, il est donc nécessaire d'introduire des hypothèses supplémentaires lorsque l'on se place dans la cadre de l'observation adaptative avec des calculs pronostiques, comme présenté dans la section suivante.

## 4.3 Méthodes pronostiques

Pour le cadre pronostique, il est nécessaire de prévoir la position des observations futures, ce qui est possible en faisant quelques hypothèses pas trop lourdes. Par contre l'anticipation des valeurs observées est beaucoup plus délicate et n'est pas réaliste. Le futur vecteur d'innovation et donc l'analyse vérifiante ne peuvent pas être utilisés pour le calcul de la fonction score ou de son gradient. C'est pourquoi la variance de $S$ est utilisée et non plus la valeur de la fonction elle même. L'objectif de l'observation adaptative n'est pas de rapprocher la prévision de l'état vrai $\boldsymbol{x}^t$ mais de réduire l'incertitude de cette prévision. On fait l'hypothèse que la prévision n'est pas biaisée. Réduire l'incertitude revient donc à améliorer la prévision en moyenne. Cette incertitude est caractérisée par la variance de $S$, $\sigma_S^2 = \mathbb{E}\left(\delta S \delta S^T\right)$. L'écart à $\boldsymbol{x}^t(t_{verif})$

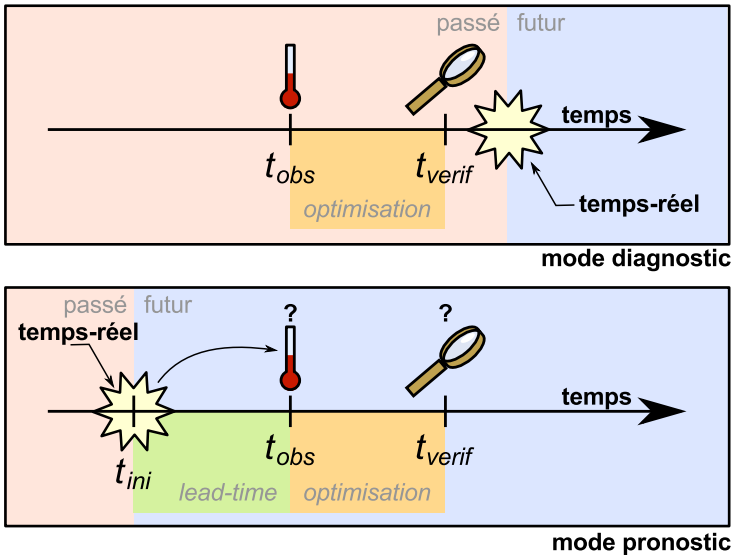



FIGURE 4.2 – Comparaison des schémas de principe des méthodes diagnostiques (impact d'observations) et des méthodes pronostiques (observation adaptative). Dans le cadre des méthodes diagnostiques, les calculs sont réalisés à posteriori de $t_{obs}$ et $t_{verif}$, tous les éléments sont donc connus (ébauche et observations à $t_{obs}$, et les analyses à $t_{obs}$ et $t_{verif}$). Pour les méthodes pronostiques, le calcul est réalisé à priori, à $t_{ini}$, de l'assimilation qui est étudiée, qui aura lien à $t_{obs}$. A $t_{ini}$ seule une ébauche issue d'une assimilation précédente est connue. Le réseau d'observation et les observations elles mêmes doivent être simulés.

n'est donc pas pris en compte pour le calcul de $S$, son écriture est alors réduite à $S(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{E}\boldsymbol{x}$. En utilisant la relation (4.1), la variance s'exprime en fonction du gradient par rapport à l'analyse et de la matrice de covariance d'erreur,

$$\sigma_S^2 = [\nabla_{\boldsymbol{x}^a} S]^T \boldsymbol{A} [\nabla_{\boldsymbol{x}^a} S]. \tag{4.9}$$

La variance de la fonction score est donc directement liée à l'incertitude sur la condition initiale. Cette incertitude est propagée le long de la trajectoire et ensuite projetée selon le gradient. La matrice $\boldsymbol{A}$ étant obtenue par l'assimilation, elle comporte l'information liée aux observations et le gradient dépend de la dynamique. En calculant la matrice $\boldsymbol{A}$ qui utilise les observations supplémentaires, il est possible de déterminer la réduction de variance due à ces observations.

L'idée de l'OA est d'adapter le réseau d'observation à la situation et aux phénomènes dont on souhaite améliorer la présivion.

Cette section présente l'observation adaptative et ses objectifs. Le formalisme du calcul de sensibilité aux observations est ensuite introduit. Enfin, la section se conclut par le développement de différentes méthodes pour calculer la sensibilité et la manière de la mettre en œuvre.

### 4.3.1 Principe de l'observation adaptative

Les observations jouent un rôle important dans la qualité des prévisions des modèles numériques. Cependant leur positionnement n'est pas toujours optimal par rapport aux phénomènes météorologiques et l'assimilation d'un grand nombre d'observation à un coût important. Le but de l'observation adaptative est de rationaliser l'impact effectif d'une observation pour permettre un calcul objectif de la localisation optimale d'une observation en fonction de l'état de l'atmosphère et du système d'assimilation (Joly, 2003). Ceci peut être effectué dans le but de déployer des observations supplémentaires ou d'optimiser le choix des observations assimilées. Ces techniques ont été développées à la fin des années 90 suite à la campagne FASTEX (Joly et al., 1999).

Le développement d'un outil pronostique impose des simplifications. En effet, le meilleur état de l'atmosphère disponible à la date de vérification $t_{verif}$ est une prévision réalisée à partir d'une analyse précédente

à $t_0$. Comme la qualité des prévisions se dégrade avec les échéances, il est délicat de choisir un instant de vérification trop éloigné de l'instant présent $t_0$. La première étape est de fixer un instant d'observation $t_{obs}$ auquel aura lieu le déploiement ($t_0 \leqslant t_{obs} \leqslant t_{verif}$).

L'objectif de l'observation adaptative est d'améliorer la prévision, en particulier la prévisibilité effective. En effet, l'état lui même de la prévision ne peut être connu car on ne sait pas à l'avance les quantités collectées par le réseau d'observation. Il est cependant possible de quantifier l'incertitude de la prévision, à partir de la variance d'une fonction score $S$ définie à $t_{verif}$. Cette variance est reliée à la matrice de covariance d'erreur de prévision $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{A}\boldsymbol{M}^T$, en supposant le modèle parfait $\boldsymbol{Q} = \boldsymbol{0}$. Les observations supplémentaires à $t_{obs}$ sont alors déterminées de manière à maximiser la réduction de variance $\sigma_S^2$. Pour ce faire il faut exprimer la variance à $t_{verif}$ en fonction des variables à $t_{obs}$. Il existe plusieurs manières de procéder dans la littérature. Ces méthodes cherchent à exprimer la matrice $\boldsymbol{A}$, dans un espace réduit (*Hessian Reduced Rank* (HRR) Leutbecher (2003)), à partir d'ensemble (*Ensemble Transform Kalman Filter* (ETKF) Bishop et al. (2001)) ou en utilisant le modèle adjoint, comme la méthode de la *Kalman Filter Sensitivity* (KFS) Bergot and Doerenbecher (2002)). Le schéma de principe de l'observation adaptative a été illustré en introduction sur la figure 1.2.

Le formalisme des équations reprend celui des méthodes diagnostiques, à la différence que l'état vrai $\boldsymbol{x}^t$ n'est pas pris en compte dans l'évaluation de $S$ (ie. $\boldsymbol{x}^t = \boldsymbol{0}$). Quatre méthodes de calcul de sensibilité sont présentées dans cette section, la *KFS* et son lien avec la *HRR*, l'application de l'ETKF à l'observation adaptative et enfin une nouvelle méthode appelée Variance Reduction Field (VRF) est proposée.

### 4.3.2 Méthode de la Kalman Filter Sensitivity

La méthode de la KFS a été introduite par Bergot and Doerenbecher (2002), les résultats sont repris à partir du formalisme présenté ci-dessus. La KFS vise à déterminer les zones géographiques où la réduction de variance de $S$ due aux nouvelles observations sera maximale. Pour cela un réseau d'observation de routine (ou conventionnel) $c$ est considéré. Ce réseau correspond à un ensemble d'observations dont on sait qu'elles seront présentes à $t_{obs}$. Le réseau d'observations supplémentaires est noté $t$ (pour *targeted* ou ciblé). Le réseau total (conventionnel et ciblé) est noté $c + t$. Les matrices de covariance d'erreur d'observation et d'opérateurs d'observation associées sont respectivement $(\boldsymbol{R}_c, \boldsymbol{H}_c)$ et $(\boldsymbol{R}_t, \boldsymbol{H}_t)$.

La matrice de covariance pour le réseau de routine $c$ s'écrit $\boldsymbol{A}_c = (\boldsymbol{B}^{-1} + \boldsymbol{H}_c^T\boldsymbol{R}_c^{-1}\boldsymbol{H}_c)^{-1}$ et la matrice pour le réseau $c+t$ : $\boldsymbol{A}_{c+t} = (\boldsymbol{B}^{-1} + \boldsymbol{H}_c^T\boldsymbol{R}_c^{-1}\boldsymbol{H}_c + \boldsymbol{H}_t^T\boldsymbol{R}_t^{-1}\boldsymbol{H}_t)^{-1}$ (le réseau supplémentaire $t$ est supposé indépendant du réseau de routine $c$). En utilisant la formule de Sherman-Morrison-Woodbury on montre que $\boldsymbol{A}_{c+t}$ peut s'écrire sous la forme

$$\boldsymbol{A}_{c+t} = \boldsymbol{A}_c - \boldsymbol{A}_c\boldsymbol{H}_t^T(\boldsymbol{R}_t + \boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{H}_t^T)^{-1}\boldsymbol{H}_t\boldsymbol{A}_c. \tag{4.10}$$

En remplaçant (4.10) dans (4.9), la réduction de variance due aux nouvelles observations s'écrit

$$\begin{aligned}(\delta\sigma_S^t)^2 &= \sigma_S^{c\,2} - \sigma_S^{c+t\,2} \\ &= (\nabla S)^T\boldsymbol{A}_c(\nabla S) - (\nabla S)^T\boldsymbol{A}_{c+t}(\nabla S) \\ &= (\nabla S)^T\boldsymbol{A}_c\boldsymbol{H}_t^T(\boldsymbol{R}_t + \boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{H}_t^T)^{-1}\boldsymbol{H}_t\boldsymbol{A}_c(\nabla S).\end{aligned} \tag{4.11}$$

Si $\boldsymbol{H}_t$ et $\boldsymbol{R}_t$ sont connus, cette relation permet de calculer directement la réduction de variance associée. Le déploiement le plus intéressant étant celui qui maximise la réduction de variance. Pour simplifier ce calcul il est nécessaire de faire des hypothèses supplémentaires pour pouvoir inverser $\boldsymbol{R}_t + \boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{H}_t^T$. C'est pourquoi pour la KFS cette formulation n'est pas utilisée directement pour obtenir un champ de sensibilité. Cependant elle permet de calculer la réduction de variance associée à des observations supplémentaires choisies. On vérifie aussi que la variance est réduite peu importe l'emplacement de l'observation supplémentaire (dans le cas optimal) puisque le terme est positif.

Dans l'application pratique de la KFS, une approche par champ de sensibilité, à la manière de la sensibilité aux observations, est effectuée comme suit.

**Champ de sensibilité de la KFS**

A partir de l'équation (4.9) il est possible de déterminer le gradient de la variance par rapport à $\nabla_{\boldsymbol{x}^a} S$, tel que $\nabla_{\boldsymbol{g}}(\sigma_S)^2 = 2\boldsymbol{A}\boldsymbol{g}$ avec $\boldsymbol{g} = \nabla_{\boldsymbol{x}^a} S$ pour simplifier les notations. En utilisant un développement de Taylor au premier ordre il est alors possible d'exprimer l'écart de variance en fonction d'une variation de $\boldsymbol{g}$ sous la forme

$$\begin{aligned} \delta(\sigma_S)^2 &= [\nabla_{\boldsymbol{g}}(\sigma_S)^2]^T \delta(\boldsymbol{g}), \\ &= 2[\boldsymbol{A}\boldsymbol{g}]^T \delta(\boldsymbol{g}). \end{aligned}$$

L'objectif de la KFS est de déterminer le lieu des observations optimales pour réduire la variance $(\sigma_S)^2$, or d'après (4.10) l'ajout d'observation diminue toujours la variance. Chercher le maximum de diminution de variance revient donc à chercher le maximum de $[\boldsymbol{A}\boldsymbol{g}]^T \delta(\boldsymbol{g})$. Une variation de gradient n'étant pas une valeur pertinente il convient de l'exprimer en fonction d'une variation d'incrément $\delta(\boldsymbol{d})$.

Le tableau ci-dessous reprend les relations utilisées pour le calcul et les approximations au premier ordre qui relient les variables.

| expression initiale | approximation au premier ordre |
|---|---|
| $\boldsymbol{x}^a(t_0) = \boldsymbol{x}^b(t_0) + \boldsymbol{K}\boldsymbol{d}$ | $\delta(\boldsymbol{x}^a) = \boldsymbol{K}\delta(\boldsymbol{d})$ |
| $\boldsymbol{x}^f(t_1) = \mathcal{M}(\boldsymbol{x}^a(t_0))$ | $\delta(\boldsymbol{x}^f) = \boldsymbol{M}\delta(\boldsymbol{x}^a)$ |
| $\nabla_{\boldsymbol{x}^f} S = \boldsymbol{E}\boldsymbol{x}^f$ | $\delta(\nabla_{\boldsymbol{x}^f} S) = \boldsymbol{E}\delta(\boldsymbol{x}^f)$ |
| $\nabla_{\boldsymbol{x}^a} S = \boldsymbol{M}^T \nabla_{\boldsymbol{x}^f} S$ | $\delta(\nabla_{\boldsymbol{x}^a} S) = \boldsymbol{M}^T \delta(\nabla_{\boldsymbol{x}^f} S)$ |

En combinant les équations la relation entre une variance de gradient $\delta(\boldsymbol{g})$ et une variation d'innovation $\delta(\boldsymbol{d})$ s'écrit $\delta(\boldsymbol{g}) = \boldsymbol{M}^T\boldsymbol{E}\boldsymbol{M}\boldsymbol{K}\delta(\boldsymbol{d})$. Ainsi rechercher le maximum de réduction de variance revient donc à chercher le maximum de

$$\begin{aligned} \max(\delta(\sigma_S)^2) &= \max\left(\langle \boldsymbol{A}\boldsymbol{g}, \boldsymbol{M}^T\boldsymbol{E}\boldsymbol{M}\boldsymbol{K}\delta(\boldsymbol{d})\rangle\right), \\ &= \max\left(\langle \boldsymbol{K}^T\boldsymbol{M}^T\boldsymbol{E}^T\boldsymbol{M}\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S, \delta(\boldsymbol{d})\rangle\right). \end{aligned}$$

A ce niveau il n'est pas possible de connaître la distribution des vecteurs d'innovation, ils sont donc supposés équiprobables. Le champ de sensibilité se résume donc au calcul du terme de gauche dans le produit scalaire. L'évaluation de ce terme est cependant assez coûteuse, il faut en effet une intégration du modèle direct et un de l'adjoint. De plus pour obtenir un résultat correct il faudrait utiliser l'adjoint du système d'assimilation ($\boldsymbol{K}^T$) qui est difficilement accessible pour prendre en compte l'adjoint complet (Trémolet, 2008), mais une estimation de $\mathbf{R}^{-1}\mathbf{H}\mathbf{A}$ reste calculable.

La pratique a montré que le champ $\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S$ contient suffisamment d'information pour être utilisé, ce qui correspond au gradient de $S$ par rapport aux observations (à projection et normalisation près). Le produit $\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S$ peut être vu comme le produit des zones d'incertitudes du modèle par les zones dynamiques, ce qui renseigne bien sur l'incertitude future de la prévision. Dans ce cas l'emplacement des observations supplémentaires est choisi comme le maximum du champ $|\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S|$. Le développement de la HRR permet de justifier aussi ce choix de champ de sensibilité.

### 4.3.3 Lien entre la KFS et la HRR

L'approche retenue par Leutbecher (2003) se rapproche d'avantage des méthodes de vecteurs singuliers. L'objectif est de calculer la condition initiale associée à une forte incertitude qui va contribuer le plus à la valeur de $S$. On définie la fonction $\mathcal{F}(\delta(\boldsymbol{x})) = \frac{\|\boldsymbol{M}\delta(\boldsymbol{x})\|^2_{\boldsymbol{E}}}{\|\delta(\boldsymbol{x})\|^2_{\boldsymbol{A}^{-1}}} = \frac{S(\delta(\boldsymbol{x}))}{\|\delta(\boldsymbol{x})\|^2_{\boldsymbol{A}^{-1}}}$ que l'on va chercher à maximiser. $\delta(\boldsymbol{x})$ est içi une perturbation par rapport à l'analyse $\boldsymbol{x}^a$. Maximiser $\mathcal{F}(\delta(\boldsymbol{x}))$ revient à chercher la perturbation de l'analyse qui est associé à une forte valeur de $S(\delta(\boldsymbol{x}))$ et à des fortes valeurs de $\boldsymbol{A}$, donc avec une forte incertitude sur la condition initiale.

Cela revient à chercher les valeurs propres de la forme,

$$\begin{aligned} \boldsymbol{M}^T\boldsymbol{E}\boldsymbol{M}\boldsymbol{v}_i &= \sigma_i^2\boldsymbol{A}^{-1}\boldsymbol{v}_i \\ \boldsymbol{A}\boldsymbol{M}^T\boldsymbol{E}\boldsymbol{M}\boldsymbol{v}_i &= \lambda_i\boldsymbol{v}_i \\ \boldsymbol{A}\nabla_{\boldsymbol{x}^a}S\left(\boldsymbol{v_i}\right) &= \lambda_i\boldsymbol{v}_i \end{aligned} \tag{4.12}$$

En posant $\lambda_1$ pour la plus grande valeur propre (en valeur absolue), $\boldsymbol{v}_1$ est alors la direction de la perturbation pour laquelle une variation de l'intensité aura un impact important sur $\mathcal{F}$. Pour la KFS, le champ de sensibilité est calculé à partir de l'ébauche $\boldsymbol{A}\nabla_{\boldsymbol{x}^a}S$. Pour une perturbation d'ébauche quelconque, on peut s'attendre à ce que $\boldsymbol{v}_1$ domine le champ de sensibilité. En plaçant des observations à la position du maximum de $\boldsymbol{A}\nabla_{\boldsymbol{x}^a}$, on va modifier les perturbations d'analyse ayant un fort impact sur $\mathcal{F}$ et donc sur $S$. Comme dans le cas optimal, l'ajout d'observation doit diminuer la variance, l'impact sera donc bénéfique.

Dans la formulation présentée par Leutbecher (2003), la HRR n'utilise pas de champ de sensibilité. La diagonalisation de la Héssienne est utilisée pour simplifier les calculs pour l'intégration des observations supplémentaires. Plusieurs réseaux d'observations supplémentaires sont alors comparés pour déterminer celui qui est le plus avantageux. Il peut s'agir de trajectoire d'avion, de ballon sonde ou de drone d'observation. Les développements sont très proches de ceux réalisés dans le cadre de l'ETKF.

### 4.3.4 Application de l'ETKF à l'observation adaptative

L'ETKF a été présenté au chapitre 2 comme un système d'assimilation d'ensemble. Celui-ci peut aussi être utilisé dans la cadre de l'observation adaptative (Bishop et al. (2001) et Majumdar et al. (2001)). En effet, la formalisme de l'ETKF permet de rajouter très simplement des observations supplémentaires à partir de l'analyse de routine.

En reprenant les équations décrites dans l'annexe C, on peut définir une matrice des perturbations d'analyse $\boldsymbol{X}^a$ à partir de l'ensemble. On calcule alors $\boldsymbol{Y}_t^a = \boldsymbol{H}_t\boldsymbol{X}^a$ et on effectue l'inversion de la Héssienne dans l'espace réduit pour tenir compte des observations supplémentaires.

On définit la fonction de score comme la différence des matrices de covariance d'erreur de prévision entre l'analyse de routine et celle contenant les observations supplémentaires :

$$\begin{aligned} \boldsymbol{S} &= \boldsymbol{B}_c(t_{verif}) - \boldsymbol{B}_{c+t}(t_{verif}) \\ &= \boldsymbol{M}\boldsymbol{A}_c\boldsymbol{M}^T - \boldsymbol{M}\boldsymbol{A}_{c+t}\boldsymbol{M}^T \\ &= \boldsymbol{M}(\boldsymbol{A}_c - \boldsymbol{A}_{c+t})\boldsymbol{M}^T \\ &= \boldsymbol{M}\boldsymbol{A}_c\boldsymbol{H}_t^T(\boldsymbol{R}_t + \boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{H}_t^T)^{-1}\boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{M}^T \end{aligned} \tag{4.13}$$

On remplace alors la matrice de covariance d'erreur d'analyse pour sa décomposition dans la base réduite. La prise en compte de la dynamique est possible grâce à l'ensemble qui fait implicitement évoluer les perturbations, $\boldsymbol{M}\boldsymbol{A_c} = 1/(N_e-1)\boldsymbol{M}\boldsymbol{X}_c^a\boldsymbol{X}_c^{aT} = 1/(N_e-1)\boldsymbol{X}_c^f\boldsymbol{X}_c^{aT}$, où $\boldsymbol{X}_c^f$ est construit à partir de l'ensemble de prévision. Le terme à inverser devient, $\boldsymbol{R}_t + \boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{H}_t^T = \boldsymbol{R}_t + 1/(N_e-1)\boldsymbol{H}_t\boldsymbol{X}_c^a(\boldsymbol{H}_t\boldsymbol{X}_c^a)^T$ qui est de la dimension de l'espace des observations supplémentaires, qui reste faible, et donc facilement inversible. On obtient, au final :

$$\boldsymbol{S} = \frac{1}{(N_e-1)^2}[\boldsymbol{H}_t\boldsymbol{X}_c^a\boldsymbol{X}_c^{fT}]^T\left(\boldsymbol{R}_t + \frac{1}{N_e-1}(\boldsymbol{H}_t\boldsymbol{X}_c^a)(\boldsymbol{H}_t\boldsymbol{X}_c^a)^T\right)^{-1}[\boldsymbol{H}_t\boldsymbol{X}_c^a\boldsymbol{X}_c^{fT}] \tag{4.14}$$

Le calcul de $\boldsymbol{S}$ se résume alors à des produits de dimension $N_e$ ou de la dimension de l'espace des observations supplémentaires. Pour obtenir un champ de réduction de variance de $S$, il suffit de prendre la diagonale de $\boldsymbol{S}$. Dans ce cas, la matrice $\boldsymbol{E}$ (associée à $S$) est supposée être l'identité.

L'utilisation des ensembles permet de simplifier les calculs. Cependant la mise en œuvre de l'ETKF pour obtenir le champ de sensibilité associé à n'importe quelle fonction score $S$ nécessite encore d'inverser la matrice $\boldsymbol{R}_t + \boldsymbol{H}_t\boldsymbol{A}_c\boldsymbol{H}_t^T$.

### 4.3.5 Variance Reduction Field

La méthode de la VRF a été développée pour proposer le calcul d'un champ de réduction de variance dans le cadre d'un système d'assimilation et de prévision d'ensemble mais aussi pour une assimilation incrémentale et un modèle déterministe.

L'hypothèse centrale de la méthode est de considérer que les observations supplémentaires se résument à ajouter une observation en un point de grille unique $i$ du modèle pour un paramètre choisi. L'opérateur d'observation $\boldsymbol{H}_{t,i}$ se résume alors à la projection sur ce point de grille. En notant $\boldsymbol{s} = \boldsymbol{A}_c[\nabla_{\boldsymbol{x}^a} S]$ le champ de sensibilité, on peut alors simplifier l'équation (4.11) pour l'observation supplémentaire en $i$,

$$(\delta\sigma^2_{S_t})_i = \frac{s_i^2}{\sigma_t^{o2} + \sigma_i^{a2}}. \tag{4.15}$$

où $s_i$ est la composante de $\boldsymbol{s}$ en $i$, $\sigma_t^{o2}$ la variance d'erreur de l'observation supplémentaire, qui peut être choisie arbitrairement, et $\sigma_i^{a2}$ la variance d'erreur d'analyse en $i$.

Il est possible de construire un champ de réduction de variance à partir de (4.15), le calcul pour chaque point de grille étant indépendant,

$$\delta\boldsymbol{\sigma}^2_{S_t} = \boldsymbol{s} \odot \boldsymbol{s} \odot (\sigma_t^{o2}\mathbf{1} + \boldsymbol{\sigma}^{\boldsymbol{a2}})^{-1}, \tag{4.16}$$

où $\mathbf{1}_i = 1$, $\forall i$ et $\boldsymbol{\sigma}^{\boldsymbol{a2}}$ est la diagonale de $\boldsymbol{A}_c$. $\odot$ représente le produit d'Hadamard (produit terme à terme des vecteurs). Comme le terme $\sigma_t^{o2}\mathbf{1}$ est choisi arbitrairement, il peut être négligé dans le calcul, cependant il peut être utilisé pour filtrer les variances d'erreur d'analyse trop faibles qui pourraient être obtenues pour $\boldsymbol{A}_c$. Si l'on prend un $\sigma_t^{o2} = 0$, avec un $\sigma_t^{a2}$ très faible on obtient des valeurs de réduction de variance aberrantes. La valeur choisie de $\sigma_t^{o2}$ n'a cependant pas d'influence sur la forme du champ de réduction de variance donc sur la position des maxima.

Le calcul du champ de VRF nécessite uniquement de connaître les deux vecteurs $\boldsymbol{s} = \boldsymbol{A}_c[\nabla_{\boldsymbol{x}^a} S]$ et $\boldsymbol{\sigma}^{\boldsymbol{a2}}$. L'implémentation en pratique de la VRF dans des systèmes d'assimilation incrémental et d'ensemble sont décrits au chapitre 7.

### 4.3.6 Développement à l'ordre 2

Les calculs pour obtenir une expression complète de la VRF au second ordre n'ont pas abouti à la rédaction de ce manuscrit. L'approche mise en œuvre est cependant décrite ici.

Le développement au second ordre pour le calcul de la VRF est réalisé à partir de l'équation (4.1) qui est de la forme $dS_1 = \Gamma_1\delta(\mathbf{x}^a)$, l'indice 1 indiquant qu'il s'agit d'un développement au premier ordre. Un second est ajouté pour prendre en compte les non-linéarités (ie. non gaussianité) lors de l'évolution de la trajectoire,

$$dS_2 = \Gamma_1^T\delta(\mathbf{x}^a) + \delta(\mathbf{x}^a)^T\Gamma_2\delta(\mathbf{x}^a). \tag{4.17}$$

L'objectif est donc d'effectuer une régression linéaire polynomiale pour déterminer les valeurs de $\Gamma_1$ et de $\Gamma_2$. Une régression linéaire polynomiale s'écrit de la forme,

$$\begin{aligned} y &= a_0 + a_1 x + a_2 x^2 + \eta \\ \boldsymbol{Y} &= \boldsymbol{M}\boldsymbol{X} \end{aligned}$$

avec $\boldsymbol{X} = [1 \; x \; x^2]$. La matrice $\boldsymbol{M}$ est alors obtenue comme $\boldsymbol{M} = \mathrm{Cov}(\boldsymbol{Y}, \boldsymbol{X})\,\mathrm{Cov}(\boldsymbol{X})^{-1}$. Dans notre cas en utilisant la convention d'Einstein pour le sommation des indices, le vecteur $\boldsymbol{X}$ s'écrit $\boldsymbol{X}^T = [\delta(x^a)_i \quad \delta(x^a)_j\delta(x^a)_k]^T$. $\boldsymbol{Y}$ est composé des valeurs $dS$ calculées à partir des membres de l'ensemble $dS_i = S(\boldsymbol{x}_i) - \mathbb{E}(S(\boldsymbol{x})$ $\Gamma_1$ et $\Gamma_2$ sont obtenus comme les termes diagonaux de $\boldsymbol{M}$.

Les calculs pour obtenir $\boldsymbol{M}$ sont détaillés dans l'annexe F. Au final on obtient,

$$\begin{aligned} \Gamma_1 &= C_1 A^{-1} \\ \Gamma_2 &= \frac{C_2}{3} A^{-2} \end{aligned}$$

avec $C_1 = \frac{1}{N_e-1}\sum_{i=1}^{N_e} \delta S_i \delta(x^a)_{j_i}$ et $C_2 = \frac{1}{N_e-1}\sum_{i=1}^{N} \delta S_i (\delta(x^a)_j \delta(x^a)_k)_i$. Ces expressions permettent de relier une variation de $dS$ à une variation $\delta(\mathbf{x}^a)$. Pour obtenir un champ de sensibilité de VRF il est nécessaire de pouvoir en déduire une expression au second ordre pour l'équation la variance de la fonction score.

A partir de l'équation (4.17), la variance de la fonction score s'écrit (voir l'annexe F),

$$\begin{aligned} \sigma_S^2 &= \mathrm{E}(dSdS^T) && (4.18)\\ &= \Gamma_1^T A \Gamma_1 + \mathrm{E}(\delta(x^a)^T \Gamma_2 \delta(x^a) \delta(x^a)^T \Gamma_2^T \delta(x^a)). && (4.19) \end{aligned}$$

Le second terme peut être simplifier en posant $\delta(\tilde{\mathbf{x}}) = \Gamma_2^{T/2}\delta(x^a)$ et en utilisant la formule de Wick (Zinn-Justin, 2005), il devient,

$$\mathrm{E}(\delta(x^a)^T \Gamma_2 \delta(x^a) \delta(x^a)^T \Gamma_2^T \delta(x^a)) = 2\left[\Gamma_2 :A\right]^2 \tag{4.20}$$

où : est l'opérateur du produit doublement contracté (ou opérateur de contraction tensorielle). L'expression de $\sigma_S^2$ au second ordre permet d'obtenir la réduction de variance $\delta\sigma_{S_t^2}$, équivalent au second ordre à l'équation (4.11). La réduction de variance associée à l'observation supplémentaire s'écrit donc,

$$\delta_{S_t^2} = \Gamma_1^T \mathbf{A}_c \Gamma_1 + 2\left[\Gamma_2 :\mathbf{A}_c\right]^2 - \Gamma_1^T \mathbf{A}_{c+t} \Gamma_1 - 2\left[\Gamma_2 :\mathbf{A}_{c+t}\right]^2. \tag{4.21}$$

On retrouve la différence $\Gamma_1^T(\mathbf{A}_c - \mathbf{A}_{c+t})\Gamma_1$ correspondant à l'équation à l'ordre 1, à laquelle s'ajoute un nouveau terme $2\left[\Gamma_2 :\mathbf{A}_c\right]^2 - 2\left[\Gamma_2 :\mathbf{A}_{c+t}\right]^2$. La différence des contractions au carré peut se ré-écrire sous la forme d'un produit de contractions,

$$[\Gamma_2 :A_c]^2 - [\Gamma_2 :A_{c+t}]^2 = [\Gamma_2 :(A_c + A_{c+t})][\Gamma_2 :(A_c - A_{c+t})]. \tag{4.22}$$

Une idée pour calculer les termes en $\mathbf{A}_c - \mathbf{A}_{c+t} = \boldsymbol{A}_c \boldsymbol{H}_t^T (\boldsymbol{R}_t + \boldsymbol{H}_t \boldsymbol{A}_c \boldsymbol{H}_t^T)^{-1} \boldsymbol{H}_t \boldsymbol{A}_c$ serait d'utiliser la méthode proposée au chapitre 3 à partir de l'estimation des longueurs de portée. Le terme en $\mathbf{A}_c + \mathbf{A}_{c+t}$ est cependant plus compliqué à calculer car il s'agit d'une matrice complète. La contraction avec $\Gamma_2$ doit donc prendre en compte l'ensemble des composantes. Une solution doit encore être trouvée pour simplifier l'équation ou avoir un moyen pratique de réaliser la contraction.

## 4.4 Utilisation des sensibilités

### 4.4.1 Déploiement d'observations supplémentaires

Le déploiement d'observations supplémentaires dans les zones de forte incertitude et potentiellement dynamique est l'objectif initial de l'observation adaptative. Elle a été mise en œuvre à partir de la campagne de mesure FASTEX et pour l'observation des cyclones tropicaux. Il y a alors deux manières de mettre en œuvre les méthodes décrites auparavant.

La première manière est de définir à l'avance un ensemble de stratégies de déploiement possible pour déterminer laquelle sera la plus profitable. Cela correspond le plus souvent à des plans de vol pour un avion d'observation. L'avantage est que cela permet de comparer des stratégies qui sont réalisables d'un point de vue logistique, à condition d'avoir un avion ou un drone disponible. De plus cela restreint le plus souvent le domaine géographique des observations supplémentaires potentielles. Un des inconvénients cependant est que l'on base souvent la décision de déclencher la mission sur la valeur prévue de sensibilité ou de réduction de variance. Pour des raisons financières, il est souhaitable d'effectuer un vol uniquement si celui-ci contribuera réellement à améliorer la prévision. Hors, comme il sera illustré au chapitre 6, la valeur elle même de la réduction de variable est très dépendante de la trajectoire du modèle. Utiliser une prévision d'ensemble et un ETKF peut cependant palier à ce défaut.

La deuxième manière, et celle qui est principalement décrite dans ces travaux, consiste à construire un champ de sensibilité. Comme précédemment la valeur du champ de sensibilité est dépendante de la trajectoire, cependant sa forme et la position du ou des maxima de sensibilité restent proches si on

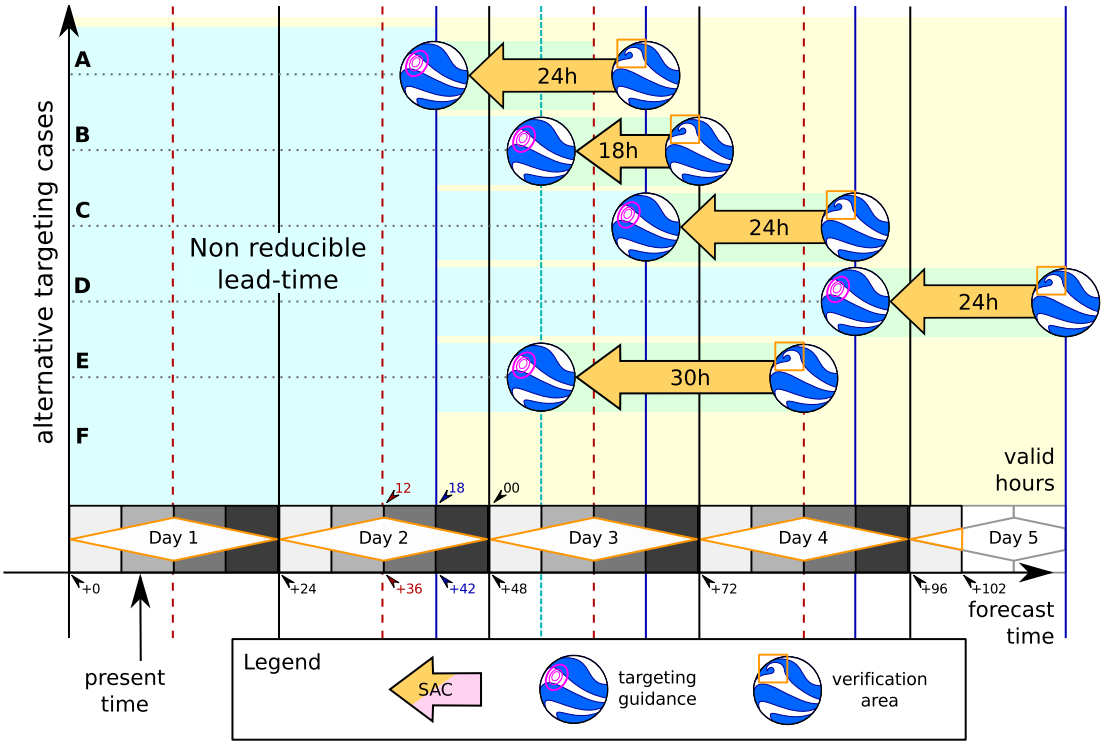



FIGURE 4.3 – Schéma de principe des calculs automatiques de champ de sensibilité pour la campagne HyMeX. Différents lead-times sont prédéterminés pour permettre aux prévisionnistes d'évaluer l'évolution des champs de sensibilité dans le temps. Ces calculs prennent en compte des prévisions allant jusqu'à $J+5$.

considère plusieurs trajectoires. Le champ de sensibilité permet d'avoir une vision globale et synthétique de notre connaissance à l'instant $t_{obs}$ de la dynamique de l'atmosphère et du système d'assimilation. Il reste cependant un problème d'ordre logistique à résoudre pour exploiter au mieux le champ de sensibilité. En effet comme il n'y a aucune contrainte géographique sur le maximum de sensibilité, celui-ci peut se trouver à un endroit où il est impossible de déployer une observation supplémentaire dans les délais.

Une autre contrainte forte qui concerne toutes les méthodes d'observation adaptative est notre capacité à prévoir le réseau de routine dans le futur. En effet comme le montre les valeurs de DFS (2.6.4), environ 50% de l'information de l'analyse provient des observations satellites et 15% des mesures avions AMDAR. Ces mesures étant intermittentes, pour disposer d'une évaluation pertinente du réseau de routine à $t_{obs}$ il faudrait donc disposer d'un modèle de prévision des observations à part entière. De plus comme la qualité des mesures satellites est très dépendante de la couverture nuageuse, il faudrait coupler ce modèle de prévision des observations à un modèle atmosphérique. Cela doit être réalisable mais non trivial.

L'observation adaptative reste donc limitée à des phénomènes de grandes échelles (dépressions et cyclones) si on conserve l'objectif de déployer des observations supplémentaires. C'est pourquoi une nouvelle manière d'utiliser le champ de sensibilité est présentée ci-dessous.

### Illustration de la KFS dans le cadre d'une expérience HyMeX

La $KFS$ a été utilisée dans le cadre de la campagne de mesure HyMeX du 5 septembre au 6 novembre 2012. L'objectif de la campagne était d'identifier l'impact des observations supplémentaires sur les prévisions de fortes précipitations sur le nord-ouest de la Méditérannée (Ducrocq et al., 2014). Cette campagne utilisait la plate-forme du DTS qui avait déjà été mise en œuvre en 2008 et 2009. Le réseau d'observations adaptatives se composait de 3 éléments, des stations de mesure de radiosondage spécifiques à la campagne de mesure (une station mobile terrestre et une station sur bateau), les radiosondages réguliers pour lesquels des lâchers supplémentaires ont été demandés à 18h UTC, et enfin les données de radiances AMSU-A et AMSU-B. Pour ces derniers les observations n'ont pas été directement modifiées mais c'est la manière de

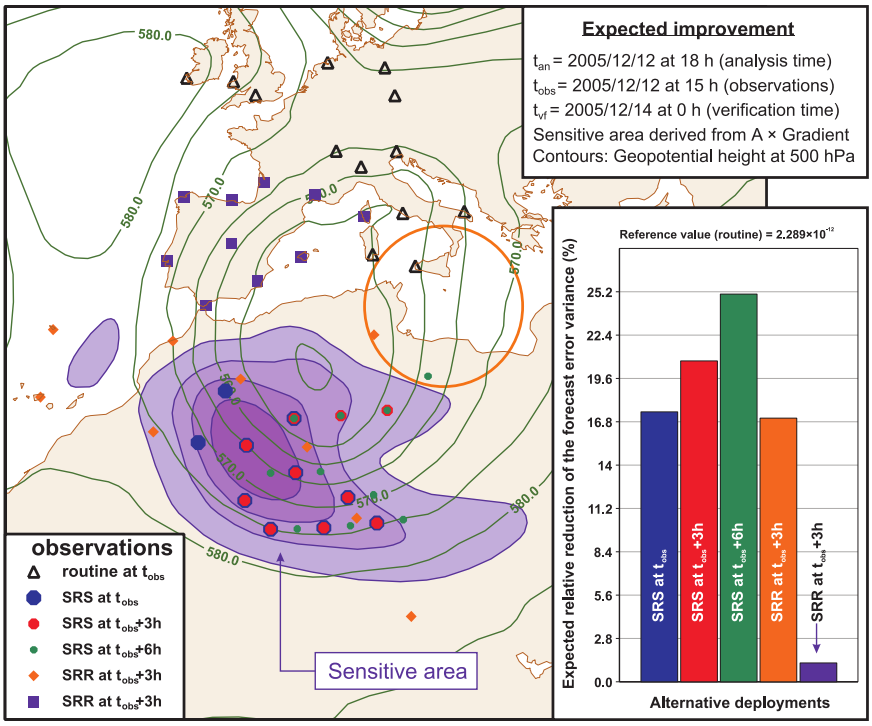



FIGURE 4.4 – Comparaison de l'impact de 5 déploiements alternatifs composés de 10 radiosondages supplémentaires. On considère 2 types de déploiements (carte). Les SRS (*Simulated Radiosoundings in Sensitive area*) sont positionnés pour échantillonner la zone sensible en 3 instants différents. Les SRR (*Simulated Routine Radiosoundings*) exploitent les positions de sites de sondage existants mais qui ne fonctionnent pas à l'instant $t_{obs}$. Le champ de sensibilité (carte) est valide à cet instant $t_{obs}$. Les déploiements simulés en bleu, orange et violet sont aussi valides à cet instant. Les déploiements simulés rouge et verte sont définis pour $t_{obs}+3h$ et $t_{obs}+6h$ à partir de l'advection du champ de sensibilité. Ces champs de sensibilité à $t_{obs}+3h$ et $t_{obs}+6h$ ne sont pas montrés. Les sondages SRS sont compliqués à faire en pratique (nécessité d'un avion ou plusieurs bateaux). Les sondages SRR sont faciles à réaliser en pratique (nécessité d'une coordination à peu de frais). On note (histogramme) l'impact plus fort des déploiements effectués en fin de fenêtre d'assimilation. À horaire identique, le déploiement SRS (bleu) a un impact équivalent au déploiement SRR (orange). Le déploiement violet (radiosondages sur l'Espagne et en dehors de la zone sensible) produit un impact très marginal.

les traiter dans l'assimilation. La résolution horizontale des données de radiance a été doublée au niveau des zones de sensibilité, en pratique cela se traduit par une diminution d'observations rejetées lors du screening (voir plus bas).

Dans ces expériences, la $KFS$ fournit un champ de sensibilité $\boldsymbol{A}_c\nabla S$ de manière régulière en tenant compte uniquement du réseau de routine. La figure 4.3 montre comment les calculs automatiques ont été configurés afin d'offrir un jeu de résultats utiles le plus tôt possible au prévisionniste. Si le prévisionniste estime que la situation est intéressante, il peut calculer de nouveaux produits et aussi élaborer un ensemble de stratégies de déploiement, comme illustré dans la figure 4.4. Pour chaque stratégie une nouvelle assimilation a lieu en prenant en compte les observations supplémentaires, et la réduction de variance est calculée comme $\delta S_t^2 = \nabla S^T \boldsymbol{A}_c \nabla S - \nabla S^T \boldsymbol{A}_{c+t} \nabla S$.

### 4.4.2 Efficacité des déploiements d'observations supplémentaires

L'exemple ci-dessus présente un dispositif très abouti et intégré pour faire de l'observation adaptative ainsi que quelques résultats de KFS. Il est naturel que le lecteur se pose la question : "Est-ce que tout

cela marche et quelle technique fonctionne le mieux ?". Pour la plupart des campagnes qui ont utilisé des dispositifs d'observation adaptative, il a été conduit des études d'impact (soit par des OSE, soit par des évaluations linéaires comme celles présentées plus tôt dans ce chapitre). Les résultats sont en général assez mitigés et ne sont pas aussi significatifs en fonctions des régions et des phénomènes météorologiques que l'on a essayé d'appréhender avec les observations supplémentaires. Pour résumer, dans le cas de l'Europe, l'amélioration des prévisions par le déploiement d'observations supplémentaires est en général difficile à montrer (Fourrié et al., 2006) (signal neutre en moyenne), sauf dans le cas de la campagne FASTEX (Desroziers et al., 2003) et sur quelques situations météorologiques particulières. Un rapport synthétique sur cette question a été publié par l'OMM (Majumdar and Co-authors, 2014) et d'une façon générale, il semble que dans les systèmes d'assimilation actuels, les déploiements que l'on peut réaliser ont un poids très faible (notamment en terme de degrés de liberté) par rapport au réseau de routine largement dominé par les données issues de satellites. En outre, il se pose la question de l'observabilité des structures très instables (baroclines) qui sont décrites par les techniques adjointes et qu'il faudrait réussir à parfaitement échantillonner dans la pratique au moment de déployer les observations pour obtenir les impacts parfois très importants (Lupu and Gauthier, 2011). C'est particulièrement le cas pour les techniques adjointes comme celles qui sont basées sur les vecteurs singuliers (Buizza and Palmer, 1995). En effet, ces approches permettent de déterminer les structures initiales à fort potentiel dynamique mais sans évaluer si celles-ci sont déjà suffisamment contraintes par des observations ou non. Elles ne prennent pas en compte le système d'assimilation et c'est pour cette raison qu'elles n'ont pas été détaillées dans ce chapitre.

### 4.4.3 Utilisation de la sensibilité aux observations pour améliorer le screening

Gratton et al. (2013) ont montré qu'il est possible d'améliorer l'efficacité de la minimisation en sélectionnant les *bonnes* observations parmi l'ensemble de celles disponibles. La fonction coût $J$ atteint une valeur plus faible en moins d'itérations, avec de fait une réduction du coût de calcul. L'algorithme proposé n'est pas applicable dans un contexte opérationnel mais l'expérience prouve qu'il y a un gain possible à optimiser la sélection des observations.

Les études récentes en observation adaptative réalisées dans le cadre de l'observation des dépressions hivernales (Hamill et al., 2013) et pour la campagne de mesure de meso-échèlle HyMeX (Campins et al., 2013) ont montré que la stratégie historique de l'observation adaptative n'est plus aussi efficace. En effet le nombre d'observations disponibles et les systèmes d'assimilation ont beaucoup évolués depuis le début des années 2000, et rajouter des observations supplémentaires à un réseau de routine n'a qu'un impact limité. L'idée est donc d'avantage d'essayer de tirer un meilleur profit des observations déjà disponibles. Comme cela a été le cas lors de la campagne HyMeX où la densité des radiances AMSU-A et AMSU-B prise en compte dans l'assimilation a été doublé au niveau des zones de sensibilité.

De plus la sensibilité aux observations est déjà utilisée dans le cadre des méthodes d'impact d'observation pour identifier les observations les plus efficaces. Les dernières études sur l'EFSO (Hotta et al. (2014) et Kalnay et al. (2012)) proposent même de rejeter les observations ayant un effet négatif sur la prévision. Il s'agit d'un contrôle de qualité proactif. Une évaluation de la fonction score à $6h$ permet déjà d'identifier les observations néfastes, ceci limite les effets de non-linéarités de la trajectoire pour le calcul de la sensibilité et le coût de calcul.

Pour pouvoir appliquer ce contrôle dans un cadre opérationnel, il faut pouvoir s'affranchir d'effectuer une seconde assimilation en soustrayant les observations identifiées comme néfastes. L'EFSO étant une méthode diagnostique, le calcul est réalisé à l'issu de l'assimilation. En pratique, la première assimilation est réalisée avec l'ensemble des observations post-screening qui permet d'appliquer l'EFSO. Comme le nombre d'observations néfastes à supprimer doit rester faible, l'analyse est directement corrigée en utilisant un développement de Taylor au premier ordre,

$$\boldsymbol{x}^{a,deny} - \boldsymbol{x}^{a} \approx -\boldsymbol{K}^{deny}\delta(\boldsymbol{y}^{o,deny}) \approx -\frac{1}{1-N_e}\boldsymbol{X}^{a}\boldsymbol{Y}^{a,deny}\boldsymbol{R}^{deny^{-1}}\delta(\boldsymbol{y}^{o,deny}). \tag{4.23}$$

$\boldsymbol{x}^{a,deny}$ est l'analyse corrigée si les observations néfastes sont écartées. $\delta(\boldsymbol{y}^{o,deny})$ contient ces observations néfastes, avec la matrice de covariance d'erreur d'observation associée $\boldsymbol{R}^{deny}$ et l'opérateur d'observation.

La matrice de gain est cependant estimée en reprenant les perturbations d'analyse $\boldsymbol{X}^a$ de la première assimilation. Il n'est alors pas nécessaire d'effectuer une seconde assimilation complète pour corriger l'analyse des observations néfastes.

Ce genre d'approche permet de faire la jonction entre l'étude d'impact des observations et l'observation adaptative et d'envisager des nouvelles méthodes de screening adaptatives.

## 4.5 Conclusion du chapitre 4

Les développements récents autour de la sensibilité aux observations se sont concentrés sur l'étude de l'impact des observations, à travers des développements au second ordre et très récemment l'utilisation d'un ensemble dans l'EFSO. Les calculs au second ordre sont rendus possibles dans les méthodes diagnostiques par la connaissance du résultat de l'assimilation de données (analyse). Ce n'est pas le cas dans la cadre pronostic de l'observation adaptative. Les méthodes historiques de l'observation adaptative ont été revisitées dans cette étude pour proposer une nouvelle approche appelée la VRF (pour *Variance Reduction Field*). Un développement au second ordre pour la VRF est tenté mais les difficultés rencontrées dans les calculs ne permettent pas d'aller jusqu'au bout. Des hypothèses simplificatrices pourraient être utilisées pour développer ces calculs plus en avant. Ceci pourra faire partie des prolongements à mener au-delà ce travail. L'implémentation de la VRF n'est pas décrite dans ce chapitre, puisqu'un chapitre lui est dédié (chapitre 7), soit à travers une implémentation de l'assimilation de données incrémentale, soit à partir d'un ensemble de prévision et d'assimilation.

On a vu que les approches diagnostiques apportent des informations complémentaires à celles apportées par d'autres diagnostics tels que les DFS (chapitre 2.6.4). Les DFS permettent de déterminer l'influence de groupes d'observations sur le système d'assimilation, cependant ils ne prennent que très marginalement en compte la dynamique des phénomènes atmosphériques qui conditionne fortement le signal dans les approches par calcul de sensibilité.

L'évolution des réseaux d'observation - avec en particulier la très forte augmentation du nombre de mesures satellites disponibles - demande à repenser l'utilisation des méthodes d'observation adaptative. De même avec la réduction des délais entre les réseaux d'assimilation, il est possible d'utiliser des techniques de calcul d'impact d'observation comme l'EFSO pour corriger en quasi-temps réel.

# Chapitre 5

# Modèles numériques utilisés

Afin d'étudier l'impact de la trajectoire sur les champs de sensibilité calculés, on souhaite simuler une dynamique simple mais réaliste et exhibant des évolutions complexes. L'objectif étant de suivre l'évolution des erreurs au cours du temps, une dynamique linéaire ne nécessiterait que la connaissance de l'instant initial. C'est pourquoi un modèle quasi-géostrophique non linéaire a été développé, le modèle de Phillips (Phillips, 1951).

Ce chapitre décrit les équations du modèle quasi-géostrophique utilisé, en précisant les hypothèses et le sens physique. L'implémentation du code est ensuite décrite ainsi que sa validation. Pour finir le comportement du modèle est illustré à l'aide de simulations.

## 5.1 Modèle Quasi-Géostrophique à deux couches

### 5.1.1 Description du modèle Quasi-Géostrophique à deux couches

#### Équations Shallow Water

Dans un premier temps l'atmosphère est assimilée à un fluide en couche mince soumis à une rotation. Le paramètre de Coriolis est noté $f = 2\Omega \sin\phi$ où $\Omega$ est la vitesse angulaire de rotation de la Terre et $\phi$ la latitude. Dans la suite les effets dus à la courbure sont négligés. A l'échelle synoptique, à partir d'une approximation en ordre de grandeur on trouve que la viscosité du fluide peut être négligée et que la vitesse verticale est faible devant la vitesse horizontale. L'accélération verticale peut aussi être négligée devant le gradient de pression, l'hypothèse hydrostatique est utilisée. L'écoulement est supposé parfait et incompressible, les équations d'Euler s'écrivent alors sous la forme

$$\frac{d\boldsymbol{u}}{dt} + f\boldsymbol{k} \times \boldsymbol{u} = -\frac{1}{\rho}\nabla P, \tag{5.1}$$

$$\frac{\partial P}{\partial z} = -\rho g, \tag{5.2}$$

$$\nabla.\boldsymbol{u} + \frac{\partial w}{\partial z} = 0, \tag{5.3}$$

avec $\boldsymbol{u} = (u, v)$ la vitesse horizontale, $w$ la vitesse verticale, $\boldsymbol{k}$ le vecteur unitaire décrivant l'axe $z$ et $g$ l'accélération de gravité. Ces équations correspondent à la conservation de la quantité de mouvement horizontale (5.1) et verticale (5.2) en négligeant $\partial_t w$, et à l'équation de continuité 5.3. La vitesse verticale n'est plus une variable pronostique mais est obtenue à partir de l'équation de la divergence. En intégrant l'équation hydrostatique (5.3) sur la verticale on obtient $P(z = h) - P(z = 0) = -\rho g h$. Le terme de droite de l'équation de conservation de quantité de mouvement horizontale (5.1) devient alors $-g\nabla h$, et on retrouve les équations de Saint Venant. Cependant l'hypothèse $\rho = cst$ sur toute la couche étant trop forte pour permettre de représenter certains phénomènes météorologiques classiques, on peut introduire une stratification sous la forme d'un écoulement multicouche.

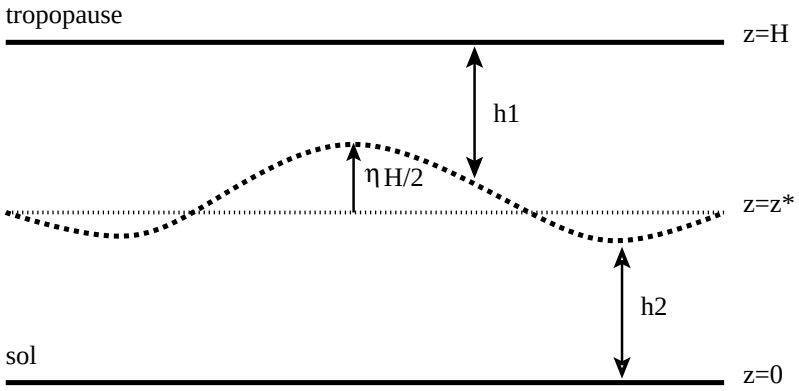



FIGURE 5.1 – Schéma d'une atmosphère à 2 couches modélisée par deux Shallow Water couplés.

L'atmosphère est donc supposé stratifié en deux couches, avec comme condition à la limite en bas un sol imperméable et en haut une tropopause imperméable. Les couches supérieure et inférieure sont respectivement numérotées 1 et 2, leurs épaisseurs respectives sont $h_1$ et $h_2$ (avec $h_1 + h_2 = H$ où $H$ est la hauteur totale de l'atmosphère) et leur interface se trouve à l'altitude $z^*$ (voir figure 5.1).

En supposant que la vitesse horizontale est homogène sur l'épaisseur de chaque couche et en intégrant l'équation de continuité (5.3) sur chaque couche on obtient

$$\begin{cases} w(z=H) - w(z=z^*) + h_1 \nabla . \boldsymbol{u}_1 = 0, \\ w(z=z^*) - w(z=0) + h_2 \nabla . \boldsymbol{u}_2 = 0, \end{cases}$$

où $w(z=z^*) = \frac{dz^*}{dt}$ est la vitesse verticale de l'interface. Les conditions d'imperméabilité donnent $w(z = H) = w(z = 0) = 0$. Par continuité à l'interface on peut écrire $\frac{dz^*}{dt} = -\frac{dh_1}{dt} = \frac{dh_2}{dt}$. En remplaçant dans les équations, on retrouve les équations Shallow Water (ou de Saint-Venant) pour chaque couche

$$\begin{cases} \frac{d\boldsymbol{u}_i}{dt} + f\boldsymbol{k} \times \boldsymbol{u}_i = -g\nabla h_i, \\ \frac{dh_i}{dt} + h_i \nabla . \boldsymbol{u}_i = 0. \end{cases} \tag{5.4}$$

En prenant le rotationnel de l'équation de quantité de mouvement horizontale projeté selon la verticale, l'équation du tourbillon s'écrit

$$\frac{d(\zeta_i + f)}{dt} + (\zeta_i + f)\nabla . \boldsymbol{u}_i = 0$$

en substituant $\nabla . \boldsymbol{u}_i$ à partir de l'équation de continuité (5.3), on obtient $\frac{d}{dt}\left(\frac{\zeta_i + f}{h_i}\right) = 0$, ce qui signifie que $\frac{\zeta_i + f}{h_i}$ est un invariant lagrangien. Il s'agit du tourbillon potentiel Shallow Water, noté $q_{SWi} = \frac{\zeta_i + f}{h_i}$. Malgré la conservation lagrangienne, ce système ne peut se réduire à une équation et il est nécessaire de résoudre le système couplé (5.4) pour construire l'évolution temporelle. On cherche maintenant à simplifier cette dynamique en éliminant les ondes de gravité.

### Approximation Quasi-Qéostrophique

On se place maintenant dans le cadre quasi-géostrophique (QG), c'est à dire que le vent horizontal s'équilibre quasi instantanément entre la force de Coriolis et le gradient de pression $f\boldsymbol{k} \times \boldsymbol{u}_{gi} = -g\nabla h_i$, ou encore $\boldsymbol{u}_{gi} = \frac{g}{f}\boldsymbol{k} \times \nabla h_i$.On pose $h_1 = \frac{H}{2}(1 - \eta)$ et $h_2 = \frac{H}{2}(1 + \eta)$. Le vent horizontal s'écrit alors

$$\boldsymbol{u}_{g1} = -\frac{gH}{2f}\boldsymbol{k} \times \nabla\eta = \boldsymbol{k} \times \nabla\psi_1,$$

$$\boldsymbol{u}_{g2} = \frac{gH}{2f}\boldsymbol{k} \times \nabla\eta = \boldsymbol{k} \times \nabla\psi_2,$$

avec $\psi_1 = -\frac{gH\eta}{2f} = -\psi_2$ la fonction de courant pour chaque niveau à constante près. Ce qui peut aussi s'écrire $\eta = \frac{f}{gH}(\psi_2 - \psi_1)$.

Sous l'hypothèse que $\eta \ll 1$, il est possible de reformuler $q_{SW}$ de manière à simplifier le système 5.4. En appliquant un développement au premier ordre en $\eta$, il vient que $1/h_1 \approx \frac{2}{H}(1+\eta)$ et $1/h_2 \approx \frac{2}{H}(1-\eta)$. Sous l'hypothèse $\zeta \ll f$, la quantité conservée peut alors s'écrire $\frac{H}{2}q_{SW1} = (\zeta_1 + f)(1+\eta) \approx \zeta_1 + f + f\eta$. Une expression similaire est trouvée pour la couche 2. Finalement en utilisant la fonction de courant et en introduisant le rayon de déformation interne de Rossby $\lambda = \frac{\sqrt{gH}}{f}$, l'expression du tourbillon potentiel QG pour chaque couche est

$$q_i = f + \Delta\psi_i - \lambda^{-2}(\psi_i - \psi_{i'})$$

avec la notation $'$ faisant référence à l'autre couche. Le tourbillon potentiel QG $q_i$ est alors un invariant lagrangien. L'intérêt majeur est que l'on a remplacé la résolution du système couplé 5.4 par une simple équation aux dérivées partielles $\frac{dq}{dt} = 0$ où seules des fonctions de $\psi$ interviennent. Ce qui permet de construire un système simple reproduisant le SW deux couches en filtrant les ondes de gravité.

Dans cette formulation discrétisée sur deux couches, on retrouve une expression similaire au tourbillon potentiel QG classique

$$q_{QG} = \Delta\psi + f + \partial_z\left(\frac{f^2}{N^2}\partial_z\psi\right), \tag{5.5}$$

avec $N$ la fréquence de Brunt-Väisälä. Le tourbillon se décompose en un terme lié à la rotation de le Terre ($f = f_0 + \beta y$), un terme de vorticité relative ($\Delta\psi$) et enfin un terme de d'étirement verticale ($\lambda^{-2}(\psi_1 - \psi_2)$).

Cette approche par développement asymptotique ne garantie pas dans le cas général que la variable ainsi approximée soit toujours un invariant lagrangien. Cependant en utilisant un développement à partir des équations de conservation on retrouve le même résultat. L'ensemble des calculs est disponible dans Pedlosky (1982).

**Equations du modèle Quasi-Géostrophique deux couches**

En reprenant la définition 5.5 du tourbillon potentiel exprimée dans un $\beta$-plan ($f = f_0 + \beta y$), le tourbillon potentiel QG sur chaque couche s'écrit finalement :

$$q_1 = f_0 + \beta y + \Delta\psi_1 - \lambda^{-2}(\psi_1 - \psi_2) \tag{5.6}$$

$$q_2 = f_0 + \beta y + \Delta\psi_2 + \lambda^{-2}(\psi_1 - \psi_2) \tag{5.7}$$

$$\frac{\partial q_i}{\partial t} = -J(\psi_i, q_i) \tag{5.8}$$

avec $\boldsymbol{u_i} = (-\partial_y\psi_i, \partial_x\psi_i)$ et $J(\psi, q)$ le jacobien de $\psi$ et $q$ (qui correspond au terme d'advection), avec $J(A,B) = \frac{\partial A}{\partial x}\frac{\partial B}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial B}{\partial x}$.

A partir du système (5.6), il est possible de calculer l'évolution temporelle du tourbillon potentiel QG sur les deux niveaux, en advectant simplement la quantité conservée par le vent déduit de la fonction de courant. La fonction de courant est ici une variable diagnostique qui est recalculée à chaque pas de temps. Le couplage des deux couches est donc traité de manière explicite.

### 5.1.2 Implémentation du modèle

Maintenant que les équations ont été obtenues, il est nécessaire de les exprimer dans un espace discrétisé pour les résoudre numériquement. Les seules variables du modèle sont $q$ sur les deux niveaux. La projection de Galerkin utilisée est dans l'espace de Fourier tronqué, les champs sont alors représentés de manière spectrale par des coefficients associés au vecteur d'onde $\boldsymbol{k} = 2\pi(\frac{k_x}{L_x}, \frac{k_y}{L_y})$ pour chaque niveaux,

$$q(\boldsymbol{x}, t) = \sum_{k_x=-T_x}^{T_x}\sum_{k_y=-T_y}^{T_y} \hat{q}_{\boldsymbol{k}}(t) e_{\boldsymbol{k}}(\boldsymbol{x}),$$

avec $T_x$ et $T_y$ les troncatures selon $x$ et $y$, $e_{\boldsymbol{k}}(\boldsymbol{x}) = \exp(i\boldsymbol{k}.\boldsymbol{x})$ et $\hat{q}_{\boldsymbol{k}}(t)$ le coefficient spectral de vecteur d'onde $\boldsymbol{k}$ à l'instant $t$. L'évolution temporelle est calculée directement pour chaque coefficient spectral.

Une troncature elliptique $\frac{k_x}{T_x}^2 + \frac{k_y}{T_y}^2 \leqslant T$ permettant d'éviter l'introduction d'anisotropie est utilisée. L'espace point de grille équivalent est un rectangle selon $(x, y)$ de dimension $(L_x, L_y)$. Le passage d'un espace à l'autre est effectué par la fonction `fft` (Fast Fourier Transform).

La projection dans l'espace de Fourier impose d'avoir des champs bi-périodiques en $x$ et $y$. Ce qui n'est pas le cas pour le terme $\beta y$ par exemple. Ce terme n'intervient que dans le terme d'advection, il peut donc être reformulé dans les équations en remplaçant $J(\psi, \beta y)$ par $\beta\partial\psi/\partial x$. Ces deux termes sont complètement équivalent mais dans ce cas $\beta\partial\psi/\partial x$ est périodique et donc projetable dans l'espace de Fourier. Le terme $\beta y$ n'est donc pas pris en compte dans la variable tourbillon potentiel du modèle mais l'équation d'évolution est remplacée par $\frac{dq_i}{dt} = -\beta\partial_x\psi_i$, où $q_i = f_0 + \Delta\psi_i - \lambda^{-2}(\psi_i - \psi'_i)$. Les notations précédentes sont cependant conservée. La partie non périodique de l'écoulement de base est traitée de la même façon.

### Introduction d'un écoulement de base

Pour pouvoir mettre en place les méthodes d'observation adaptative il est nécessaire de simuler un phénomène dont la durée de vie est de quelques jours, comme la cyclogénèse. La cyclogénèse intervient lorsqu'une anomalie de tourbillon potentiel d'échelle synoptique interagit avec une zone barocline de plus grande échelle, associée un cisaillement vertical caractérisé par un vent fort en altitude (jet). Pour représenter ce phénomène l'écoulement est scindé en un écoulement haute fréquence (l'anomalie) qui est intégré dans le temps, et un écoulement basse fréquence (le jet) rendu stationnaire par un terme de forçage $F$. Les champs sont décomposés en un écoulement de base $(\overline{q}, \overline{\psi})$ et une perturbation $(q', \psi')$. L'équation de conservation pour l'écoulement de base s'écrit alors

$$\frac{\partial\overline{q}}{\partial t} + J(\overline{\psi}, \overline{q}) + \beta\partial_x\overline{\psi} = F$$

En posant $F = J(\overline{\psi}, \overline{q}) + \beta\partial_x\overline{\psi}$, on obtient $\frac{\partial\overline{q}}{\partial t} = 0$, conduisant à un écoulement de base stationnaire. L'équation d'évolution de la perturbation s'écrit alors

$$\frac{\partial q'}{\partial t} + J(\overline{\psi}, q') + J(\psi', \overline{q}) + J(\psi', q') = D(q') - \beta\partial_x\psi.$$

Cette relation s'applique pour les deux couches, avec des forçages différents. Le terme $D(q')$ est un terme de dissipation ajouté pour supprimer les petites échelles et assurer la stabilité à long terme du modèle. L'écoulement de base choisi ainsi que l'ensemble des paramètres utilisés lors des simulations sont donnés en section 5.1.4.

### Couplage tourbillon potentiel - fonction de courant

Le passage de $q$ à $\psi$ se fait en inversant les relations 5.6. En prenant la somme et la différence des équations on obtient

$$\begin{cases} q_1 + q_2 = \Delta(\psi_1 + \psi_2), \\ q_1 - q_2 = (\Delta - 2\lambda^{-2})(\psi_1 - \psi_2) \end{cases}$$

d'où les champs $\psi$ recherchés

$$\psi_1 = \frac{A+B}{2} \text{ et } \psi_2 = \frac{A-B}{2},$$

avec $A = \Delta^{-1}(q_1 + q_2)$ et $B = (\Delta - 2\lambda^{-2})^{-1}(q_1 - q_2)$.

La dernière étape pour coder le modèle consiste à discrétiser la dérivée temporelle.

**Schéma numérique temporel**

Le schéma en temps utilisé est un *leap-frog*, avec un premier pas de temps d'initialisation d'Euler explicite,

$$\frac{q^{n+1}-q^{n-1}}{2dt} = -J(\psi^n, q^n) + D\left(\frac{q^{n+1}+q^{n-1}}{2}.\right)$$

Les modes numériques issus du *leap-frog* sont filtrés par un filtre d'Asselin (Asselin, 1972) modifié, la formulation de Williams (Williams, 2009) est utilisée ($\varepsilon = 0.01$ et $\alpha = 0.53$),

$$\begin{aligned} q^n &= q^n + \frac{1}{2}\varepsilon\alpha(q^{n-1} + q^{n+1} - 2q^n), \\ q^{n+1} &= q^{n+1} - \frac{1}{2}\varepsilon(1-\alpha)(q^{n-1} + q^{n+1} - 2q^n). \end{aligned}$$

Le terme de dissipation choisi est un tri-laplacien avec des viscosités différentes selon $x$ et $y$ ($D = \nu_x\partial_x^6 + \nu_y\partial_y^6$). Ce terme permet de dissiper l'énergie dans les petites échelles sans modifier les structures de grandes échelles. Un pas de temps constant est utilisé, lorsque les perturbations croissent la condition CFL peut être mise en défaut et créer des instabilités à partir des petites échelles. Pour filtrer ces artefacts numériques il est nécessaire de paramétrer correctement la diffusion ($\nu_x$,$\nu_y$) pour garantir une stabilité sans détériorer la partie importante de signal. La mise en place d'un terme de forçage sans terme de dissipation pure oblige à limiter l'intégration du modèle à quelques jours, ce qui convient pour cette étude.

### 5.1.3 Validation du modèle non linéaire

Le modèle ainsi codé étant assez complexe, il est nécessaire de valider son comportement dans un cas simple, ainsi que les équations du modèle tangent et son adjoint. La validation du modèle se fait dans un cas de cisaillement vertical où la solution analytique est connue. La validation du tangent et de l'adjoint se fait en revenant à leurs définitions.

**Instabilité de Eady**

Pour valider le code il est intéressant de se ramener à un cas d'étude simple. L'écoulement de base est réduit a un vent zonal uniforme cisaillé verticalement ($U_1$ et $U_2$, $\overline{q} = 0$), dans un $f$-plan ($\beta = 0$), les termes de dissipation sont annulés (ce qui limite la durée des simulations). L'écoulement de base s'écrit donc $\overline{\psi} = -Uy$ et $\overline{q} = -\lambda^{-2}U'y$ avec $U$ le vent sur la couche considérée et $U'$ le cisaillement vertical. Le modèle se simplifie donc sous la forme

$$\frac{\partial q'}{\partial t} = -\lambda^{-2}U'\frac{\partial \psi'}{\partial x} - U\frac{\partial q'}{\partial x}. \tag{5.9}$$

La solution étant identique sur les deux niveaux, un seul résultat est présenté dans la suite.

Ces conditions sont celles du problème d'Eady, il est possible de montrer que des modes instables existent dans l'écoulement et de calculer leurs taux de croissance. La perturbation est prise sous la forme d'une onde sinusoïdale selon $x$ (invariant selon $y$), $u(\boldsymbol{x}) = A(t)\cos(kx)$ où $A$ est l'amplitude du mode. Les modes instables se développent dans la direction du cisaillement, la condition d'instabilité théorique le nombre d'onde $k$ est $k^2 \leq 2\lambda^{-2}$. Dans le cas d'étude ($Lx = 28000km$ et $\lambda = 500km$), le nombre d'onde limite théorique est $12.5k_0$ (avec $k_0 = 2\pi/L_x$). Cette valeur se retrouve dans le modèle, voir figure 5.2.

Pour calculer ces taux de croissance exponentiel une intégration suffisamment longue est nécessaire, un terme de dissipation faible est rajouté pour assurer la stabilité numérique. Les calculs des taux de croissance se faisant par rapport des amplitudes entre deux instants, $\sigma = \frac{1}{T}log\left(\frac{A(t+T)}{A(t)}\right)$. Les valeurs réelles de $\sigma$ ne sont pas directement calculées selon l'unité standard. C'est pourquoi les valeurs normalisées sont présentées, la forme de la courbe est conservée.

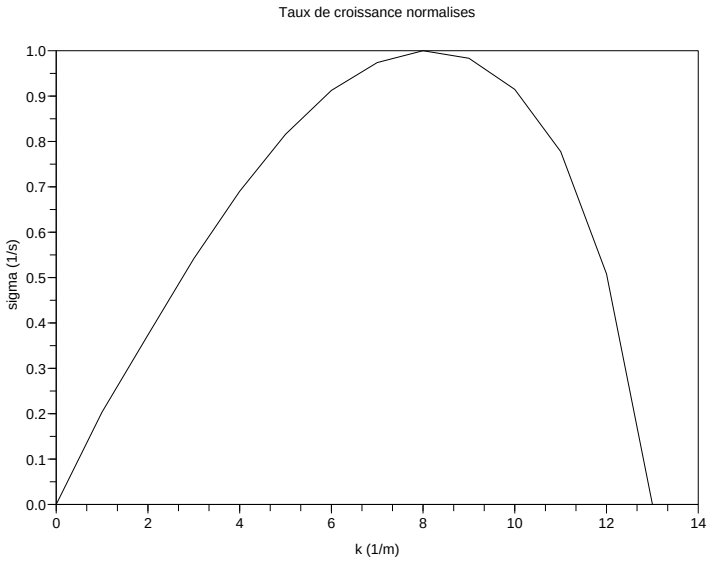



FIGURE 5.2 – Taux de croissance normalisés par le mode le plus instable en fonction du nombre d'onde normalisé par $k_0$. Pour $k > 12$, $\sigma$ devient complexe.

### Validation du linéaire tangent et de l'adjoint

En plus du modèle d'intégration non linéaire, il a été nécessaire de coder son linéaire tangent et son adjoint pour l'assimilation 4D-var (section 2.2.2) et le calcul du gradient de fonction score pour la KFS (section 4.3.2). On ne s'intéresse ici qu'à la validation de ces deux modèles.

**Modèle linéaire tangent** Le modèle linéaire tangent est défini à partir du développement de Taylor du modèle non linéaire $\mathcal{M}$,

$$\mathcal{M}(\boldsymbol{x} + \alpha\boldsymbol{u}) = \mathcal{M}(\boldsymbol{x}) + \alpha\boldsymbol{M}_{\boldsymbol{x}}\boldsymbol{u} + \mathcal{O}(\alpha^2),$$

avec $\boldsymbol{x}$ la trajectoire et $\boldsymbol{u}$ un vecteur arbitraire fixé. L'écart $\varepsilon$ entre la trajectoire non linéaire et le développement à l'ordre 1 peut s'écrire

$$\varepsilon = \frac{\mathcal{M}(\boldsymbol{x} + \alpha\boldsymbol{u}) - \mathcal{M}(\boldsymbol{x})}{\alpha} - \boldsymbol{M}_{\boldsymbol{x}}\boldsymbol{u} = \mathcal{O}(\alpha). \tag{5.10}$$

Ainsi en calculant $\varepsilon$, à $\boldsymbol{x}$ et $\boldsymbol{u}$ fixés, en faisant varier la valeur de $\alpha$, on doit retrouver un comportement linéaire en $\alpha$. Ceci a été observé dans notre cas validant le modèle linéaire (non montré ici).

**Modèle adjoint** Comme dans le cas du modèle tangent, la définition est utilisée pour valider le modèle adjoint. Celui-ci est défini par

$$\langle \boldsymbol{u} | \boldsymbol{M}\boldsymbol{v} \rangle_{\mathcal{T}_1} = \langle \boldsymbol{M}^T\boldsymbol{u} | \boldsymbol{v} \rangle_{\mathcal{T}_0}, \tag{5.11}$$

pour tout $(\boldsymbol{u}, \boldsymbol{v}) \in \mathcal{T}_1 \times \mathcal{T}_0$. $\mathcal{T}_i$ représente l'espace du vecteur d'état à l'instant $i$, $\langle .|. \rangle_{\mathcal{T}_i}$ est le produit scalaire euclidien dans cet espace, $\boldsymbol{M}^T$ désigne l'adjoint de $\boldsymbol{M}$. $\mathcal{T}_i$ représente l'espace des vecteurs d'états à $t_i$, dans le cas de l'adjoint du modèle d'intégration. Pour valider l'adjoint, $\boldsymbol{u}$ et $\boldsymbol{v}$ sont deux vecteurs aléatoires. Si $\boldsymbol{M}^T$ est bien l'adjoint de $\boldsymbol{M}$, l'écart entre les deux valeurs doit être de l'ordre de l'erreur machine ($10^{-16}$ dans notre cas).

Cette méthode se généralise à tout opérateur adjoint en calculant le produit scalaire dans l'espace correspondant. L'ensemble des opérateurs intervenant dans le code a été validé avec une erreur de l'ordre de la précision machine.

| paramètre | symbole | valeur |
|---|---|---|
| dimension en longitude | $L_x$ | $16000.10^3 m$ |
| dimension en latitude | $L_y$ | $8000.10^3 m$ |
| troncature en longitude | $T_x$ | 64 |
| troncature en latitude | $T_y$ | 32 |
| pas de temps | $dt$ | $450s$ |
| rayon de déformation de Rossby | $\lambda$ | $300.10^3 m$ |
| $\beta$ à 45 degrés | $\beta$ | $1.610^{-11} m^{-1} s^{-1}$ |
| viscosité tri-laplacien selon $x$ | $\nu_x$ | $6.10^{22} m^6 s^{-1}$ |
| viscosité tri-laplacien selon $y$ | $\nu_y$ | $6.10^{20} m^6 s^{-1}$ |

TABLE 5.1 – Liste des paramètres utilisés dans le modèle.

Plus d'informations sur l'interprétation du modèle adjoint et la manière de l'implémenter sont données dans le cours de Blayo and Nodet (2009).

### 5.1.4 Comportement du modèle

Le modèle QG deux couches est capable de représenter des interactions baroclines et barotropes, il est donc possible de simuler une cyclogénèse sur quelques jours. Dans la suite, la valeur des paramètres utilisés est donnée dans le tableau 5.1. Avant de décrire l'écoulement de base et l'anomalie utilisés, on se propose d'analyser le mécanisme de l'interaction barocline.

#### Modèle conceptuel de l'interaction barocline

Lors d'une cyclogénèse il faut distinguer deux phases, l'apparition des anomalies et leur croissance. Le modèle QG à deux couches étant un cas limite du Shallow-Water, les raisonnements physiques peuvent être appliqués sur le tourbillon $q_{SW}$. Considérons une anomalie positive de $q_1 = \frac{\zeta_1 + f}{h_1}$, qui peut être interprétée comme une élévation locale de l'interface entre les deux couches, $h_1$ diminue alors (voir 5.3). L'étirement de la couche inférieure entraîne la formation de vorticité $\delta\zeta_2 > 0$ pour assurer la conservation du tourbillon potentiel $q_2$. L'écoulement associé à ce tourbillon advecte alors le tourbillon potentiel de l'écoulement de base. Il apparaît une anomalie positive de $q_2$. Une fois les deux anomalies en place la cyclogénèse se développe.

Le modèle conceptuel de la cyclogénèse est rappelé dans la figure 5.4. Le cisaillement vertical de vent associé au jet d'altitude engendre un gradient de tourbillon potentiel dirigé vers le nord en altitude et vers le sud en basses couches. L'advection du tourbillon potentiel de l'écoulement de base (indiquée par les flèches) accentue l'anomalie positive vers l'aval en basse couche et vers l'amont à la tropopause, ce qui crée un déphasage vertical entre les deux niveaux. La croissance de $q_2$ correspond à un étirement et donc une élévation de l'interface au niveau de l'anomalie de basse couche. Cette élévation renforce l'anomalie négative en haut créée par advection. Il y a croissance barocline. Comme l'anomalie d'altitude est aussi transportée par le jet, le phasage des anomalies peut être conservé. On obtient donc des anomalies de plus en plus positives aux deux niveaux, ce qui accentue l'étirement du tourbillon en bas et l'anomalie d'épaisseur en haut.

Pour que cette interaction barocline ait lieux il est donc nécessaire de construire un environnement similaire, ce qui est décrit maintenant.

#### Choix de l'écoulement de base

Le cisaillement dans le modèle est obtenue par un vent zonal uniforme sur chaque couche (mais de valeurs différentes). A ceci se rajoute un jet localisé oscillant suivant $x$ qui va crée un champ de déformation plus complexe. Le profil de vitesse selon $y$ pour le jet est pris gaussien. Le caractère oscillant du jet est

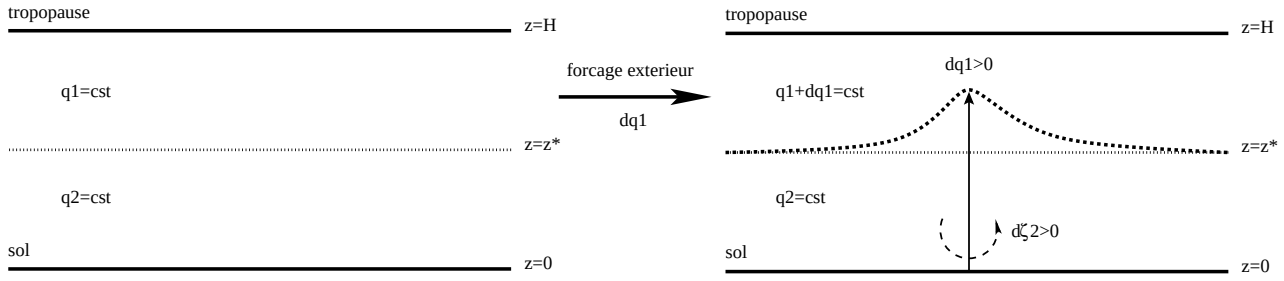



FIGURE 5.3 – Anomalie positive de tourbillon potentiel associée à une élévation locale de l'interface. L'étirement de la couche inférieure crée un tourbillon positif en bas pour conserver le tourbillon potentiel.

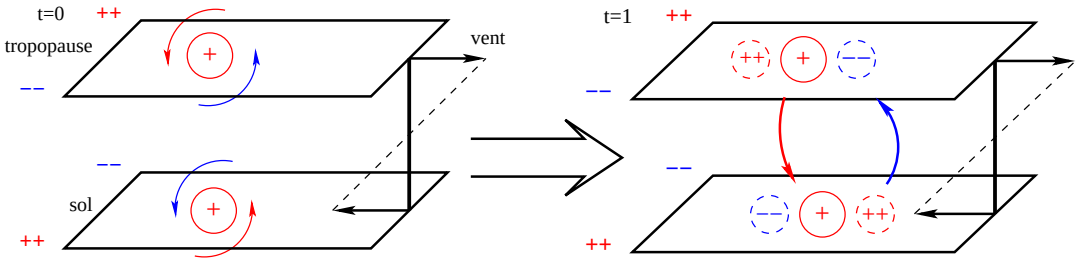



FIGURE 5.4 – Schéma conceptuel d'interaction barocline pour deux instants successifs. Les cercles représentent des anomalies de tourbillon potentiel, en rouge une anomalie positive et en bleu une négative. Les flèches représentent l'advection. Les signes $++$ et - - donnent la direction du gradient de tourbillon potentiel de l'écoulement de base. Le cisaillement de vent vertical est indiqué sur la droite des schémas.

choisi pour limiter le développement d'instabilité de type Eady (cf. 5.1.3), mais aussi en accord avec des résultats antérieurs (Gilet et al., 2009). La croissance barocline est favorisée selon la direction du champ de déformation. Les oscillations permettent donc de diversifier les scénarios selon la position initiale de l'anomalie en longitude et en latitude. Le champ de vent de l'écoulement de base est montré figure (5.5). Dans la suite, la forme analytique du jet est

$$U(x,y) = \overline{U} + \mu \exp\left\{ \frac{y - \frac{L_y}{2} - \delta_y \cos\left(\frac{4\pi}{L_x} x\right)}{\delta_y^2} \right\},$$

avec $\mu$ le terme d'amplitude et $\delta_y = L_y/12$.

### Création de l'anomalie

L'anomalie gaussienne est créée sur les deux niveaux. La gaussienne du bas est légèrement déphasée vers l'aval de l'écoulement ce qui correspond au cas d'interaction le plus fort entre les deux niveaux.

$$q'_k = A \exp\left( \frac{(x - x_0 - \delta_{k2} d)^2}{\sigma^2} + \frac{(y - y_0)^2}{\sigma^2} \right)$$

$(x_0, y_0)$ étant le centre de la gaussienne et $\sigma$ caractérise sa largeur et $d$ est le déphasage entre les deux niveaux. Les valeurs suivants ont été utilisées, $A = 2.10^{-5} s^{-1}$, $\sigma = 700km$ et $d = 500km$. Ce choix d'anomalie initialement déphasée permet d'avoir une croissance immédiate de celle-ci et ainsi limiter le temps d'intégration nécessaire pour obtenir une cyclogénèse.

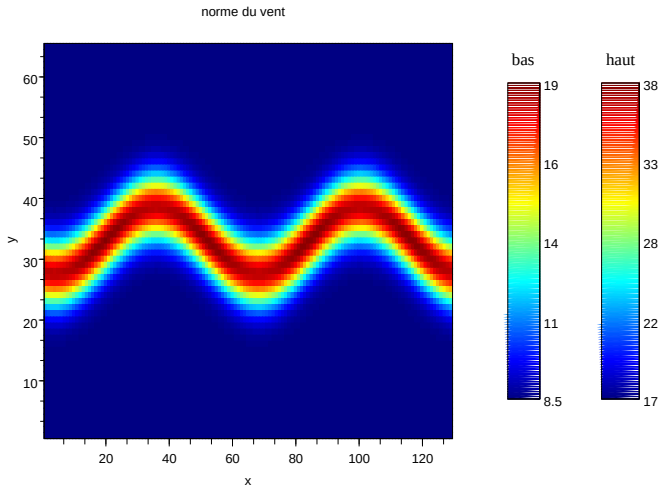



FIGURE 5.5 – Intensité du jet sur les deux niveaux. La forme est identique mais les valeurs sont différentes, échelle de gauche pour la couche inférieure et de droite pour la couche supérieure.

**Simulation numérique**

Une intégration de 4 jours à partir de cette anomalie permet d'obtenir une cyclogénèse bien établie avec la présence de fronts. La figure 5.6 montre l'évolution d'une anomalie sur 4 jours.

L'anomalie initiale est gaussienne sur chaque niveau. Un développement aval apparaît le premier jour. A 48h on observe l'apparition de gradients de tourbillon associés aux fronts chauds et froids. Une fois les fronts en place, la perturbation s'amplifie en puisant dans la zone barocline du jet. A quatre jours l'amplitude est telle que le modèle devient instable, les ondes de petites échelles sont le début d'une instabilité numérique.

La trajectoire obtenue est contrainte par le jet et par l'amplitude de l'anomalie initiale. Cette configuration permet d'avoir une variabilité importante entre les différentes trajectoires issues d'anomalies perturbées. Pour limiter les calculs une intégration de deux jours uniquement est utilisée pour les calculs de la KFS.

## 5.2 Modèle de Lorenz

Le deuxième modèle utilisé pour les simulations numériques, le modèle dit de Lorenz 96, est beaucoup plus simple que le modèle quasi-géostrophique à 2 couches. Sa simplicité permet d'effectuer des calculs explicites à moindre coût des valeurs que l'on souhaite estimer, comme la réduction de variance de la fonction score $S$.

### 5.2.1 Equation du modèle de Lorenz

Le modèle de Lorenz a été introduit par Lorenz and Emanuel (1998). Il est très utilisé en météorologie car il permet de représenter un caractère chaotique de manière simple, avec l'association de termes non linéaires et d'un forçage. Sa dynamique est décrite par l'équation suivante,

$$\frac{dX_i}{dt} = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F, \tag{5.12}$$

pour $i = 1..N$ avec des conditions limites périodiques. Pour les simulations, la dimension du vecteur d'état choisie est généralement de $N = 60$. $F$ est le terme de forçage. Sa valeur est fixée à 4 pour obtenir un système stationnaire (la dissipation de l'énergie dans tout le système est égale à celle apportée par le

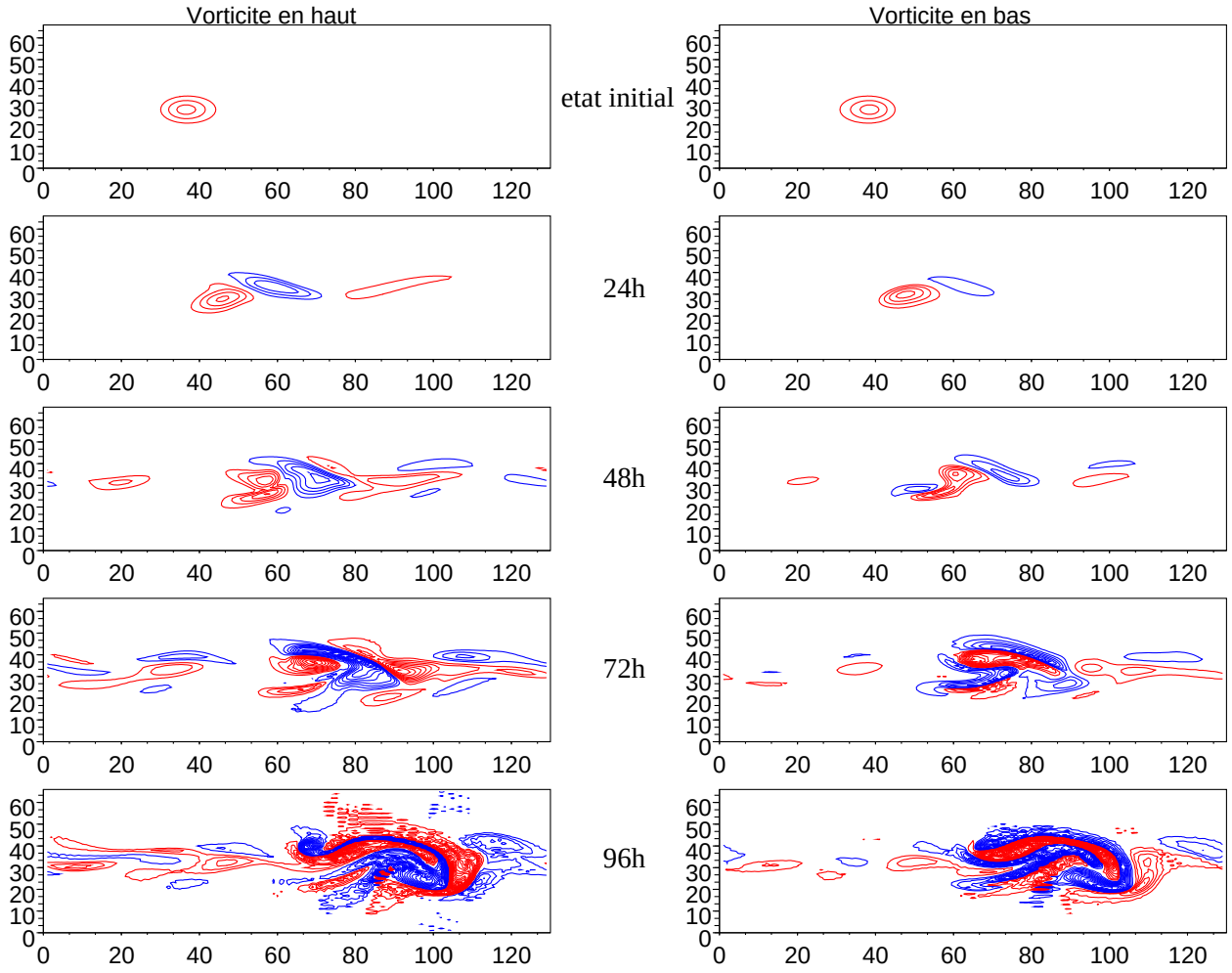



FIGURE 5.6 – Évolution d'une anomalie de tourbillon sur 4 jours. La colonne de gauche montre la couche supérieure, celle de droite la couche inférieure. Valeurs positives en rouge, négatives en bleu, intervalle $2.10^{-5}$.

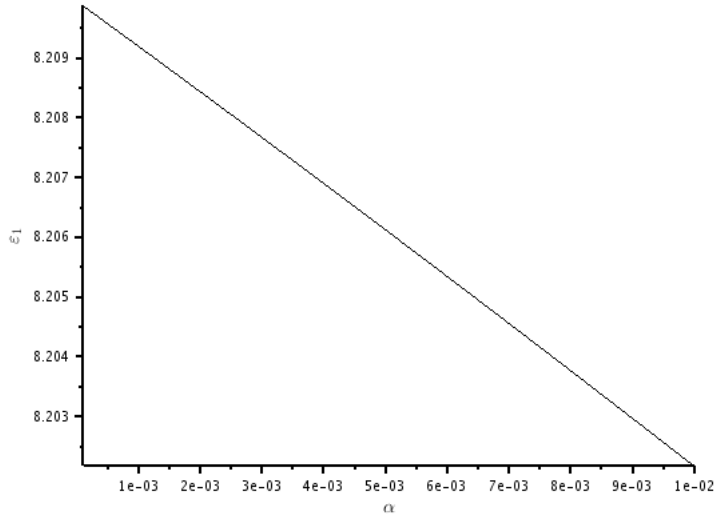



FIGURE 5.7 – Evolution de l'erreur $\varepsilon = \frac{\mathcal{M}(\boldsymbol{x}+\alpha\boldsymbol{u})-\mathcal{M}(\boldsymbol{x})}{\alpha} - \boldsymbol{M}_{\boldsymbol{x}}\boldsymbol{u}$ en fonction de $\alpha$ pour la première composante.

forçace, voir Lorenz and Emanuel (1998)). Pour retrouver les taux de croissances de l'atmosphère, le pas de temps est choisi égal à $dt = 0.05$, ce qui correspond à $6h$.

Le schéma temporelle utilisé est un Runge-Kutta d'ordre 4. Ce schéma a l'avantage de proposer un développement à un ordre élevé sans nécessiter le calcul des dérivées successives de $\frac{dX_i}{dt}$. Ce schéma est stable même sans filtrage supplémentaire pour les paramètres utilisés dans les expériences.

### 5.2.2 Implémentation du modèle tangent linéaire et de l'adjoint dans Lorenz

Compte tenu de la faible dimension du vecteur d'état, il est possible de construire explicitement la matrice du modèle tangent linéaire $\boldsymbol{M}$. Pour se faire, on estime chacune de ses composantes à partir du calcul du gradient dans chaque direction,

$$m_{ij}(\boldsymbol{x}, h) = \left(\frac{\mathcal{M}(\boldsymbol{x} + h\boldsymbol{e}_j) - \mathcal{M}(\boldsymbol{x})}{h}\right)_i \tag{5.13}$$

avec $\boldsymbol{e}_j$ le vecteur de la base canonique. Pour $h$ suffisamment petit, on utilise $h = 10^{-5}|\overline{\boldsymbol{x}}|$, on obtient une bonne estimation du gradient. Le modèle adjoint est ainsi directement obtenu comme la transposée de $\boldsymbol{M}$.

#### Validation des modèles tangent linéaire et de l'adjoint

Comme pour le modèle QG, une validation des modèles tangent linéaire et de l'adjoint est réalisée. La mise en œuvre est beaucoup plus aisée dans ce modèle de petite dimension mais ce sont les mêmes critères de validation qui sont utilisés ((5.10) et (5.11)).

La figure 5.7 montre l'évolution de $\varepsilon = \frac{\mathcal{M}(\boldsymbol{x}+\alpha\boldsymbol{u})-\mathcal{M}(\boldsymbol{x})}{\alpha} - \boldsymbol{M}_{\boldsymbol{x}}\boldsymbol{u}$ en fonction de $\alpha$ pour $\alpha$ décrivant $[10^{-4}\ 10^{-2}]$ pour la première composante. La décroissance étant linéaire, le résidu est bien un terme d'ordre 2 en $\alpha$. Ce qui valide le modèle linéaire au premier ordre. Comme le modèle adjoint est construit directement comme la transposée de la matrice du tangent linéaire, on retrouve bien l'égalité des produits scalaires en prenant 2 vecteurs quelconques (non illustrée).

## 5.3 Bilan

Pour les besoins de l'étude de l'influence de la trajectoire de la KFS, un modèle suffisamment réaliste mais simple à mettre en place est nécessaire. Le modèle QG à deux couches satisfait à ces contraintes. De plus lors de la mise en place de calcul de champ de sensibilité, le cadre synoptique est supposé bien représenté par le modèle. Ceci correspond à l'écoulement de base, stationnaire, la KFS est ici utilisée pour optimiser les observations par rapport à la trajectoire de l'anomalie. Ce qui est présenté dans le chapitre 6.

Le modèle de Lorenz est un modèle couramment utilisé en PNT pour la validation de nouveaux algorithmes. Sa dynamique est suffisamment complexe pour que les solutions ne soient pas triviales mais le vecteur d'état de petite dimension permet d'effectuer tous les calculs de manière explicite. Il est utilisé dans le chapitre 7 pour illustrer l'implémentation de la VRF.

# Chapitre 6

# Etude de la sensibilité à la trajectoire

Comme indiqué en introduction (chap. 1), l'objectif premier de ce travail de thèse est de documenter l'incertitude liée à la trajectorie de référence lorsque l'on effectue un calcul de sensibilité aux observations en utilisant un modèle adjoint. Cette étude est réalisée en utilisant des ensembles de champ de sensibilité de KFS, la méthode d'observation adaptative utilisée à Météo-France et présentée au chapitre 4.3.2. Trois expériences ensemblistes différentes ont été construites pour documenter la variabilité des statégies de déploiement proposées par la KFS (exp. I, 6.2.1), mais aussi pour proposer une méthode permettant de tirer profit de l'information issue de l'ensemble (exp. II, 6.2.2). Une validation du formalisme et des hypothèses est ensuite envisagée (exp. III, 6.2.3).

Pour analyser les limites de la KFS, il est nécessaire de construire un système d'assimilation pour le modèle QG (décrit en section 5.1.4). La problématique de l'étude est d'abord rappelée et le système d'assimilation utilisé est présenté 6.1.2. Ensuite le protocole expérimental, les différentes configurations implémentées et les expériences menées sont introduites. Finalement les résultats de ces expériences sont montrés et discutés. Les résultats présentés dans ce chapitre ont fait l'objet d'une publication dans la revue Quarterly Journal of the Royal Meteorological Society d'avril 2012, (Oger et al., 2012), cet article figure dans l'annexe H.

### Problématique de l'étude

L'écoulement de base du modèle QG correspond à une situation synoptique de grande échelle que l'on suppose certaine. La KFS est ici mise en place pour déterminer le déploiement d'observations supplémentaires associé à la trajectoire non linéaire d'une perturbation de cette écoulement, comme une anomalie d'altitude par exemple. L'objectif de l'étude est d'analyser la dépendance et la variabilité de la KFS à cette trajectoire. Pour ce faire une méthode basée sur des ensembles est utilisée. Chaque membre de l'ensemble représente donc une réalisation possible de cette perturbation. L'incertitude initiale sur cette perturbation est représentée par la matrice de covariance d'erreur d'ébauche $\boldsymbol{B}$. Suivant la position initiale de la perturbation on s'attend à obtenir des trajectoires différentes, et donc des solutions diverses pour le déploiement des observations. Le rôle du système d'assimilation n'est cependant pas négligeable, en effet les opérateurs d'assimilation interviennent directement dans les calculs de la KFS. Une étude purement dynamique n'est donc pas envisageable. Le calcul de la KFS est réalisé pour chaque membre, ce qui permet de représenter directement la variabilité à laquelle est soumise la KFS. Un récapitulatif du calcul d'un champ de sensibilité est présenté ci-dessous.

### Calcul du champ de sensibilité

Le champ de sensibilité est donné par $\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S$ (voir (4.3.2)). Il est donc nécessaire d'intégrer l'état initial $\boldsymbol{x}^b(t_{obs})$ de $t_{obs}$ à $t_{verif}$ par le modèle non linéaire, $\boldsymbol{x}^b(t_{verif}) = \mathcal{M}(\boldsymbol{x}^b(t_{obs}))$. On en déduit le gradient de $S$ à $t_{verif}$, $\nabla_{\boldsymbol{x}^f} S$, qui est ramené à l'instant d'assimilation par le modèle adjoint, $\nabla_{\boldsymbol{x}^a} S = M^T \nabla_{\boldsymbol{x}^f} S$. En parallèle une assimilation est effectuée pour obtenir une approximation de la matrice $\boldsymbol{A}_c$ par l'algorithme

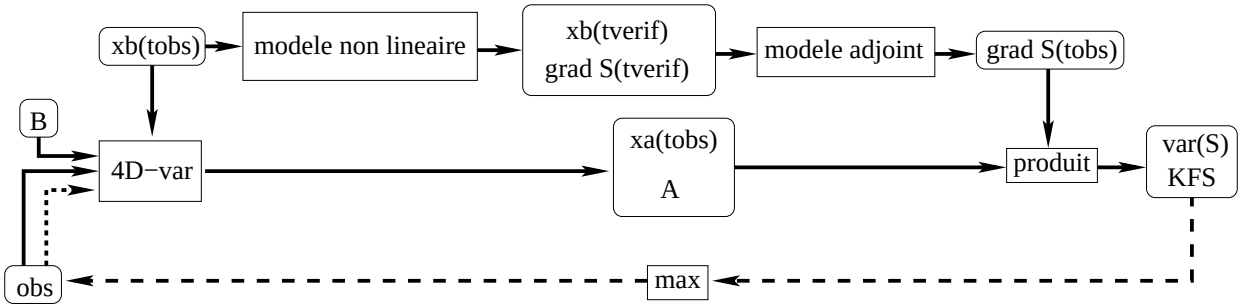



FIGURE 6.1 – Algorithme élémentaire de calcul de la KFS.

de Lanczos (voir annexe B). Pour finir le produit est calculé, $KFS = \boldsymbol{A}\nabla_{\boldsymbol{x}^a} S$. La variance est calculée en même temps par $\sigma^2_{S_c} = (\nabla_{\boldsymbol{x}^a} S)^T \boldsymbol{A}_c \nabla_{\boldsymbol{x}^a} S$. Les observations supplémentaires sont positionnées aux maxima de champ de sensibilité sur les deux niveaux. Une nouvelle assimilation est calculée en utilisant ces observations supplémentaires, ce qui donne une estimation de $\boldsymbol{A}_{c+t}$ et la variance associée $\sigma^2_{S_{c+t}}$. On en déduit une estimation de la réduction de variance associée à ces observations supplémentaires.

La figure 6.1 résume les calculs à effectuer. Le terme *obs* comprend les valeurs des observations $\boldsymbol{y}^o$, les opérateurs $\boldsymbol{H}$ et les erreurs associées $\boldsymbol{R}$. La boucle de retour entre le champ de KFS et les matrices $\boldsymbol{H}$ et $\boldsymbol{R}$ indique que l'on peut reproduire le calcul de manière récursive en prenant en compte les observations supplémentaires déjà introduites.

Pour démarrer ce calcul, il est donc nécessaire de construire la condition initiale $\boldsymbol{x}^b(t_{obs})$, mais aussi le système d'assimilation $\boldsymbol{B}$, $\boldsymbol{H}_c$ et $\boldsymbol{R}_c$. La trajectoire est calculée sur deux jours afin d'obtenir le gradient de fonction score à $t_{verif}$. La fonction score choisie est l'enstrophie dans la couche inférieure, son gradient est donc la vorticité $\zeta_2 = \Delta\psi_2$.

## 6.1 Protocole et cadre expérimental

Cette section introduit les différentes configurations des simulations et du système d'assimilation utilisées. Les différentes expériences réalisées à partir des ensembles sont ensuite présentées.

### 6.1.1 Choix de la condition initiale

Trois types de conditions initiales ont été utilisées pour créer les ensembles. La première, la plus simple, consiste a déphaser l'anomalie initiale, introduite en section 5.1.4, en latitude et longitude de manière aléatoire. Son amplitude est aussi modifiée de manière aléatoire. Une trajectoire de contrôle est définie comme l'état non perturbé (pas de déphasage ), les champs associés sont montrés en exemple 6.1.4. Le déphasage génère des perturbations très distinctes les unes des autres, ce qui permet une grande variabilité des trajectoires.

Ce premier cas étant trop simple, il n'est pas possible de déterminer la correction à la trajectoire lié aux observations. Pour obtenir un calcul plus réaliste un second type de conditions initiales est envisagé. Une partie de l'écoulement de base est introduit dans la variable d'état du modèle. Le jet n'est alors plus stationnaire et son évolution est prise en compte. De plus il n'est plus considéré comme certain, et est lui aussi déphasé de manière aléatoire pour chaque membre. Cela permet de lui associer une incertitude.

Pour ces deux types de conditions initiales, une matrice $\boldsymbol{B}$ est construite de manière à ce que le champ de variance corresponde à la distributions des perturbations (voir 6.1.2). Le dernier type de condition initiale est lui construit de manière à ce que la distribution des erreurs soit gaussienne et compatible avec le $\boldsymbol{B}$ de l'assimilation (voir annexe D). Les perturbations sont obtenus comme $\boldsymbol{x}^b_i = \overline{\boldsymbol{x}}^b + \varepsilon^b$, où $\overline{\boldsymbol{x}}^b$ est l'anomalie de contrôle et $\varepsilon^b$ une erreur d'ébauche.

Pour chaque type de condition initiale, une perturbation est choisie comme trajectoire de référence $\boldsymbol{x}^t$.

Les observations sont alors construites en perturbant le champ de tourbillon potentiel $q$, $\boldsymbol{y}^o = \boldsymbol{H}\boldsymbol{x}^t + \varepsilon^o$, avec $\varepsilon^o$ compatible avec $\boldsymbol{R}$. A chaque condition initiale correspond aussi une matrice $\boldsymbol{B}$.

### 6.1.2 Système d'assimilation de données

En parallèle des conditions initiales il est nécessaire de choisir un système d'assimilation de données. Pour l'ensemble des simulations, le système utilisé est un 4D-var. La fenêtre d'assimilation est limitée de l'instant initial $t_{obs}$ à une échéance de 3h, seuls ces deux instants sont observés. Le critère d'arrêt pour la minimisation est un nombre maximum de 15 itérations, ou la norme du gradient $\nabla J < 10^{-9}$. La matrice $\boldsymbol{B}$ et le réseau d'observation de routine sont choisis en fonction de l'expérience.

#### Modélisation de B

Comme présenté dans le chapitre 2 consacré à l'assimilation de données, la matrice de covariance d'erreur d'ébauche $\boldsymbol{B}$ représente la structure spatiale de la distribution des erreurs d'ébauche. Elle est représentative des zones d'incertitude dans le modèle. Une zone de forte incertitude correspond à des fortes valeurs de variance (la diagonale de $\boldsymbol{B}$) et inversement des zones de confiance correspondent à des faibles valeurs. La partie extra-diagonale de $\boldsymbol{B}$, les covariances, propage l'information des observations pour corriger l'ébauche. Pour simplifier la représentation en machine de $\boldsymbol{B}$, il est classique de la modéliser sous la forme d'un produit d'opérateurs. Une première décomposition permet de séparer la variance des corrélations,

$$\boldsymbol{B} = \Sigma \boldsymbol{C} \Sigma^T, \tag{6.1}$$

où $\Sigma$ est la matrice d'écart-type d'ébauche en chaque point de grille et $\boldsymbol{C}$ la matrice des corrélations (Courtier et al., 1998). La fonction de corrélation est supposée homogène en espace, ce qui correspond à une décorrélation des nombres d'ondes. Plus d'informations sur ce que représentent les matrices $\Sigma$ et $\boldsymbol{C}$ et la manière de les construire sont données en annexe A.

Les champs de variance utilisés pour chaque expérience sont montrés figure 6.2, de part la manière dont sont crées les ensembles on trouve une zone de forte variance au niveau de la position de l'état de contrôle. La variance est dehors de cette zone est faible, les observations auront donc un impact limité à la région de l'anomalie initiale.

#### Réseau de routine

Le réseau d'observation de routine est supposé constant dans le temps et est identique sur les deux niveaux du modèle QG. Deux réseaux d'observation sont envisagés (voir figure 6.3). Le premier réseau (*ResD*) est choisi de manière à très bien observer les zones de fortes variances visibles sur la figure 6.2. Les expériences ont aussi montrées que cette dans zone le système est aussi le plus sensible aux conditions initiales (exp I, 6.2.1). Le second réseau (*ResH*) est lui choisi homogène de manière à moins contraindre les champs de sensibilité.

### 6.1.3 Description des expériences réalisées

A partir des différents éléments introduits précédemment, plusieurs expériences ont été réalisées afin de diagnostiquer le comportement de la KFS par rapport aux trajectoires.

- La première expérience (exp I) consiste à évaluer la variabilité du champ de KFS en fonction des composants du système d'assimilation utilisé. Pour ce faire trois ensembles de trajectoires sont utilisés à partir des perturbations de l'anomalie seule. Pour chacun des ensembles la matrice $\boldsymbol{B}$ ou le réseau d'observation est modifié. Le jet est ici supposé parfaitement connu.
- Une seconde expérience (exp II) a été effectuée en incluant le jet dans la condition initiale, ce qui permet de lui attribuer une incertitude. A partir de cet ensemble une stratégie objective de déploiement d'observations supplémentaires est proposée 6.2.2.

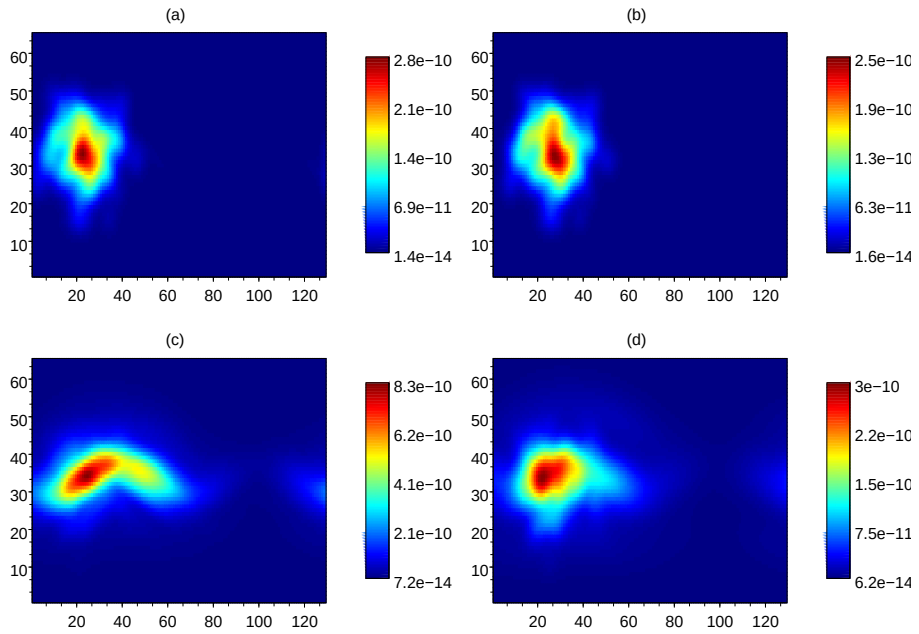



FIGURE 6.2 – Variance d'erreur d'ébauche pour l'anomalie seule perturbée en haut, et pour l'anomalie et le jet en bas. Le colonne de gauche correspond au niveau supérieur et de droite au niveau inférieur.

– La dernière expérience (exp III) consiste à évaluer la capacité de la KFS à anticiper la réduction de variance d'erreur de prévision au travers de la fonction score $S$. Pour se faire les variances $\sigma_S^2$ estimées par la KFS sont comparées à celles obtenues de manière effective par l'ensemble.

### 6.1.4 Exemple de résultats préliminaires

Avant de décrire les résultats obtenus à partir des ensembles, deux exemples sont utilisés pour illustrer les champs manipulés lors du calcul de la KFS. Le premier montre l'évolution de l'incertitude entre l'ébauche et l'analyse par la prise en compte des observations. Le second est le calcul de la KFS pour la trajectoire de contrôle (non perturbée).

#### Réduction de variance due à l'assimilation

Le but de l'assimilation de données est de combiner l'information issue de l'ébauche à celle des observations pour réduire l'incertitude et obtenir un état plus proche de la réalité. Dans le cas du 4D-var cette analyse est obtenue par minimisation de la fonction coût $J$. Lors de la minimisation une estimation de la matrice de covariance $\boldsymbol{A}$ est obtenue par l'algorithme de Lanczos. A partir de cette estimation il est possible de déterminer la carte de variance associée à $\boldsymbol{A}$.

La figure 6.4 compare la variance d'erreur d'ébauche et l'estimation de la variance d'erreur d'analyse pour les deux niveaux dans le cas du réseau d'observation *ResD*. Une première remarque est que la valeur maximale de variance d'erreur a fortement diminué après l'assimilation. Pour la couche supérieure par exemple le maximum passe de $2.8\,10^{-10}$ à $8.1\,10^{-11}$. De manière générale l'analyse est donc plus certaine que l'ébauche. En observant la champ de variance analysée entre les indices 30 et 40 en $y$, le rôle des observations est clairement visible. La région de forte incertitude sur $\boldsymbol{B}$ a été supprimée par la bande d'observation dense. La matrice $\boldsymbol{A}$ intervenant dans le calcul de la KFS, cela permet de comprendre le rôle du réseau d'observation.

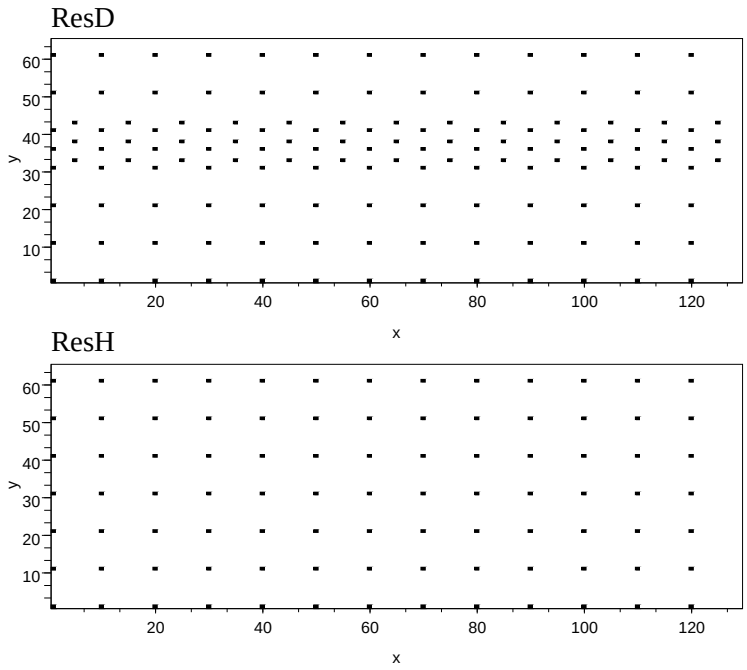



FIGURE 6.3 – Réseau d'observation de routine, dense en haut et homogène en bas.

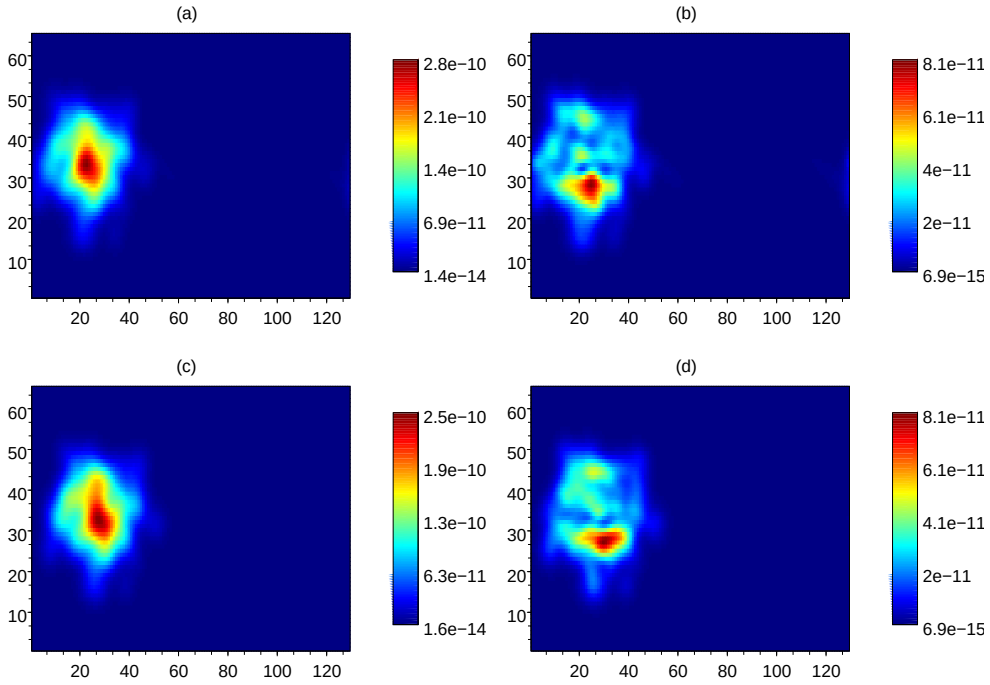



FIGURE 6.4 – Variances d'erreur d'ébauche (a) et (c) comparées à l'estimation des variances d'erreur d'analyse (b) et (d), en haut sur la couche supérieure et en bas sur l'inférieure.

**Exemple du calcul de champ de sensibilité**

Les différents champs utilisés lors du calcul d'un champ de sensibilité sont maintenant illustrés. Le calcul est ici effectué en considérant une anomalie placée au centre du jet.

L'évolution à deux jours (figure 6.5) montre la croissance de l'anomalie et sa déformation. Le champ de vorticité est semblable à celui de tourbillon potentiel. Une fois propagé de manière rétrograde par l'adjoint, on obtient le gradient de $S$ à l'instant initial (figure 6.6). Le réseau d'observation *ResD* est utilisé pour ce calcul. Le rôle des observations se retrouve en comparant les champs $\boldsymbol{B}\nabla S$ et $\boldsymbol{A}\nabla S$, les fortes valeurs du gradient entre 30 et 40 en $y$ ont disparu. Il ne subsiste des fortes valeurs que dans la région sud du gradient qui est moins observée.

A partir de ce type de champ il n'est pas possible d'évaluer directement le rôle de la trajectoire. C'est pourquoi des ensembles de trajectoires sont générés pour étudier la variabilité de la KFS. Tous calculs ci-dessus sont effectués pour chaque membre.

## 6.2 Résultats et discussion

Cette section présente le développement des résultats associés aux expériences présentées précédemment. Une discussion est effectuée à la fin de chaque section.

Les expériences utilisant des configurations différentes un récapitulatif est donné dans le tableau suivant.

| | type de perturbation | réseau d'observation | observations perturbées | nombre de membres |
|---|---|---|---|---|
| exp I | anomalie seule | variable | non | 50 |
| exp II | anomalie et jet | *ResD* | oui | 100 |
| exp III | compatible avec $\boldsymbol{B}$ | *ResD* | oui | 100 |

Pour les expériences II et III, pour que la dispersion des ensembles assimilés soit représentative de l'erreur d'observation, celles-ci sont perturbées pour chaque membre par un bruit blanc compatible avec $\boldsymbol{R}$ (Burger et al., 1998).

### 6.2.1 Exp I : Variabilité du champ de KFS à la trajectoire

L'exemple présenté montre l'importance du système d'assimilation dans le champ de KFS. L'objectif de cette expérience est de tester la variabilité des champs de KFS à la trajectoire en fonction du système d'assimilation. Pour ce faire des ensembles de champs de KFS sont calculés à partir de réseau d'observation et de champs de variance d'ébauche différents. Le tableau ci-dessous décrit les conditions pour les trois expériences.

| | réseau d'observation | $\sigma^{b2}$ |
|---|---|---|
| exp I.a | dense | spatialisée |
| exp I.b | homogène | spatialisée |
| exp I.c | homogène | uniforme |

Les réseaux d'observations *ResD* et *ResH* correspondent à ceux présentés figure 6.3. La variance spatialisée correspond au champ (a) et (b) de la figure 6.2, dans le cas uniforme la valeur est choisie est la moyenne du champ spatialisé. Chacune des expériences contient 50 membres.

Les champs étudiés ici sont les champs moyens pour chaque ensemble, ainsi que la localisation des maxima et leurs valeurs pour les différentes expériences.

**Analyse des champs moyens**

Les figures 6.7, 6.8 et 6.9 montrent les différents champs moyennés sur l'ensemble pour respectivement les expériences I.a, I.b et I.c. Pour les trois expériences on retrouve un maximum de gradient $\nabla_{\boldsymbol{x}^a} S$ au niveau de la position moyenne de l'anomalie, ce qui correspond aussi au maximum de $\sigma^b$. Le champ $\boldsymbol{B}\nabla S$

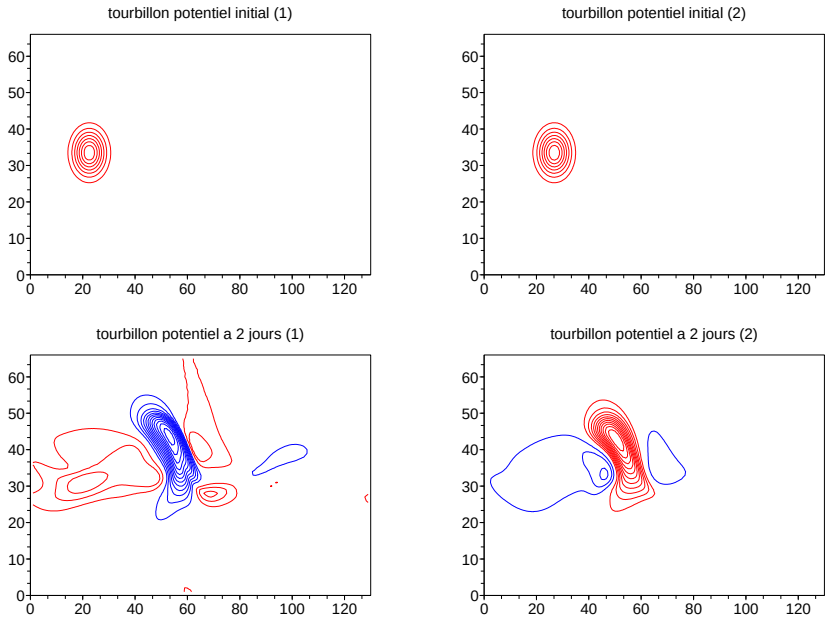



FIGURE 6.5 – Tourbillon potentiel QG à l'instant initial (haut) et à deux jours (bas) pour la trajectoire de contrôle, à gauche sur la couche supérieure et à droite sur la couche inférieure. Valeurs positives en rouge, négatives en bleu. Intervalle à l'instant initial $5.10^{-6}$ et à deux jours $2.10^{-5}$

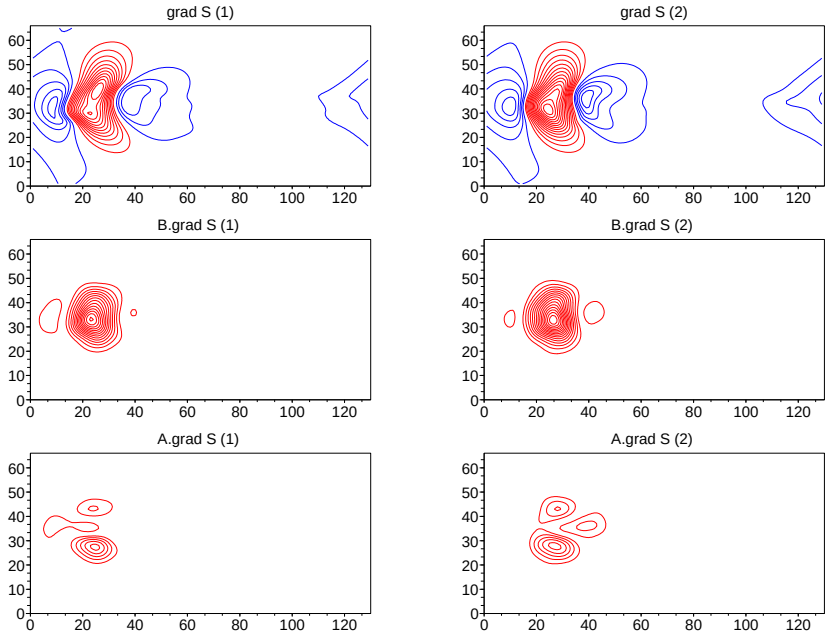



FIGURE 6.6 – Gradient de fonction score ($\nabla S$) en haut, $|\boldsymbol{B}\nabla S|$ au centre et $|\boldsymbol{A}\nabla S|$ en bas, à gauche sur la couche supérieure et à droite sur la couche inférieure. Intervalle $1.10^{-8}$ pour le gradient et $1.10^{-12}$ pour les champs de sensibilité.

moyen est donc fortement corrélé au maximum de variance de $\boldsymbol{B}$. Ces résultats sont cohérents étant donné que la dynamique et les perturbations pour les trois ensembles sont identiques. Des différences interviennent cependant sur les champs de sensibilité. Dans le cas I.a, lorsque le réseau *ResD* est utilisé on retrouve le comportement de l'exemple, la zone de sensibilité est associée à la marge sud du gradient. En relâchant la contrainte du réseau d'observation, expérience I.b avec le réseau *ResH*, le signal de sensibilité est ramené sur la zone de forte valeur du gradient et de $\sigma^b$. Celui-ci se limite tout de même à la bande non observée au centre du domaine (voir réseau d'observation figure 6.3). Le maximum de valeur moyenne est encore peu étalé et fortement corrélé à $\sigma^b$. En prenant un champ de $\sigma^b$ uniforme, expérience I.c, le champ moyen de sensibilité est beaucoup plus étalé et correspond à la forme du gradient.

Au niveau des variances de KFS pour les deux expériences (I.a et I.b) les champs sont sensiblement semblables à ceux de la moyenne, ce qui indique que les membres sont en accord en ce qui concerne la position mais que la valeur du champ est elle très variable. Pour l'expérience I.c en revanche un zone de plus forte variance est visible autour de (40,25). Celle est reliée aux valeurs fortes de KFS données par plusieurs membres et visibles figure 6.10.

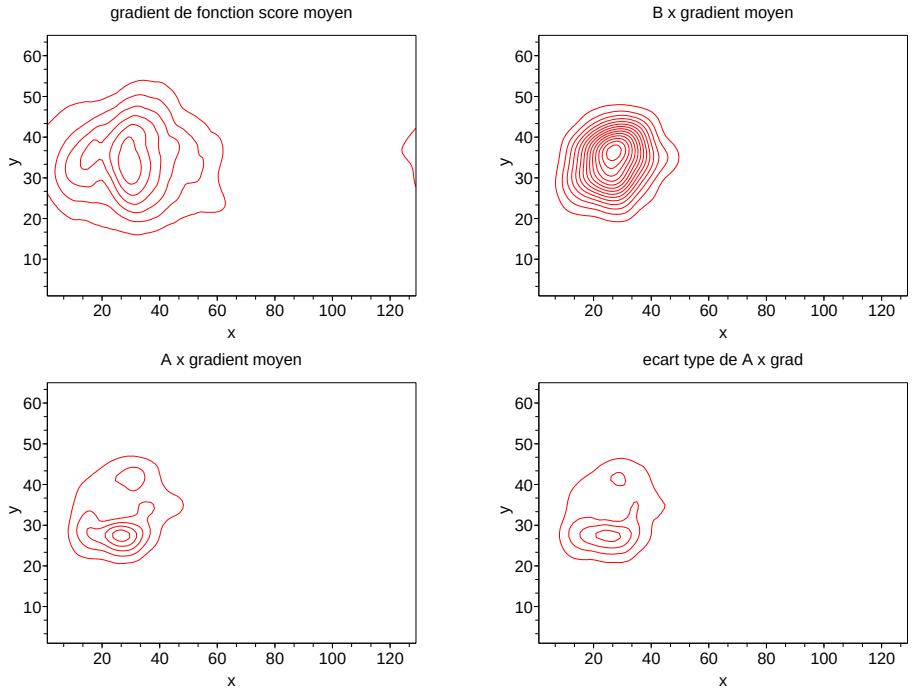



FIGURE 6.7 – Champs composites, pour l'expérience I.a, des deux niveaux calculés à partir de l'ensemble. Haut gauche : moyenne du gradient de fonction score $\nabla S$ à $t_{obs}$. Haut droite : moyenne de $\boldsymbol{B}\nabla S$. Bas gauche : moyenne de $\boldsymbol{A}\nabla S$. Bas droite : écart type de $\boldsymbol{A}\nabla S$. Intervalle isolignes, pour le gradient $1.10^{-8}$, pour les autres champs $5.10^{-9}$.

### Distribution des maxima

La figure 6.10 montre la position et la valeur du maximum de champ de KFS pour les 50 membres pour les trois différentes expériences. Chaque point correspond donc à l'observation supplémentaire pour le membre correspondant. Les valeurs maximales pour les expériences I.a et I.b sont équivalentes et surtout positionnées dans les zones peu observées (le sud du jet pour a et le centre pour b), ce qui est cohérent. En modifiant les valeurs de $\sigma^b$ dans l'expérience c, on réduit aussi les valeurs de KFS, la région de forte incertitude ayant disparu. Pour l'expérience I.a, on observe peu des fortes valeurs (rouge) par rapport à l'expérience I.b et les valeurs de maxima entre les membres sont peu dispersées géographiquement. Pour I.b, la dispersion géographique est toujours faible, les valeurs sont alignées dans l'axe absent d'observation,

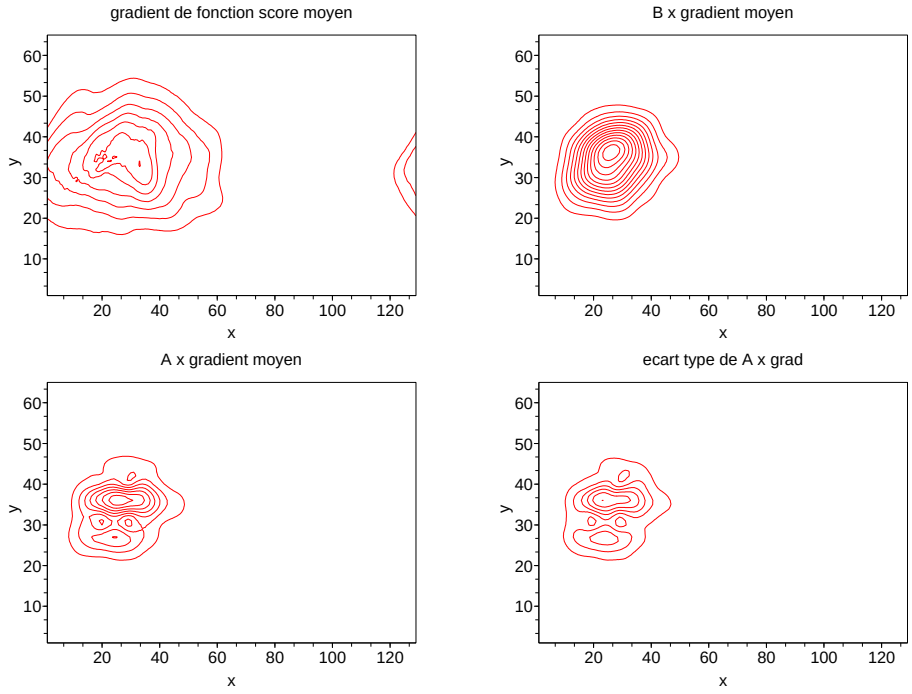



FIGURE 6.8 – Champs composites pour l'expérience I.b, idem que 6.7

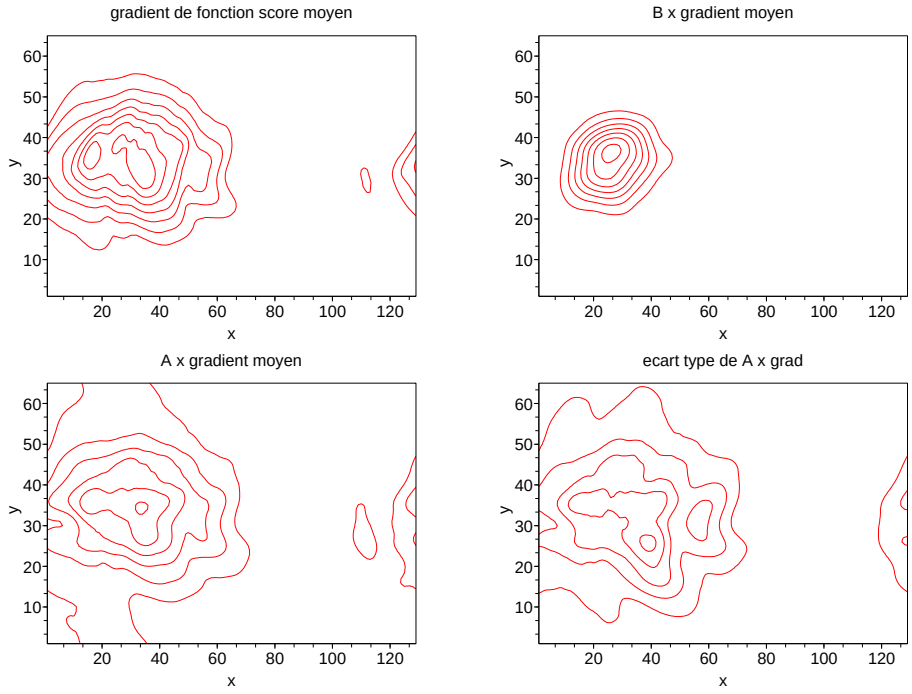



FIGURE 6.9 – Champs composites pour l'expérience I.c, idem que 6.7

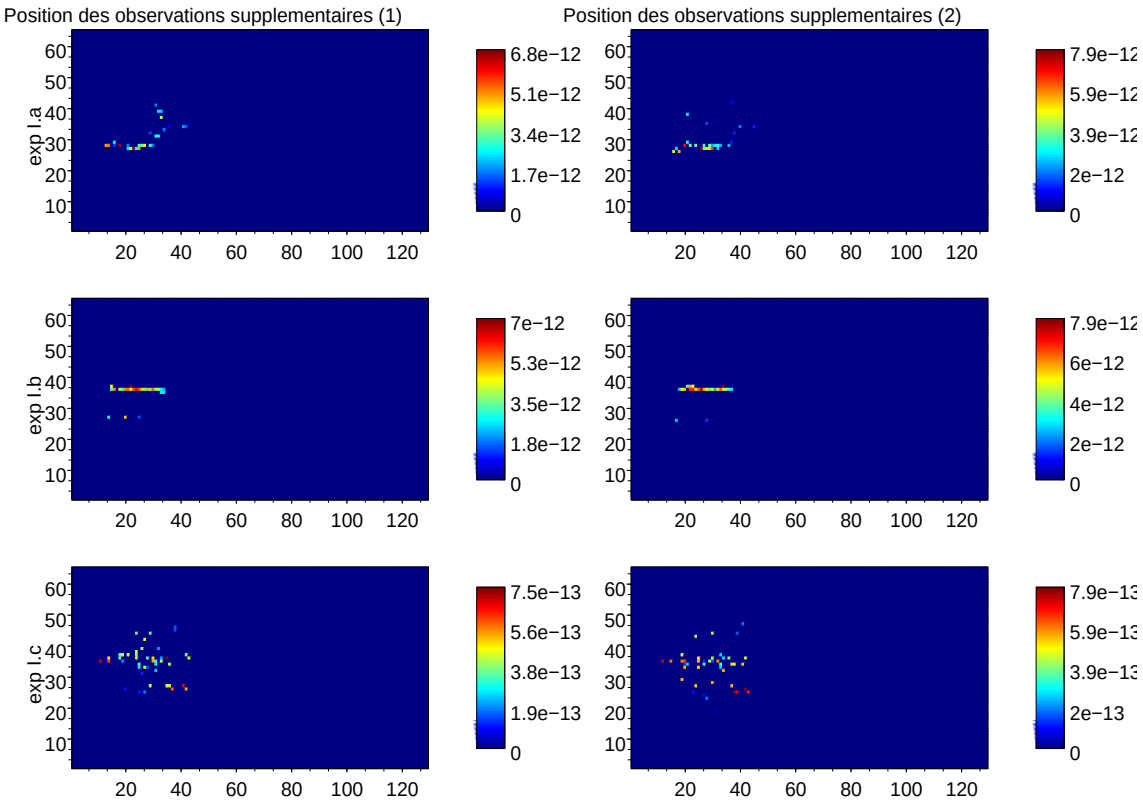



FIGURE 6.10 – Position des maxima de KFS sur les deux niveaux pour l'ensemble des membres. La valeur indiquée est la moyenne des maxima en ce point de grille (voir texte pour les différentes expériences).

cependant les écarts de valeur sont grands. Dans le cas de I.c, les valeurs sont alors beaucoup plus dispersées, et il est beaucoup moins évident de spécifier une région de forte valeur. On retrouve la bande centrale mais aussi la région (40,25) où plusieurs membres ont des fortes valeurs. En contraignant d'avantage le système d'assimilation on réduit donc la part d'incertitude liée à la trajectoire.

Lors des calculs de KFS, la variance $\sigma_S^2 = [\nabla_{\boldsymbol{x}^a} S]^T \boldsymbol{A} [\nabla_{\boldsymbol{x}^a} S]$ est aussi calculée en utilisant la matrice associée au réseau de routine $\boldsymbol{A}_c$ et celle incluant les observations supplémentaires $\boldsymbol{A}_{c+t}$. La figure 6.11 montre la variance estimée de la fonction score et la réduction de variance $\delta\sigma^2_{S_t} = \sigma^2_{S_{c+t}} - \sigma^2_{S_c}$ due à l'observation supplémentaire, pour chaque membre, en fonction de la valeur maximale de KFS, pour l' expérience I.c. On trouve une croissance de la variance $\sigma_S^2$ en fonction du maximum de KFS $\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S$, ce qui correspond au formalisme, de même pour la réduction. En considérant un membre unique, une observation positionnée dans une zone de forte valeur de KFS, réduira d'autant plus la variance.

### Discussion et résultats de l'exp I

A partir des différentes expériences (a, b et c), l'influence du réseau d'observation et de la carte de variance $\sigma^{b2}$ sont clairement identifiées. Ainsi la prédiction du réseau d'observation de routine peut contraindre fortement le champ de sensibilité obtenu et donc le choix du déploiement. De même la carte de variance doit être suffisamment représentative de l'incertitude de la situation donnée. **L'influence de la trajectoire est d'autant plus faible sur la forme des champs que les contraintes du système d'assimilation sont fortes.**

En pratique lorsqu'un calcul de KFS est effectué le déploiement d'observations supplémentaires est conditionné à la valeur estimée de la réduction de variance. A partir des figures 6.10 et 6.11, il vient que dans le cas où l'on ne dispose que d'une unique simulation, et donc une unique valeur de KFS, il n'est pas

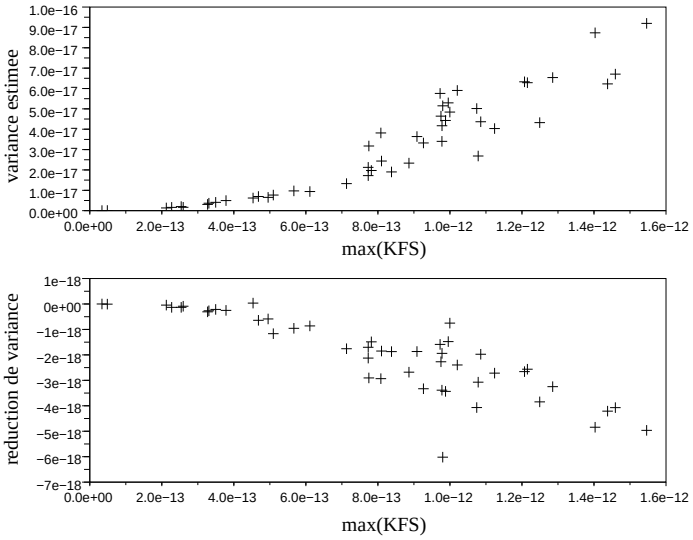



FIGURE 6.11 – Corrélation de la somme des maxima de KFS pour les deux niveaux par rapport la variance estimée en haut et la réduction de variance estimée en bas. Chaque croix donne les valeurs pour un membre de l'ensemble.

possible de discerner si une région correspond à une forte sensibilité aux conditions initiales uniquement par la valeur du maximum. Si une région de quelques points de grille est choisie comme zone de déploiement, suivant le membre considéré la réduction de variance associée peut fortement varier. **L'impact estimé d'une observation dépend autant de la position de cette observation que de la trajectoire considérée.**

En analysant les résultats de l'expérience I, il s'est avéré que le gradient de fonction score, et donc les zones de sensibilité, est gouverné par le jet. Les positions d'observations supplémentaires ne correspondent donc pas que au vecteur d'état utilisé, qui est ici uniquement une anomalie, si bien qu'il n'est pas possible de quantifier l'apport de ces observations supplémentaires sur la dispersion de l'ensemble. C'est pourquoi une partie de l'écoulement de base est introduite dans le vecteur d'état pour l'expérience II.

### 6.2.2 Exp II : Valorisation de l'ensemble pour optimiser le déploiement

L'expérience I a montré une grande variabilité de valeur des champs de KFS en fonction des trajectoires lorsque le système d'assimilation n'est pas trop contraignant. L'expérience II vise maintenant à proposer une stratégie pour utiliser l'information issue de cet ensemble de champs de sensibilité. Les conditions de l'expérience II sont une perturbation de l'anomalie et du jet, et un réseau d'observation *ResH*. L'analyse se porte sur la dispersion de l'ensemble en fonction de la stratégie de déploiement utilisée, en supposant le nombre d'observations supplémentaires disponibles limité et fixé à une par niveau içi.

Pour se faire la variance totale de l'ensemble est calculée comme la somme de la variance de tourbillon potentiel pour tous les points de grille ($\sigma_q^2$). Cette variance est calculée à l'instant initial et à deux jours pour trois *états* de l'ensemble, après assimilation du réseau de routine, après assimilation avec les observations ciblées de chaque membre et après l'assimilation avec des observations supplémentaires identiques pour tout l'ensemble.

L'objectif est ici d'utiliser l'information issue de l'ensemble pour choisir des observations supplémentaires pertinentes de manière à prendre en compte l'incertitude sur la trajectoire. Ce choix d'observation est appelé *observations consensus*. En réécrivant l'expression de la variance de fonction score en prenant

| réduction de variance en % | $\sigma_c^2$ | $\sigma_{c+t}^2$ | $\sigma_{c+m}^2$ | $\sigma_{c+r}^2$ |
|---|---|---|---|---|
| instant initial | 1 | 0.95 | 0.90 | 1.01 |
| prévision à deux jours | 1 | 0.97 | 0.94 | 1.009 |

TABLE 6.1 – Comparaison des stratégies de déploiement. Les valeurs sont des pourcentages de réduction par rapport à la variance sans observation supplémentaire. $t$ correspond à l'observation ciblée, $m$ au consensus et $r$ à l'aléatoire.

en compte le caractère probabiliste de la trajectoire, on obtient

$$\begin{aligned} \sigma_S^2 &= \mathbb{E}\left\{\mathbb{E}(\delta S\ \delta S^T|\boldsymbol{x}^a)\right\} \\ &= \mathbb{E}\left\{[\nabla_{\boldsymbol{x}^a} S]^T \mathbb{E}(\delta \boldsymbol{x}^a\ \delta \boldsymbol{x}^{aT})[\nabla_{\boldsymbol{x}^a} S]\right\}, \\ &= \mathbb{E}\left\{[\nabla_{\boldsymbol{x}^a} S]^T \boldsymbol{A}[\nabla_{\boldsymbol{x}^a} S]\right\}. \end{aligned}$$

Dans l'expression classique de la KFS, on suppose que $\nabla_{\boldsymbol{x}^a} S$ est certain et il n'y a donc pas de fonction de distribution associée. L'utilisation de l'ensemble permet d'estimer la distribution associé à la trajectoire par une méthode de Monte-Carlo. L'espérance est alors vue comme une moyenne d'ensemble.

En posant le problème de cette manière, on associe donc la réduction de variance due à des observations supplémentaires au champ $\mathbb{E}(\boldsymbol{A}\nabla_{\boldsymbol{x}^a} S)$, c'est à dire à la moyenne des champs de KFS. En effet une anomalie à forte croissance positionnée dans une région de forte incertitude donne de fortes valeurs de KFS. En effectuant la moyenne de tous les champs on retrouve donc ces zones sensibles, mais sans accorder une pleine confiance à une trajectoire unique. On peut donc définir de manière objective les positions des observations supplémentaires comme les maxima de moyenne pour chaque niveau. Il peut cependant être intéressant de pouvoir identifier des trajectoires de fortes valeurs extrêmes, qui seraient filtrées lors de la moyenne. Le calcul de la variance du champ de KFS permet de retrouver ces membres extrêmes de manière subjective. Le déploiement n'étant pas effectué de manière automatique, les champs de sensibilité sont soumis à une expertise objective et subjective.

Pour évaluer le choix des observations consensus, la réduction de variance de l'ensemble due aux observations ciblées pour chaque membre est comparée à celle due aux observations consensus sur chaque niveau. Une expérience de référence a aussi été effectuée en rajoutant une observation supplémentaire aléatoire sur chaque niveau. Les variances totales des ensembles sont données dans la tableau 6.1. Dans les deux cas d'observation supplémentaire une réduction de la dispersion de l'ensemble est obtenue. La réduction est plus important lorsque l'observation consensus est choisie. Dans le cas d'une observation prise de manière aléatoire, celle-ci n'a pas d'influence sur la dispersion de l'ensemble. L'observation consensus semble donc être un choix raisonnable pour le déploiement d'une observation supplémentaire.

#### Discussion et résultats de l'exp II

Plusieurs stratégies de déploiement ont été comparées, dans l'expérience II, pour tirer profit de l'information issue de l'ensemble. **Le choix de l'observation consensus est la manière la plus efficace de réduire la variance de l'ensemble, comparé à un choix aléatoire et à une stratégie où chaque membre utilise son observation ciblée.**

### 6.2.3 Exp III : Validation du formalisme

Le formalisme de la KFS est destiné à prendre en compte de manière implicite l'incertitude d'une prévision. Or en utilisant un ensemble de trajectoire il est possible de calculer explicitement l'incertitude. Le score $S$ de chaque trajectoire est calculé à partir de la prévision, en supposant l'ensemble non biaisé. La valeur de l'état vrai n'influence pas la dispersion des valeurs de score. A partir de cet ensemble de réalisations de $S$, on peut en déduire $\sigma_S^2$ et comparer cette valeur à celle obtenue par la KFS.

**Comparaison des calculs de variance de score**

Pour que les deux calculs soient équivalents il faut que la matrice de covariance de l'ensemble soit semblable à celle utilisée dans l'assimilation de données et donc par la KFS. C'est pourquoi les perturbations sont créés de manière cohérente avec $\boldsymbol{B}$ (comme défini précédemment). L'observation supplémentaire utilisée est la consensus afin que la trajectoire de chacun soit une réalisation du même déploiement d'observation (et donc pouvoir comparer les valeurs). Les valeurs de variance de fonction score obtenues sont données dans le tableau 6.2.

| variance de fonction score ($10^{15}s^{-2}$) | $\sigma^2_{S_c}$ | $\sigma^2_{S_{c+m}}$ | $\delta\sigma^2_{S_m}$ |
|---|---|---|---|
| KFS | 4588 | 3639 | -948 |
| ensemble | 8 | 3 | -5 |

TABLE 6.2 – Comparaison de la variance calculée par la KFS et par l'ensemble. $\sigma^2_{S_c}$ correspond à la variance lorsque uniquement le réseau de routine est utilisé, $\sigma^2_{S_{c+m}}$ tient compte de l'observation consensus. Les valeurs indiquées pour l'ensemble correspondent à la moyenne sur l'ensemble.

On observe une grande différence de valeur entre les deux calculs, avec une surestimation importante par la KFS. Celle-ci calcule une estimation de la variance sous la forme $\sigma^2_S = [\nabla_{\boldsymbol{x}^a} S]^T \boldsymbol{A} [\nabla_{\boldsymbol{x}^a} S]$. L'écart peu donc être du à une erreur dans l'estimation de $\boldsymbol{A}$ par l'algorithme de Lanczos ou au calcul de gradient par l'adjoint le long de la trajectoire.

**Analyse du champs de variance d'analyse**

La figure 6.12 montre la variance d'erreur d'analyse calculée à partir de l'ensemble comparée à l'estimation obtenue par le Lanczos. Les écarts de valeur sont faibles, même si la distribution spatiale des variances n'est pas exactement égale cela n'explique pas la différence de variance de fonction score (tableau 6.2). Cela permet aussi de valider l'utilisation de l'algorithme de Lanczos pour estimer $\boldsymbol{A}$.

**Étude du transport linéaire**

Lors du calcul du gradient à partir du modèle adjoint la trajectoire de l'ébauche (non assimilée) est utilisée. Or dans le formalisme de la KFS en écrivant $\delta S = [\nabla_{\boldsymbol{x}^a} S]^T \delta \boldsymbol{x}^a$ cela suppose que la trajectoire utilisée est celle de l'analyse (qui ne peut pas être connue dans un contexte temps réel). Une seconde hypothèse est le transport de petites perturbations qui est nécessaire à l'utilisation de l'adjoint. Pour vérifier cette hypothèse, un nouveau calcul de variance de $S$ est effectué à partir des ensembles. On considère maintenant chaque membre comme une perturbation d'un état moyen $\boldsymbol{x} = \overline{\boldsymbol{x}} + \varepsilon$. L'état moyen est calculé à partir de tous les membres, on en déduit la perturbation $\varepsilon$ pour chaque membre. Ce sont ces perturbations que l'on va faire évoluer le long de la trajectoire $\overline{\boldsymbol{x}}$ , $\boldsymbol{x}(t_{verif}) = \mathcal{M}(\overline{\boldsymbol{x}}(t_{obs})) + \boldsymbol{M}\varepsilon(t_{obs})$. Le modèle linéaire tangent peut alors être défini par rapport à la trajectoire de l'état moyen $\overline{\boldsymbol{x}}$ ou de la trajectoire du membre considéré $\boldsymbol{x}$. La figure 6.13 montre le champ de tourbillon potentiel à deux jours pour un membre suivant la trajectoire non linéaire utilisée. L'ordre de grandeur des perturbations est de celui des champs issus des trajectoires non linéaires. Bien que les deux trajectoires soient assez proche, les perturbations obtenues à deux jours ne sont pas identiques, la perturbation (d) à des structures plus fines et plus marquées. Cela correspond aux structures tourbillonnaires crées par la trajectoire du membre perturbé où des non linéarités en amont et en aval du centre de l'anomalie se sont développées.

A partir de l'intégration des perturbations il est ensuite possible de calculer la variance $\sigma^2_S$ de l'ensemble. Les valeurs de variance sont données dans le tableau 6.3. Les valeurs sont plus proches de celle obtenues à partir des trajectoires non linéaires mais sont supérieures. L'hypothèse du transport des petites perturbations semble donc être mise en défaut. Il n'y a pas de croissance linéaire de l'incertitude $\boldsymbol{B} = \boldsymbol{M}\boldsymbol{A}\boldsymbol{M}^T$ pour deux jours d'intégration.

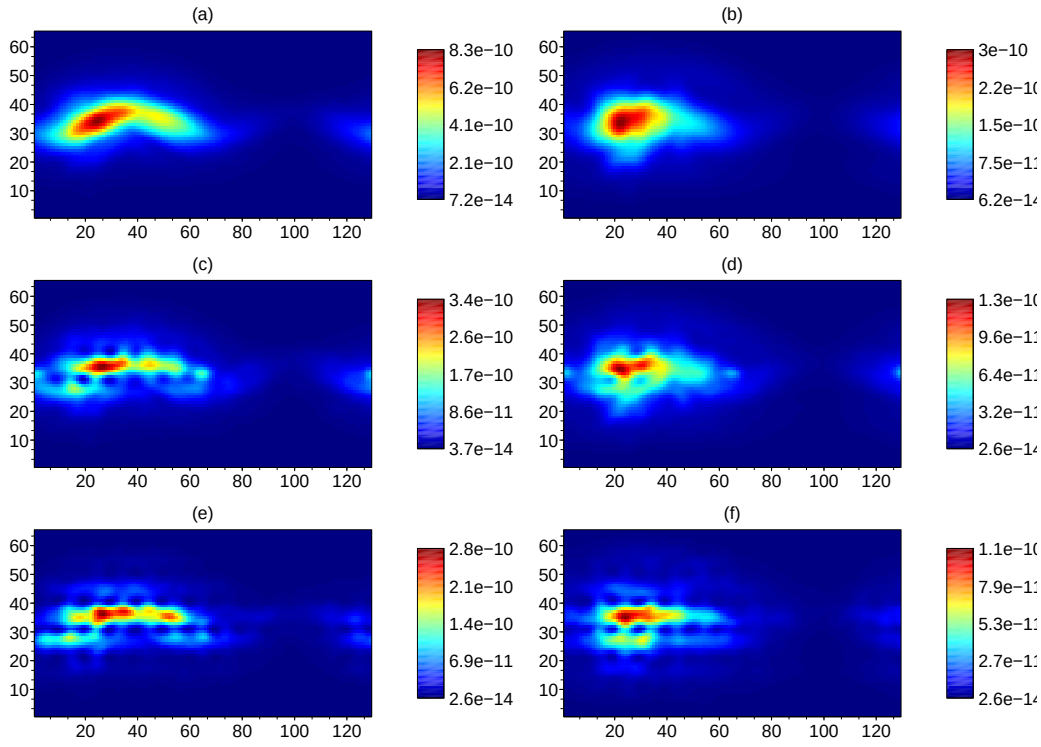



FIGURE 6.12 – En haut (a) et (b) variances d'erreur d'ébauche données au système. Au centre (c) et (d) variances d'erreur d'analyse estimées par l'algorithme de Lanczos. En bas (e) et (f) variances d'erreur d'analyse calculées à partir de l'ensemble. La colonne de gauche donne les valeurs pour la couche supérieure, celle de droite la couche inférieure.

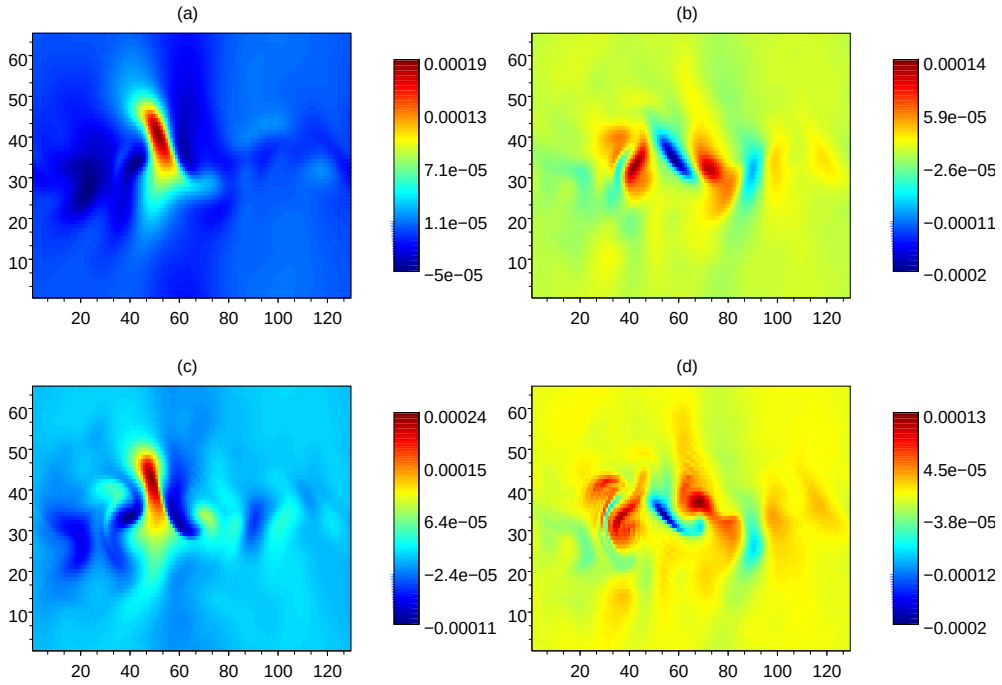



FIGURE 6.13 – Champ de tourbillon potentiel à deux jours sur la couche inférieure. A gauche les champs sont issus des trajectoires non linéaires $\boldsymbol{x}(t_{verif})$ de l'état analysé moyen (a) et de l'analyse du membre (c). Les perturbations $\varepsilon(t_{verif})$ propagées le long de ces trajectoires sont indiquées à droite (b) et (d).

| variance de fonction score $(10^{15}s^{-2})$ | $\sigma^2_{S_c}$ | $\sigma^2_{S_{c+m}}$ | $\delta\sigma^2_{S_m}$ |
|---|---|---|---|
| trajectoire moyenne | 22 | 5 | -17 |
| trajectoire de chaque membre | 128 | 42 | -86 |

TABLE 6.3 – Comparaison de la variance calculée en propageant linéairement les perturbations le long de la trajectoire moyenne ou le long de la trajectoire non linéaire du membre considéré.

### Approche ensembliste du champ de sensibilité

Le paragraphe précédent a montré que l'estimation de la variance par la KFS est surestimée par l'utilisation du modèle adjoint. Il semble donc intéressant de vérifier les champs de KFS ($\boldsymbol{A}\nabla_{\boldsymbol{x}^a}S$) eux-mêmes. Le calcul du champ de KFS est sensé indiquer les zones où une variation d'analyse engendre une forte diminution de valeur de fonction score. En utilisant le même ensemble que pour la section précédente il est possible de faire le lien directement entre une variance d'erreur d'analyse $\varepsilon^a$ et une valeur de fonction score $\delta S$ de manière statistique. Le $\varepsilon^a_i$ est encore défini comme l'écart du membre à la moyenne des analyses, et $\delta S_i$ est l'écart de fonction score par rapport à la moyenne sur tous les membres. La régression s'écrit alors sous la forme

$$\delta S = \Gamma^T \varepsilon^a$$

avec $\Gamma^T = \mathrm{Cov}(\delta S, \varepsilon^a)\,\mathrm{Cov}(\varepsilon^a)^{-1}$ où Cov est la covariance. On peut alors écrire la régression sous la forme, $\Gamma^T = \boldsymbol{u}^T\boldsymbol{A}^{-1}$, soit $\Gamma = \boldsymbol{A}^{-1}\boldsymbol{u}$, avec $\boldsymbol{u}^T = \mathrm{Cov}(\delta S, \varepsilon^a)$ et $\mathrm{Cov}(\varepsilon^a)^{-1} = \boldsymbol{A}^{-1}$. $\boldsymbol{u}$ peut être calculé à partir de l'ensemble. La matrice de covariance est obtenue à partir de l'algorithme de Lanczos. Cette méthode, appelée *Ensemble Sensitivity*, a été proposée par Ancell et Hakim (2006)Ancell and Hakim (2007). Cependant ils ne considère alors que la partie diagonale de $\boldsymbol{A}^{-1}$. L'approche variationnelle permet d'obtenir $\boldsymbol{A}^{-1}$, couplée à un ensemble il est donc possible de calculer l'ensemble des termes de la régression.

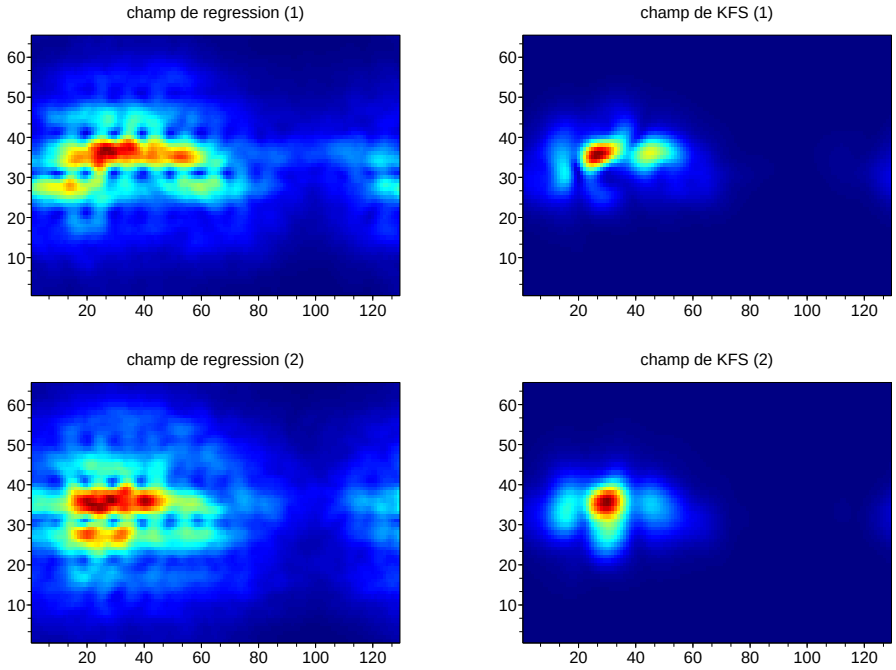



FIGURE 6.14 – Champs de régression $\mathrm{Cov}(\delta S, \varepsilon^a)$ sur les deux niveaux à gauche comparés aux champs de sensibilité obtenus par la KFS. Le réseau d'assimilation est celui de routine.

Les champs de régression considérés se limitent donc à $\mathrm{Cov}(\delta S, \varepsilon^a)$. Ceux-ci sont comparés aux champs de KFS sur les deux niveaux sur la figure 6.14. Les valeurs ne sont pas indiquées, les champs n'étant pas exprimés dans la même unité.

Le champ de régression est plus étalé que celui de KFS. En effet les perturbations $\varepsilon^a$ sont distribuées dans les zones de forte variance de $\boldsymbol{A}$, visibles figure 6.12. Dans ce cadre d'étude où la totalité du domaine est considéré pour calculer la fonction score, il ne se dégage pas de région privilégiée de forte valeur de $S$. On retrouve le résultat précédent que la zone sensible est le jet en lui-même. Pour le calcul des champs de régression les perturbations étant distribuées sur sa longueur il n'apparaît pas de signal particulier. Le champ de KFS moyen est lui cependant moins étalé car il est calculé en utilisant la trajectoire non assimilé pour prendre en compte la dynamique. Des amplitudes plus fortes de perturbations $\varepsilon^b$ sont alors présentes au niveau du maximum de variance de $\boldsymbol{B}$. Ce sont ces perturbations, combinées à la présence du jet, qui contribuent fortement à la valeur du gradient $\nabla_{\boldsymbol{x}^a} S$ et donc au champ de KFS.

La figure 6.15 montre les mêmes champs que la figure 6.14 mais calculés après que l'observation consensus a été déployée. On retrouve la diminution de sensibilité à l'endroit où l'observation a été déployée, aux alentours de (30,35). Un nouveau maximum de sensibilité est mis en évidence plus en aval et ceci pour les deux calculs (surtout visible sur la couche supérieure). Cette zone correspond au chevauchement du champ de variance d'erreur d'ébauche avec le jet de l'écoulement de base. Dans ce cas, la zone de sensibilité ne correspond plus à la zone de plus forte variance d'analyse mais semble dépendre d'avantage de la dynamique. Ceci est d'autant plus visible au niveau des champs de KFS qui en prennent directement compte. Comme pour la cas précédent les champs de corrélation sont bruités par l'étalement des structures de $\varepsilon^a$.

### Discussion et résultats de l'exp III

**En comparant l'estimation de la variance par la KFS à son équivalent calculé à partir d'un ensemble, on obtient que l'estimation de $\sigma_S^2$ par la KFS est fortement surestimée.** L'utilisation de l'algorithme de Lanczos pour obtenir une approximation de $\boldsymbol{A}$ a cependant été validée et n'est pas la

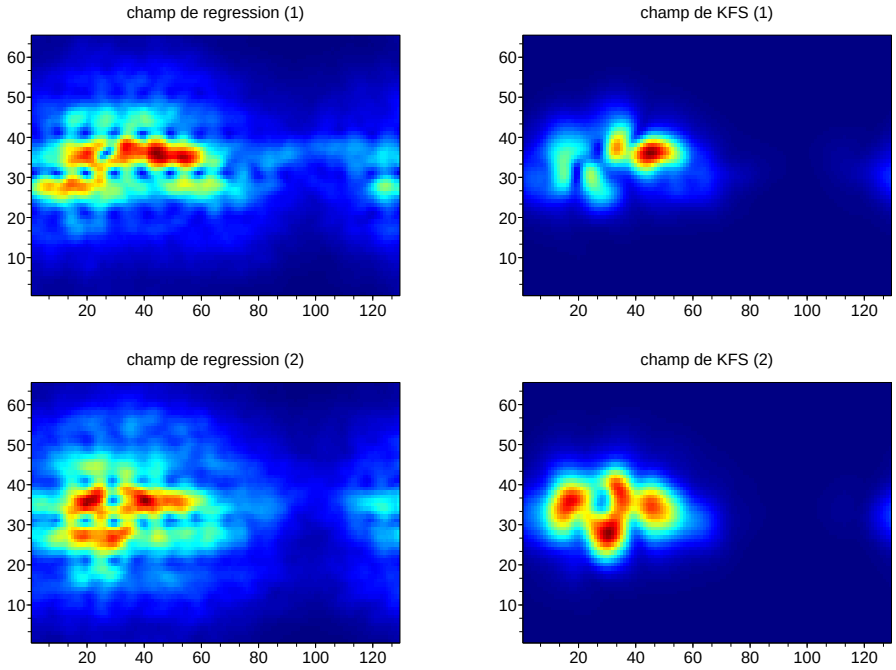



FIGURE 6.15 – Idem champs que 6.14 mais en tenant compte des observations consensus dans l'assimilation.

cause de cet écart. L'hypothèse de transport linéaire par le modèle adjoint et le linéaire tangent semble être la cause de cet écart. En effet les champs transportés sont du même ordre de grandeur que la variable d'état et ne correspondent donc pas à des petites perturbations.

Une approche similaire a été utilisée pour évaluer les champs obtenus par la KFS $\boldsymbol{A}\nabla_{\boldsymbol{x}^a}S$ à partir de champs de régression $\mathrm{Cov}(\delta S, \varepsilon^a)$. **Les champs obtenus par la KFS et la régression sont assez similaire et représentent la même information**. La prise en compte explicite de la dynamique par la KFS semble apporter des résultats plus déterminants. Cette question sera étudiée de plus près avec l'implémentation de la VRF.

## 6.3 Conclusion des expériences

L'objectif de quantifier l'influence de la trajectoire sur la KFS a été traité dans l'expérience I. Les résultats ont permis de montrer le rôle important du système d'assimilation pour limiter la variabilité à la trajectoire. Dans un cadre où la zone de sensibilité est déjà bien observée, toute perturbation initiale associée à cette zone aura peu d'influence sur la prévision. En revanche dans le cas où la zone de sensibilité est peu observée, les positions de déploiement proposées par la KFS varient en fonction du membre. Cela implique qu'une attention toute particulière doit être apportée lors de la prédiction du réseau de routine mais aussi dans le choix des cartes de variance $\sigma^{b2}$.

De plus l'estimation de la réduction de variance, et donc de l'impact d'une observation supplémentaire, est directement corrélée à la valeur de KFS. Cependant ces valeurs sont très variables selon les membres. Ainsi, à une même région géographique peut être associée une forte réduction de variance de score ou une faible selon la trajectoire considérée. Appliquer un critère de seuil pour décider du déploiement d'observations supplémentaires à partir d'un unique champ de KFS semble donc difficile à mettre en œuvre. Utiliser un ensemble de KFS même de dimension faible semble donc un choix plus judicieux.

L'expérience II propose une méthode pour synthétiser l'information issue d'un ensemble de KFS. En considérant un nombre de déploiements possibles limité, utiliser le maximum du champ moyen s'est

montré plus efficace pour réduire la variance de l'ensemble à deux jours, par rapport à un choix aléatoire. Cela permet de valider les apports des techniques d'observation adaptative mais aussi de montrer qu'une méthode simple comme la moyenne permet de tirer partie d'un ensemble. Dans le cadre où un nombre plus important d'observations peuvent être déployées, d'autres critères peuvent être utilisées, comme la position du membre à plus forte valeur ou un choix par rapport au champ de variance de KFS.

Enfin une tentative de validation des estimations de variance de score a été réalisée dans l'expérience III. Les résultats montrent que les valeurs estimées par la KFS sont très supérieures aux valeurs effectives de variance calculées par l'ensemble. Les investigations ont permis de valider l'utilisation de l'algorithme de Lanczos pour estimer $\boldsymbol{A}$, cependant la raison de cette surestimation n'est pas encore entièrement élucidée. Les hypothèses de transport de petites perturbations semblent mises en défaut lorsqu'elles sont appliquées à des trajectoires de deux jours, mais cela n'explique pas toute la différence entre les valeurs obtenues. Une approche d'ordre supérieure, comme utilisée pour étudier l'impact des observations, pourrait être élaborée mais nécessite un développement théorique préalable.

# Chapitre 7

# Implémentation de la VRF dans le modèle de Lorenz

Ce chapitre illustre les deux manières d'implémenter le calcul du champ de réduction de variance issu de la VRF (voir section 4.3.5 pour les développements théoriques). Les développements théoriques de la VRF et les simulations réalisées dans le modèle de Lorenz 96 ont fait l'objet d'un article soumis à la revue Monthly Weather Review.

Les principaux éléments de cette article sont repris ci-dessous. L'article en lui même est présenté en 7.5.

## 7.1 Implémentation de la VRF

Les développements théoriques de la section 4.3.5 montrent que pour effectuer le calcul de la VRF, il est nécessaire de connaître deux quantités, le champ de sensibilité $\boldsymbol{s} = \boldsymbol{A}_c[\nabla_{\boldsymbol{x}^a} S]$, et la variance d'erreur d'analyse en chaque point de grille $\sigma_c^{a2}$. Les paragraphes ci-dessous résument comment calculer ces quantités avec un modèle déterministe et un VDAS, d'une part et avec un ensemble, d'autre part.

Pour un système de PNT déterministe avec assimilation variationnelle, le champ de sensibilité $\boldsymbol{s}$ est calculé exactement comme pour le champ de sensibilité de la KFS : Le gradient de $S$ par rapport à l'analyse est obtenu par le modèle adjoint. Pour effectuer le produit par $\mathbf{A}$ et obtenir la variance d'erreur d'analyse, on utilise l'algorithme de Lanczos lors de la minimisation (dans l'assimilation). Il permet d'avoir une estimation des valeurs propres et vecteurs propres de $\mathbf{A}^{-1}$, et d'en déduire $\mathbf{A}$ et $\mathbf{A}^{1/2}$, voir les annexes B et D.

Dans un système ensembliste le calcul est plus direct. A partir de l'ensemble de prévision, on construit l'ensemble des perturbations $dS_i$ de la fonction score. A partir de l'ensemble d'assimilation, on construit l'ensemble des perturbations $\varepsilon_i^a$ de l'analyse. La variance d'erreur d'analyse est obtenue comme $\boldsymbol{\sigma^{a2}} = \frac{1}{N-1}\sum_{k=1}^{N} \boldsymbol{\varepsilon}_k^a \odot \boldsymbol{\varepsilon}_k^a$, où $\odot$ est le produit d'Hadamard. En reprennant l'équation (4.1), on peut écrire,

$$\begin{aligned} \mathbb{E}(\boldsymbol{\varepsilon}^a\, \delta S) &= \mathbb{E}(\boldsymbol{\varepsilon}^a \boldsymbol{\varepsilon}^{aT} [\nabla_{\boldsymbol{x}^a} S]) \\ &= \boldsymbol{A}_c[\nabla_{\boldsymbol{x}^a} S]. \end{aligned} \tag{7.1}$$

On obtient donc directement $\boldsymbol{s}$ à partir des ensembles $dS_i$ et $\varepsilon_i^a$.

Le calcul de la VRF en lui-même est ensuite réalisé en suivant l'équation 4.16,

$$\delta \boldsymbol{\sigma}_{S_t}^2 = \boldsymbol{s} \odot \boldsymbol{s} \odot (\sigma_t^{o2}\mathbf{1} + \boldsymbol{\sigma^{a2}})^{-1}. \tag{7.2}$$

La valeur de $\sigma_t^o$ peut être choisie arbitrairement.

## 7.2 Cadre expérimental des équations

Le modèle utilisé pour réaliser les simulations est le modèle de Lorenz 96, décrit au chapitre 5.2. La dimension du vecteur d'état est de 60, et le pas de temps de $dt = 0.05$, équivalent à $6h$ en terme de croissance de l'erreur pour un modèle géophysique. La taille de l'ensemble est fixée à 60 membres. Le calcul de la fonction score est réalisé à $t_{verif} = 1$ jours

Dans ce contexte très simplifié, il est possible d'utiliser directement l'équation du *BLUE* (2.4) pour obtenir l'analyse et la matrice de covariance d'erreur d'analyse. Il est choisi d'observer un point de grille sur deux. La matrice de covariance d'erreur d'ébauche $B$ est construite à partir de l'ensemble. Comme la matrice $\mathbf{A}$ est pleinement connue, il n'est pas nécessaire d'utiliser l'algorithme de Lanczos pour l'estimer, et on ne fait donc pas d'assimilation variationnelle.

L'adjoint du modèle de Lorenz est déterminé en reprenant la définition du modèle tangent linéaire $\mathbf{M}$. Il est calculé à partir de différences finies dans chaque direction pour obtenir la matrice $\mathbf{M}$ complète. L'adjoint est alors $\mathbf{M}^T$.

Pour comparer les champs de VRF calculés par la méthode déterministe et la méthode ensembliste, un calcul explicite de la réduction de variance est réalisé : C'est-à-dire que pour chaque point de grille, une assimilation supplémentaire est réalisée avec l'observation supplémentaire et ce pour chaque membre de l'ensemble. Une prévision est ensuite calculée pour chaque analyse afin d'obtenir la variance de la fonction score ${\sigma_{t,i}^S}^2$.

## 7.3 Résultats

Pour l'expérience présentée dans l'article, les trois méthodes de calcul mènent à la même localisation pour le maximum de réduction de variance, et donc à la même stratégie de déploiement. Compte tenu du grand nombre d'hypothèses prises en compte, les trois champs de sensibilité ne sont pas strictement égaux. Des écarts de réduction de variance peuvent apparaître au niveau des pics, là où il n'y a pas d'observation de routine. Cependant les champs sont globalement cohérents entre eux.

Des expériences supplémentaires, non illustrées dans l'article, ont été réalisées en augmentant le *optimization-time*. Le modèle s'écarte alors de son régime linéaire et les différences entre les trois méthodes sont beaucoup plus visibles. La VRF calculée à partir de l'adjoint tend alors à minimiser la réduction de variance effective, ce qui est moins vrai pour le calcul via l'ensemble.

## 7.4 Discussion sur l'implémentation de la VRF

Les aspects théoriques de la VRF ont été développés au chapitre 4.3.5. Ce chapitre décrit et illustre les deux manières d'implémenter la VRF, à partir d'un système déterministe avec l'utilisation du modèle adjoint et un système d'assimilation variationnel, et à partir d'un ensemble. Les simulations réalisées dans le cadre idéalisé du modèle de Lorenz 96 prouvent que l'algorithme fonctionne, lorsque l'on se place dans un contexte suffisamment linéaire et avec un ensemble de grand taille (pour les simulations nous avons $N = N_e$).

Le calcul de la VRF à partir d'un ensemble est très simple à mettre en place. Outre la difficulté à écrire un développement au second ordre, la méthode souffre d'un défaut de conception. En effet, tel que le calcul est réalisé en prenant $\mathbb{E}(\boldsymbol{\varepsilon}^a\, \delta S)$ l'ensemble du vecteur d'état est pris en compte pour la sensibilité. Par exemple, on se place dans une situation où deux dépressions similaires se trouvent une sur l'Atlantique et l'autre sur le Pacifique, produisant deux régions de forte incertitude, mais déconnectées entre elles. Avec un calcul de la fonction score sur l'Europe, la VRF va indiquer que rajouter des observations au niveau

de la dépression sur le Pacifique va réduire l'incertitude sur la dépression arrivant depuis l'Atlantique. Ce qui n'est pas très cohérent.

Une localisation semble donc nécessaire pour limiter l'impact des corrélations à très longue distance qui pourraient apparaître dans un ensemble de petite taille. Cette localisation n'est cependant pas triviale car la fonction score qui est calculée est définie pour la totalité du vecteur d'état à $t_{verif}$. Il n'est pas possible d'identifier la contribution de chaque point de grille de $t_{obs}$ sur la trajectoire à $t_{verif}$. Les méthodes de localisation récentes proposent d'utiliser la vitesse de groupe (Kalnay et al., 2012) ou une advection issue du transport Lagrangien (Desroziers et al., 2015) pour transporter les fonctions de localisation dans le temps. Ces méthodes sont développées pour des transports sur des fenêtres d'assimilation (environ $6h$) qui sont plus courtes que les *optimization-times* couramment utilisés en observation adaptative. Une adaptation pourrait donc être nécessaire.

Une mise en œuvre de la VRF dans un modèle opérationnel de PNT n'a pas pu être réalisée dans le cadre de cette thèse. Néanmoins, le chapitre suivant présente comment la VRF peut être implémentée dans l'environnement quasi-opérationnel OOPS.

## 7.5 Forecast variance reduction field, a new tool for adaptive observation : ensemble versus adjoint computation

# Forecast variance reduction field a new tool for adaptive observation: ensemble versus adjoint computation

N. Oger, O. Pannekoucke, A. Doerenbecher and P. Arbogast
Météo-France CNRM/GMAP, 42, Av. Coriolis, 31057 Toulouse, France

April 27, 2015

ABSTRACT

Adaptive observation seeks to build an optimal frame for an observation network that accounts for the dynamical aspect of the atmosphere and for the data assimilation system. This note describes a new statistical tool for observation targeting called Variance Reduction Field (VRF). This technique computes the expected variance reduction of a forecast score function. This forecast score variance reduction is linked to a test probe assimilated for each model grid point. The VRF allows to discriminate regions of interest for targeted observations. The VRF is composed of two aspects: a sensitivity information and the variance of the analysis error distribution, both of which must be predicted. To compute these two fields, two approaches are proposed. The first one is based on a deterministic simulation. The second approach consists of using an ensemble data assimilation and prediction system.

## 1 Introduction

In meteorology, the goal of data assimilation is to seek the most likely state of the atmosphere at a given time. For this purpose observations $\boldsymbol{y}^o$ and a priori information, called the background state $\boldsymbol{x}^b$ are used. It is generally a previous forecast given by the Numerical Weather Prediction (NWP) system. A data assimilation system also accounts for the uncertainty associated to both sources of information, i.e. $\boldsymbol{y}^o$ and $\boldsymbol{x}^b$. It is usually assumed that errors on the observations and the background state are unbiased and Gaussian. Thus, the best estimate of the most likely atmosphere state (also called analysis $\boldsymbol{x}^a$) can be expressed as the solution of the cost function minimization $J$,

$$J(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{x}^b\|^2_{\boldsymbol{B}^{-1}} + \frac{1}{2}\|\mathcal{H}(\boldsymbol{x})-\boldsymbol{y}^o\|^2_{\boldsymbol{R}^{-1}}, \quad (1)$$

$$\boldsymbol{x}^a = \operatorname{Arg\,min} J, \quad (2)$$

where $\boldsymbol{B}$ and $\boldsymbol{R}$ are respectively the background and observation error covariance matrices, and $\mathcal{H}$ the observation operator which maps the model state into observations. When $\mathcal{H}$ is linear, and denoted $\boldsymbol{H}$, the problem is quadratic and the unique solution is given by the Best Linear Unbiased Estimator (BLUE) $\boldsymbol{x}^a = \boldsymbol{x}^b + \boldsymbol{K}\boldsymbol{d}$, where $\boldsymbol{K} = (\boldsymbol{B}^{-1} + \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{R}^{-1} = \boldsymbol{A}\boldsymbol{H}^T\boldsymbol{R}^{-1}$ is the Kalman gain matrix, and $\boldsymbol{d} = \boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x}^b$ the innovation vector. $\boldsymbol{A}$ is the analysis error covariance matrix and the inverse of the Hessian of the cost function $J''$.

The Data Assimilation Systems (DAS), which consists of minimizing of a cost function to compute the analysis, are called variational (VDAS). By using the observation operator $\mathcal{H}$, VDAS can deal with time distributed observations and non-linear system, using an incremental formulation (such as 4D-var, Courtier *et al.* 1994). Another type of DAS, operationally used in NWP, is based on ensemble forecasts, such as the Ensemble Kalman Filter (EnKF) (Houtekamer and Mitchell 1998) or ensemble square root filters (Tippett *et al.* 2003, Hunt *et al.* 2007).

As observations are sparsely distributed in time and space, the observation network is not always effective in describing the state of the atmosphere. The Adaptive Observation (AO) strives to locate sensitive regions where more observations must be deployed (called targeted observation), at time $t_{obs}$, to improve the weather forecast, at time $t_{verif}$. Several methods of AO have been developed in National Meteorological Services (NMSs). As for DAS, two methods are available. The first is based on a deterministic approach, using the adjoint model and a VDAS, like the Hessian singular vectors (Leutbecher 2003) or the Kalman Filter Sensitivity (KFS, Bergot and Doerenbecher 2002). The other is based on an ensemble, like the Ensemble Transform Kalman Filter (ETKF, Bishop et al. 2001) or the Ensemble Sensitivity (Ancell and Hakim 2006).

In this study a new method to predict the sensitive area, named VRF, is proposed. The method is derived from the Kalman Filter Sensitivity (KFS, Doerenbecher and Bergot 2001) and can be used either with a deterministic approach or an ensemble approach. It simulates the assimilation of additional observation (or probe) at each model grid points.

The note is organized as follows. Section 2 introduces the adaptive observation, its notations and presents the development of the Variance Reduction Field (VRF). Section 3 proposes to compute the VRF, using a deterministic and an ensemble approach. The last section 5 provides a summary and conclusion.

## 2 Adaptive observation

### 2.1 Forecast aspect and notations

In order to assess the impact of the targeted observations, a score function should be chosen. The forecast score function $S$ is defined at the verification time $t_{verif}$. This function needs no particular properties apart from being related to the observed meteorological event and being quadratic, for example $S$ can be the enstrophy . It is usually computed in a specific area called the verification region.

The theoretical development is based on the observation sensitivity (Baker and Daley 2000), which includes the AO and the observation impact. In observation impact studies (Cardinali 2008, Trémolet 2008) the computation is done in a diagnostic way. The function $S$ is defined as the distance between the forecast from $t_{obs}$ and the verifying analysis at $t_{verif}$, which is known. However, for the prognostic computation of AO, the analysis at $t_{verif}$ is not yet available. Therefore, the variance of $S$ ($\sigma_S^2$), which measures the uncertainty of the forecast, is used here. The AO methods link this variance $\sigma_S^2$, or its reduction, to the location of the deployment of additional observations.

In order to determine the location of additional observation deployment, a variation $\delta S$ at $t_{verif}$ is expressed as a function of the analysis $\boldsymbol{x}^a$ at $t_{obs}$, using a first order Taylor expansion.

$$\begin{aligned} \delta S &= S(\boldsymbol{x}^f + \boldsymbol{\varepsilon}^f) - S(\boldsymbol{x}^f), \\ &= [\nabla_{\boldsymbol{x^f}} S]^T \boldsymbol{\varepsilon}^f, \\ &= [\nabla_{\boldsymbol{x^f}} S]^T \boldsymbol{M}_{t_{obs} \to t_{verif}} \boldsymbol{\varepsilon}^a, \\ &= [\nabla_{\boldsymbol{x^a}} S]^T \boldsymbol{\varepsilon}^a, \end{aligned} \tag{3}$$

where $[\nabla_{\boldsymbol{x^a}} S] = \boldsymbol{M}^T_{t_{obs} \to t_{verif}} [\nabla_{\boldsymbol{x^f}} S]$ is the gradient field at $t_{obs}$ and $\boldsymbol{M}^T_{t_{obs} \to t_{verif}}$ is the adjoint model. It propagates the perturbation backward, from $t_{verif}$ to $t_{obs}$.

As the perturbation $\boldsymbol{\varepsilon}^a$ is distributed according to the analysis error probability density function, the analysis covariance matrix can be written as $\boldsymbol{A} = \mathbb{E}(\boldsymbol{\varepsilon}^a \boldsymbol{\varepsilon}^{aT})$. The variance of $S$ can be written as follows :

$$\sigma_S^2 = \mathbb{E}(\delta S \, \delta S^T) = [\nabla_{\boldsymbol{x^a}} S]^T \boldsymbol{A} [\nabla_{\boldsymbol{x^a}} S]. \tag{4}$$

The forecast uncertainty at $t_{verif}$ is directly related to the distribution of the analysis error at $t_{obs}$, propagated by the model $\boldsymbol{M}$. The analysis error covariance matrix $\boldsymbol{A}$ includes the information of the a priori uncertainty on the background state by $\boldsymbol{B}$, and the information issued form the observations by $\mathcal{H}$ and $\boldsymbol{R}$. The AO method aims to tune the observation operator $\mathcal{H}$ in order to reduce $\sigma_S^2$.

### 2.2 Forecast score variance reduction equations

The computation of Eq. (4) could be done with every possible observation network. This would allow to compare the deployment strategies. In order to keep the computational cost affordable, the observation network is usually divided into a routine (or conventional) network, called $c$, and a targeted network, called $t$. The respective analysis-error covariance matrices are $\boldsymbol{A}_c$ and $\boldsymbol{A}_{c+t}$, depending on the assimilation or non assimilation of the targeted observations. The goal of the method is to assess when the targeted network maximizes the reduction of the variance $\sigma_S^2$. $\boldsymbol{A}_{c+t}$ can be expressed as a function of $\boldsymbol{A}_c$,

$$\boldsymbol{A}_c = (\boldsymbol{I} - \boldsymbol{K}_c \boldsymbol{H}_c)\boldsymbol{B} \tag{5}$$

$$\boldsymbol{A}_{c+t} = (\boldsymbol{I} - \boldsymbol{K}_{c+t} \boldsymbol{H}_t)\boldsymbol{A}_c \tag{6}$$

where, using the Sherman-Morrison-Woodbury formula, $\boldsymbol{K}_c = (\boldsymbol{B}^{-1} + \boldsymbol{H}_c^T \boldsymbol{R}_c^{-1} \boldsymbol{H}_c)^{-1} \boldsymbol{H}_c^T \boldsymbol{R}_c^{-1}$ and $\boldsymbol{K}_{c+t} = (\boldsymbol{A}_c^{-1} + \boldsymbol{H}_t^T \boldsymbol{R}_t^{-1} \boldsymbol{H}_t)^{-1} \boldsymbol{H}_t^T \boldsymbol{R}_t^{-1}$. This leads to the first order reduction of forecast score variance due to the additional observations:

$$\begin{aligned} \delta\sigma_{S_t}^2 &= \sigma_{S_c}^2 - \sigma_{S_{c+t}}^2 \\ &= (\boldsymbol{H}_t \boldsymbol{A}_c [\nabla_{\boldsymbol{x^a}} S])^T \left(\boldsymbol{R}_t + \boldsymbol{H}_t \boldsymbol{A}_c \boldsymbol{H}_{\boldsymbol{t}}^{\boldsymbol{T}}\right)^{-1} (\boldsymbol{H}_t \boldsymbol{A}_c [\nabla_{\boldsymbol{x^a}} S]) \, . \end{aligned} \tag{7}$$

This expression ensures that when the optimal solution of the data assimilation is reached, additional observations reduce the variance of the forecast score. The notation $\boldsymbol{s} = \boldsymbol{A}_c [\nabla_{\boldsymbol{x^a}} S]$, for sensitivity field, is used hereafter.

Eq. (7) allows to compute the forecast score variance reduction for any targeted observations. Computing the optimal additional network will lead to a minimization problem on the observation operator $\boldsymbol{H}_t$, which is rather difficult to undertake. In order to make the computation cost affordable, the additional observation is supposed to be expressed by a model variable at a given grid point.

Now, each model grid point can supposedly be a deployment location. This simulates the sampling from a test probe. $\boldsymbol{H}_{t,i}$ is built as the restriction of the whole field to the grid point $i$.

This simplifies even more the Eq. (7), the sensitivity field contribution becomes $\boldsymbol{H}_{t,i} \boldsymbol{A}_c [\nabla_{\boldsymbol{x}^a} S] = \boldsymbol{s}_i$, where $\boldsymbol{s}_i$ is the $i$th component of $\boldsymbol{s}$. The matrix $\boldsymbol{R}_t$ is a scalar noted $\sigma_t^{o\,2}$, as well as $\boldsymbol{H}_{t,i} \boldsymbol{A}_c \boldsymbol{H}_{t,i}^T = \sigma^{a\,2}_i$, with $\sigma^{a\,2}_i$ the analysis error variance at the grid point $i$. The inverse $\left(\boldsymbol{R}_t + \boldsymbol{H}_t \boldsymbol{A}_c \boldsymbol{H}_{\boldsymbol{t}}^{\boldsymbol{T}}\right)^{-1}$ is then straightforward to compute. Eq. (7) can be rewritten:

$$(\delta\sigma_{S_t}^2)_i = \frac{s_i^2}{\sigma_t^{o\,2} + \sigma^{a\,2}_i}. \tag{8}$$

Eq. (8) allows to compute the variance reduction associated to a test probe at the grid point $i$. To compute $(\delta\sigma_{S_t}^2)_i$, the sensitivity field $\boldsymbol{s}_i$ and the analysis error variance $\sigma^{a\,2}_i$ are both required. As all grid points

are independent, the VRF can be expressed using the Hadamard product $\odot$,

$$\delta\boldsymbol{\sigma}^2_{S_t} = \boldsymbol{s} \odot \boldsymbol{s} \odot (\sigma_t^{o2}\mathbf{1} + \boldsymbol{\sigma}^{\boldsymbol{a2}})^{-1}, \tag{9}$$

where $\mathbf{1}_i = 1$, $\forall i$. The VRF provides an evaluation of the impact of any additional observation location. The field is expressed in the model space and thus is easily handled. Since $\sigma_t^{o2}$ can not be associated to any effective observation, the term $\sigma_t^{o2}\mathbf{1}$ can be neglected. This field is uniform and we are interested here in the spatial distribution of $\delta\boldsymbol{\sigma}^2_{S_t}$. The shape of the field, rather than its values, is to be considered.

Two fields have to be computed in order to get the VRF, the sensitivity field $\boldsymbol{s} = \boldsymbol{A}_c[\nabla_{\boldsymbol{x}^a} S]$ and the analysis error variance map (i.e. the diagonal of $\boldsymbol{A}_c$). The method to compute them is now explained.

# 3 Two methods for the sensitivity field computation

The sensitivity field and the variance map can be obtained using a deterministic or ensemble approach. The computation of each field could be done separately or by combining both methods. It is assumed here that the deterministic approach only use a single model forecast and that the adjoint model is available to perform the VDAS. The ensemble approach only needs a DAS and an ensemble of forecasts.

## 3.1 Using a deterministic NWP system

Using the adjoint model $\boldsymbol{M}^T$, as showed in Eq. (3), $\nabla_{\boldsymbol{x}^a} S$ can be computed backward in time from $t_{verif}$.

$$\nabla_{\boldsymbol{x}^a} S = \boldsymbol{M}^T \nabla_{\boldsymbol{x}^f} S \tag{10}$$

This needs two time integrations, one forward with $\boldsymbol{M}$ and one backward in time. The use of an adjoint model also limits the sensitivity field which must be valid for the reference trajectory only. Oger *et al.* (2011) showed that the reference trajectory can have a strong influence on the sensitivity field.

The sensitivity field $\boldsymbol{s}$ needs a product by $\boldsymbol{A}_c$ in order to be computed. As mentioned before, in a VDAS the analysis is obtained as the minimum of the cost function $J$ (Eq. (1)), and the covariance matrix $\boldsymbol{A}_c$ is not directly available. An approximation of the analysis error covariance matrix is obtained following Fisher and Courtier (1995), as described now. The classical preconditioning (change of variable) $\omega = \boldsymbol{B}^{-1/2}\boldsymbol{x}$ is firstly applied to $J(\boldsymbol{x})$. The Hessian of the pre-conditioned cost function $\tilde{J}$ is $\tilde{\boldsymbol{J}}'' = \boldsymbol{I} + \boldsymbol{B}^{T/2}\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H}\boldsymbol{B}^{1/2}$, with $\boldsymbol{I}$ the identity matrix. The use of a combined conjugate gradient (CG) and Lanczos algorithm during the minimization of the cost function allows to get an approximation of the pre-conditioned Hessian, $\widetilde{\boldsymbol{A}}^{-1}$ (Fisher and Courtier 1995). It is approximated by a tri-diagonal matrix expressed in the reduced rank base spanned by the descent vectors. The algorithm is shown in the appendix.

As the dimension of the tri-diagonal is small, it is cheap to diagonalize it. The Hessian can then be approximated by

$$\widetilde{\boldsymbol{A}}^{-1} = \boldsymbol{I} + \sum_{r=1}^{D} (\lambda_r - 1)\boldsymbol{u}_r\boldsymbol{u}_r^T, \tag{11}$$

where $D$ is the dimension of the reduced base and $(\lambda_r, \boldsymbol{u}_r)$ the singular value and singular vector couples expressed in the pre-conditioned space. The identity matrix is added here to have a full ranked approximation. Trémolet (2008) proposed the same approximation without the identity matrix in order to computed the VDAS adjoint. The inversion of Eq. (11) is direct and gives:

$$\widetilde{\boldsymbol{A}} = \boldsymbol{I} + \sum_{r=1}^{D} (1/\lambda_r - 1)\boldsymbol{u}_r\boldsymbol{u}_r^T \tag{12}$$

By multiplying by $\boldsymbol{B}^{1/2}$ in the left hand-side and $\boldsymbol{B}^{T/2}$ in the right hand-side of Eq. (12), we get the approximation of the analysis error covariance matrix, to account for the change of variable,

$$\boldsymbol{A} = \boldsymbol{B} + \sum_{r=1}^{D} (\frac{1}{\lambda_r} - 1)(\boldsymbol{B}^{1/2}\boldsymbol{u}_r)(\boldsymbol{B}^{1/2}\boldsymbol{u}_r)^T. \tag{13}$$

The Lanczos algorithm allows to compute this approximation with little additional computational cost compared to the total computational cost of the minimization. Once the assimilation is done and the gradient $\nabla_{\boldsymbol{x}^a} S$ computed, the sensitivity is available.

Yet, the analysis error variance field is not directly available from the Lanczos algorithm. In order to compute the diagonal of $\boldsymbol{A}_c$, a stochastic method is used. An ensemble of realizations $\hat{\boldsymbol{\varepsilon}}^a_k = \boldsymbol{A}^{1/2}\boldsymbol{\eta}_k$ are built, with $\boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0}, \boldsymbol{I})$. $\mathcal{N}(\boldsymbol{M}, \boldsymbol{V})$ acounting for a Gaussian distribution with mean $\boldsymbol{M}$ et covariance matrix $\boldsymbol{V}$. The covariance matrix of this ensemble is $\mathbb{E}(\hat{\boldsymbol{\varepsilon}}^a\hat{\boldsymbol{\varepsilon}}^{aT}) = \boldsymbol{A}^{1/2}\boldsymbol{A}^{T/2} = \boldsymbol{A}$. The square root matrix can also be obtained by the Lanczos algorithm,

$$\boldsymbol{A}^{1/2} = \boldsymbol{B}^{1/2} + \sum_{k=1}^{r} (\frac{1}{\sqrt{\lambda_k}} - 1)\boldsymbol{B}^{1/2}\boldsymbol{u}_k\boldsymbol{u}_k^T. \tag{14}$$

Given enough realizations $\hat{\boldsymbol{\varepsilon}}^a$ and by computing only the variance of the ensemble, this method offers a good approximation of the analysis error variance field. The VRF can then be computed using Eq. (9).

The use of the combined CG-Lanczos algorithm is the main tool for this computation. Its implementation can be done in any VDAS. For VDAS, the adjoint $\boldsymbol{M}^T$ is needed, leading in pratice to huge developments and maintenance in the numerical prediction weather system. These cost could be reduced thanks to the VRF since it can be performed without any adjoint model by the use of an ensemble. This is now described.

### 3.2 Using an ensemble NWP system

Some reformulation is needed in order to write the VRF with an ensemble approach. As for the previous subsection the first goal is to get the sensitivity field $\boldsymbol{s} = \boldsymbol{A}_c \nabla_{\boldsymbol{x}^a} S$. From Eq. (3), $\delta S = \boldsymbol{\varepsilon}^{aT}[\nabla_{\boldsymbol{x}^a} S]$, by multiplying by $\boldsymbol{\varepsilon}^a$ and taking the expectation, we get:

$$\begin{aligned} \mathbb{E}(\boldsymbol{\varepsilon}^a \, \delta S) &= \mathbb{E}(\boldsymbol{\varepsilon}^a \boldsymbol{\varepsilon}^{aT}[\nabla_{\boldsymbol{x}^a} S]) \\ &= \boldsymbol{A}_c[\nabla_{\boldsymbol{x}^a} S]. \end{aligned} \tag{15}$$

So, $\boldsymbol{s}$ is computed from the ensemble as $\boldsymbol{s} = \frac{1}{N}\sum_{k=1}^{N} \boldsymbol{\varepsilon}_k^a \delta S_k$, where $N$ is the ensemble size. Assuming that the ensemble is unbiased, $\delta S_k$ is defined as $\delta S_k = S(\boldsymbol{x}_k^f) - S(\overline{\boldsymbol{x}^f})$ and $\boldsymbol{x}_k^a = \overline{\boldsymbol{x}^a} + \boldsymbol{\varepsilon}_k^a$, the overline accounting for the ensemble average. The computation of the sensitivity field $\boldsymbol{s}$ is thus straightforward with an ensemble. It only needs a serie of couples of $\widetilde{\delta S}$ and $\boldsymbol{\varepsilon}^a$.

The time propagation is here implicit in the correlation. The analysis error variance map is computed from the ensemble as the diagonal of the matrix $\boldsymbol{A}_c = 1/N \sum_{k=1}^{N} \boldsymbol{\varepsilon}_k^a \boldsymbol{\varepsilon}_k^{aT}$,

$$\boldsymbol{\sigma}^{\boldsymbol{a2}} = \frac{1}{N-1} \sum_{k=1}^{N} \boldsymbol{\varepsilon}_k^a \odot \boldsymbol{\varepsilon}_k^a. \tag{16}$$

It can be seen that, if an ensemble data assimilation and prediction system are available, the VRF method is easy to implement. This method provides an advanced adaptive observation tool which can be used to tune the observation network or to forecast deployments for measurement campaigns. The next section shows some results of VRF fields for the two methods in an simplified framework.

# 4 Application of the VRF: experimental setup

This section illustrates the theoretical developments implemented on a toy model used by Lorenz and Emmanuel (1998), with a BLUE data assimilation system. An ensemble of trajectories is created to compute the VRF using both methods. The toy model is first introduced, then the data assimilation system and finally some results are shown.

### 4.1 The Lorenz model

The toy model used here is the one used by Lorenz and Emmanuel (1998). it has been firstly introduced in 1996 an is commonly called 'Lorenz 96'. This model is widely used as a basic simulation of mid-latitude zonal circle of the global atmosphere. The equation is nonlinear and uses a forcing term leading to various dynamics depending on the values chosen. The time evolution follows,

$$\frac{dX_i}{dt} = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F, \tag{17}$$

for $i = 1..n_{lorenz}$, with cyclic bondaries. The value of the forcing term $F = 4$ is taken to build chaotic time variations (see Lorenz and Emmanuel 1998). The members of the ensemble are built as $\boldsymbol{x}_i^b = \boldsymbol{x}_{ref}^b + \varepsilon_i^b$ with $\varepsilon_i^b \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{B})$.

In these simulations, the state vector dimension $n_{lorenz}$ and the ensemble size $N$ are taken equal to 60. The time propagation scheme is a 4-order Runge-Kutta with a time step of $dt = 0.05$ time unit, this is equivalent to a $6h$ time step for a geophysical model in terms of error growth.

### 4.2 BLUE assimilation system

The computation of the VRF as been introduced, in section 3, using NWP data assimilation systems (4Dvar and ensemble square root filter), but it can be applied to any data assimilation system. Considering the few dimensions of the model, the BLUE can be used without expensive computational costs. The analysis state is computed following $\boldsymbol{x}_c^a = \boldsymbol{x}^b + \boldsymbol{K}_c \boldsymbol{d}_c$. The observations are generated from the reference state, $\boldsymbol{y}^o = \boldsymbol{x}_{ref}^b + \varepsilon^o$ with $\varepsilon^o \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$. One grid point out of two is observed in the first half of the domain, one out of five for the second half. The quality of each observation is supposed the same and they are uncorreleted (*ie.* $\boldsymbol{R} = \sigma_o \boldsymbol{I}$).

The adjoint method uses directly the $\boldsymbol{B}$ matrix from the computation of $\boldsymbol{A}$. On the contrary the ensemble method does not use the $\boldsymbol{B}$ matrix, the correlation are implicit in the computation of (15). Thus in order to avoid a biais between the two methods, the background error covariance matrix $\boldsymbol{B}$ used for the BLUE is computed from the ensemble, $\boldsymbol{B} = E(\varepsilon^b \varepsilon^{bT})$.

Furthermore to fully account for the observation error correlation in the assimilation proccess of an ensemble, the observations have to be perturbed according to $\boldsymbol{R}$ (Burgers *et al.* 1998). The values assimilated as the observation for each membres follow $\tilde{\boldsymbol{y}^o} = \boldsymbol{y}^o + \varepsilon_i^o$, $\varepsilon_i^o$ with $\varepsilon_i^o \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$.

### 4.3 Explicit computation of the variance reduction

In order to assess the accuracy of the two methods proposed for the VRF, and since the computational costs are cheap, the explicit variance reduction due to the additional observations is computed. To do so, for each initial state vector and for each grid point, the analysis state using an additional observation at that grid point is computed. It is obtained using the BLUE relation from the previous assimilation state $\boldsymbol{x}_c^a$,

$$\boldsymbol{x}_{t,i}^a = \boldsymbol{x}_c^a + \boldsymbol{K}_{t,i}(\boldsymbol{y}^o - \boldsymbol{H}\boldsymbol{x}_c^a), \tag{18}$$

with $\boldsymbol{K}_{t,i} = (\boldsymbol{A}_c^{-1} + \boldsymbol{H}_{t,i}^T \boldsymbol{R}_{t,i}^{-1} \boldsymbol{H}_{t,i})^{-1} \boldsymbol{H}_{t,i}^T \boldsymbol{R}_{t,i}^{-1}$, the subscript $t$ meaning targeted and $i$ the grid point of the additional targeted observation.

All analysis states ($\boldsymbol{x}_c^a$ and $\boldsymbol{x}_{t,i}^a$ $\forall i$) are evolved into $t_{verif}$, making a forecast ensemble for each observation

network $\boldsymbol{H}_{t,i}$. The variance of the score function $S$ is the computed from these $N_{lorenz}+1$ ensembles, $\sigma_S^2 = \mathrm{E}((S-\overline{S})^2)$. Using $(\delta\sigma^2_{S_t})_i = \sigma^2_{S_c} - \sigma^2_{S_{c+i}}$, it is thus computed explicitly.

The figure 1 shows a summary of the three ways to compute the VRF, using the adjoint, an ensemble or the explicit assimilation of the additional observations. The results of the implementation of these methods are now discussed.

### 4.4 Preliminary results

The framework has been introduced in the previous sections and the computation is summarised in Fig. 1. Before showing the values of $VRF$, let's take a look at some by-products of the computation. The figure 2 shows the forecast error variance before the assimilation proccess compared to the analysis error variance. Two analysis error variances are ploted, $\boldsymbol{A}_{BLUE}$ is computed from the BLUE equation (eq. (5)) and $\boldsymbol{A}_{ens}$ is computed from the ensemble of $\boldsymbol{x}^a$. Both analysis error variances are not equal due to the lack of consistancy between the $\boldsymbol{R}$ matrix used for the BLUE, which is built diagonal, and the pratical $\boldsymbol{R}$ used in the ensemble data assimilation system, where different observations are correlated due to sampling noise.

It can be also noticed that the reduction of variance between the forecast error ($\boldsymbol{B}$) and the analysis error ($\boldsymbol{A}$) has spred at all grid points, not only for the being observed (illustradted by the $+$). The propagation of the error variance is due to the correlation between grid points in $\boldsymbol{B}$ (not shown). This is a limitation with the use of a toy model like Lorenz 95 where the dynamic is not able to build structures and thus there is spurious correlation between grid points. Furthermore, from Eq. 9, the interpretation of the non linear equation between the variance, the gradient and the $VRF$ is not straightforward.

### 4.5 VRF Results

This part illustrates the three methods of computing the $VRF$. The first, shown in Fig. 3, is the explicit computation. It is chosen as a reference since no hypothetis has been used. The additional targeted observation proposed by the $VRF_{expl}$ is located at grid point 31.

The Fig. 4 shows the $VRF$ for each computation (ie. explicit, using the ensemble and using the gradient). All three fields are similar. They lead to the same optimal deployment strategy at grid point 31. The differences between $VRF_{expl}$ and the two other methods can be explained by the hypothesis introduced for the computation. It has been shown in part 4.4.4 that the ensemble does not fully represent the variability and correlation between the observation, leading to variation in analysis error variances and covariances.

On the over side, the gradient method uses the $\boldsymbol{A}$ matrix from the BLUE (eq. (5)), which does not provide the same variances as $\boldsymbol{A}_{ens}$ (see figure 2). Furthermore the backward time integration for the gradient needs the adjoint model, under the hypothesis of linear transport of small perturbations. The lorenz 95 model is far from linear. However on a 1 day forecast, the gradient method still proves as effective as the others, which is not the case if we make a longer time integration (not shown here).

## 5 Conclusion

The theoretical aspects of the variance reduction field (VRF) have been developed in this note. This adaptive observation tool aims to build objective fields of sensitivity to the observations. The only assumption needed, on the form or the location of the targeted observations (or probe) simulated here, is that it is expressed in the model grid. Two approaches are proposed for the computation, the first one using a deterministic system, and the second one using a forecast ensemble.

The deterministic method uses the adjoint model to propagate in time the information. The analysis error covariance matrix $\boldsymbol{A}$ is got as a by-product of the minimization of the variational data assimilation system. The Lanczos algorithm, combined with a conjugate gradient minimizer, allows to compute an approximation of the Hessian of the functional, here $\boldsymbol{A}^{-1}$, in a reduced rank base, and then $\boldsymbol{A}$ is obtained. The sensitivity field and the error variance map are computed using this approximated $\boldsymbol{A}$.

The implementation of the VRF using an ensemble is easier. An ensemble of analyzed states and their respective forecast is needed. The sensitivity field, including the time propagation, are computed by a linear regression. The error variance map is diagnosed from the forecast ensemble. Thus the VRF needs no particular development when applied to an operational ensemble forecast system.

The simple experiments of the last section show that both methods lead to the same targeting strategy in an optimal case, i.e. when the ensemble is large enough to describe the full variance and when the linear transport hypothetis is satisfied. The values of the VRF are also consistent with the explicit computation of the variance reduction du to the additional observations.

This note only presents the computation of a VRF for a single additional observation. A series of additional observations can also be dealt with by assimilating the previous one in a sequential way. The VDAS needs to be re-run with the previous $\boldsymbol{A}^{-1}$ approximation as preconditioner. As most ensemble assimilation systems are sequential, this only needs to assimilate the additional observation already targeted.

The theoretical aspects of the VRF are shown here and two ways of implementation are proposed. Further investigations have to be done in order to assess the efficiency of each approach, and compare them to the existing methods in adaptive observation.

*Acknowledgments*
The authors acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant Geo-FLUIDS (ANR- 09-SYSC-005-01) "Analyse et simulation d'écoulements fluides à partir de séquences d'images : application à l'étude d'écoulements géophysiques". They also thank Marie Boisserie for her comments on an earlier version of the paper.

# Combined Conjugate Gradient and Lanczos algorithm

This appendix shows the combined CG-Lanczos algorithm. It allows to compute an approximation of the Hessian of the functional $J$ to minimize during the computation. Initialization: $\delta\boldsymbol{x}_0 = \boldsymbol{0}$, $\boldsymbol{d}_0 = \boldsymbol{g}_0 = -\nabla J(\delta\boldsymbol{x}_0)$

- $\alpha_k = \frac{\boldsymbol{g}_k^T \boldsymbol{g}_k}{\boldsymbol{d}_k^T \boldsymbol{J''} \boldsymbol{d}_k}$
- $\delta\boldsymbol{x}_{k+1} = \delta\boldsymbol{x}_k + \alpha_k d_k$
- $\boldsymbol{g}_{k+1} = -\nabla J(\delta\boldsymbol{x}_{k+1})$
- $\gamma_k = \frac{\boldsymbol{g}_{k+1}^T \boldsymbol{g}_{k+1}}{\boldsymbol{g}_k^T \boldsymbol{g}_k}$
- $\boldsymbol{d}_{k+1} = \boldsymbol{g}_{k+1} + \gamma_k \boldsymbol{d}_k$
- $\delta_{k+1} = \frac{1}{\alpha_{k+1}} + \frac{\gamma_k}{\alpha_k}$
- $\beta_{k+1} = -\frac{c_k}{\alpha_k c_{k+1}}$

with $c_k = \|\boldsymbol{g}_k\|^{-1}$.

The normalized descend vectors, $\boldsymbol{q}_k = \frac{\boldsymbol{g}_k}{\|\boldsymbol{g}_k\|}$, satisfy the relation, $\boldsymbol{J''}\boldsymbol{q}_k = \beta_{k+1}\boldsymbol{q}_{k+1} + \delta_k \boldsymbol{q}_k + \beta_k \boldsymbol{q}_{k-1}$. $\boldsymbol{J''}$ can thus be approximated by $\boldsymbol{T}_D = \boldsymbol{Q}_D^T \boldsymbol{J''} \boldsymbol{Q}_D$, where $D$ is the number of iteration, and the rank of $\boldsymbol{T}_D$. $\boldsymbol{T}_D$ is a symmetric tridiagonal matrix, with the $\delta_k$ coefficients on its diagonal and the $\beta_k$ coefficients above and under. $\boldsymbol{Q}_D$ is the matrix which columns are the vectors $\boldsymbol{q}_k$.

# References

Ancell, B. and G. J. Hakim, 2007: Comparing Adjoint- and Ensemble-Sensitivity Analysis with Applications to Observation Targeting. *Mon. Wea. Rev.*, **135**, 4117–4134.

Bergot, T. and A. Doerenbecher, 2002: A study on the optimization of the deployment of targeted observations using adjoint-based methods. *Quart. J. Roy. Meteor. Soc.*, **128**, 1689–1712.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform kalman filter. part i: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436.

Burger, V. P., G. and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724.

Cardinali, C., 2009: Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.*, **135**, 239–250.

Courtier, P., J. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, **120**, 1367–1387.

Doerenbecher, A. and T. Bergot, 2001: Sensitivity to observations applied to FASTEX cases. *Nonlinear Processes in Geophysics*, **8**, 467–481.

Fisher, M. and P. Courtier, 1995: Estimating the covariance matrix of analysis and forecast error in variational data assimilation. Tech. rep., ECMWF.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2006: Efficient Data Assimilation for Spatiotemporal Chaos: a Local Ensemble Transform Kalman Filter. *Physica D*, **230**, 112–126.

Leutbecher, M., 2003: A reduced rank estimate of forecast error variance changes due to intermittent modifications of the observing network. *J. Atmos. Sci.*, **60 (5)**, 729–742, doi:10.1175/1520-0469(2003)060<0729:ARREOF>2.0.CO;2.

Lorenz, E. N. and K. A. Emanuel, 1998: Optimal Sites for Supplementary Weather Observations: Simulation with a Small Model. *J. Atmos. Sci.*, **55**, 399–414.

N. Oger, A., O. Pannekoucke and P. Arbogast, 2012: Assessing the trajectory influence in adaptive observation Kalman filter sensitivity method. *Quart. J. Roy. Meteor. Soc.*, **138**, 813–825, doi:10.1002/qj.950.

NL., B. and D. R, 2000: Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quart. J. Roy. Meteor. Soc.*, **126**, 1431–1454.

PL Houtekamer, H. M., 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131**, 1485–1490.

Trémolet, Y., 2008: Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus*, **60**, 964–978.

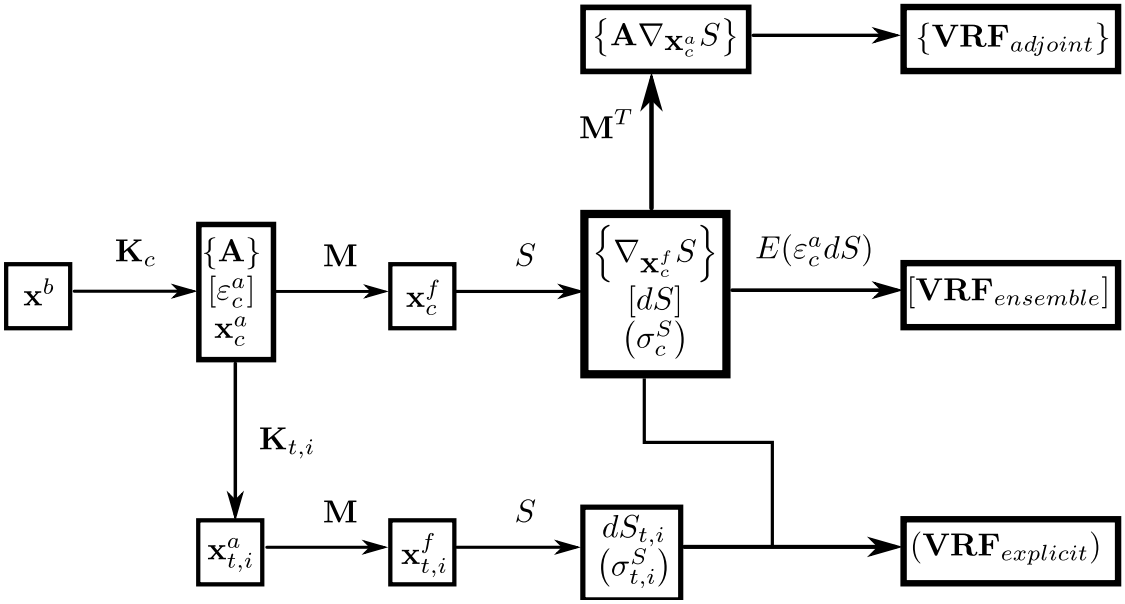



Figure 1: Summary of the two methods to compute the VRF fields and the explicit variance reduction. { } shows the terms used for the adjoint method, [ ] for the ensemble method and ( ) for the explicit computation. Some shortcuts have been made to make the picture readable. The different integrations for each member of the ensemble have not been shown.

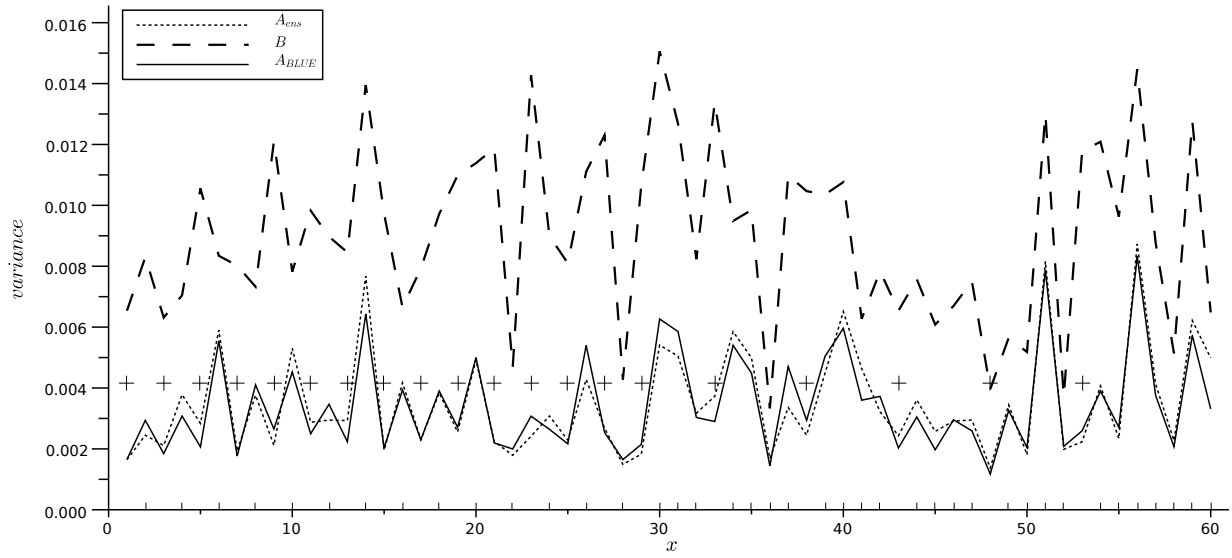



Figure 2: Error variance reduction between the forecast and the analysis. $\boldsymbol{A}_{BLUE}$ is the analysis error variance computed from the BLUE equation (eq. (5)). $\boldsymbol{A}_{ens}$ is computed from the ensemble of $\boldsymbol{x}^a$. The $+$ show the location of the observation sites.

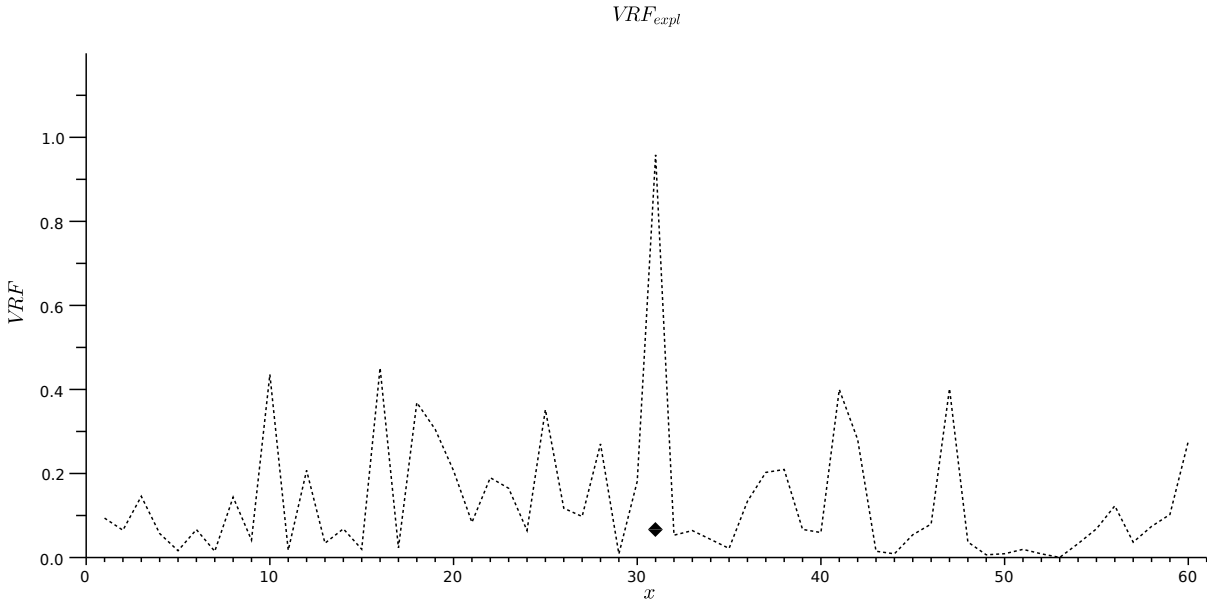



Figure 3: One day variance reduction expectation from $VRF$ explicit computation. The maximum value of $VRF$ is emphasized by the diamond.

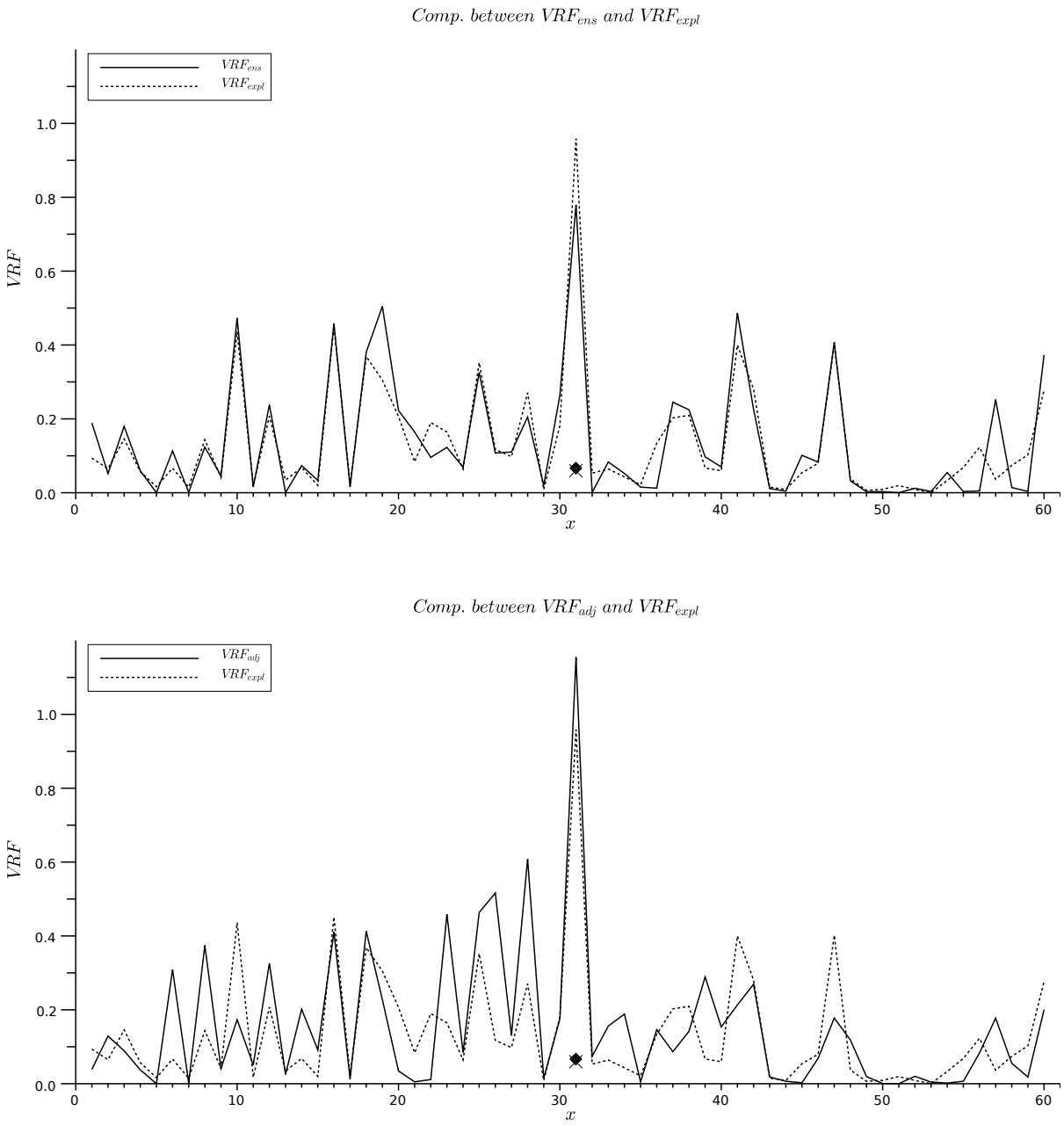



Figure 4: One day variance reduction expectation using the three methods. The upper picture shows the comparison between the $VRF$ computed from the ensemble ($VRF_{ens}$) and the $VRF$ computed explicitly ($VRF_{expl}$). The lower picture shows the comparison between the $VRF$ computed from the ensemble ($VRF_{ens}$) and the $VRF$ computed explicitly ($VRF_{expl}$). The three methods lead to the same deployment strategy at grid point 31.

# Chapitre 8

# Implémentation de la VRF dans OOPS

Le chapitre précédent a introduit un nouvel outil pour l'observation adaptative, la VRF. L'implémentation et la validation de cette approche ont été réalisées dans un cadre académique, celui du modèle de Lorenz 1995 avec une implémentation avec le logiciel Scilab. Pour que cette démonstration soit pleinement convaincante et satisfaisante pour la prévision numérique, il est nécessaire d'apporter la preuve de l'intérêt de la VRF pour la PNT opérationnelle. Pour cela nous avons besoin de développer l'algorithme de la VRF dans le code ARPEGE/IFS.

Malgré son apparente simplicité algorithmique, l'implémentation de la VRF dans un système opérationnel représente un travail très difficile. Pour parvenir à tester la VRF dans un contexte opérationnel, une possibilité est de s'appuyer sur l'environnement OOPS (*Object-Oriented Prediction System*) en cours de développement et de validation à la fois au CNRM et au CEPMMT. Ainsi, en implémentant la VRF dans OOPS et en la testant dans un modèle simplifié (le QG est déjà disponible sous OOPS), il sera possible d'étudier l'organisation des composants à mettre en place tout en proposant une solution aboutie utilisable dans le code opérationnel. Une motivation supplémentaire quant au choix de l'implémentation sous OOPS est de pouvoir partager plus facilement les développements logiciel, par exemple en partageant des algorithmes avec d'autres laboratoires partenaires, comme le CERFACS.

Ce chapitre décrit l'environnement OOPS et l'implémentation de la VRF dans ce cadre. Le projet OOPS vise à révolutionner l'organisation des codes informatiques des systèmes d'assimilation pour la prévision numérique du temps à Météo-France (ARPEGE et AROME) et au CEPMMT (IFS). La formalisation dans le code d'objets communément manipulés dans les écritures mathématiques de l'assimilation doit permettre une flexibilité accrue de ce code pour l'implémentation de nouveaux algorithmes d'assimilation. Une couche de code en langage C++ est donc ajoutée au code historiquement en FORTRAN lui-même révisé pour gagner en modularité.

Le contenu présenté ici est très technique et n'a pas vocation à ajouter des nouveautés scientifiques à ce manuscrit. Son objectif est de décrire le contenu et le fonctionnement d'une implémentation OOPS, en reprenant dans un premier temps l'existant (réaliser une intégration temporelle, faire une assimilation de données, générer un ensemble) pour finir sur une illustration de la manière d'implémenter la VRF dans cet environnement.

## 8.1 Contexte et objectifs de OOPS

Les systèmes de prévision numérique du temps à Météo-France sont de plus en plus complexes. Depuis la mise en place du 4D-Var au début des années 2000, le développement de systèmes de prévision ensemblistes et leur interconnexion avec les systèmes déterministes eux-mêmes en évolution (modèle AROME dès 2008) résultent en un ensemble cohérent de systèmes interdépendants. En outre, le schéma d'assimilation varie en fonction du modèle, ainsi que les observations utilisées. Le 3D-Var de AROME assimile des données de réfléctivité radar non exploitées dans le 4D-Var ARPEGE. Il faut donc faire évoluer des spécificités algorithmiques dans un code unique, de surcroît partagé avec le CEPMMT (modèle IFS et son

4D-Var) et les partenaires ALADIN (autre modèle à aire limitée qui fonctionne sur la base des librairies ARPEGE/IFS).

Au fur et à mesure des nouvelles versions des codes informatiques, la correction des bugs et la validation de nouvelles configurations prend de plus en plus de temps. Par ailleurs, ces codes, dédiés aux approches variationnelles sont devenus assez difficiles à faire évoluer vers de nouveaux algorithmes d'assimilation, notamment ceux exploitant pleinement les ensembles.

Le projet OOPS a été conçu pour transformer le code ARPEGE/IFS de telle sorte qu'il se rapproche de la manière dont sont écrites les équations algébriques de l'assimilation, voir Trémolet and Fisher (2012). En effet, les algorithmes de l'assimilation de données reposent sur un ensemble d'opérateurs et d'états qu'il est possible d'organiser différemment suivant le type de formulation à implémenter. Lagarde et al. (2001) proposent ainsi une visualisation graphique des différentes variables, opérateurs et espaces qui interviennent lors de l'analyse. Ces illustrations permettent de mettre en évidence les points communs et variations entre les différentes méthodes variationnelles. L'étape suivante consiste à pouvoir construire un code en suivant la philosophie des schémas proposés.

Passer par un langage orienté objet, ici le C++, permet d'exploiter les notions de polymorphisme et d'abstraction. Nous avons alors uniquement besoin de savoir quel objet est manipulé (un vecteur d'état, une observation, une matrice de covariance d'erreur) sans devoir décrire la manière dont il est représenté en mémoire (dimension du tableau par exemple). L'architecture du code avec OOPS est découpée en trois niveaux de spécialisation, qui sont décrits dans la section suivante.

## 8.2 Architecture générale du code

Le code OOPS a été construit en trois niveaux, du plus abstrait au plus proche des données. Le premier niveau correspond aux classes génériques applicables à l'ensemble des simulations, le deuxième niveau regroupe les classes spécifiques à un modèle et le troisième niveau contient le code FORTRAN qui implémente les objets du deuxième niveau. Ces trois niveaux sont définis par des fichiers d'interface qui permettent de passer de l'un à l'autre en spécifiant les objets. La définition des `Traits` relie les classes génériques du premier niveau à celles plus spécifiques du deuxième. Le fichier `*Fortran.h` définie les liens entre les fonctions C++, du deuxième niveau, et les fonctions FORTRAN du troisième niveau.

### 8.2.1 Classes génériques : 1$^{er}$ niveau

Le premier niveau contient les classes génériques (un modèle, une observation, un ensemble, etc.), c'est le plus haut niveau d'abstraction. Ces classes sont basées sur la notion de *template* du C++, elles sont indépendantes du modèle utilisé (ARPEGE, AROME, QG 2 couches ou Lorenz 95) et à ce niveau le structure des données en mémoire est inconnue. Certaines méthodes génériques sont cependant directement codées à ce niveau, comme les algorithmes de minimisation.

La classe `Forecast` est, par exemple, définie par la ligne suivante, dans le fichier `Forecast.h` :

```
template <typename MODEL> class Forecast {}
```

où `typename MODEL` introduit un type générique `MODEL`. Ce type est spécifié lors de l'instanciation de la classe. La déclaration suivante utilise deux niveaux de template pour définir l'instance de l'objet `run`,

```
qg::RunQg< oops::Forecast<qg::QgTraits> > run(argc, argv);
```

avec `qg::QgTraits` le type associé au template `Forecast` et `oops::Forecast<qg::QgTraits>` le type associé au template `RunQg`.

La classe `Run`, dont `RunQg` est une spécialisation, est la classe de plus haut niveau dans l'architecture OOPS. C'est elle qui est définie dans la fonction `main()` du programme principal qui sera exécuté. Elle regroupe l'ensemble des objets de la simulation. Son type indique le choix de la simulation qui est réalisée,

`Forecast` pour une prévision, `Variational` pour une assimilation variationnelle ou `GensEnsPertB` pour la génération d'un ensemble, par exemple.

La figure 8.1 illustre les liens entre les différentes classes de niveau 1. Pour chacune de ces classes un type non défini `MODEL` est utilisé. Ces classes, leurs attributs et leurs méthodes sont valables pour n'importe quel type de modèle utilisé, que ce soit le QG ou encore le Lorenz 95 .

### 8.2.2 Classes spécifiques : 2$^{e}$ niveau

Le 2$^{e}$ niveau décrit les spécialisations des classes du 1$^{er}$ niveau propre à un modèle, c'est-à-dire qu'en fonction du `MODEL` passé en argument du template du 1$^{er}$ niveau le code va utiliser un objet spécifique. Par exemple la classe `State` est définie de manière abstraite au 1$^{er}$ niveau et elle peut être spécialisée en un objet `StateL95` pour le modèle de Lorenz ou un objet `QgState` pour le QG.

C'est le type `qg::QgTraits`, décrit dans l'exemple précédent, qui fait l'interface vers le 2$^{e}$ niveau. Ici il indique que c'est le modèle Quasi-Géostrophique qui va être utilisé. La structure `QgTraits`, dans le fichier `QgTraits.h`, regroupe les liens entre les classes du 1$^{er}$ niveau et du 2$^{e}$,

```
struct QgTraits {
   typedef qg::QgState               State;
   typedef qg::QgModel               ModelConfiguration;
   typedef qg::QgIncrement           Increment;
   typedef qg::QgTLM                 LinearModel;

   typedef qg::QgErrorCovariance     ModelSpaceCovariance;
   typedef oops::NullModelAux        ModelAuxControl;
   typedef oops::NullModelAux        ModelAuxIncrement;

   typedef qg::QgObservation         ObsOperator;
   typedef qg::ObsTrajQG             ObsOperatorLinearizationTrajectory;
   typedef oops::NullObsAux          ObsAuxControl;
   typedef oops::NullObsAux          ObsAuxIncrement;
   typedef qg::ObsVecQG              ObsVector;

   typedef qg::GomQG                 ModelAtLocations;
   typedef qg::LocQG                 Locations;

   typedef qg::LocalizationMatrixQG  LocalizationMatrix;
};
```

où l'ensemble des éléments définis comme `qg::Qg*` (à gauche) sont les classes de 2$^{e}$ niveau, et les classes à droite sont celles du 1$^{er}$ niveau.

Les classes de 2$^{e}$ niveau sont spécifiques au modèle QG, des classes équivalentes sont définies pour le modèle de Lorenz 95 dans le fichier `L95Traits.h`. Pour une implémentation dans un autre modèle (ARPEGE ou AROME), il est nécessaire d'identifier ces même types de classe dans un fichier `*Traits.h`.

Ces classes sont spécifiques à chaque modèle car elles implémentent des fonctions propres au modèle, comme par exemple

```
QgFields::QgFields(const QgGeometry & geom, const VarQG & vars, const
util::DateTime & time):
   geom_(new QgGeometry(geom)), vars_(new VarQG(vars)), time_(time) {
   qg_field_create_f90(&flds_, geom_->toFortran(), vars_->toFortran());
}
```

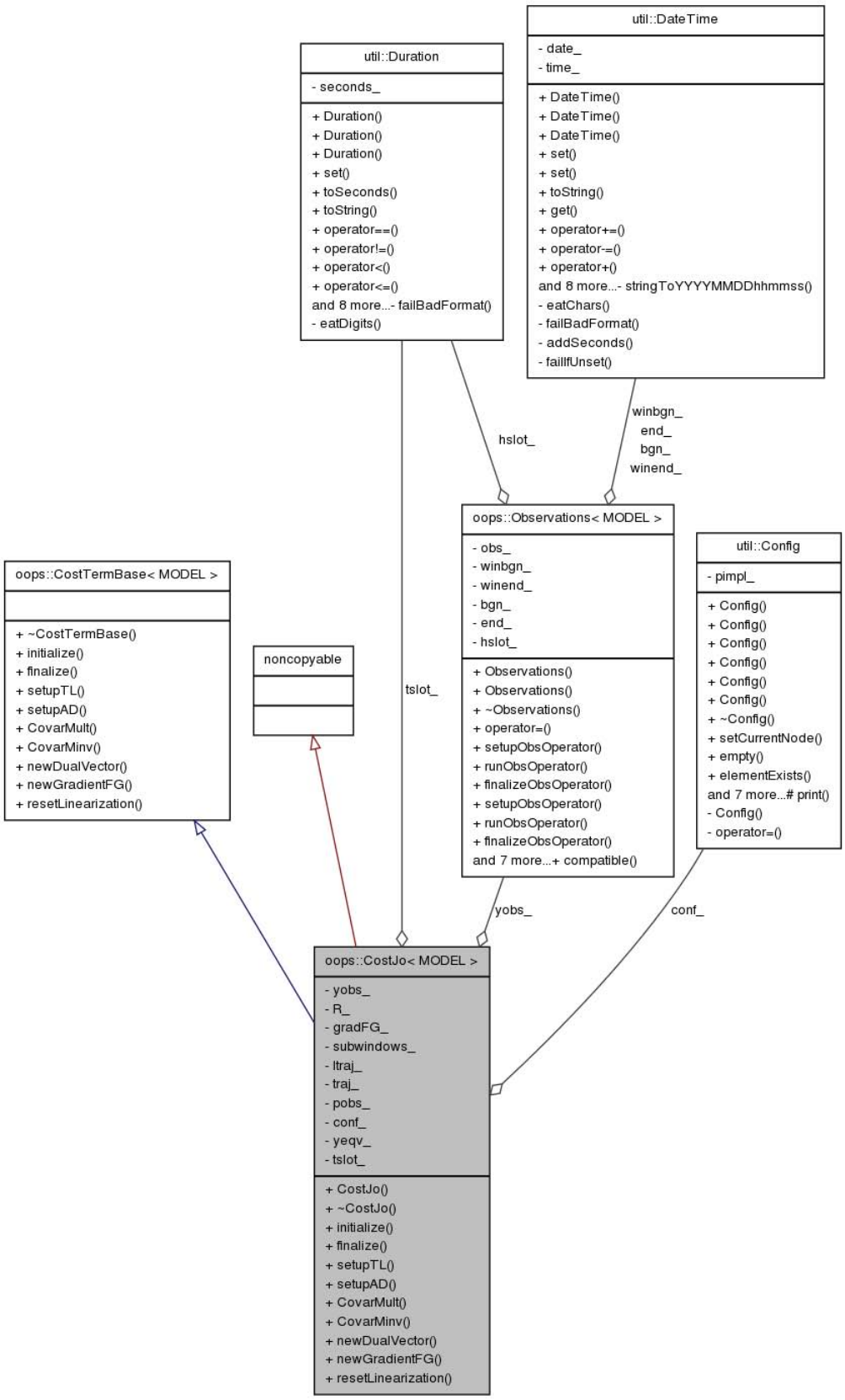



FIGURE 8.1 – Diagramme de classe pour `CostJo`. Les attributs de la classe sont listés dans le 1er rectangle sous le nom de la classe, ses méthodes sont listées dans le 2e rectangle. Les flèches symbolisent l'héritage entre les classes. La classe `CostJo` est une spécialisation de `CostTermBase`. Elle est composée d'un ensemble d'`Observations` qui contiennent elles-mêmes des données. Les classes `Duration` et `Config` sont des classes *outils* permettant l'implémentation.

qui génère un champ pour le modèle QG à partir de sa géométrie (`geom_`, nombre de points de grille en latitude et longitude) et du type de variable (`vars_`). La fonction `qg_field_create_f90` est ici une fonction C++, alors que comme le laisse penser son suffixe `f90` le code à l'intérieur est écrit en FORTRAN.

Le 3$^{e}$ niveau fait le lien entre les fonctions définies au niveau 2 et le code FORTRAN qui va effectivement réaliser les calculs. Ce troisième niveau permet de conserver l'ensemble des optimisations de code de bas niveau et de reprendre le code FORTRAN existant (légèrement réusiné). Les objets en mémoire sont transmis du niveau 2 au niveau 3 par des pointeurs.

### 8.2.3 Lien entre le C++ et le FORTRAN : 3$^{e}$ niveau

La majorité les traitements sur les valeurs en mémoire sont codés en FORTRAN, ce qui définit les fonctions du 3$^{e}$ niveau. Les translations entre le code C++ et le FORTRAN sont réalisées à l'aide de deux mécanismes, l'utilisation de librairies FORTRAN lors de la compilation du C++ et la conversion des pointeurs (dans les deux sens).

#### Implémentation d'une fonction FORTRAN dans un code C++

Pour illustrer le mécanisme, un exemple tiré du fichier FORTRAN `qg_fields_interface.f90` est utilisé.

```
subroutine qg_field_create_c (c_self,c_geom,c_vars) bind(c,name='qg_field_create_f90')
```

La fonctionnalité ISO C Binding de FORTRAN 2003 permet de convertir la fonction `qg_field_create_c` codée en FORTRAN en une fonction C++ appelée `qg_field_create_f90`. La liste des fonctions FORTRAN utilisées est définie dans le fichier `QgFortran.h`, on y retrouve la déclaration de la fonction `qg_field_create_f90`,

```
void qg_field_create_f90(F90flds **, const F90geom * const *, const F90vars * const *);
```

L'ISO C Binding permet de faire le lien entre les librairies de calcul en FORTRAN et le code C++. L'exécutable qui est lancé au niveau du système est donc obtenu par la compilation du code C++. Cette compilation génère des fichiers binaires qui sont généralement appelés `.x` et auxquels on passe en argument un fichier XML de paramètres. Par exemple pour calculer une trajectoire de référence,

```
./${model}_forecast.x ${model}/scripts/truth.XML
```

où $model peut être égal à L95 (pour le modèle de Lorenz 95), QG (pour le modèle quasi-géostrophique), ARPEGE ou AROME. Le déclenchement des exécutables binaires avec leurs fichiers de configuration sont rassemblés dans des scripts (ou job) qui sont soumis au serveur de calcul.

#### Illustration graphique des trois niveaux

Le schéma 8.2 illustre les niveaux présentés précédemment, à partir de l'exécutable binaire jusqu'à la lecture du fichier de données contenant les valeurs de l'objet `Field` à créer. Un développement similaire pourrait être fait pour les objets `QgModel` et `VarQG` mais alourdirait le schéma. De nombreuses représentations sont possibles pour illustrer les liens entre les objets, celle retenue pour ce schéma est de suivre la déclaration des objets dans les différents fichiers.

#### Passage des variables en mémoire

Les variables sont à la fois manipulées par le code C++ et par le code FORTRAN. Il faut donc utiliser une représentation en mémoire qui soit compatible avec ces deux langages. L'exemple au-dessus de création d'un champ pour le QG est repris pour illustrer la manière dont la variable est représentée en mémoire.

L'objet `qg_field` dispose d'une structure en FORTRAN définie dans le fichier `qg_fields.f90`,

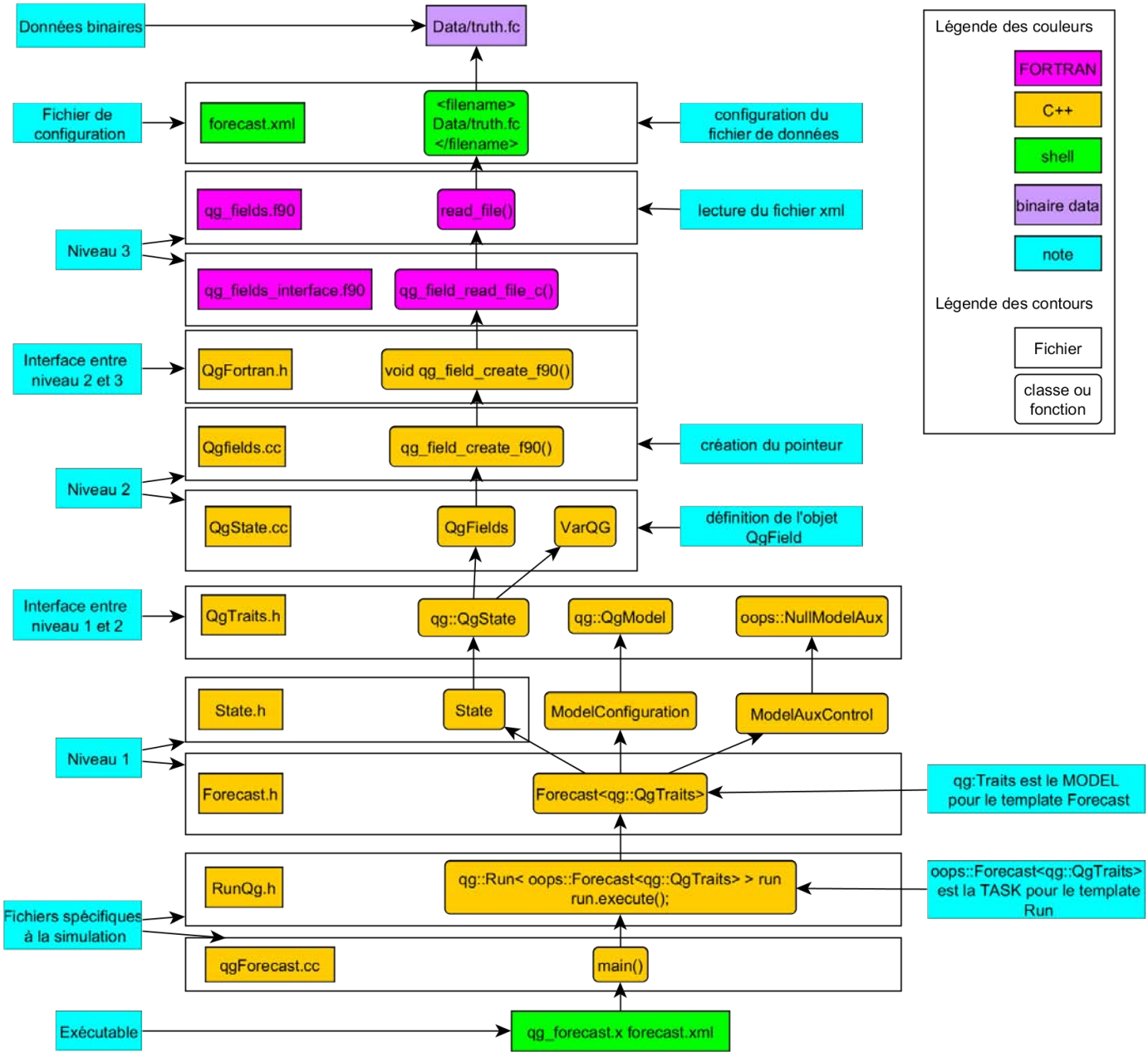



FIGURE 8.2 – Illustration de l'implémentation à 3 niveaux dans OOPS. La lecture du schéma se fait du bas vers le haut, en partant de l'exécutable jusqu'à la mise en mémoire des valeurs. Seule l'arborescence pour la création de la variable `QgField` est représentée.

```
type :: qg_field
   integer :: nx                     !< Zonal grid dimension
   integer :: ny                     !< Meridional grid dimension
   integer :: nl                     !< Number of levels
   integer :: nf                     !< Number of fields
   logical :: lbc                    !< North-South boundary is present
   real(kind=kind_real), pointer :: gfld3d(:,:,:)    !< 3D fields
   real(kind=kind_real), pointer :: x(:,:,:)         !< Stream function
   real(kind=kind_real), pointer :: q(:,:,:)         !< Potential vorticity
   real(kind=kind_real), pointer :: u(:,:,:)         !< Zonal wind
   real(kind=kind_real), pointer :: v(:,:,:)         !< Meridional wind
   real(kind=kind_real), pointer :: xbound(:)        !< Streamfunction on walls
   real(kind=kind_real), pointer :: x_north(:)       !< Streamfunction on northern wall
   real(kind=kind_real), pointer :: x_south(:)       !< Streamfunction on southern wall
   real(kind=kind_real), pointer :: qbound(:,:)      !< PV on walls
   real(kind=kind_real), pointer :: q_north(:,:)     !< PV on northern wall
   real(kind=kind_real), pointer :: q_south(:,:)     !< PV on southern wall
   character(len=1), allocatable :: fldnames(:) !< Variable identifiers
end type qg_field
```

L'ensemble des éléments de la structure, hors paramètres de la grille, sont des pointeurs de dimension dynamique. La translation des pointers C++ vers le FORTRAN est réalisée par la fonction `c_f_pointer(c_ptr,f_ptr` où `c_ptr` est le pointeur C++ et `f_ptr` le pointeur en FORTRAN. Lors de l'appel à la fonction FORTRAN `qg_field_create_c`, les arguments `c_geom` et `c_vars` sont des pointeurs C++, ils sont convertis avant d'être repris en FORTRAN.

```
call c_f_pointer(c_geom,geom)
call c_f_pointer(c_vars,vars)
allocate(qg_field)
call create(self, geom, vars)
c_qgfield = c_loc(qg_field)
```

La dernière fonction utilisée, `c_loc`, permet de retourner l'adresse du pointeur FORTRAN au format C++. Ainsi le pointeur renvoyé vers le code C++ par le `return` du FORTAN est un pointeur C++.

La figure 8.3 illustre le passage en argument des différents pointeurs et l'organisation en mémoire. L'enchaînement des appels de fonctions lors de la création d'un objet `QgFields` est montré au niveau du code, à gauche de l'image. Les fichiers `QgFields.h`, `QgFields.cc` et `QgFortan.h` sont codés en C++, le fichier `qg_fields_interface.f90` est codé en FORTRAN. Le constructeur de la classe `QgFiedls` appelle la fonction `qg_field_create_f90`. Le fichier `QgFortan.h` fait le lien entre `qg_field_create_f90` et `qg_field_create_c`. Les objets en mémoire sont représentés sur la droite de l'image. Les pointeurs `geom_` et `vars_` sont passés du C++ au FORTRAN par l'appel `toFortran()` qui correspond à `c_f_pointer`. Le pointeur vers la mémoire correspondant au tableau de valeurs `flds_` est retourné par la fonction `c_loc`.

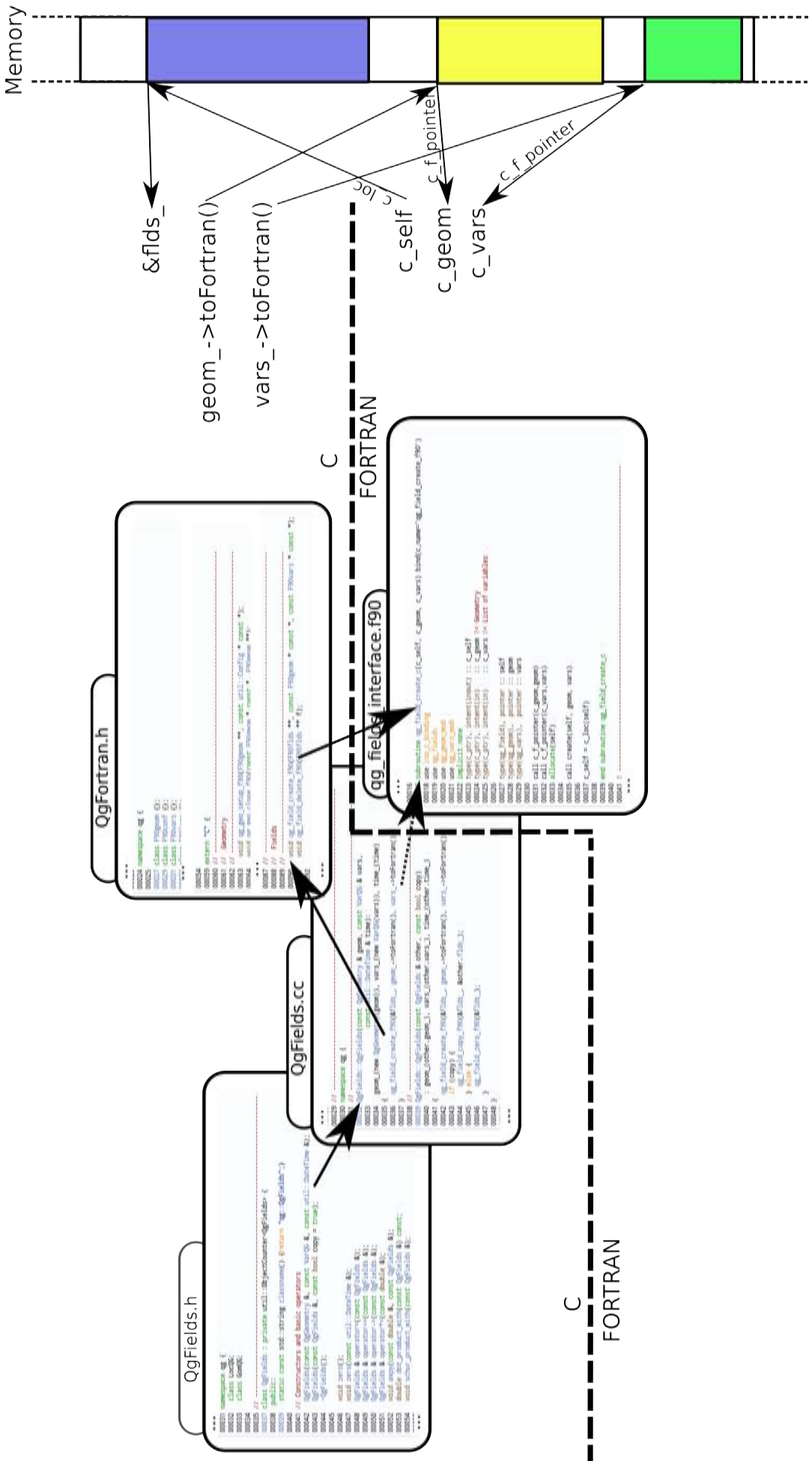



FIGURE 8.3 – Voir le texte pour les commentaires.

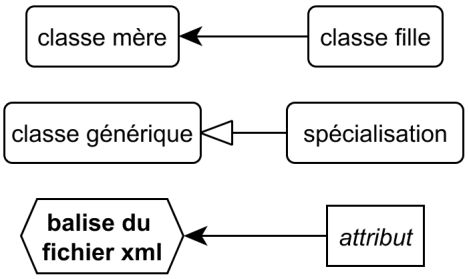



FIGURE 8.4 – Principes de représentation graphique utilisés dans les figures suivantes. Les rectangles arrondis servent à identifier les classes. Les hexagones en gras représentent les balises contenues dans le fichier XML de configuration. Et les rectangles en italique les attributs définis à l'intérieur des balises. Une flèche simple symbolise une relation de composition, la classe mère est composé (ou a besoin) de la classe fille. La flèche en pointe (vide) symbolise une spécialisation, une classe spécialisée reprend l'ensemble des attributs de la classe générique.

## 8.3 Description des objets utilisés dans OOPS

La section précédente a présenté l'architecture du code, l'utilisation des *templates* et les interfaces entre le code FORTRAN et le C++. Les classes d'objet codées dans OOPS sont maintenant décrites. Les classes sont organisées par `namespaces` C++, qui regroupent les classes de même usage. Le `namespace` est indiqué devant la classe lors d'un appel par la syntaxe `Namespace::Classe`. Il y a trois types de `namespaces`,

- les `namespaces` utilitaires qui peuvent être utilisés dans tous les codes, avec `logger` pour l'écriture dans des fichiers de log, et `util` pour la manipulation des dates et des objets `Config`,
- le `namespace oops` qui regroupe toutes les classes de niveau 1,
- les `namespaces` spécifiques à un modèle, qui regroupent les classes de niveau 2, `lorenz95` et `qg`.

Dans cette section seul le `namespace oops` est décrit pour réaliser une description exhaustive des différents objets de niveau 1 utilisés dans chaque type de simulation. Comme présenté sur la figure 8.2, l'objet de plus haut niveau est de type `Run`. Dans l'exemple de la figure, la classe `RunQg` est une spécialisation de la classe `Run`. Cette classe `Run` récupère en argument le fichier XML qui sera utilisé pour définir l'ensemble des objets `Config` (voir 8.3.2).

### 8.3.1 Organisation des classes dans OOPS

Une description des classes de niveau 1 peut être trouvée dans la documentation de Trémolet and Fisher (2012). Une description textuelle des différentes classes n'est pas reprise dans ce manuscrit, mais nous nous intéressons plutôt à l'organisation entre les classes et comment celles-ci sont définies à partir des fichiers XML passés en argument.

Les figures suivantes reprennent une présentation simplifiée de schémas UML pour illustrer quels sont les liens entre les différentes classes. Les principaux des éléments (classe, balise, etc.) utilisés dans les figures 8.5, 8.6, 8.7, 8.8, 8.9 et 8.10, et les principes de représentations qui les identifient sont schématisés dans la figure 8.4. Pour simplifier les diagrammes, l'ensemble des interactions entre classes n'est pas toujours exhaustif sur chaque figure. Si une classe a été décrite en détail dans un schéma, tous les liens ne seront pas repris dans les figures suivantes.

Les schémas présentés ici sont réalisés avec une configuration correspondant à un modèle QG et une minimisation par un algorithme de Lanczos. De même, les différents vecteurs d'état sont configurés pour être lus à partir de fichier déjà existant, il n'y a donc pas de configuration pour la création d'un vecteur d'état. Les paramètres relatifs au modèle, le type d'observation et les éléments du variationnel peuvent varier suivant le type de modèle choisi. Cependant l'organisation entre les classes est indépendante du modèle choisi ou du type d'assimilation.

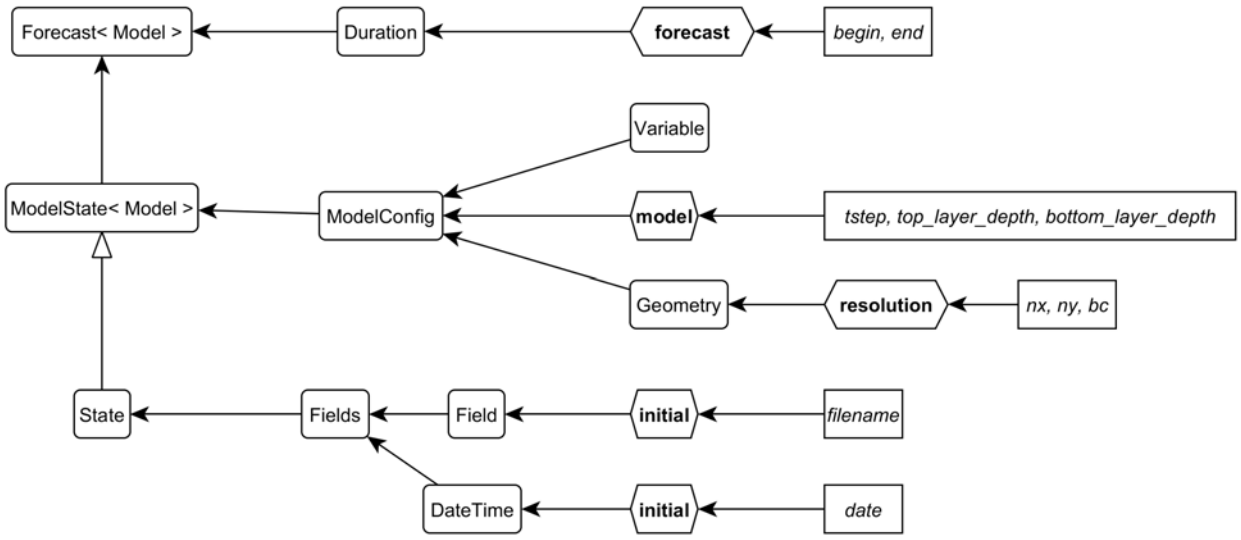



FIGURE 8.5 – Interactions entre les classes autour de `State`. La classe `State` représente le vecteur d'état du modèle, elle implémente la fonction `forecast`. Le vecteur d'état est composé de la classe `Fields` et reprend les paramètres de `ModelConfig`. `ModelConfig` regroupe la définition du type de variable (température, vent, etc,) qui est non spécifié pour le QG dans l'exemple ici, de l'objet `Geometry` qui décrit la grille utilisée, et des paramètres spécifiques au modèle (pas de temps et hauteurs des 2 couches du QG). L'objet `Fields` regroupe la date de validité du vecteur d'état et l'objet `Field` qui contient les pointeurs vers le tableau en mémoire. La balise `initial` indique dans cette exemple que le vecteur d'état est construit à partir d'un fichier déjà existant.

### 8.3.2 Configuration d'une simulation

Certaines balises utilisées pour définir les paramètres passés en argument au fichier binaire pour construire l'objet `Config` général ont été présentées dans les figures précédentes. Tous ces paramètres sont définis dans des fichiers XML. Ces fichiers de configuration XML remplacent les *namelists* couramment utilisés en PNT.

Dans le code C++, les paramètres définis dans ces fichiers sont repris dans les objets de type `Config`. La classe `Config` intègre un mécanisme pour extraire une partie de la configuration de l'objet général pour ne pas avoir à passer l'ensemble de la configuration à toutes les classes, par la syntaxe `config($valeur)`. Où `$valeur` correspond aux différentes balises du XML lui-même, `resolution`, `forecast` ou `observations`, par exemple.

La configuration d'un ficher XML est à adapter aux besoins de la simulation et est forcément évolutive. Une description exhaustive des paramètres possibles n'est pas donc réalisée ici. Les extraits de fichier XML présentés ci-dessous permettent de replacer les éléments des figures du paragraphe 8.3.1 dans le contexte d'une simulation. Voici un exemple de XML utilisé pour réaliser une prévision sur le modèle QG.

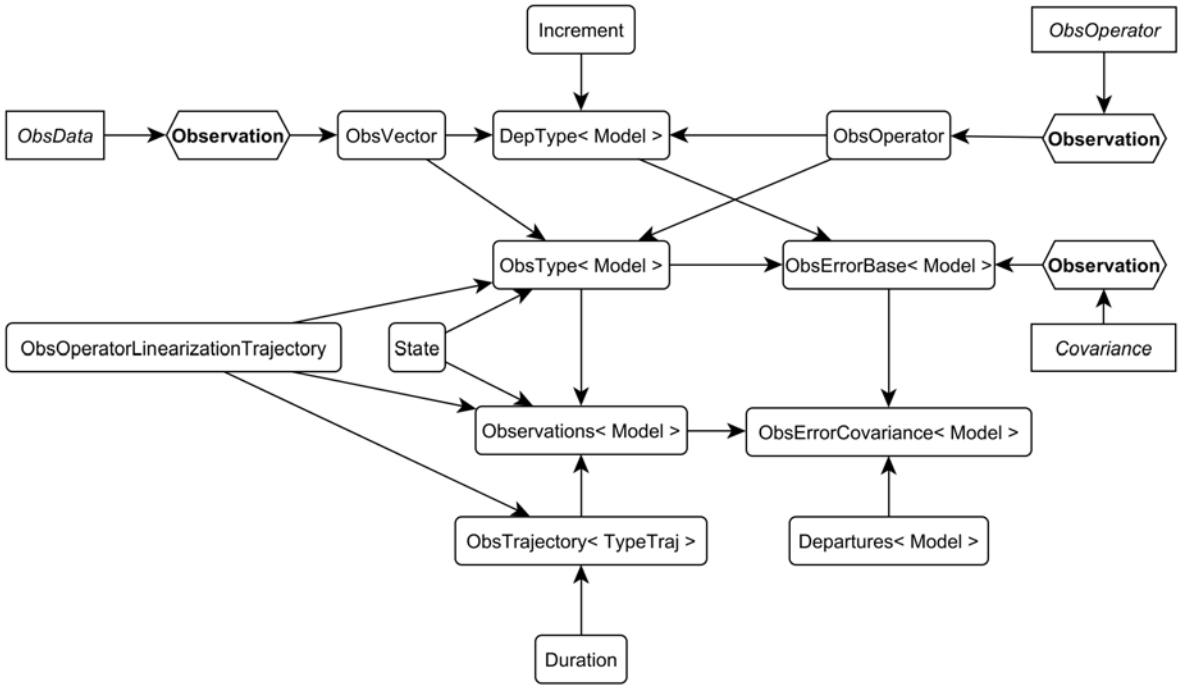



FIGURE 8.6 – Interactions entre les classes autour d'`Observations`. Les observations sont regroupées par type d'observation (`Stream`, `Wind`, ou `WSpeed` pour le QG). Cela se traduit par un objet `ObVector` (qui correspond à $\mathbf{y}^o$) et un objet `ObsOperator` (qui correspond à $\mathbf{H}$) par objet `ObsType`. Toutes les observations sont définies par rapport à un vecteur d'état, de manière à ce que l'ensemble des fonctions relatives aux observations (calcul du vecteur d'innovation par exemple) soient directement accessible via la classe `Observations`. Cela inclut aussi le transport linéaire éventuel pour un 4D-Var (via `ObsTrajectory`). La matrice de covariance d'erreur $\mathbf{R}$ est aussi décomposée en plusieurs objets `ObsErrorBase` selon le type d'observation au sein de la classe `ObsErrorCovariance`.

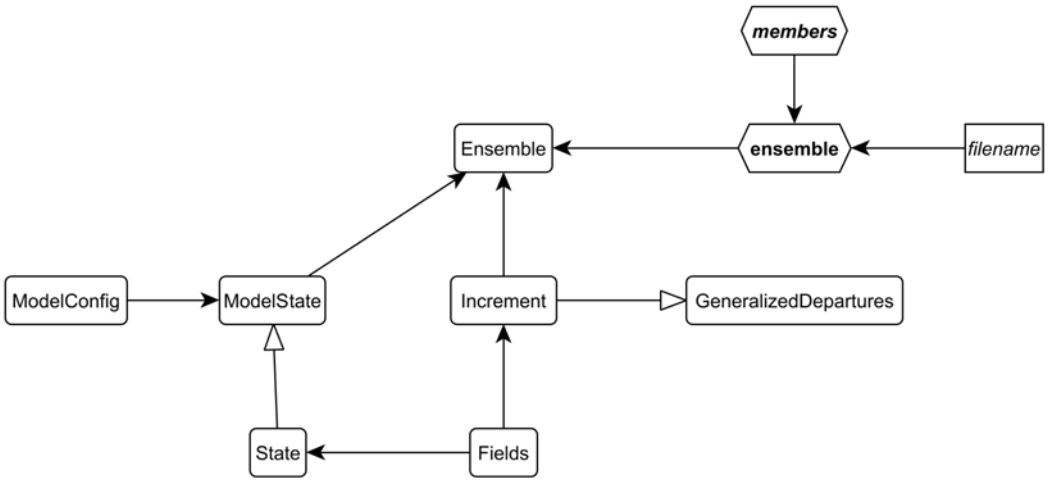



FIGURE 8.7 – Interactions entre les classes autour d'`Ensemble`. Un ensemble est composé d'un vecteur d'état (`ModelState`) et d'une collection d'`Increment`. `GeneralizedDepartures` est une classe générique pour représenter des écarts entre 2 objets du même type, dans l'espace du modèle ou des observations. Ici la définition de l'ensemble se fait à partir d'une liste de fichiers contenant les vecteurs d'état.

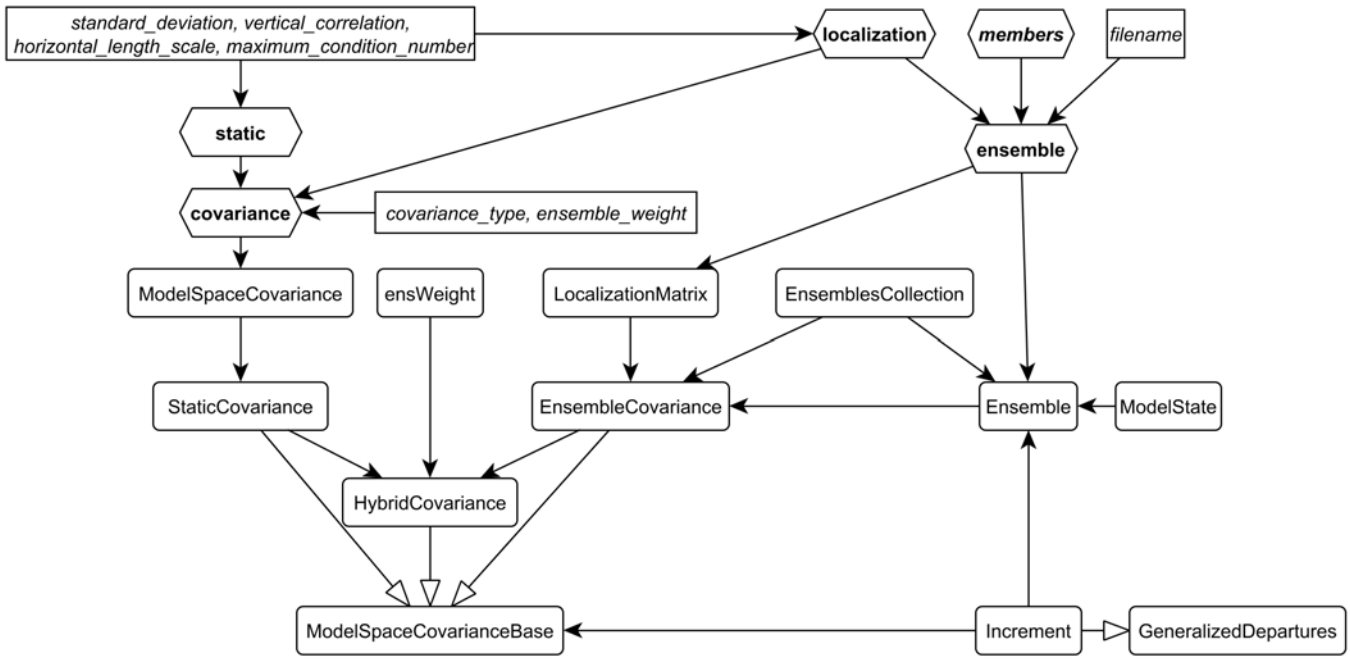



FIGURE 8.8 – Interactions entre les classes autour de `ModelSpaceCovariance`. 3 types de matrice de covariance d'erreur sont possibles dans OOPS pour $\mathbf{B}$ et $\mathbf{A}$ : (i) une matrice statique pour laquelle le choix des fonctions de corrélation et des écart-types est défini dans le fichier XML ; (ii) une matrice de covariance déterminée à partir d'un ensemble, auquel est associée une matrice de localisation ;(iii) une matrice hybride qui est une moyenne pondérée entre une matrice statique et une matrice issue d'un ensemble. La matrice de covariance d'erreur peut être utilisée pour générer un $\varepsilon^b$ ou $\varepsilon^a$ sous la forme d'un `Increment`.

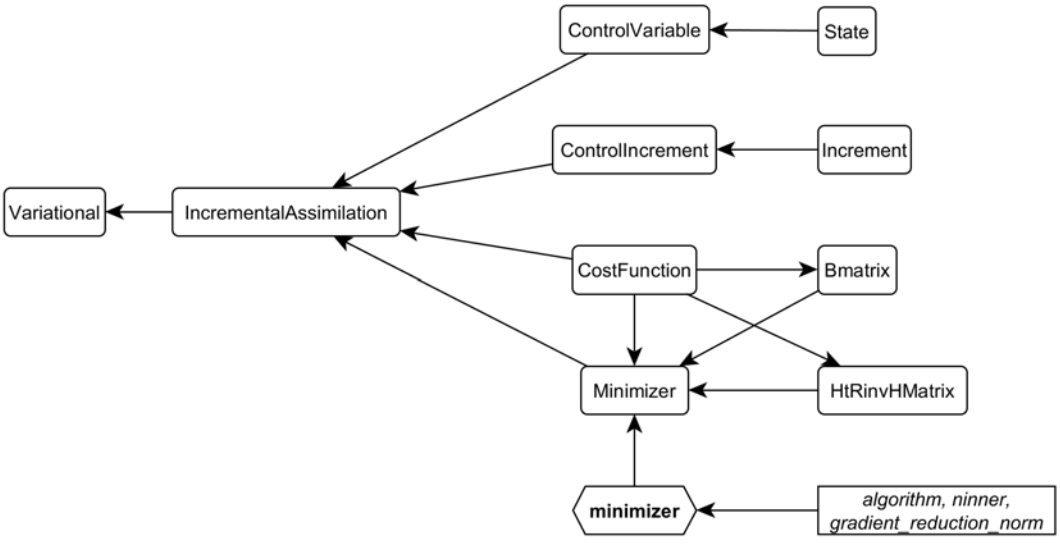



FIGURE 8.9 – Interactions entre les classes autour de `Variational`. L'assimilation variationnelle incrémentale est la combinaison de la variable de contrôle, de l'incrément, de la fonction coût et d'un algorithme de minimisation. L'algorithme de minimisation, `Minimizer` inclus le pré-conditionnement à travers `Bmatrix` obtenu via la fonction coût. Dans l'exemple de l'algorithme de Lanczos représenté ici, la multiplication par $\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ est aussi nécessaire.

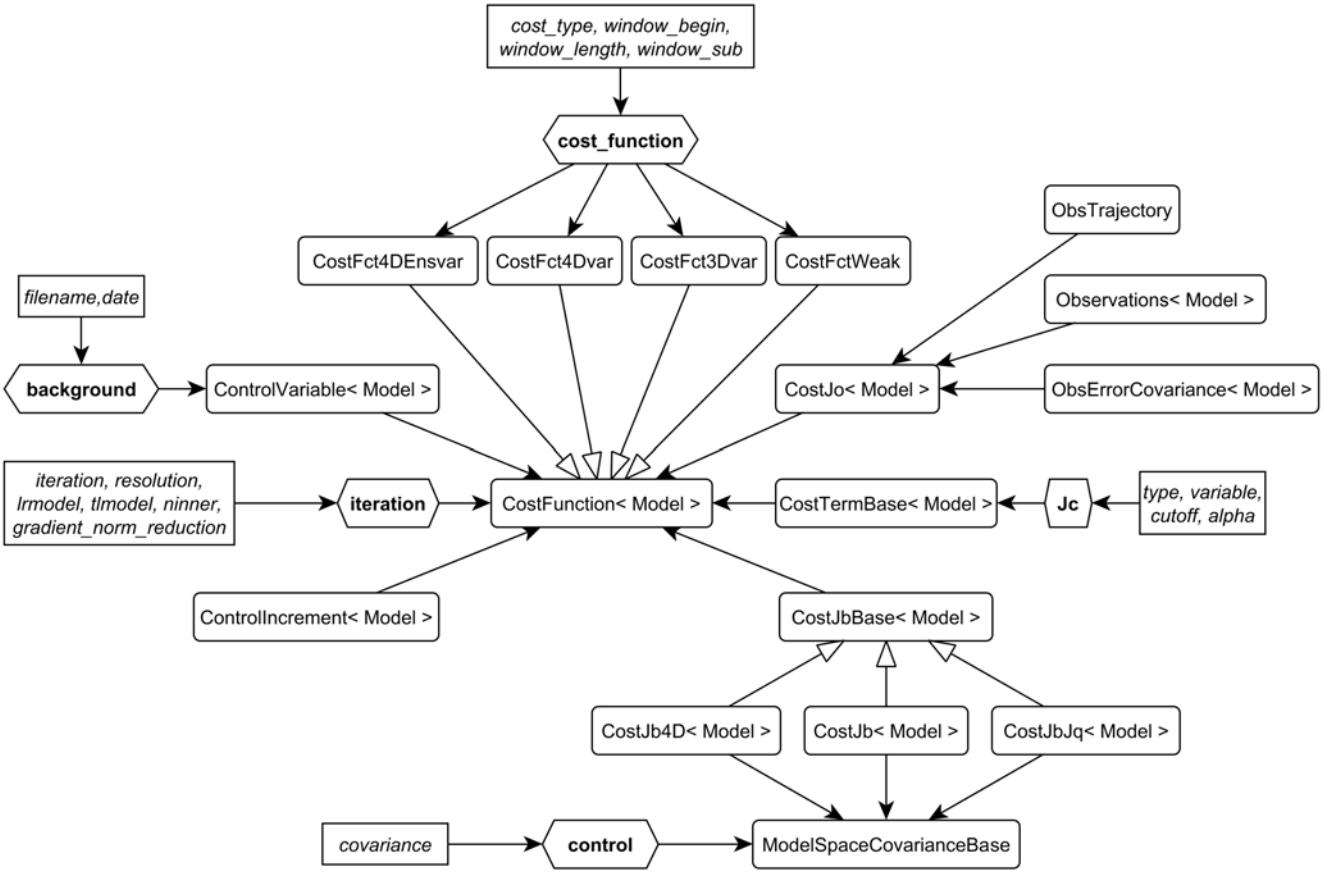



FIGURE 8.10 – Interactions entre les classes autour de `CostFunction`. La fonction de coût est définie par la variable de contrôle et l'incrément (qui peuvent être dans l'espace du modèle ou des observations), et par plusieurs termes qui contribuent à la valeur et au gradient, $J_b$ via `CostJbBase`, $J_o$ via `CostJo` et $J_c$ via `CostTermBase`. 4 types de fonctions coûts sont implémentées dans OOPS , 3D-Var, 4D-Var, 4D-En-Var et 4D-Var en contrainte faible. Le choix de la fonction va déterminer la forme du $J_b$ (extension temporelle pour le 4D-Var et prise en compte de l'erreur modèle pour le 4D-Var en contrainte faible).

Paramètres de la grille :

```
<resolution>
  <nx>40</nx>
  <ny>20</ny>
  <bc>1</bc>
</resolution>
```

Paramètres spécifiques au modèle QG :

```
<model>
  <tstep>PT1H</tstep>
  <top_layer_depth>
    5500.0
  </top_layer_depth>
  <bottom_layer_depth>
    4500.0
  </bottom_layer_depth>
</model>
```

Et la description d'une assimilation de données,

Définition de la condition initiale :

```
<initial>
  <filename>
    Data/truth.fc.P16D
  </filename>
  <date>
    2009-12-31T00:00:00Z
  </date>
</initial>
```

Paramètres pour la prévision :

```
<forecast>
  <begin>
    2009-12-31T00:00:00Z
  </begin>
  <end>
    2010-01-02T00:00:00Z
  </end>
</forecast>
```

Définition de la fonction coût :

```
 <cost_function>
  <cost_type>3D-Var</cost_type>
  <window_begin>
    2009-12-31T00:00:00Z
  </window_begin>
  <window_length>PT6H</window_length>
</cost_function>
```

Choix de l'algorithme de minimisation :

```
<minimizer>
  <algorithm>DIPCG</algorithm>
</minimizer>
```

Définition de la matrice **B** et de l'ébauche :

```
 <control>
  <covariances>
   <covariance>
    <covariance_type>
      QgError
    </covariance_type>
    <standard_deviation>
      0.8
    </standard_deviation>
    <vertical_correlation>
      0.2
    </vertical_correlation>
    <horizontal_length_scale>
      1e6
    </horizontal_length_scale>
    <maximum_condition_number>
      1.0e6
    </maximum_condition_number>
    <date>
      2009-12-31T00:00:00Z
    </date>
   </covariance>
  </covariances>
</control>

<background>
  <states>
   <state>
    <filename>
      Data/truth.fc.P16D
    </filename>
    <date>
      2009-12-31T00:00:00Z
    </date>
  </state>
 </states>
</background>
```

Définition des observations, de l'opérateur d'observation **H** et de **R** :

```
<Observations>
  <Observation>
    <ObsType>Stream</ObsType>
    <ObsOperator>OnTheFly</ObsOperator>
    <ObsData>
      <ObsHandler>ObsHelpQG</ObsHandler>
      <ObsFileIn>
        Data/qg.truth3d.obt
      </ObsFileIn>
      <ObsFileOut>
        Data/qg.3dvar.obt
      </ObsFileOut>
      <obsvalue>ObsVal</obsvalue>
      <obserror>ObsErr</obserror>
    </ObsData>
    <Covariance>diagonal</Covariance>
  </Observation>
  <Observation>
    <ObsType>Wind</ObsType>
    <ObsOperator>OnTheFly</ObsOperator>
    <ObsData>
      <ObsHandler>ObsHelpQG</ObsHandler>
      <ObsFileIn>
        Data/qg.truth3d.obt
      </ObsFileIn>
      <ObsFileOut>
        Data/qg.3dvar.obt
      </ObsFileOut>
      <obsvalue>ObsVal</obsvalue>
      <obserror>ObsErr</obserror>
    </ObsData>
    <Covariance>diagonal</Covariance>
  </Observation>
  <Observation>
    <ObsType>WSpeed</ObsType>
    <ObsOperator>OnTheFly</ObsOperator>
    <ObsData>
      <ObsHandler>ObsHelpQG</ObsHandler>
      <ObsFileIn>
        Data/qg.truth3d.obt
      </ObsFileIn>
      <ObsFileOut>
        Data/qg.3dvar.obt
      </ObsFileOut>
      <obsvalue>ObsVal</obsvalue>
      <obserror>ObsErr</obserror>
    </ObsData>
    <Covariance>diagonal</Covariance>
  </Observation>
</Observations>
```

Les éléments décrits ci-dessus dans le fichier XML sont repris dans les objets `Config` pour être passés en argument pour l'instanciation des objets plus complexes. La plupart des classes passent par des constructeurs classiques pour l'instanciation, comme dans l'exemple du `QgFields` précédent,

```
QgFields::QgFields(const QgGeometry & geom, const VarQG & vars,
 constutil::DateTime & time):
 geom_(new QgGeometry(geom)), vars_(new VarQG(vars)), time_(time)
{
  qg_field_create_f90(&flds_, geom_->toFortran(), vars_->toFortran());
}
```

Certaines classes plus complexes (observations, matrices de covariance ou fonctions de coût) requièrent des méthodes de fabrication abstraite, qui sont appelées `Factory`. Ces `Factory` sont instanciées lorsque le code souhaite créer un nouvel objet complexe. La `Factory` définit alors dynamiquement l'objet à instancier en fonction des éléments contenus dans le XML. Cette mécanique permet de simplifier la lecture du code pour ne pas avoir à définir, élément par élément, les différents types d'observation.

## 8.4 Algorithme de la VRF dans OOPS

L'analyse présentée ci-dessous a été réalisée avec la version du code OOPS disponible fin 2013 à Météo-France. Des évolutions du code ont eu lieu depuis, avec entre autres des développements pour le 4D-En-Var à Météo-France pour ARPEGE et un 3D-Var pour AROME. Vouloir implémenter la VRF dans OOPS est très stimulant : l'environnement OOPS est encore très jeune et non adapté par défaut à l'algorithme de la VRF. Ainsi, introduire de nouvelle problématique dans OOPS contribue à enrichir la réflexion sur l'environnement et son développement. Un autre aspect séduisant quant aux développements dans OOPS est que le travail réalisé dans cet environnement pourra être utilisé à la fois pour ARPEGE, IFS et AROME ; voir même exporté vers d'autres laboratoires comme le CERFACS.

L'implémentation de la VRF dans OOPS est proposée en suivant les algorithmes présentés au chapitre 7. L'implémentation des deux cas de figure, en utilisant le modèle adjoint ou en utilisant un ensemble, peut être réalisée dans une unique classe `RunVRF`. Le traitement doit être différent suivant le paramètre `<VRF_algo>` du fichier XML de configuration. La valeur `adjoint` est associée à une implémentation utilisant le modèle adjoint et l'assimilation variationnelle. La valeur `ensemble` est associée à l'implémentation à partir des ensembles de prévision et d'assimilation. Comme les éléments nécessaires aux calculs sont différents pour ces deux implémentations, la description ne se limite pas uniquement à la classe `RunVRF` mais à l'ensemble des étapes de la simulation.

Les différentes étapes pour chaque implémentation sont rassemblées dans un script qui exécute les binaires au fur et à mesure. La structure du script et les éléments spécifiques des fichiers de configuration XML utilisés pour chaque implémentation sont décrits dans les sous-sections suivantes. Les fichiers XML ayant déjà été illustrés dans la section 8.3.2, seuls les éléments principaux pour chaque étape sont repris ci-après.

### 8.4.1 Cadre des simulations

Les deux scripts se placent dans le cadre d'une OSSE où l'état vrai, les observations et les ébauches sont construites directement lors de la simulation. Dans un contexte opérationnel, il suffirait d'exécuter le binaire correspondant à la VRF en indiquant les données d'entrée issues de la chaîne opérationnelle dans le fichier de configuration XML. Les implémentations sont illustrées avec les exemples de fichiers binaires et de fichier de configuration XML du modèle QG.

Pour la simulation les paramètres suivants sont choisis,

- Date de l'état de référence : $t_{obs}$=`2014-01-19T00:00:00Z`,
- Système d'assimilation : 4D-Var,

- Longueur de la fenêtre d'assimilation : de $t_{obs}$ jusqu'à `2014-01-19T03:00:00Z` avec des observations toutes les heures,
- Paramètre observé : uniquement le vent (`Wind`),
- *Lead-time* : 48h soit $t_{verif}$ =`2014-01-21T00:00:00Z`,
- Nombre de membres de l'ensemble : 20.

La création de l'état de référence et des observations est identique pour la simulation déterministe ou la simulation utilisant l'ensemble. Elles sont donc d'abord présentées avant de décrire les particularités de chaque implémentation.

### Création d'un état de référence

Comme la simulation est réalisée dans un contexte OSSE, l'état de référence est créé artificiellement. L'élément principal du fichier `truth.xml` est défini dans la balise `initial` :

```
<initial>
 <date>2014-01-19T00:00:00Z</date>
 <read_from_file>0</read_from_file>
 <top_layer_depth>6000.0</top_layer_depth>
 <bottom_layer_depth>4000.0</bottom_layer_depth>
 <perturb>0</perturb>
</initial>
```

La valeur `read_from_file` à 0 indique que l'objet `QgState` sera crée lors de la simulation et non lu à partir d'un fichier comme c'était le cas sur la figure 8.5 où la balise `initial` contient l'argument `filename`.

Il n'est pas nécessaire d'effectuer l'intégration temporelle de l'état de référence à ce niveau. En effet la classe `MakeObs`, qui sera utilisée pour générer les observations, effectue elle même le calcul de la trajectoire sur l'ensemble de la fenêtre d'assimilation. Les balises `forecast` et `output` doivent tout de même être renseignées pour disposer des fonctions d'écriture des fichiers de sortie.

```
<forecast>
  <begin>2014-01-19T00:00:00Z</begin>
  <end>2014-01-19T01:00:00Z</end>
</forecast>

<output>
  <frequency>PT1H</frequency>
  <datadir>Data</datadir>
  <exp>truth</exp>
  <date>2014-01-19T03:00:00Z</date>
  <type>fc</type>
</output>
```

Cette configuration va générer deux fichiers de sortie `Data/truth.fc.2014-01-19T00:00:00Z.$time`, avec `$time` valant `PT0S` (pour l'instant initial) et `PT1H` (pour la prévision à $1h$ d'échéance). Ce second fichier n'est pas utilisé.

Le code est exécuté par le binaire `qg_forecast.x`,

```
qg_forecast.x truth.xml > truth.fc.out
```

### Création des observations

La trajectoire de référence décrite au dessus est récupérée pour créer les observations,

```
<initial>
   <filename>Data/truth.fc.2014-01-19T00:00:00Z.PT0S</filename>
   <date>2014-01-19T00:00:00Z</date>
</initial>
```

La longueur de la fenêtre d'assimilation, la fréquence et la répartition des observations sont ensuite spécifiées dans le XML,

```
<assimilation_window>
  <begin>2014-01-19T00:00:00Z</begin>
  <end>2014-01-19T03:00:00Z</end>
</assimilation_window>

<observer>
  <frequency>PT1H</frequency>
</observer>

<Observations>
  <Observation>
    <ObsType>Wind</ObsType>
    <ObsOperator>OnTheFly</ObsOperator>
    <ObsData>
      <ObsHandler>ObsHelpQG</ObsHandler>
      <ObsFileOut>Data/qg.truth4d.obt</ObsFileOut>
      <obserror>ObsErr</obserror>
    </ObsData>
    <Generate>
      <begin>PT0S</begin>
      <obs_period>PT1H</obs_period>
      <obs_density>100</obs_density>
      <nx>40</nx>
      <ny>20</ny>
    </Generate>
    <Covariance>diagonal</Covariance>
  </Observation>
<Observations>
```

Cette configuration va générer un fichier de données `Data/qg.truth4d.obt` contenant l'ensemble des observations. L'exécution de la commande dans le script est,

```
qg_makeobs.x makeobs4d.xml > qg.makeobs4d.out
```

Pour la suite, le contenu du script va dépendre de l'implémentation retenue (déterministe ou via un ensemble).

### 8.4.2 Avec un adjoint et une assimilation variationnelle

Les différentes étapes pour le calcul en utilisant une assimilation variationnelle et le modèle adjoint sont décrites ci-dessous,

1. Création de l'état de référence (décrit précédemment),
2. Création des observations à partir de l'état de référence (décrit précédemment),
3. Création de l'état initial,

4. Analyse variationnelle en utilisant l'algorithme Lanczos pour obtenir une estimation de $\mathbf{A}$,
5. Calcul de la trajectoire jusqu'à $T_{verif}$, du gradient de la fonction score à $T_{verif}$ $\nabla_{\mathbf{x}^f} S$ et du gradient de la fonction score à $T_{obs}$ via l'adjoint $\nabla_{\mathbf{x}^a} S$,
6. Estimation de la variance d'erreur d'analyse,
7. Calcul du champ de sensibilité $VRF$ à partir de $\boldsymbol{s} = \mathbf{A}\nabla_{\mathbf{x}^a} S$.

Le script déroulant les différents binaires pour réaliser la simulation est de la forme,

```
# Création de l'état de référence
qg_forecast.x truth.xml > truth.fc.out

# Génération des observations 4D à partir de l'état de référence
qg_makeobs.x makeobs4d.xml > qg.makeobs4d.out

# Création de la condition initiale
qg_forecast.x state.xml > state.fc.out

# Assimilation de données
qg_4dvar.x 4dvar.xml > qg.4dvar.out

# Calcul de la VRF
qg_VRF.x VRF_adjoint.xml > VRF.out
```

La description de la configuration de chaque état est réalisée ci-dessous.

### Création de l'état initial

Le fichier `invent_state.f90` qui est utilisé pour générer l'état initial est construit de manière à produire toujours le même état de référence, sauf si le paramètre `perturb` est différent de 0. La condition initiale n'est donc pas directement reprise à partir du fichier `truth` mais construite de la même manière que l'état de référence.

```
<initial>
  <date>2014-01-19T00:00:00Z</date>
  <read_from_file>0</read_from_file>
  <top_layer_depth>   6000.0</top_layer_depth>
  <bottom_layer_depth>4000.0</bottom_layer_depth>
  <perturb>20</perturb>
</initial>
```

L'assimilation 4D-Var et la VRF effectuent eux-mêmes l'intégration temporelle dont ils ont besoin. Comme pour l'état de référence, les balises `forecast` et `output` sont renseignées pour disposer des fonctions d'écriture des fichiers de sortie.

```
<forecast>
  <begin>2014-01-19T00:00:00Z</begin>
  <end>2014-01-19T01:00:00Z</end>
</forecast>

<output>
  <frequency>PT1H</frequency>
  <datadir>Data</datadir>
  <exp>state</exp>
```

```
    <date>2014-01-19T03:00:00Z</date>
    <type>fc</type>
  </output>
```

Le fichier de sortie produit est `Data/state.fc.2014-01-19T00:00:00Z.PT0S`, le second fichier est encore inutile. La commande exécutée dans le script est :

```
qg_forecast.x state.xml > state.fc.out
```

### Assimilation variationnelle

Pour effectuer l'assimilation, l'état initial créé précédemment est défini comme l'ébauche (`background`) et le fichier d'observation commun est spécifié en `ObsFileIn`.

```
  <background>
    <states>
     <state>
      <filename>Data/state.fc.2014-01-19T00:00:00Z.PT0S</filename>
      <date>2014-01-19T03:00:00Z</date>
     </state>
    </states>
  </background>

  <Observations>
    <Observation>
      <ObsType>Wind</ObsType>
      <ObsOperator>OnTheFly</ObsOperator>
      <ObsData>
        <ObsHandler>ObsHelpQG</ObsHandler>
        <ObsFileIn>Data/qg.truth4d.obt</ObsFileIn>
        <ObsFileOut>Data/qg.4dvar.obt</ObsFileOut>
        <obsvalue>ObsVal</obsvalue>
        <obserror>ObsErr</obserror>
      </ObsData>
      <Covariance>diagonal</Covariance>
    </Observation>
  <Observations>
```

Les configurations relatives à la variable de contrôle, à la fonction coût et au 4D-Var ne sont décrites ici. Le seul élément spécifique pour le calcul de la VRF est le choix du minimiseur qui doit être l'algorithme de Lanczos :

```
  <minimizer>
    <algorithm>DPLanczos</algorithm>
  </minimizer>
```

C'est un élément important pour le calcul de la VRF, Cependant, il est encore absent dans la version du code OOPS que nous utilisons ici. L'algorithme du Lanczos doit permettre de construire une estimation de la matrice de covariance $\mathbf{A}$ à partir des termes `alpha_` et `beta_` du code calculés lors de la minimisation, mais l'export de ces valeurs n'est pas codé. Il s'agit d'une évolution à réaliser pour que l'implémentation déterministe de la VRF soit possible.

**Calcul de la VRF via l'adjoint**

Pour effectuer le calcul de la VRF, il est nécessaire de disposer du vecteur de la condition initiale pour le calcul du gradient de la fonction score et des éléments issus de l'algorithme de Lanczos pour modéliser la matrice $\mathbf{A}$. Le choix de la fonction score doit aussi être défini dans le fichier XML.

```
<initial>
  <date>2014-01-19T00:00:00Z</date>
  <filename>Data/state.fc.2014-01-19T00:00:00Z.PT0S</filename>
</initial>
```

Le calcul de la trajectoire jusqu'à $t_{verif}$ est réalisé à ce niveau en utilisant la méthode `forecast` associée à l'objet `QgState`.

```
<forecast>
  <begin>2014-01-19T00:00:00Z</begin>
  <end>2014-01-21T00:00:00Z</end>
</forecast>
```

Il n'existe pas de classe pour le calcul d'une fonction score ou de son gradient. Dans le cas d'une implémentation pour le QG, on contourne le problème en choisissant l'enstrophie comme fonction score, car son gradient est la vorticité. La vorticité est obtenue comme le laplacien de la fonction courant (la fonction en FORTRAN est `laplacian_2d`).

Pour ramener le gradient à $t_{obs}$, on passe par la classe `QgTLM` pour le QG, qui regroupe les modèles tangent linaire et l'adjoint. L'intégration temporelle de l'adjoint se fait via le méthode `stepAD` appliquée au gradient de $S$ à $t_{verif}$.

La construction de la matrice $\mathbf{A}$ sous forme d'opérateur est décrite dans l'annexe B et le calcul de la variance d'erreur d'analyse associé dans l'annexe D. Ces calculs peuvent être codés directement en C++. Comme $\mathbf{A}$ est construit sous la forme $\boldsymbol{A} \approx \boldsymbol{B} + \sum_{k=1}^{r}(\lambda_k^{-1} - 1)(Lu_k)(Lu_k)^T$, le produit matrice vecteur entre le gradient et $\mathbf{A}$ peut donc être réalisé à partir de la fonction `dot_product` et d'additions.

```
<VRF>
   <VRF_algo>adjoint</VRF_algo>
   <ScoreFunction>enstrophy</ScoreFunction>
   <ALanczosIn>qg.4dvar.out</ALanczosIn>
</VRF>
```

Le champ de réduction de variance ainsi obtenu doit être sauvegardé dans un fichier. L'écriture des fichiers est traitée par un `PostProcessor`.

### 8.4.3 Avec un ensemble

En suivant la même logique que pour le cas du calcul déterministe, l'implémentation de la VRF via un ensemble est maintenant décrite.

- Création de l'état de référence (décrit précédemment),
- Création des observations à partir de l'état de référence (décrit précédemment),
- Génération d'un ensemble et intégration temporelle,
- Analyse pour chaque membre de l'ensemble à partir d'observations perturbées,
- Calcul de la $VRF$ (calcul de la fonction score pour chaque membre, de $\boldsymbol{s}$ et de $\sigma_c^{a2}$.

Le script déroulant les différents binaires pour réaliser la simulation est de la forme :

```
# Création de l'état de référence
qg_forecast.x truth.xml > truth.fc.out
```

```
# Génération des observations 4D à partir de l'état de référence
qg_makeobs.x makeobs4d.xml > qg.makeobs4d.out

# Génération des membres de l'ensemble et intégration temporelle
qg_genpert.x genenspert.xml > VRF.ens.out

# Assimilation variationnelle pour chaque membre de l'ensemble
for membre `seq 1 $Ne;
do
  sed "s/Ne_/${membre}/" 4dvar.xml.ori > 4dvar.xml
  qg_4dvar.x 4dvar.xml > qg.4dvar.out
  mv Data/state.an.2014-01-19T00:00:00Z Data/state.an.${membre}.2014-01-19T00:00:00Z
  rm -f 4dvar.xml
done

# Calcul de la VRF
qg_VRF.x VRF_ensemble.xml > VRF.out
```

La description de la configuration de chaque étape est réalisée ci-dessous.

### Génération de l'ensemble et intégration temporelle

La génération de l'ensemble est réalisée à travers la classe `GenEnsPertB`. Les différents membres de l'ensemble sont créés à partir de l'état initial et de la matrice de covariance d'erreur d'ébauche $\mathbf{B}$ spécifiée dans le fichier XML. Pour chaque membre, la trajectoire est calculée jusqu'à $t_{verif}$.

```
   <initial>
    <filename>Data/state.fc.2014-01-19T00:00:00Z.PT0S</filename>
    <date>2014-01-19T00:00:00Z</date>
  </initial>

  <covariance>
    <covariance_type>QgError</covariance_type>
    <standard_deviation>0.8</standard_deviation>
    <vertical_correlation>0.2</vertical_correlation>
    <horizontal_length_scale>1e6</horizontal_length_scale>
    <maximum_condition_number>1.0e6</maximum_condition_number>
    <date>2014-01-19T00:00:00Z</date>
  </covariance>

  <members>20</members>

  <forecast>
    <begin>2014-01-19T00:00:00Z</begin>
    <end>2014-01-21T00:00:00Z</end>
  </forecast>
```

Le binaire produit deux fichiers de sortie par membre de la forme `state.ens.$Ne.2014-01-19T00:00:00Z.$ech`, avec `$Ne` le numéro de membre et `$ech` valant `PT0S` ou `P2D`.

```
qg_genpert.x genenspert.xml > VRF.ens.out
```

### Assimilation d'ensemble

Seule l'assimilation variationnelle est disponible dans la version de OOPS avec laquelle nous avons travaillé. Une assimilation variationnelle doit donc être réalisée pour chaque membre de l'ensemble à partir des observations perturbées. Pour que le binaire `qg_4dvar.x` soit exécuté pour chaque membre de l'ensemble, il est nécessaire d'adapter le fichier de configuration XML via la commande linux *sed* à chaque appel du binaire. L'assimilation produit un fichier `state.an.2014-01-19T00:00:00Z` qu'il est nécessaire de renommer.

```
for membre ‘seq 1 $Ne;
do
   sed "s/Ne_/${membre}/" 4dvar.xml.ori > 4dvar.xml
   qg_4dvar.x 4dvar.xml > qg.4dvar.out
   mv Data/state.an.2014-01-19T00:00:00Z Data/state.an.${membre}.2014-01-19T00:00:00Z
   rm -f 4dvar.xml
done
```

La seule modification par rapport au fichier `4dvar.xml` utilisé précédemment est la ligne `<filename>Data/state.ens.Ne_.2014-01-19T00:00:00Z.PT0S</filename>` du `background` pour permettre à la commande *sed* de remplacer `Ne_` par la bonne valeur.

### VRF d'ensemble

Les données d'entrée pour le calcul de la VRF via l'ensemble sont les fichiers de l'ensemble de prévision et d'assimilation et le choix de la fonction score. Pour que le calcul de l'espérance soit correct, il est nécessaire que les indices des membres de l'ensemble de prévision correspondent avec ceux de l'ensemble d'analyse.

```
<VRF>
  <VRF_algo>ensemble</VRF_algo>
  <members>20</members>
  <VRFForecastEnsemble>
    <members>20</members>
    <filename>Data/state.ens.1.2014-01-21T00:00:00Z</filename>
    ...
    <filename>Data/state.ens.20.2014-01-21T00:00:00Z</filename>
  </VRFForecastEnsemble>
  <VRFAnalysisEnsemble>
    <members>20</members>
    <filename>Data/state.an.1.2014-01-19T00:00:00Z</filename>
    ...
    <filename>Data/state.an.20.2014-01-19T00:00:00Z</filename>
  </VRFAnalysisEnsemble>
  <ScoreFunction>enstrophy</ScoreFunction>
  <VarR>0.6</VarR>
</VRF>
```

Deux `Ensembles` sont utilisés pour contenir les valeurs des prévisions d'ensemble (balise `VRFForecastEnsemble`) et des analyses (balise `VRFAnalysisEnsemble`). Comme nous l'avons vu avec l'implémentation basée sur un adjoint, il n'existe pas de fonction native pour le calcul de la valeur d'une fonction score. A partir de la valeur de la fonction score de chaque membre, on en déduit la moyenne et les perturbations $dS_i$ (à sauvegarder dans un tableau C++).

Le calcul des $\varepsilon_i^a$ peut aussi être réalisé directement en C++. Le calcul de la variance de l'ensemble à partir des $\varepsilon_i^a$ permet de connaître $\sigma^a$ en chaque point de grille. Le champ de sensibilité est obtenu comme $\boldsymbol{s} = 1/N_e \sum_1^{N_e} dS_i \varepsilon_i^a$.

Il suffit ensuite d'inverser $\sigma^{a2} + \sigma_t^{o2}$ (cf. équation 7.2, où $\sigma_t^{o2}$ est la variance liée à l'observation supplémentaire) et d'utiliser la fonction `dot_product` pour réaliser les produits d'Hadamard. Comme pour l'implémentation avec l'adjoint, on sauvegarde le champ de réduction de variance dans un fichier.

Une grande partie des fonctionnalités nécessaires à ce calcul sont déjà disponibles dans la version d'OOPS utilisée ici. Une ébauche de code C++ pour cette implémentation est présentée dans l'annexe I.

## 8.5 Commentaires sur l'implémentation dans OOPS

OOPS est un environnement de développement orienté objet développé au CEPMMT et à Météo-France pour l'assimilation de données. L'objectif initial d'OOPS est de proposer un cadre permettant de simplifier la construction des algorithmes d'assimilation de données en travaillant sur une organisation de haut niveau. La VRF n'est pas directement un algorithme d'assimilation de données, aussi le développement de la VRF dans OOPS est particulièrement stimulant puisqu'il contribue à ouvrir le champ d'application initial d'OOPS et participe ainsi à la maturation de l'outil.

Dans ce chapitre, les composants intervenant dans l'environnement OOPS ont été décrits pour une version du code datant de 2013. Ainsi la description qui a été faite, ne reflète pas parfaitement les caractéristiques de la dernière version de l'environnement mais permet cependant d'en présenter les principaux ingrédients, tels les différents niveaux du code entre le C++ et le FORTRAN. Le choix d'une granularité fine dans la description de OOPS s'explique par la nécessité que nous avons ressentie de bien comprendre l'ensemble des interactions au sein du code afin de pouvoir y implémenter la VRF

Une mise en œuvre concrète de la VRF dans le cadre d'OOPS n'a pas pu être réalisée. Cependant l'analyse des éléments nécessaires au codage a été faite, et ce pour les deux méthodes (déterministe et ensembliste). Outre le codage de la classe `RunVRF` en elle-même, plusieurs fonctionnalités seront à développer dans OOPS. Par exemple, la définition d'une fonction score, le calcul de sa valeur en un point et le calcul de son gradient. Pour la formulation déterministe, utilisant l'adjoint et l'assimilation variationnelle, l'algorithme de Lanczos est disponible pour la minimisation mais il faudra développer une manière de sauvegarder les facteurs lors du calcul. Ceux-ci sont nécessaires pour l'estimation de la matrice $\mathbf{A}$ sous forme d'un opérateur. Pour le développement de la méthode d'ensemble, il serait intéressant de développer une méthode d'assimilation non incrémentale, telle que l'ETKF par exemple. L'implémentation proposée avec une assimilation variationnelle pour chaque membre de l'ensemble n'est sûrement pas très optimisée pour les coûts de calcul, mais ceci pourra constituer un point à développer dans la réalisation pratique.

Ainsi, la réflexion sur la faisabilité de la VRF dans OOPS a permis d'identifier des axes de développement contribuant à l'évolution et l'enrichissement des composants sous OOPS. Une fois l'implémentation de ces composants réalisés, le code de la VRF en lui-même sera entièrement implémenté en C++. Il sera alors possible d'appliquer l'algorithme de la VRF à n'importe quel modèle de PNT (ARPEGE, AROME).

# Chapitre 9

# Conclusion et perspectives

Ce chapitre résume les différents éléments de discussions et de conclusions abordés dans les chapitres précédents.

## 9.1 Retour sur la problématique de la thèse

L'application au cadre pronostique de la sensibilité de la prévision aux observations a permis de diversifier les méthodes d'observation adaptative de seconde génération. Ces approches en sensibilité sont très dépendantes des développements liés au modèle adjoint (prévision et assimilation). Cette génération de méthodes d'observation adaptative (HRR, KFS et ETKF) a été développée suite à la campagne FASTEX et certaines ont pu être implémentées dans le cadre de campagnes de terrain sous l'égide du programme WWRP Thorpex (2003-2013, dont les campagnes NA-TreC, T-PARC, Medex, Majumdar and Co-authors (2014)). En parallèle, l'approche en sensibilité a permis d'importantes avancées pour les études d'impact des observations comme outil de diagnostic du système d'assimilation variationnel. Ces développements théoriques ont permis d'obtenir des expressions au second ordre par rapport aux perturbations à l'instant d'analyse et ainsi de réduire les erreurs d'estimation liées aux hypothèses de transport linéaire par le modèle. Une revue des différentes méthodes de sensibilité aux observations (diagnostiques et pronostiques) est réalisée au chapitre 4.

En parallèle des développements propres à l'assimilation variationnelle déterministe, les approches basées sur les ensembles ont connu un fort essor ces dernières années. La prévision d'ensemble apporte un éclairage nouveau sur les notions de prévisibilité et le caractère déterministe de la PNT. L'utilisation des ensembles permet aussi de documenter les matrices de covariance d'erreur d'ébauche $\mathbf{B}$ *du jour*, et construire des ensembles d'analyse prenant en compte toutes incertitudes associées à l'analyse. Les travaux récents sur le 4D-En-Var permettent de faire converger les méthodes basées sur le filtre de Kalman d'ensemble et l'approche variationnelle de l'assimilation de données.

Les méthodes de sensibilité aux observations n'ont cependant pas profité de ces évolutions. C'est dans ce contexte que se déroule ce travail de thèse. L'objectif de ce manuscrit est de répondre aux deux questions suivantes :

- Quelle est la part d'incertitude liée à la trajectoire de référence dans le calcul de la sensibilité aux observations ?
- Quelle information supplémentaire peut on tirer d'un ensemble ?

Pour répondre à cette double problématique, un état des lieux des méthodes de sensibilité aux observations a été réalisé, à la fois en ce qui concerne les méthodes d'observations adaptatives mais aussi de calcul d'impact d'observation. Ces dernières ont en effet bénéficié d'un fort intérêt ces dernières années. L'étude montre que l'ensemble des avancées obtenues dans le cadre des études d'impact d'observations ne sont pas transposables directement à l'observation adaptative. En particulier, les différents développements au second ordre, qui ont sont considérés comme essentiels pour le calcul de l'impact des observations, nécessi-

tent de connaître des valeurs qui sont inconnues pour l'observation adaptative. L'analyse à $t_{verif}$ est prise comme la vérité pour le calcul de l'erreur de prévision, et l'analyse à $t_{obs}$ elle même est utilisée. Les observations fictives utilisées en observation adaptative ne permettent pas de manipuler l'analyse directement, mais seulement l'erreur associée.

## 9.2 Analyse de la sensibilité de la KFS à trajectoire de référence

L'objet initial du travail de thèse était de documenter l'incertitude du champ de sensibilité de la KFS à la trajectoire et aux composants du système d'assimilation de données. Cette étude a été réalisée en utilisant un ensemble de prévisions dans le cadre idéalisé du modèle Quasi-Géostrophique à deux couches, et d'un système d'assimilation 4D-Var. Pour chaque expériences présentées au chapitre 6, un ensemble de prévisions est construit et un champ de sensibilité est calculé pour chaque membre de l'ensemble.

L'expérience I a permis de mettre en évidence la dispersion des solutions proposées par la KFS en fonction de la trajectoire retenue. L'influence de la trajectoire est d'autant plus faible sur la forme des champs de sensibilité que les contraintes du système d'assimilation sont fortes. Ces résultats peuvent s'appliquer à n'importe quelle méthode de sensibilité aux observations mettant en œuvre un transport de l'information par le modèle adjoint. Les méthodes d'observation adaptative sont d'autant plus sensibles à cette incertitude qu'elles reposent sur des hypothèses fortes pour la prévision du réseau de routine.

L'expérience II a été conçue pour évaluer l'intérêt d'utiliser un ensemble de prévision pour limiter l'impact de cette incertitude sur la trajectoire de référence. La stratégie proposée est de considérer le déploiement associé au maximum du champ moyen de sensibilité. Cette stratégie appelée *observation consensus* se montre plus efficace pour réduire la variance d'erreur à $t_{verif}$ que l'assimilation pour chaque membre d'une observation au maximum de son propre champ de sensibilité. Cependant cette méthode n'est pas applicable en pratique puisque le calcul d'un ensemble de champ de sensibilité déterministe, et donc une assimilation variationnelle pour chaque membre de l'ensemble, n'est pas réaliste vu les coûts de calcul induits.

Ces expériences mettent en évidence l'intérêt de développer une méthode de calcul de sensibilité reposant sur des techniques uniquement ensemblistes.

## 9.3 Sensibilité aux observations à l'aide d'un ensemble

Le cadre théorique de la KFS ne permet pas de répondre à cette seconde problématique de la thèse. C'est pourquoi une nouvelle méthode, appelée *Variance Reduction Field* (VRF) a été présentée au chapitre 4. Cette méthode présente la particularité de pouvoir être implémentée soit dans un système de prévision déterministe, avec un modèle adjoint et une assimilation variationnelle, soit dans un système utilisant un ensemble de prévisions et d'assimilations. Ces deux implémentations sont illustrées au chapitre 7 dans un modèle de Lorenz 96. Ces expériences ont montrées que dans un contexte idéalisé, le calcul de la réduction de variance à travers les deux implémentations de la VRF sont très proches du calcul explicite. On observe aussi que l'utilisation d'un ensemble semble plus robuste que le transport linéaire lorsque le temps d'optimisation augmente. Cependant l'analyse de la dynamique du modèle est limitée par la nature du modèle de Lorenz.

Bien que la VRF n'est pas pu être validée dans un contexte opérationnel dans le cadre de cette thèse, les résultats obtenus en parallèle avec l'EFSO par Hotta et al. (2014) sont très encourageant pour l'utilisation d'un ensemble pour effectuer des calculs de sensibilité aux observations. De plus, le dernier chapitre de ce manuscrit, consacré à la présentation de l'implémentation de la VRF dans l'environnement OOPS, illustre bien la facilité à implémenter cette méthode dans un système d'assimilation et de prévision d'ensemble. Pour un système d'assimilation incrémentale, l'implémentation est plus complexe avec la nécessité de disposer de l'algorithme de Lanczos pour la minimisation et du modèle adjoint pour le transport du gradient de la fonction score. Par cette étude préparatoire d'implémentation, le travail réalisé contribue

donc au développement d'OOPS, puisque elle en élargit le domaine d'application essentiellement orienté vers l'algorithmie de l'assimilation.

## 9.4 Perspectives de travail à la suite de cette thèse

Cette thèse a aussi été l'occasion de remettre en question la finalité de l'observation adaptative. Les objectifs initiaux très ambitieux de déploiement d'observations mobiles hérités de la campagne FASTEX ne sont encore envisagés que dans le cadre des campagnes hivernales du NCEP et NOAA (Winter Storm Reconnaissance Program) et pour la prévision des cyclones tropicaux. Avec l'augmentation très importante de la couverture et du nombre des observations satellites, la question n'est plus tant de disposer d'une observation au niveau de la zone sensible mais de faire le meilleur usage de celles disponibles. Dans les campagnes de DTS (comme la campagne de mesure HyMeX) les observations supplémentaires sont directement issues du réseau de radiosondage de routine, avec des lâchers supplémentaires en dehors des heures principales. L'évaluation de l'impact d'un tel système adaptatif très contraint révèle une amélioration très faible des prévisions. Une option envisagée mais non appliquée lors de la campagne est aussi d'augmenter la densité des observations de satellites défilants a proximité de la zone de sensibilité. Cela rejoint les résultats de l'EFSO concernant la détection des observations néfastes pour les écarter de l'analyse. La sensibilité aux observations ne doit plus être seulement considérée pour effectuer un tri des observations mais d'avantage comme un outil de paramétrisation du système d'assimilation. Les méthodes de calcul d'impacts d'observation permettant de configurer les matrices $\mathbf{R}$ et $\mathbf{B}$, les méthodes d'observation adaptative permettant d'adapter l'intégration des observations à travers la matrice $\mathbf{H}$. Avec l'augmentation de la fréquence des cycles d'assimilation (le modèle AROME est assimilé toutes les heures), la nécessité de distinguer les méthodes diagnostiques et pronostiques est de moins en moins présente.

Dans le cadre de la thèse, un diagnostic synthétique de la réduction de variance associée à l'assimilation d'une observation a été introduit au chapitre 3. Son calcul repose sur la description locale de la fonction de corrélation d'erreur de prévision au voisinage de l'observation. En pratique, la caractérisation locale passe par le calcul du tenseur de métrique (généralisation de la notion de longueur de portée) estimé à l'aide des méthodes d'ensemble. L'avantage de ce nouveau diagnostic est de pouvoir estimer l'intérêt d'une observation, sans réaliser d'assimilation, et donc de sélectionner les observations importantes pour le système de prévision (étape de screening dans l'assimilation de observations). Il serait intéressant de tenter l'implémentation de ce diagnostic dans un système de prévision quasi-opérationnel afin d'évaluer son bénéfice. Cet axe pourra être suivi dans le prolongement des travaux décris ici.

Les grandes avancées en PNT des dernières années sont en partie dues à l'amélioration des systèmes d'assimilation de données (mise en place du 4D-Var, assimilation de nouveaux satellites comme Metop ou Megha-Tropiques) et à l'augmentation de la résolution des modèles qui suit l'augmentation de la puissance de calcul. Les nouvelles approches comme le 4D-En-Var permettent d'avancer dans la représentation des matrices de covariance d'erreur d'ébauche du jour, qui est un point faible du 4D-Var classique. En parallèle il permet aussi de simplifier la maintenance du code avec la possibilité de disposer d'une assimilation variationnelle sans la nécessité de conserver un modèle tangent linéaire et un adjoint.

Les observations restent encore cependant à la marge de ces évolutions. La prise en compte des termes non diagonaux dans la matrice d'erreur d'observation est encore un sujet peu documenté. Et les critères de sélection des observations pour l'assimilation restent relativement empiriques. Des méthodes comme la VRF ou l'EFSO permettent d'ajouter des critères objectifs à cette sélection en fonction de l'impact attendu des observations. Dans un soucis toujours constant d'optimiser les coûts de calcul, l'efficience est peut être maintenant à chercher dans la rationalisation de cette sélection.

# Bibliographie


B. Ancell and G. J. Hakim. Comparing Adjoint- and Ensemble-Sensitivity Analysis with Applications to Observation Targeting. *Mon. Wea. Rev.*, 135 :4117–+, 2007.

R. Asselin. Frequency filter for time integrations. *Mon. Wea. Rev.*, 100 :487–490, 1972.

H. Avron and S. Toledo. Randomized algorithms for estimating the trace of an implicit symmetric positive semi-definite matrix. *J. ACM*, 58 :8 :1–8 :34, 2011.

M. Belo Pereira and L. Berre. The use of an ensemble approach to study the background error covariance in a global nwp model. *Mon. Wea. Rev.*, 134 :2466–2489, 2006.

Th. Bergot. Adaptive observations during fastex : A systematic survey of upstream flights. *Quart. J. Roy. Meteor. Soc.*, 125 :3271–3298, 1999.

Th. Bergot and A. Doerenbecher. A study on the optimization of the deployment of targeted observations using adjoint-based methods. *Quart. J. Roy. Meteor. Soc.*, 128 :1689–1712, July 2002.

L. Berre, O. Pannekoucke, G. Desroziers, S. Stefanescu, B. Chapnik, and L. Raynaud. A variational assimilation ensemble and the spatial filtering of its error covariances : increase of sample size by local spatial averaging. In ECMWF, editor, *ECMWF Workshop on Flow-dependent aspecyts of data assimilation, 11-13 June 2007.*, pages 151–168, Reading, UK, 11-13 June 2007.

C. H. Bishop and Z. Toth. Ensemble transformation and adaptive observations. *J. Atmos. Sci.*, 56 : 1748–1765, 1999.

C. H. Bishop, B. J. Etherton, and S. J. Majumdar. Adaptive sampling with the ensemble transform kalman filter. part i : Theoretical aspects. *Mon. Wea. Rev.*, 129(3) :420–436, 2001.

E. Blayo and M. Nodet. *Introduction à l'assimilation de données variationnelle*. Université de Grenoble, 2009.

N. Boullot, F. Rabier, R. Langland, R. Gelaro, C. Cardinali, V. Guidard, P. Bauer, and A. Doerenbecher. Observation impact over the southern polar area during the concordiasi field campaign. *Quart. J. Roy. Meteor. Soc.*, 2015.

F. Bouttier, B. Vié, O. Nuissier, and L. Raynaud. Impact of stochastic physics in a convection-permitting ensemble. *Mon. Wea. Rev.*, 140 :3706–3721, 2012.

M. Buehner and M. Charron. Spectral and spatial localization of background error correlations for data assimilation. *Quart. J. Roy. Meteor. Soc.*, 133 :615–630, 2007.

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

G. Burger, P. J. Van Leewen, and G. Evensen. Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, 126 :1719–1724, 1998.

J. Campins, B. Navascués, C. Santos, and A. Amo-Baladrón. Influence of targeted observations on short-term forecasts of high-impact weather events in the mediterranean. *Nat Hazards Earth Syst. Sci.*, 13 : 2891–2910, 2013.

C. Cardinali. Monitoring the observation impact on the short-range forecast. *Quart. J. Roy. Meteor. Soc.*, 135 :239–250, 2009.

M. Clerc and S. Mallat. Estimating deformations of stationary processes. *Annals Statistics*, 31 :1772–1821, 2003.

Ph. Courtier and J.-F. Geleyn. A global numerical weather prediction model with variable resolution : Application to the shallow-water equations. *Quart. J. Roy. Meteor. Soc.*, 114 :1321–1346, 1988.

Ph. Courtier, C. Freydier, J.F. Geleyn, F. Rabier, and M. Rochas. The arpege project at météo-france. In *ECMWF Seminar Proceedings, 9–13 sept. 1991*, pages 193–231. ECMWF, 1991.

Ph. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.*, 120 :1367–1387, July 1994.

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

D. N. Daescu. On the Sensitivity Equations of Four-Dimensional Variational (4D-Var) Data Assimilation. *Mon. Wea. Rev.*, 136 :3050–+, 2008.

D. N. Daescu and R. Todling. Adjoint estimation of the variation in model functional output due to the assimilation of data. *Mon. Wea. Rev.*, 137 :1705–1716, 2009.

R. Daley. *Atmospheric data analysis*. 1991.

P. Del Moral. *Feynman-Kac Formulae : Genealogical and Interacting Particle Systems With Applications*. Probability and Its Applications. Springer-Verlag New York Inc., 2004.

L. Descamps, C. Labadie, A. Joly, E. Bazile, Ph. Arbogast, and P. Cebron. Pearp, the météo-france short-range ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, 2014.

G. Desroziers. A coordinate change for data assimilation in spherical geometry of frontal structures. *Mon. Wea. Rev.*, 125 :3030–3038, 1997.

G. Desroziers, G. Hello, and J.N. Thépaut. A 4d-var re-analysis of the fastex experiment. *Quart. J. Roy. Meteor. Soc.*, 129 :1301–1316, 2003.

G. Desroziers, J.-T. Camino, and L. Berre. 4denvar : link with 4d state formulation of variational assimilation and different possible implementations. *Quart. J. Roy. Meteor. Soc.*, 140 :2097–2110, 2014.

G. Desroziers, E. Arbogast, and L. Berre. Development of 4denvar at météo-france. In *Theoretical aspects of ensemble data assimilation for the Earth system, Les Houches, April 5-10 2015*, pages 0–0. Ecole de physique des Houches, 2015.

A. Doerenbecher and Th. Bergot. Sensitivity to observations applied to fastex cases. *Nonlinear Processes in Geophysics*, 8 :467–481, 2001.

V. Ducrocq, Braud I., Davolio S., Ferretti R., Flamant C., Jansa A., Kalthoff N., Richard E., Taupier-Letage I., Ayral P.-A., Belamari S., Berne A., Borga M., Boudevillain B., Bock O., Boichard J.-L., Bouin M.-N., Bousquet O., Bouvier Ch., Chiggiato J., Cimini D., Corsmeier U., Coppola L., Cocquerez Ph., Defer E., Delanoë J., Di Girolamo P., Doerenbecher A., Drobinski Ph., Dufournet Y., Fourrié N.,

Gourley J., Labatut L., Lambert D., Le Coz J., Marzano F. S., Molinié G., Montani A., Nord G., Nuret M., Ramage K., Rison W., Roussot O., Said F., Schwarzenboeck A., Testor P., Van Baelen J., Vincendon B., Aran M., and Tamayo J. Hymex-sop1 : The field campaign dedicated to heavy precipitation and flash flooding in the northwestern mediterranean. *Bull. Amer. Meteor. Soc.*, 95 :1083–1100, 2014.

E. S. Epstein. Stochastic dynamic prediction. *Tellus A*, 16 :739–759, 1969.

R. M. Errico. What Is an Adjoint Model ? *Bull. Amer. Meteor. Soc.*, 78 :2577–2591, November 1997.

R. M. Errico. Interpretations of an adjoint-derived observational impact measure. *Tellus A*, 59A :273–276, 2007.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error. *J. Geophys. Res.*, 99 :10 143–10 162, 1994.

G. Evensen. The ensemble kalman filter : theoretical formaulation and practical implementation. *Ocean Dynamics*, 53 :343–367, 2003.

C. Fischer. An overview and illustration of sensitivity of forecasts with respect to initial conditions. *Tellus A*, 66, 2014.

M. Fisher. Background error covariance modelling. In *Seminar on Recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003*, pages 45–64. ECMWF, 2003.

N. Fourrié, D. Marchal, F. Rabier, B. Chapnik, and G. Desroziers. Impact study of the 2003 North Atlantic THORPEX Regional Campaign. *Quart. J. Roy. Meteor. Soc.*, 132 :275–295, 2006.

G. Gaspari and S. Cohn. Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, 125 :723–757, 1999.

R. Gelaro and Y. Zhu. Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus A*, 61 :179–193, March 2009.

J.-B. Gilet, M. Plu, and G. Rivière. Nonlinear Baroclinic Dynamics of Surface Cyclones Crossing a Zonal Jet. *J. Atmos. Sci.*, 66 :3021–+, 2009.

N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. *Radar Signal Processing, IEE Proceedings F*, 140 :107–113, 1993.

S. Gratton, M. Rincon-Camacho, and Ph. L. Toint. Adaptive observations and multilevel optimization in data assimilation. Technical report, University of Namur, 2013.

T. M. Hamill, F. Yang, C. Cardinali, , and S. J. Majumdar. Impact of targeted winter storm reconnaissance dropwindsonde data on midlatitude numerical weather predictions. *Mon. Wea. Rev.*, 141 :2058–2065, 2013.

D. Hotta, E. Kalnay, Y. Ota, and T. Miyoshi. Ensemble forecast sensitivity to observations (efso) and proactive quality control. World Weather Open Science Conference (WWOSC), Montreal, Canada, 2014.

P.L. Houtekamer and H.L. Mitchell. A sequential ensemble kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, 129 :123–137, 2001.

B. R. Hunt, E. J. Kostelich, and I. Szunyogh. Efficient Data Assimilation for Spatiotemporal Chaos : a Local Ensemble Transform Kalman Filter. *Physica D*, 230 :112–126, Dec 2006.

K. Ide, Ph. Courtier, M. Ghil, and A. C. Lorenc. Unified notation for data assimilation : Operational, sequential and variational. *Meteor. Soc. Japan*, 75 :181–189, 1997.

A. Joly. Du meilleur endroit pour placer le thermomètre : L'observation adaptative. *La Météorologie*, 43 : 22–36, 2003.

A. Joly, K. A. Browning, P. Bessemoulin, G. Caniaux J.-P. Cammas, J.-P. Chalon, S. A. Clough, R. Dirks, K. A. Emanuel, L. Eymard, R. Gall, T. D. Hewson, P. H. Hildebrand, D. Jorgensen, F. Lalaurette, R. H. Langland, Y. Lemaître, P. Mascart, J. A. Moore, P. O. G. Persson, F. Roux, M. A. Shapiro, C. Snyder, Z. Toth, and R. M. Wakimoto. Overview of the field phase of the Fronts and Atlantic Storm-Track Experiment (FASTEX) project. *Quart. J. Roy. Meteor. Soc.*, 125 :3131–3164, 1999.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal Basic Engineering*, 82 :35–45, 1960.

E. Kalnay, Y. Ora, and T. Miyoshi. A simpler formulation of forecast sensitivity to observations : application to ensemble kalman filters. *Tellus A*, 64, 2012.

Th. Lagarde, A. Piacentini, and O. Thual. A new representation of data-assimilation methods : The palm flow-charting approach. *Quart. J. Roy. Meteor. Soc.*, 127 :189–207, 2001.

R. H. Langland and N. L. Baker. Estimation of observation impact using the nrl atmospheric variational data assimilation adjoint system. *Tellus A*, 56A :189–201, 2004.

F.-X. Le Dimet, I. M. Navon, and D. N. Daescu. Second-order information in data assimilation. *Mon. Wea. Rev.*, 130 :629–648, 2002.

M. Leutbecher. A reduced rank estimate of forecast error variance changes due to intermittent modifications of the observing network. *J. Atmos. Sci.*, 60(5) :729–742, 2003.

E. N. Lorenz and K. A. Emanuel. Optimal Sites for Supplementary Weather Observations : Simulation with a Small Model. *J. Atmos. Sci.*, 55 :399–414, 1998.

C. Lupu and P. Gauthier. Observability of flow-dependent structure functions for use in data assimilation. *Mon. Wea. Rev.*, 139 :713–725, 2011.

S. J. Majumdar and Co-authors. Targeted observation for improving numerical weather prediction. Rapport WWRP 15, Organisation Météorologique Mondiale (OMM), 2014.

S. J. Majumdar, C. H. Bishop, B. J. Etherton, I. Szunyogh, and Z. Toth. Can an ensemble transform kalman filter predict the reduction in forecast-error variance produced by targeted observations ? *Quart. J. Roy. Meteor. Soc.*, 127 :2803–2820, 2001.

Y. Michel. Estimating deformations of random processes for correlation modelling in a limited area model. *Quart. J. Roy. Meteor. Soc.*, 139 :534–547, 2013a.

Y. Michel. Estimating deformations of random processes for correlation modelling : methodology and the one-dimensional case. *Quart. J. Roy. Meteor. Soc.*, 139 :771–783, 2013b.

I. Mirouze and A. T. Weaver. Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Quart. J. Roy. Meteor. Soc.*, 136(651) :1421–1443, Jul 2010.

N. Oger, O. Pannekoucke, A. Doerenbecher, and Ph. Arbogast. Assessing the trajectory influence in adaptive observation Kalman filter sensitivity method. *Quart. J. Roy. Meteor. Soc.*, 138 :813–825, 2012.

N. Oger, O. Pannekoucke, A. Doerenbecher, and P. Arbogast. Forecast variance reduction field a new tool for adaptive observation : ensemble versus adjoint computation. *Mon. Wea. Rev.*, X :x–x, soumis en 2014.

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

O. Pannekoucke. Heterogeneous correlation modeling based on the wavelet diagonal assumption and on the diffusion operator. *Mon. Wea. Rev.*, 137 :2995–3012, 2009.

O. Pannekoucke and S. Massart. Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation. *Quart. J. Roy. Meteor. Soc.*, 134 :1425–1438, 2008.

O. Pannekoucke, L. Berre, and G. Desroziers. Filtering properties of wavelets for local background-error correlations. *Quart. J. Roy. Meteor. Soc.*, 133 :363–379, 2007.

O. Pannekoucke, L. Berre, and G. Desroziers. Background error correlation length-scale estimates and their sampling statistics. *Quart. J. Roy. Meteor. Soc.*, 134 :497–508, 2008.

O. Pannekoucke, E. Emili, and O. Thual. Modelling of local length-scale dynamics and isotropizing deformations. *Quart. J. Roy. Meteor. Soc.*, 140 :1387–1398, 2014.

Joseph Pedlosky. *Geophysical fluid dynamics*. Springer-Verlag, 1982.

N. A. Phillips. a Simple Three-Dimensional Model for the Study of Large-Scale Extratropical Flow Patterns. *J. Atmos. Sci.*, 8 :381–394, December 1951.

C. Prates, D. Richardson, and C. Sahin. Final report of the preview observation data targeting system (dts). Tech. Memo 581, ECMWF, 2009.

R. J. Purser, W.-S. W., D. F. Parrish, and N. M. Roberts. Numerical aspects of the application of recursive filters to variational statistical analysis. part i : Spatially homogeneous and isotropic gaussian covariances. *Mon. Wea. Rev.*, 131 :1524–1535, 2003a.

R. J. Purser, W.-S. W., D. F. Parrish, and N. M. Roberts. Numerical aspects of the application of recursive filters to variational statistical analysis. part ii : Spatially inhomogeneous and anisotropic general covariances. *Mon. Wea. Rev.*, 131 :1536–1548, 2003b.

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

L. Raynaud and O. Pannekoucke. Sampling properties and spatial filtering of ensemble background-error length-scales. *Quart. J. Roy. Meteor. Soc.*, 139 :784–794, 2013.

Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The arome-france convective scale operational model. 139 :976–991, 2011.

J. R. Shewchuk. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain*. School of Computer Science Carnegie Mellon University, 1994.

Y. Trémolet. Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus A*, 60 :964–978, October 2008.

Y. Trémolet and M. Fisher. Object-oriented prediction system overview and design principles. Technical report, ECMWF, 2012.

P. J. van Leeuwen. A variance-minimizing filter for large-scale applications. *Mon. Wea. Rev.*, 131 : 2071–2084, 2003.

H. Varella, L. Berre, and G. Desroziers. Diagnostic and impact studies of a wavelet formulation of background-error correlations in a global model. *Quart. J. Roy. Meteor. Soc.*, 137 :1369–1379, 2011.

A. T. Weaver and Ph. Courtier. Correlation modelling on the sphere using a generalized diffusion equation (tech. memo. ecmwf, num. 306). *Quart. J. Roy. Meteor. Soc.*, 127 :1815–1846, 2001.

A. T. Weaver and I. Mirouze. On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Quart. J. Roy. Meteor. Soc.*, 139(670) :242–260, Jan 2013.

P. D. Williams. A Proposed Modification to the Robert-Asselin Time Filter. *Mon. Wea. Rev.*, 137 : 2538–+, 2009.

J. Zinn-Justin. *Transitions de phase et groupe de renormalisation.* EDP SCIENCES, 2005. ISBN 2868837905.

# Annexe A

# Modélisation de B

L'opérateur $\boldsymbol{B}$ utilisé est modélisé sous la forme

$$\boldsymbol{B} = \Sigma \boldsymbol{C} \Sigma^T,$$

où $\Sigma$ est la matrice d'écart-type d'ébauche en chaque point de grille et $\boldsymbol{C}$ la matrice des corrélations (Courtier et al., 1998). $\Sigma^2$ ainsi définie correspond à la diagonale de $\boldsymbol{B}$ en point de grille et correspond à la variance. L'hypothèse d'homogénéité de la fonction de corrélation implique une décorrélation des nombres d'onde dans l'espace spectral, donc un opérateur diagonal. On peut donc écrire $\boldsymbol{C}$ sous la forme, $\boldsymbol{C} = \boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S}^{-T}$, où $\boldsymbol{S}$ est la matrice de passage de l'espace point de grille à l'espace spectral. $\boldsymbol{D}$ correspond au spectre d'énergie en spectral, c'est-à-dire à la tranformée de Fourier de la fonction de corrélation d'après le théorème de Wiener-Khinchin.

L'opérateur $\boldsymbol{B}$ en spectral se met donc sous la forme

$$\boldsymbol{B} = \boldsymbol{S}\Sigma\boldsymbol{S}^{-1}\boldsymbol{C}\boldsymbol{S}^{-T}\Sigma^T\boldsymbol{S}^T.$$

La matrice $\boldsymbol{B}$ ainsi obtenue permet de prendre en compte les zones d'incertitude de l'écoulement et un calcul explicite des corrélations sans nécessiter un stockage mémoire très important.

Cette décomposition permet de construire le préconditionnement du gradient conjugué sous la forme suivante : $\boldsymbol{B}^{1/2} = \boldsymbol{S}\Sigma\boldsymbol{S}^{-1}\boldsymbol{C}^{1/2}$.

### Construction du champ de variance

Le champ de variance est obtenu à partir d'un ensemble de conditions initiales perturbées (voir 6.1.1). La variance pour chaque point de grille est calculée à partir de cet ensemble et fournit la carte des écart-types utilisés pour la diagonale de $\Sigma$.

### Construction du spectre

La fonction de corrélation horizontale choisie est une gaussienne. Le spectre d'énergie obtenu en prenant la transformée de Fourier est donc aussi gaussien. Les corrélations entre les variables d'état sur les deux niveaux de modèles ne sont pas directement prises en compte dans $\boldsymbol{B}$.

# Annexe B

# Algorithme combiné du gradient conjugué et de Lanczos

Le formulation variationnelle de l'assimilation de données consiste à minimiser une fonction $J$. Plusieurs algorithmes de minimisation existent dans le cadre de problèmes de grandes dimensions. On utilise ici l'algorithme du gradient conjugué car il permet d'obtenir une estimation de la Hessienne de la fonction coût $\boldsymbol{B}^{-1} + \boldsymbol{H}^{\boldsymbol{T}}\boldsymbol{R}^{-\mathbf{1}}\boldsymbol{H} = \boldsymbol{A}^{-1}$, un ingrédient important pour le calcul de la KFS (comme décrit en section 4.3.2). Sa convergence peut être grandement améliorée grâce à un pré-conditionnement adéquat. Une description détaillée de l'algorithme de gradient conjugué est donnée dans Shewchuk (1994).

**Algorithme combiné du gradient conjugué et de Lanczos**

La méthode du gradient conjugué est un algorithme classique de minimisation de fonction quadratique utilisant une orthogonalisation des directions de descente à chaque itération. On veut minimiser $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}^{-\mathbf{1}}\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x}$, ce qui revient à résoudre $\nabla f(x) = \boldsymbol{A}^{-\mathbf{1}}\boldsymbol{x} + \boldsymbol{b} = 0$.

L'algorithme de Lanczos permet une estimation rang réduit des valeurs propres et vecteurs propres de la Hessienne $\boldsymbol{M}$. Cette approximation est calculée dans une base réduite constituée des $m$ vecteurs calculés par le gradient conjugué et qui sont associés à la matrice de Lanczos tridiagonale $\boldsymbol{T}_m = \boldsymbol{V}_m^T\boldsymbol{M}\boldsymbol{V}_m$, par la relation $\boldsymbol{M}\boldsymbol{v}_k = \beta_{k+1}\boldsymbol{v}_{k+1} + \delta_k\boldsymbol{v}_k + \beta_k\boldsymbol{v}_{k-1}$, où $\delta$ sont les coefficients sur la diagonale et $\beta$ les coefficients sur les sous et sur-diagonales. Il est donc possible de calculer une estimation de l'image d'un vecteur par $\boldsymbol{M}$ en le projetant dans la base réduite des vecteurs de Lanczos. La matrice obtenue dans l'espace réduit est facile à diagonaliser, et donc à inverser. En utilisant les différentes matrices de passage on peut donc obtenir est bonne estimation de l'inverse de la Hessienne.

| | |
|---|---|
| position courante au cours de la minimisation | $\boldsymbol{x}_i$ |
| gradient de la fonction coût au point $\boldsymbol{x}_i$ | $\boldsymbol{g}_i$ |
| direction de descente à l'itération $i$ | $\boldsymbol{d}_i$ |
| facteurs diagonaux de $\boldsymbol{T}$ | $\delta$ |
| facteurs sur et sous diagonaux de $\boldsymbol{T}$ | $\beta$ |

**Algorithme du gradient conjugué et de Lanczos combiné**

- Initialisation $\delta\boldsymbol{x}_0$, $\boldsymbol{d}_0 = \boldsymbol{g}_0 = -\nabla f(\delta\boldsymbol{x}_0)$
- BOUCLE
  Minimisation
  1. Calcul du pas optimal : $\alpha_k = \frac{\boldsymbol{g}_k^T\boldsymbol{g}_k}{(\boldsymbol{M}\boldsymbol{d}_k.\boldsymbol{d}_k)}$
  2. Mise à jour de la position : $\delta\boldsymbol{x}_{k+1} = \delta\boldsymbol{x}_k + \alpha_k d_k$
  3. Calcul du gradient : $\boldsymbol{g}_{k+1} = -\nabla f(\delta\boldsymbol{x}_{k+1})$
  4. Orthogonalisation de $\boldsymbol{d}_{k+1}$ : $\gamma_k = \frac{\boldsymbol{g}_{k+1}^T\boldsymbol{g}_{k+1}}{\boldsymbol{g}_k^T\boldsymbol{g}_k}$

5. Mise à jour de la direction de descente : $\boldsymbol{d}_{k+1} = \boldsymbol{g}_{k+1} + \gamma_k \boldsymbol{d}_k$
6. Calcul du coefficient diagonal de $\boldsymbol{T}$ : $\delta_{k+1} = \frac{1}{\alpha_{k+1}} + \frac{\gamma_k}{\alpha_k}$
7. Calcul du coefficient sur-diagonal : $\beta_{k+1} = -\frac{c_k}{\alpha_k c_{k+1}}$

FIN BOUCLE

avec $c_k = \|\boldsymbol{g}_k\|^{-1}$

Comme $\boldsymbol{M}$ est quadratique, $\nabla f(x_{k+1}) = \nabla f(x_k) + M(x_{k+1} - x_k) + \mathcal{O}(|x_{k+1} - x_k|^2)$ et comme $g$ est initialisé à $g_0 = -\nabla f(x_0)$ et $x_{k+1} - x_k = \alpha_k d_k$, on retrouve que $\forall k\ g_{k+1} = g_k - \alpha_k M d_k$, l'autre écriture classique du gradient conjugué.

L'efficacité de cet algorithme est liée au conditionnement [1] de la Hessienne, c'est pourquoi un changement de variable est effectué avant la minimisation.

### Pré-conditionnement

Pour améliorer la convergence, un changement de variable est réalisé pour se ramener à un problème dont le conditionnement est plus faible. Il est classique d'introduire, notamment dans le 4D-var, le pré-conditionnement défini par le changement de variable suivant : $\boldsymbol{w} = \boldsymbol{L}^{-1}\delta\boldsymbol{x}$ avec $\boldsymbol{L} = \boldsymbol{B}^{1/2}$.

La fonction coût modifiée devient alors

$$\tilde{J}(\boldsymbol{w}) = \frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + \frac{1}{2}\sum_{i=0}^{T_f}(d_i - \boldsymbol{H}_i\boldsymbol{M}_{0\to i}\boldsymbol{L}\boldsymbol{w})^T\boldsymbol{R}_i^{-1}(d_i - \boldsymbol{H}_i\boldsymbol{M}_{0\to i}\boldsymbol{L}\boldsymbol{w})$$

et la Hessienne

$$\tilde{J}'' = \boldsymbol{I} + \sum_{i=0}^{T_f}\boldsymbol{L}^T\boldsymbol{M}_{0\to i}^T\boldsymbol{H}_i^T\boldsymbol{R}_i^{-1}\boldsymbol{H}_i\boldsymbol{M}_{0\to i}\boldsymbol{L}$$

On obtient la somme d'une matrice identité et d'une matrice de rang plus faible (au maximum la dimension de l'espace des observations). Une grande partie des valeurs propres sera alors proche de 1, et la Hessienne peut être approximée dans la base $\boldsymbol{w}$ par

$$\tilde{J}'' = \boldsymbol{I} + \sum_{k=1}^{r}(\lambda_k - 1)\boldsymbol{u}_k\boldsymbol{u}_k^T$$

avec $(\lambda_k, \boldsymbol{u}_k)$ $r$ valeurs et vecteurs propres sélectionnés de $\tilde{J}''$.

Une estimation de la matrice de covariance $\boldsymbol{A}$ est obtenue en inversant les valeurs propres puis en revenant dans l'espace du modèle.

$$\boldsymbol{A} \approx \boldsymbol{B} + \sum_{k=1}^{r}(\lambda_k^{-1} - 1)(Lu_k)(Lu_k)^T$$

Il suffit de diagonaliser la matrice exprimée dans la base réduite obtenue par l'algorithme de Lanczos pour avoir une estimation des valeurs propres et vecteurs propres de $\boldsymbol{A}^{-1}$. $r$ étant de l'ordre de quelques dizaines l'inversion se fait à un coût numérique raisonnable.

1. Le conditionnement d'une matrice est le quotient de la plus grande valeur propre de $\boldsymbol{A}^{-1}$ sur sa plus petite.

# Annexe C

# Algorithme de l'ETKF

L'algorithme présenté ici vise à minimiser la fonctionnelle $\widetilde{J}$ suivante (Hunt et al., 2006) : $\widetilde{J}(\boldsymbol{w}) = \boldsymbol{w}^T\boldsymbol{w} + (\boldsymbol{y}^o - \boldsymbol{y}^b - \boldsymbol{Y}^b\boldsymbol{w})^T\boldsymbol{R}^{-1}(\boldsymbol{y}^o - \boldsymbol{y}^b - \boldsymbol{Y}^b\boldsymbol{w})$ avec $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$

1. ensemble à priori, perturbations espace modèle :$\boldsymbol{X}^b = \frac{1}{\sqrt{N_e - 1}}[\boldsymbol{x}_1^b ... \boldsymbol{x}_{N_e}^b]$
2. ensemble à priori, perturbations espace des obs. : $\boldsymbol{Y}^b = \boldsymbol{H}\boldsymbol{X}^b$
3. calcul de $\boldsymbol{C} = {\boldsymbol{Y}^b}^T\boldsymbol{R}^{-1}$
4. calcul de la matrice à inverser de dim $N_e \times N_e$ : $\widetilde{\boldsymbol{A}}^{-1} = (\boldsymbol{I} + \boldsymbol{C}\boldsymbol{Y}^b)$
5. diagonalisation de $\widetilde{\boldsymbol{A}}^{-1} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^T$
6. mise à 0 de la valeur propre de la valeur propre supplémentaire pour avoir un rang de $N_e - 1$
7. correction de la moyenne $\overline{\boldsymbol{w}^a} = \widetilde{\boldsymbol{A}}\boldsymbol{C}(\boldsymbol{y}^o - \boldsymbol{y}^b)$
8. calcul des analyses $\boldsymbol{x}_i^a = \overline{\boldsymbol{x}^a} + (\boldsymbol{I} - \boldsymbol{X}^b\widetilde{\boldsymbol{K}}\boldsymbol{H})\varepsilon_i^b$

Comme par construction $\sum_n \boldsymbol{x}_n^b = 0$, une des valeurs propres de $\widetilde{\boldsymbol{A}}^{-1}$ est nulle, ou très proche de 0 selon les erreurs d'arrondis. Avant d'inverser $\boldsymbol{D}$, il est donc nécessaire d'identifier toutes les valeurs propres très faibles qui donneraient des valeur aberrantes une fois $\boldsymbol{D}$ inversée. Toutes les valeurs propres inférieures à un seuil choisi sont donc mises à 1 avant l'inversion, puis mises à 0.

# Annexe D

# Création d'une distribution d'erreur à partir d'un opérateur

Lorsque les matrices de covariances n'existent que sous la forme d'opérateurs, il n'est pas possible de calculer explicitement les caractéristiques de distributions issues de ces matrices. De plus dans le cadre de l'étude (au chapitre 6), il est nécessaire de manipuler des ensembles dont la distribution d'erreur correspond à la matrice de covariance.

On souhaite dans ce cas que la matrice de covariance associée à l'ensemble soit $\boldsymbol{B}$. C'est-à-dire que $\mathbb{E}(\boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^{bT}) = \boldsymbol{B}$, avec $\boldsymbol{\varepsilon}^b = \boldsymbol{x}^b - \overline{\boldsymbol{x}}^b$ l'écart d'un membre à la moyenne de l'ensemble.

En utilisant des $\boldsymbol{\varepsilon}^b$ de la forme

$$\boldsymbol{\varepsilon}^b = \boldsymbol{B}^{1/2}\eta$$

avec $\eta \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$, la matrice de covariance obtenue est bien

$$\begin{aligned}
\mathbb{E}(\boldsymbol{\varepsilon}^b \boldsymbol{\varepsilon}^{bT}) &= \mathbb{E}\left(\boldsymbol{B}^{1/2}\eta)(\boldsymbol{B}^{1/2}\eta)^T\right) \\
&= \boldsymbol{B}^{1/2}\mathbb{E}\left(\eta\,\eta^T\right)\boldsymbol{B}^{T/2} \\
&= \boldsymbol{B}^{1/2}\boldsymbol{I}\boldsymbol{B}^{T/2} \\
&= \boldsymbol{B}
\end{aligned}$$

avec $\mathcal{N}(0, I)$ un vecteur gaussien de moyenne nulle et de matrice de covariance identité.

Suivant la modélisation de $\boldsymbol{B}$ choisie (voir section 6.1.2), la mise en place de l'opérateur $\boldsymbol{B}^{1/2}$ est aisée. Cette méthode permet donc de créer des perturbations distribuées de manière cohérente avec la matrice utilisée pour l'assimilation de données.

A l'inverse, cette technique peut aussi être utilisée pour calculer une estimation de la variance d'erreur d'analyse. En effet avec $\boldsymbol{A} \approx \boldsymbol{B} + \sum_{k=1}^{r}(\lambda_k^{-1} - 1)(Lu_k)(Lu_k)^T$ obtenue par l'algorithme de Lanczos (cf. annexe B), on en déduit $\boldsymbol{A}^{1/2} = \boldsymbol{B}^{1/2} + \sum_{k=1}^{r}\frac{1}{\|Lu_k\|}(\sqrt{(}\lambda_k)^{-1} - 1)(Lu_k)(Lu_k)^T$.

En générant un ensemble de réalisations de $\boldsymbol{\varepsilon}^a = \boldsymbol{A}^{1/2}\eta$, une estimation de $\mathrm{diag}\left(\mathbb{E}(\boldsymbol{\varepsilon}^a \boldsymbol{\varepsilon}^{aT})\right)$ peut être obtenue pour un coût numérique faible, la variance d'erreur d'analyse est calculée ainsi.

# Annexe E

# Calcul du gradient de la fonction score par régression linéaire

Le gradient permet de calculer au premier ordre la variation $dS$ associée à une perturbation $\boldsymbol{dx}$,

$$dS = \boldsymbol{dx}^T \nabla S. \tag{E.1}$$

Le gradient de la fonction score à l'instant initial est généralement calculé en utilisant le modèle adjoint. Il est cependant possible de calculer une estimation du gradient $\nabla S$ à partir d'une régression linéaire multivariée.

On considérant un ensemble de perturbations $\boldsymbol{dx}_i$ et de variations $dS_i$ de taille $N_e$, on construit la matrice $[\boldsymbol{dx}]$ dont les colonnes sont les différents vecteurs $\boldsymbol{dx}_i$ et $[dS]$ le vecteur colonne où $[dS]_i = dS_i \ \forall i$. L'équation (E.1) peut être réécrite sous la forme,

$$[dS] = \widetilde{[\boldsymbol{dx}]}^T \widetilde{\nabla S}, \tag{E.2}$$

avec $\widetilde{[\boldsymbol{dx}]} = [\underline{\boldsymbol{dx}}]^T [\boldsymbol{dx}]$ la matrice des vecteurs de l'ensemble exprimée dans leur propre base (le souligné indique que les vecteurs sont normés) et $\widetilde{\nabla S}$ est le gradient exprimé dans l'espace réduit des perturbations.

Avec les notations de statistiques classiques, $\boldsymbol{Y} = [dS]$, $\widetilde{[\boldsymbol{dx}]} = \boldsymbol{X}^T$, $\widetilde{\nabla S} = \boldsymbol{a}$ et $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{a}$. L'estimateur de $\boldsymbol{a}$ à partir de l'ensemble est

$$\hat{\boldsymbol{a}} = (\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X} \boldsymbol{Y}, \tag{E.3}$$

avec $\boldsymbol{X}^T \boldsymbol{X}$ de dimension $N_e \times N_e$ donc facilement diagonalisable et inversible.

Le gradient exprimé dans l'espace réduit se calcule donc comme

$$\widetilde{\nabla S} = ([\widetilde{\boldsymbol{dx}}][\widetilde{\boldsymbol{dx}}]^T)^{-1} [\widetilde{\boldsymbol{dx}}][dS]. \tag{E.4}$$

Par construction des perturbations de l'ensemble, la somme des $\boldsymbol{dx}_i$ est nulle. Lorsque l'on diagonalise $[\widetilde{\boldsymbol{dx}}][\widetilde{\boldsymbol{dx}}]^T$, il est nécessaire de filtrer la valeur propre nulle avant d'inverser la matrice.

L'expression du gradient dans l'espace physique est obtenue par

$$\nabla S = [\underline{\boldsymbol{x}^b}] \widetilde{\nabla S} \tag{E.5}$$

En choisissant $\boldsymbol{dx} = \boldsymbol{\varepsilon}^{\boldsymbol{a}}$, le gradient de $S$ par rapport à l'analyse est donc disponible.

# Annexe F

# Développement de la VRF au second ordre à partir d'un ensemble

Cette annexe décrit les calculs relatifs au développement de la VRF au second ordre à partir d'une régression linéaire polynomiale.

La variation de la fonction score au second ordre est définie par

$$dS_2 = \Gamma_1^T \delta(\mathbf{x}^a) + \delta(\mathbf{x}^a)^T \Gamma_2 \delta(\mathbf{x}^a), \tag{F.1}$$

avec $\Gamma_1$ et $\Gamma_2$ les inconnues que l'on souhaite calculer.

## F.1 Résolution de la régression linéaire polynomiale

On cherche d'abord à résoudre la régression linéaire polynomiale suivante,

$$\boldsymbol{Y} = \boldsymbol{M}\boldsymbol{X}, \tag{F.2}$$

avec $\boldsymbol{Y}_i = dS_i$, $\boldsymbol{X}^T = [\varepsilon_i^a \;\; \varepsilon_j^a \varepsilon_k^a]^T$ et

$$\boldsymbol{M} = \begin{pmatrix} \Gamma_1 & \boldsymbol{M}_{1,2} \\ \boldsymbol{M}_{2,1} & \Gamma_2 \end{pmatrix}. \tag{F.3}$$

La matrice $\boldsymbol{M}$ est alors obtenue comme $\boldsymbol{M} = \mathrm{Cov}(\boldsymbol{Y}, \boldsymbol{X})\,\mathrm{Cov}(\boldsymbol{X})^{-1}$.

On commence par le calcul de $\mathrm{Cov}(X)$,

$$\mathrm{Cov}(X) = \begin{bmatrix} \mathrm{E}(\varepsilon_i^a \varepsilon_j^a) = A_{ij} & \mathrm{E}(\varepsilon_i^a \varepsilon_j^a \varepsilon_k^a) = 0 \\ \mathrm{E}(\varepsilon_i^a \varepsilon_j^a \varepsilon_k^a) = 0 & \mathrm{E}(\varepsilon_i^a \varepsilon_j^a \varepsilon_k^a \varepsilon_l^a) \end{bmatrix} \tag{F.4}$$

L'espérance des termes d'ordre impair en $\varepsilon^a$ est nulle, on trouve donc que $\boldsymbol{M}_{1,2} = \boldsymbol{M}_{2,1} = 0$. Le terme $\mathrm{E}(\varepsilon_i^a \varepsilon_j^a \varepsilon_k^a \varepsilon_l^a)$ est décomposé en utilisant le théorème de Wick,

$$\mathrm{E}(\varepsilon_i^a \varepsilon_j^a \varepsilon_k^a \varepsilon_l^a) = \mathrm{E}(\varepsilon_i^a \varepsilon_j^a)\,\mathrm{E}(\varepsilon_k^a \varepsilon_l^a) + \mathrm{E}(\varepsilon_i^a \varepsilon_k^a)\,\mathrm{E}(\varepsilon_j^a \varepsilon_l^a) + \mathrm{E}(\varepsilon_i^a \varepsilon_l^a)\,\mathrm{E}(\varepsilon_j^a \varepsilon_k^a). \tag{F.5}$$

On en déduit $\mathrm{Cov}(X)$,

$$\mathrm{Cov}(X) = \begin{bmatrix} A_{ij} & 0 \\ O & A_{ij}A_{kl} + A_{ik}A_{jl} + A_{il}A_{jk} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & 3A^2 \end{bmatrix} \tag{F.6}$$

Pour inverser cette matrice, il est nécessaire de disposer de $\mathbf{A}^{-1}$ et $\mathbf{A}^{-2}$. L'inversion de $\mathbf{A}$ est réalisée dans l'espace des perturbations de l'ensemble, en suivant ce qui est fait pour la réalisation de l'ETKF. On pose $\boldsymbol{X}^a$ le vecteur des perturbations de l'ensemble. La restriction de $\mathbf{A}$ à l'espace défini par les membres de l'ensemble s'écrit alors $\widetilde{\mathbf{A}} = \frac{1}{N_e - 1}\boldsymbol{X}^a \boldsymbol{X}^{aT}$. On exprime alors $\widetilde{\mathbf{A}}$ dans la base des vecteurs de l'ensemble

pour obtenir une matrice de dimension $N_e \times N_e$. Dans cet espace réduit, il est possible d'appliquer un algorithme de calcul des vecteurs propres pour diagonaliser la matrice réduite.

Pour obtenir $\widetilde{\mathbf{A}}^{-1}$, il suffit donc d'inverser la matrice diagonale obtenue et d'appliquer les différentes matrices de changement de base pour retrouver l'espace physique. On effectue de même pour $\widetilde{\mathbf{A}}^{-2}$ mais en prenant le carré de l'inverse de la matrice diagonale.

Le calcul du second terme pour $\boldsymbol{M}$ est plus simple. $Cov(Y,X)$ est un vecteur ligne dont les composantes sont

$$Cov(S,(x^a;x^{a2})) = \frac{1}{N-1}\left[\sum_{i=1}^{N}\delta S_i\varepsilon^a_{ki};\sum_{i=1}^{N}\delta S_i(\varepsilon^a_k\varepsilon^a_l)_i\right] = [C_1;C_2] \tag{F.7}$$

$k$ et $l$ parcourant les points de grille.

On obtient finalement,

$$M = [C_1 \quad C_2]\begin{bmatrix}A^{-1} & 0\\ 0 & \frac{1}{3}A^{-2}\end{bmatrix} = [C_1A^{-1} \quad \tfrac{C_2}{3}A^{-2}] \tag{F.8}$$

d'où

$$\begin{aligned}\Gamma_1 &= C_1A^{-1}\\ \Gamma_2 &= \frac{C_2}{3}A^{-2}.\end{aligned}$$

Ceci nous permet d'avoir une expression de $dS_2$ au second ordre. Afin de pouvoir utiliser cette expression dans un calcul de VRF, il est nécessaire de déduire la réduction de variance de la fonction score associée.

## F.2 Calcul de la variance de la fonction score au second ordre

L'expression de la variance devient alors

$$\begin{aligned}\sigma_S^2 &= \mathrm{E}(dS_2dS_2^T)\\ &= \mathrm{E}\left((\Gamma_1^T\varepsilon^a+\varepsilon^{aT}\Gamma_2\varepsilon^a)(\Gamma_1^T\varepsilon^a+\varepsilon^{aT}\Gamma_2\varepsilon^a)^T\right)\\ &= \Gamma_1^T\,\mathrm{E}(\varepsilon^a\varepsilon^{aT})\Gamma_1+\mathrm{E}(\Gamma_1^T\varepsilon^a\varepsilon^{aT}\Gamma_2^T\varepsilon^a)+\mathrm{E}(\varepsilon^{aT}\Gamma_2\varepsilon^a\Gamma_1^T\varepsilon^a)+\mathrm{E}(\varepsilon^{aT}\Gamma_2\varepsilon^a\varepsilon^{aT}\Gamma_2^T\varepsilon^a)\\ &= \Gamma_1^T\mathbf{A}\Gamma_1+\mathrm{E}(\varepsilon^{aT}\Gamma_2\varepsilon^a\varepsilon^{aT}\Gamma_2^T\varepsilon^a)\end{aligned} \tag{F.9}$$

Comme précédemment l'espérance des produits d'ordre impaire est nulle. On retrouve alors le terme du premier ordre $\Gamma_1^T\mathbf{A}\Gamma_1 = \nabla S^T\mathbf{A}\nabla S$ auquel s'ajoute un deuxième terme.

Comme $\varepsilon^{aT}\Gamma_2\varepsilon^a$ est un scalaire, on a $(\varepsilon^{aT}\Gamma_2\varepsilon^a) = (\varepsilon^{aT}\Gamma_2\varepsilon^a)^T$. Le deuxième terme de $\sigma_S^2$ se réécrit alors,

$$\begin{aligned}\mathrm{E}(\varepsilon^{aT}\Gamma_2\varepsilon^a\varepsilon^{aT}\Gamma_2^T\varepsilon^a) &= \mathrm{E}\left((\varepsilon^{aT}\Gamma_2\varepsilon^a)(\varepsilon^{aT}\Gamma_2\varepsilon^a)^T\right)\\ &= \mathrm{E}\left((\varepsilon^{aT}\Gamma_2\varepsilon^a)(\varepsilon^{aT}\Gamma_2\varepsilon^a)\right).\end{aligned}$$

Pour simplifier ce deuxième terme, on suppose que la matrice $\Gamma_2^{1/2}$ est définie et on pose $\tilde{\varepsilon} = \Gamma_2^{1/2}\varepsilon$. En utilisant à nouveau le théorème de Wick on obtient,

$$\begin{aligned}\mathrm{E}(\tilde{\varepsilon}^T\tilde{\varepsilon}\tilde{\varepsilon}^T\tilde{\varepsilon}) &= \mathrm{E}(\tilde{\varepsilon}^T\tilde{\varepsilon})\,\mathrm{E}(\tilde{\varepsilon}^T\tilde{\varepsilon})+\mathrm{E}(\tilde{\varepsilon}^T\tilde{\varepsilon})\,\mathrm{E}(\tilde{\varepsilon}^T\tilde{\varepsilon})\\ &= 2\,\mathrm{E}(\tilde{\varepsilon}^T\tilde{\varepsilon})^2\\ &= 2\,\mathrm{E}(\varepsilon^{aT}\Gamma_2\varepsilon^a)^2\\ &= 2\,\mathrm{E}(\Gamma_{2ij}\varepsilon^a_i\varepsilon^a_j)^2\\ &= 2\left[\Gamma_{2ij}\,\mathrm{E}(\varepsilon^a_i\varepsilon^a_j)\right]^2\\ &= 2\left[\Gamma_2 :A\right]^2\end{aligned}$$

où : est la double contraction tensorielle. On note qu'il n'est pas nécessaire de connaître $\Gamma_2^{1/2}$ pour la suite du calcul.

## F.3 Réduction de variance de la fonction au second ordre

Comme pour le calcul au premier ordre, on reprend l'équation (4.11), $(\delta\sigma_S^t)^2 = {\sigma_S^c}^2 - {\sigma_S^{c+t}}^2$ (où la mention $c$ fait référence ou réseau d'observation de routine et $c+t$ fait référence au réseau d'observation de routine enrichi de l'observation supplémentaire $t$, en utilisant cette fois-ci l'expression de $\sigma_S$ au second ordre. On a,

$$\begin{aligned} {\sigma_S^c}^2 &= \Gamma_1^T \mathbf{A}_c \Gamma_1 + 2[\Gamma_2 : \mathbf{A}_c]^2 \\ {\sigma_S^{c+t}}^2 &= \Gamma_1^T \mathbf{A}_{c+t} \Gamma_1 + 2[\Gamma_2 : \mathbf{A}_{c+t}]^2. \end{aligned}$$

D'où

$$\delta_{S_t^2} = \Gamma_1^T \mathbf{A}_c \Gamma_1 + 2\left[\Gamma_2 : \mathbf{A}_c\right]^2 - \Gamma_1^T \mathbf{A}_{c+t} \Gamma_1 - 2\left[\Gamma_2 : \mathbf{A}_{c+t}\right]^2. \tag{F.10}$$

# Annexe G

# Calcul de la trace d'une matrice par une méthode de Monte-Carlo

Plusieurs méthodes existent pour estimer la trace d'une matrice symétrique à partir d'une méthode de Monte-Carlo, voir Avron and Toledo (2011).

**Méthode d'Hutchinson**

Soit $\mathbf{A}$, une matrice symétrique de dimension $N$ et de trace non nulle, $\mathrm{Trace}(\mathbf{A}) \neq 0$.
Pour réduire le coût numérique associé à la génération d'une variable Gaussienne, et sachant que seul l'estimation des moments est nécessaire ici, une astuce souvent introduite consiste à remplacer la variable Gaussienne centrée réduite $z$ par une variable de Bernoulli centrée réduite. En effet, pour générer une variable de Bernoulli, il suffit de générer un échantillon suivant une loi uniforme, peu coûteuse à générer. $z$ est alors un vecteur aléatoire tel que $\Pr(z_i = \pm 1) = 1/2$, alors

$$\mathrm{E}(\boldsymbol{z}^T \mathbf{A} \boldsymbol{z}) = \mathrm{Trace}(\mathbf{A}) \tag{G.1}$$

$$\mathrm{Var}(\boldsymbol{z}^T \mathbf{A} \boldsymbol{z}) = 2\left( \|\mathbf{A}\|_F^2 - \sum_{i=1}^{N} a_{ii}^2 \right) \tag{G.2}$$

Il suffit d'être capable de calculer le produit d'un vecteur $z$ par la matrice $\mathbf{A}$ pour avoir sa trace. Même si $\mathbf{A}$ n'est pas connue explicitement, si construite sous la forme d'un opérateur par exemple.

**Lien entre la double contraction tensorielle et la trace d'un produit de matrice**

Soient $\mathbf{A} = [a_{ij}]$ et $\mathbf{B} = [b_{ij}]$ deux matrices symétriques de dimension $N$. La double contraction tensorielle de $\mathbf{A}$ par $\mathbf{B}$ est définie par,

$$\mathbf{A} : \mathbf{B} = \sum_{i}^{N} \sum_{j}^{N} a_{ij} b_{ij} \tag{G.3}$$

Le produit de matrices est donné par $(\mathbf{AB})_{ij} = \sum_k^N a_{ik} b_{kj}$. Donc la trace s'écrit,

$$\mathrm{Trace}(\mathbf{AB}) = \sum_{i}^{N} \sum_{k}^{N} a_{ik} b_{ki}. \tag{G.4}$$

Comme $\mathbf{A}$ et $\mathbf{B}$ sont symétriques, on obtient :

$$\mathbf{A} : \mathbf{B} = \mathrm{Trace}(AB) \tag{G.5}$$

# Annexe H

# Assessing the trajectory influence in adaptive observation Kalman filter sensitivity method.

L'article qui suit a été publié en avril 2012 dans la revue *Quarterly Journal of the Royal Meteorological Society* (Oger et al., 2012).

RMetS
Royal Meteorological Society

# Assessing the influence of the model trajectory in the adaptive observation Kalman Filter Sensitivity method

N. Oger,* O. Pannekoucke, A. Doerenbecher and P. Arbogast
*Météo-France, CNRM-GAME URA 1357, Toulouse, France*

*Correspondence to: N. Oger, Météo-France CNRM/GMAP/RECYF, 42 av. G. Coriolis, 31057 Toulouse Cedex, France. E-mail: niels.oger@meteo.fr

**The purpose of adaptive observation strategies is to design optimal observation networks in a prognostic way. The implementation of such strategies is based on adaptive observation numerical techniques that provide guidance on where to deploy future additional observations. Most advanced techniques account for the dynamical aspects of the atmosphere and the data assimilation system (DAS). This study aims to assess the influence of the model trajectory on the Kalman Filter Sensitivity (KFS) method used at Météo-France. KFS is an adjoint-based method identifying sensitive areas by means of a forecast score variance. Targeted observations are designed to reduce this score variance. In its first version, KFS was not able to deal with trajectory uncertainties. An ensemble-based approach is undertaken to investigate if it is possible to account for these uncertainties. We assess the robustness of the method regarding trajectory errors and propose a practical solution. To avoid high computational costs, a simplified framework is used. We perform experiments with a two-layer quasi-geostrophic model in an incremental 4D-Var system. KFS is computed to target a potential vorticity anomaly interacting with a simplified jet stream. Numerical experiments show that the impact of additional observations depends equally on the suboptimal formulation of the DAS components and the trajectory errors. When the KFS sensitivity fields are constrained by the DAS components, the sensitivity values are determined by the selection of the reference trajectory. A targeting strategy based on the ensemble mean sensitivity field is proposed. Experiments show this to be more efficient than a random targeting strategy.** Copyright © 2011 Royal Meteorological Society

*Key Words:* data assimilation; ensemble method; observation sensitivity; Observing System Simulation Experiment

*Received 20 December 2010; Revised 23 July 2011; Accepted 19 September 2011; Published online in Wiley Online Library 14 October 2011*

*Citation:* Oger N, Pannekoucke O, Doerenbecher A, Arbogast P. 2012. Assessing the influence of the model trajectory in the adaptive observation Kalman Filter Sensitivity method. *Q. J. R. Meteorol. Soc.* **138**: 813–825. DOI:10.1002/qj.950

## 1. Introduction

The number of observations available for data assimilation has been continually increasing. Recent studies have been carried out on observation sensitivity and observation impact, and have presented methods to verify the effectiveness of the observations used in the assimilation process (Daescu, 2008; Trémolet, 2008; Gelaro and Zhu, 2009; Cardinali, 2009; Gelaro *et al.*, 2010). The Data Targeting System (DTS) experiment, currently hosted by the European Centre for Medium-Range Weather Forecasts (ECMWF), aims to add routine observations at non-standard times to help study high-impact weather events over Europe. It is based on adaptive observation methods (singular vector and ensemble) which

Copyright © 2011 Royal Meteorological Society

estimate the area where a forecast is sensitive to additional observations.

Adaptive observation strategies consist of the design, at the current time, $t_0$, of a future observation network at time $t_{\rm obs}$. This design generally considers deploying a few additional observations. The assimilation of the enhanced observation network aims to reduce the uncertainty of the subsequent forecast at a predetermined range ($t_{\rm verif} > t_{\rm obs}$) and over a predetermined area. The reduction of uncertainty is expressed with respect to the forecast without any added observations. Several (so-called targeting) techniques for computing deployments have been developed at national meteorological services: the Ensemble Transform Kalman Filter (ETKF, Bishop *et al.*, 2001); singular vectors (Palmer *et al.*, 1998); the Kalman Filter Sensitivity (KFS, Bergot and Doerenbecher, 2002). The last technique has been developed at Météo-France and implemented in the numerical model ARPÈGE in research mode. It is a second-generation targeting tool, which is similar to techniques that estimate observation sensitivity. These techniques mix dynamics and assimilation using adjoint-based computations. This method links the uncertainty of a forecast to the variance of a score function computed at $t_{\rm verif}$. The score is computed over an area of interest and is related to a critical aspect in the prediction, e.g. high-impact weather. The reduction of the forecast score variance due to additional targeted observations at time $t_{\rm obs}$ is then computed. The variance is connected to the covariance matrix of the analysis errors **A** and to the gradient of the score function at $t_{\rm obs}$. The adjoint model must compute the gradient from $t_{\rm verif}$ backwards. The time interval between the start of the procedure $t_0$ and the time of the observations $t_{\rm obs}$ and the verification $t_{\rm verif}$ is typically several days. The KFS computes the variance of the score function, an estimation of its reduction due to supplementary observations, and also a sensitive area for targeting the observations.

The gradient of the score function at the verification time and the adjoint model are based on a single trajectory. This trajectory is usually a previous deterministic forecast of the model ($t_0 - t_{\rm obs}$) of about a few dozen hours. The purpose of this study is to evaluate the influence of the trajectory and of the data assimilation system (DAS) on the computation of the optimal network, made of supplementary observations on top of the routine observing network. The sensitivity to the trajectory also applies to the other targeting techniques that use the adjoint technology, such as adjoint sensitivity and singular vectors.

When focusing on the quality of any forecast, it is important to know the uncertainties associated with the forecast. When assessing forecast uncertainties, ensemble forecasting is a natural approach. Within an ensemble, all forecasts (or trajectories) are equally likely, even if they may significantly differ in terms of meteorological features. So when considering the uncertainty on the initial state, it seems reasonable to consider a probabilistic approach through an ensemble of trajectories.

To reduce the prohibitive computational cost of a large ensemble, and the several experiments needed in a context of large-dimensional operational numerical weather prediction (NWP) system, a two-layer quasi-geostrophic (QG) toy model (Phillips, 1951) is implemented. The data assimilation system used is an incremental four-dimensional variational assimilation (4D-Var; Courtier *et al.*, 1994), and all experiments take place in a totally simulated (including the truth) but fully consistent framework, similar to Observing System Simulation Experiments (OSSEs).

The article is organized as follows. Section 2 reviews the data assimilation formalism and particularly the incremental 4D-Var before introducing the KFS method. Section 3 describes the framework of the simulation, including a description of the two-layer QG model used, of the synoptic framework of the simulations and the OSSE components. The results are shown in section 4. The conclusion is presented in section 5.

## 2. Data assimilation and adaptive observation formalism

### 2.1. *Incremental variational data assimilation*

The notation hereafter follows Ide *et al.*(1997). In a variational data assimilation system (VDAS), the analysis $\mathbf{x}^{\rm a}$ is the result of the minimization of a cost function

$$J(\mathbf{x}) = \frac{1}{2}\|\mathbf{x}-\mathbf{x}^{\rm b}\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|H(\mathbf{x})-\mathbf{y}^{\rm o}\|^2_{\mathbf{R}^{-1}}, \quad (1)$$

$$\mathbf{x}^{\rm a} = \mathrm{Arg\,min}\, J, \quad (2)$$

where $\|\cdot\|_{\mathbf{C}}$ is the inner product with respect to the **C** matrix metric, $\mathbf{x}^{\rm b}$ is the background state estimate, **B** and **R** are respectively the background- and observation-error covariance matrices, and $H$ is the observation operator which maps the model state into observations. When $H$ is linear (denoted **H**), we are faced with a quadratic problem whose unique solution is given by the Best Linear Unbiased Estimator (BLUE) $\mathbf{x}^{\rm a} = \mathbf{x}^{\rm b} + \mathbf{K}\mathbf{d}$, where

$$\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^{\rm T}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\rm T}\mathbf{R}^{-1} = \mathbf{A}\mathbf{H}^{\rm T}\mathbf{R}^{-1}$$

is the Kalman gain matrix and $\mathbf{d} = \mathbf{y}^{\rm o} - \mathbf{H}\mathbf{x}^{\rm b}$ the innovation vector. **A** is the analysis-error covariance matrix and also the inverse of the Hessian of the cost function, $J''$.

When $H$ contains nonlinearities, the assimilation system considers a series of state variables $\mathbf{x}_j$ along which $H$ may be linearized. This method is called incremental VDAS and was introduced by Courtier *et al.*(1994). The first state variable is taken as the background state $\mathbf{x}^{\rm b} = \mathbf{x}_0$; at iteration $j$ the cost function is

$$J(\delta\mathbf{x}) = \frac{1}{2}\|\delta\mathbf{x}-\mathbf{b}_j\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\|\mathbf{H}_j\delta\mathbf{x}-\mathbf{d}_j\|^2_{\mathbf{R}^{-1}}, \quad (3)$$

where $\delta\mathbf{x} = \mathbf{x}_j - \mathbf{x}_{j-1}$ is the result of the minimization, $\mathbf{b}_j = \mathbf{x}^{\rm b} - \mathbf{x}_{j-1}$, $\mathbf{d}_j = \mathbf{y}^{\rm o} - H(\mathbf{x}_{j-1})$ and $\mathbf{H}_j$ is the observation operator linearized near the state estimate $\mathbf{x}_{j-1}$. In order to get the absolute minimum and not a local one, the first guess (background) needs to be close enough to the truth. Furthermore, in the 4D-Var formulation, $H$ includes the time propagation of the model state from the assimilation time to the time of observation. This is achieved through the dynamical model $M$ or its linear tangent model **M**.

In such a case, it may be necessary to repeat the minimization several times. For each minimization, the previous assimilated state is taken as the first guess. The update of the trajectory $\mathbf{x}_j$ during assimilation is called the outer loop. The inner loop refers to the iteration of the minimization algorithm.

The classic pre-conditioning $\chi = \mathbf{B}^{1/2}\delta\mathbf{x}$ (Gilbert and Lemaréchal, 1989) is used. The minimization is done through a combined conjugate-gradient (CG) Lanczos algorithm.

Copyright © 2011 Royal Meteorological Society

### 2.2. Adaptive observation

A measure (or function) has to be chosen to quantitatively evaluate the quality of the forecast of interest. Let $S(\mathbf{x})$ be this function that is defined at $t_{\text{verif}}$ and may be arbitrary. However, the interpretation of the final results is simplified when $S(\mathbf{x})$ is related to relevant meteorological features (i.e. wind, vorticity or humidity depending on the nature of the weather event). In our QG system, $S(\mathbf{x})$ is the total enstrophy in a verifying area located at the lower model level, $S = \sum_i \zeta_i^2$. Unlike in observation impact studies, the value of the forecast score is irrelevant here; instead, the goal is to evaluate the variance of the function, i.e. the uncertainty on the forecast, and to link it to the initial state at $t_{\text{obs}}$. Though a probabilistic approach can be found in Bergot and Doerenbecher (2002), the one presented here follows the development of Ancell and Hakim (2007).

If $\varepsilon^{\text{f}}$ is a small forecast error at $t_{\text{verif}}$, inducing a variation $\delta S$ on the forecast score, then, by a first-order Taylor expansion, $\delta S$ can be expressed as a function of a small perturbation of the analysis $\varepsilon^{\text{a}}$ at $t_{\text{obs}}$:

$$\begin{aligned} \delta S &= S(\mathbf{x}^{\text{f}} + \varepsilon^{\text{f}}) - S(\mathbf{x}^{\text{f}}) \\ &= [\nabla_{\mathbf{x}^{\text{f}}} S]^{\text{T}} \varepsilon^{\text{f}} \\ &= [\nabla_{\mathbf{x}^{\text{f}}} S]^{\text{T}} \mathbf{M}_{t_{\text{obs}} \to t_{\text{verif}}} \varepsilon^{\text{a}} \\ &= [\nabla_{\mathbf{x}^{\text{a}}} S]^{\text{T}} \varepsilon^{\text{a}}, \end{aligned} \tag{4}$$

where $[\nabla_{\mathbf{x}^{\text{a}}} S] = \mathbf{M}^{\text{T}}_{t_{\text{obs}} \to t_{\text{verif}}} [\nabla_{\mathbf{x}^{\text{f}}} S]$ is the gradient field at $t_{\text{obs}}$, referred as $\nabla S$ hereafter. As the perturbation $\varepsilon^{\text{a}}$ is distributed according to the analysis-error probability density function, it is straightforward to see that $\mathbf{A} = \mathbb{E}(\varepsilon^{\text{a}} \varepsilon^{\text{a}\text{T}})$. Assuming that the gradient $\nabla S$ is statistically independent of $\varepsilon^{\text{a}}$, the variance of the forecast score can be written

$$\sigma_S^2 = \mathbb{E}(\delta S \, \delta S^{\text{T}}) = \nabla S^{\text{T}} \mathbf{A} \nabla S \,. \tag{5}$$

The uncertainty on the forecast score is here expressed directly as a function of the uncertainty on the initial state $\mathbf{A}$. Assuming that the model is perfect, Eq. (5) can also be written $\sigma_S^2 = [\nabla_{\mathbf{x}^{\text{f}}} S]^{\text{T}} \mathbf{P} [\nabla_{\mathbf{x}^{\text{f}}} S]$, where $\mathbf{P} = \mathbf{M}\mathbf{A}\mathbf{M}^{\text{T}}$, the time-integrated error covariance matrix.

In Eq. (5), it has been assumed that the gradient is statistically independent from the trajectory. This leads to a deterministic view of the forecast score variance. The purpose of this study is to test whether this hypothesis is valid and, if not valid, to assess how to modify it.

### 2.3. Kalman Filter Sensitivity

The KFS method aims at finding the location of supplementary observations that will lead to the greatest reduction of the variance given by Eq. (5); the details of mathematical developments can be found in Bergot and Doerenbecher (2002). The observation network is supposed to be composed of a routine, or conventional, network (denoted by suffix c) and of the targeted component (denoted by suffix t). The respective analysis-error covariance matrices are $\mathbf{A}_{\text{c}}$ and $\mathbf{A}_{\text{c+t}}$ depending on whether the targeted observation are assimilated or not. Using the Sherman-Morrison-Woodbury formula, $\mathbf{A}_{\text{c+t}}$ can be expressed as a function of $\mathbf{A}_{\text{c}}$:

$$\left.\begin{aligned} \mathbf{A}_{\text{c}} &= (\mathbf{I} - \mathbf{K}_{\text{c}} \mathbf{H}_{\text{c}}) \mathbf{B}, \\ \mathbf{A}_{\text{c+t}} &= (\mathbf{I} - \mathbf{K}_{\text{c+t}} \mathbf{H}_{\text{t}}) \mathbf{A}_{\text{c}} \end{aligned}\right\} \tag{6}$$

$$\begin{aligned} &\text{where} \quad \mathbf{K}_{\text{c}} = (\mathbf{B}^{-1} + \mathbf{H}_{\text{c}} \mathbf{R}_{\text{c}}^{-1} \mathbf{H}_{\text{c}}^{\text{T}})^{-1} \mathbf{H}_{\text{c}}^{\text{T}} \mathbf{R}_{\text{c}}^{-1} \\ &\text{and} \quad \mathbf{K}_{\text{c+t}} = (\mathbf{A}_{\text{c}}^{-1} + \mathbf{H}_{\text{t}} \mathbf{R}_{\text{t}}^{-1} \mathbf{H}_{\text{t}}^{\text{T}})^{-1} \mathbf{H}_{\text{t}}^{\text{T}} \mathbf{R}_{\text{t}}^{-1}. \end{aligned}$$

This leads to an expression for the reduction of forecast score variance due to the additional observations:

$$\begin{aligned} \delta\sigma_{S_{\text{t}}}^2 &= \sigma_{S_{\text{c}}}^2 - \sigma_{S_{\text{c+t}}}^2 \\ &= [\nabla S]^{\text{T}} \mathbf{A}_{\text{c}}^{\text{T}} \mathbf{H}_{\text{t}}^{\text{T}} (\mathbf{R}_{\text{t}} + \mathbf{H}_{\text{t}} \mathbf{A}_{\text{c}} \mathbf{H}_{\text{t}}^{\text{T}})^{-1} \mathbf{H}_{\text{t}} \mathbf{A}_{\text{c}} [\nabla S]. \end{aligned} \tag{7}$$

This expression ensures that when the optimal solution of the data assimilation problem is reached, the addition of any observations reduces the variance of the forecast score.

The minimization of the variance, or maximization of $\delta\sigma_{S_{\text{t}}}^2$, is a matrix optimization problem with $\mathbf{H}_{\text{t}}$ as a control variable. Advanced mathematical tools are required to solve this problem. As computation costs would be too high, a direct minimization is therefore discarded.

Instead, the KFS uses $\mathbf{A}_{\text{c}} \nabla S$, which is a sensitivity field derived from $\sigma_S^2$. At this stage there is no direct link between the additional observation impact and the value of the sensitivity. However, targeting becomes possible since the experience indicates that $\mathbf{A}_{\text{c}}$ gives enough information about VDAS and $\nabla S$ about the dynamics of the atmosphere. A more accurate expression of the sensitivity field is under investigation which will allow the computation of the variance reduction for each grid point without having to assimilate these additional observations.

The field $\mathbf{A}\nabla S$ is close to the expression of the observation sensitivity $\nabla_{\mathbf{y}^{\text{o}}} S = \mathbf{K}^{\text{T}} \nabla_{\mathbf{x}^{\text{a}}} S$ (Baker and Daley, 2000). As $\mathbf{K}^{\text{T}} = \mathbf{R}^{-1}\mathbf{H}\mathbf{A}$, the computation of the observation sensitivity from the KFS sensitivity field only needs projection into the observation space and weighting by the inverse of $\mathbf{R}$.

Examples of sensitivity fields computed with the KFS during the DTS Mediterranean Experiment (MEDEX) in Autumn 2009 are shown in Figures 1 and 2. The meteorological conditions of 9 November 2009 (strong winds over Catalonia and Corsica and heavy precipitation over the Algiers region) were used to compute all the sensitive areas $\mathbf{A}_{\text{c}} \nabla S$. This high-impact weather was related to cyclogenesis in the northwest of the Mediterranean Sea due to a cut-off low. The low was located over the British Isles 30 h prior to the cut off, and the 30 h optimized sensitive areas are located upstream of this low, in the middle of the Atlantic Ocean. This region matches the tip of the strong ridge which triggers the cut off in a wave-breaking process (G. Rivière, personal communication). The shapes of the sensitivity areas are determined by the dynamical processes, but also from the routine observation network and the $\mathbf{B}$ matrix used in ARPEGE. Thus, the location of the sensitivity maximum should not be interpreted in terms of dynamics only.

Despite very few differences between the two forecasts of the $Z500$ field depicted, the KFS computations do not yield the same result. The first sensitivity field (Figure 1) computed for the 42 h forecast shows the optimal targeted location near 20°W, close to the European coast. The 18 h forecast for the same targeting time gives an optimal location further west in the middle of the ocean, near 36°W. In Figure 2,The whole sensitivity area is shifted westward with respect to Figure 1. This longitude shift is shown in the upper-right panels of Figures 1 and 2; these panels depict vertical cross-sections of the 3D sensitivity field, overlaid with the potential vorticity

Copyright © 2011 Royal Meteorological Society

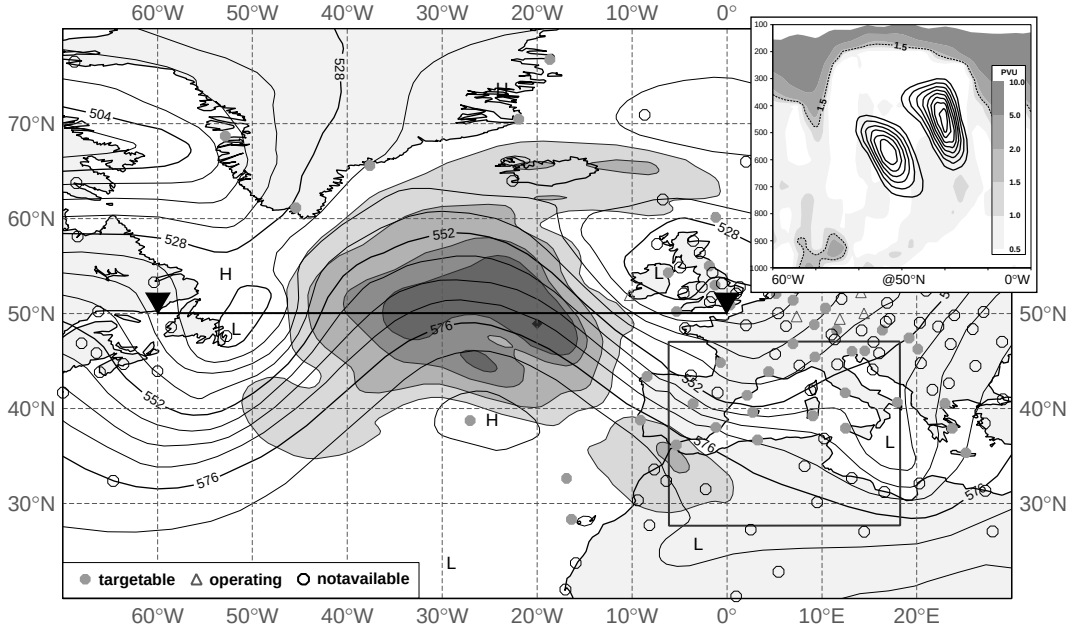



**Figure 1.** Optimised sensitive areas ($\mathbf{A}_c \nabla S$) computed at Météo-France/CNRM for the DTS-MEDEX project (case 1093). The targeting time is 1800 UTC on 7 November 2009, the 42 h forecast being based on the 0000 UTC analysis on 6 November. The shading denotes vertically integrated total energy and encloses areas of 8, 4, 2 and $1 \times 10^6$ km$^2$. The contours show the $Z500$ (gpdm) valid at the targeting time. The square box over the western Mediterranean shows the verification region valid 30 h later at 0000 UTC on 9 November. The various symbols show targetable, operating and non-available radiosonde sites. The diamond at 49°N, 20°W shows the position of maximum sensitivity. The bold horizontal line between the black triangles shows the axis of the vertical cross-section (longitude–pressure) in the upper-right panel. This panel shows sensitivity contours and the potential vorticity (shading, pvu) in the trajectory valid at the same time.

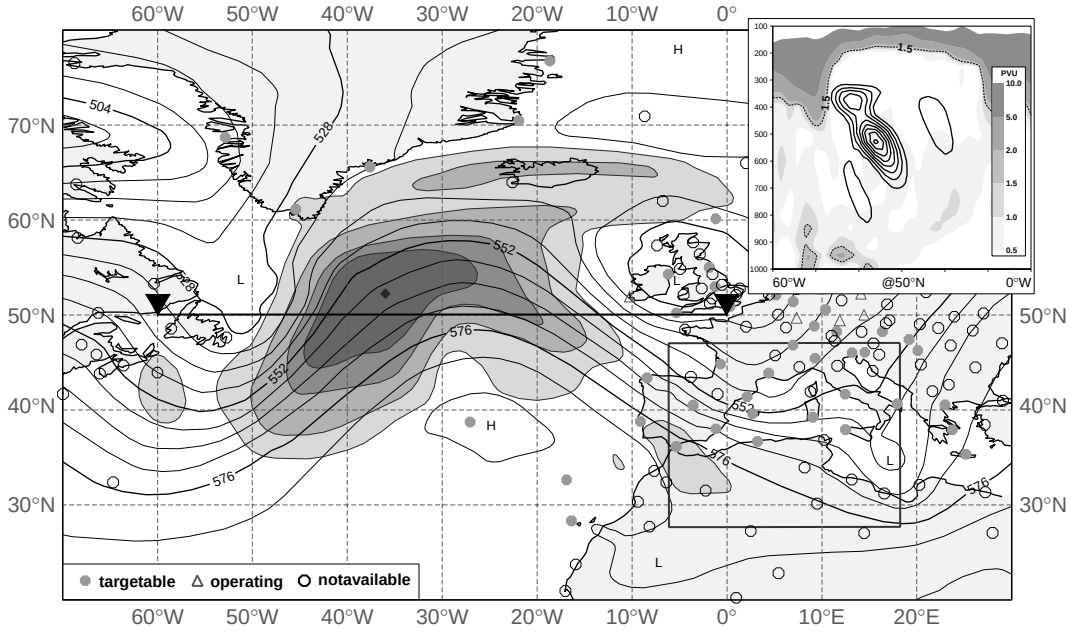



**Figure 2.** As Figure 1, but with a sensitivity computed from the 18 h forecast, based on the 0000 UTC analysis on 7 November 2009. The targeting time is 1800 UTC on 7 November 2009.

in the trajectory. The sensitivity field is made of two nodes. In the shorter-range forecast, the upstream node (on the left) has the highest amplitude (Figure 2), whereas in Figure 1 the downstream node has the highest amplitude. The dramatic differences exhibited by this example are the results of some complex interactions between the dynamics of the trajectory with both the data assimilation structure functions and the coverage of the assimilated routine observations (not shown). This interaction may enhance the single effect of the trajectory, but the consequence is that the location of the targeted observation is seen to be sensitive to the trajectory in this illustration.

In practice $S$ is defined within a verifying area focused on the event of interest. The adjoint model links the perturbations at $t_{\text{verif}}$ to their initial state at $t_{\text{obs}}$. The computation is made along the nonlinear trajectory defining the forecast. Thus the sensitive area is only effective for this particular forecast. This also holds for other techniques that use adjoint models to compute sensitive areas, such as the singular vector-based technique. By dealing only with the

Copyright © 2011 Royal Meteorological Society 

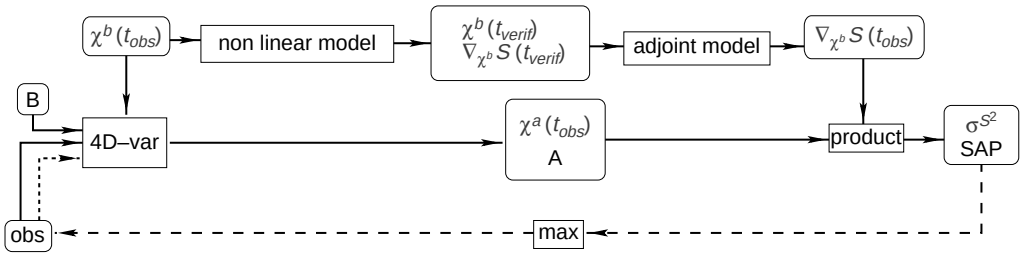



**Figure 3.** Steps for a sensitivity field computation (see text for details).

dynamical aspects, the sensitive area could already be well observed by routine observations. The purpose of weighting $\nabla S$ by the analysis-error covariance matrix $\mathbf{A}_c$ is to highlight areas subject to strong uncertainty, i.e. high values of error variance, where no observations are available. In the same way, areas well observed are associated with low values of analysis-error variance. The analysis-error covariance matrix can be written $\mathbf{A}_c = (\mathbf{I} - \mathbf{K}\mathbf{H}_c)\mathbf{B}$, observations reducing the variance near their locations.

To run the KFS method, it is necessary to compute $\mathbf{A}_c \nabla S$ and $\sigma^2_{S_c}$, followed by $\sigma^2_{S_{c+t}}$ once the targeted observation is chosen. The steps followed to achieve this are illustrated in Figure 3. $\nabla S$ is obtained by integrating backward the gradient of $S$ using the adjoint model along the initial forecast. Since this is a prognostic method, the values of the observations at $t_{\mathrm{obs}}$ cannot be known, nor can the routine network. Thus all these values are simulated, and the VDAS is only used to get an approximation of $\mathbf{A}_c$; the analysis field itself is not used. The sensitivity field to the initial conditions is thus computed from the background trajectory $\nabla_{\mathbf{x}^b} S$ instead of the one issuing from the analysis $\nabla_{\mathbf{x}^a} S$. The matrix $\mathbf{A}_c$ is obtained as a by-product during the minimization with the Lanczos algorithm (Fisher and Courtier, 1995).

The Lanczos algorithm gives an estimate of the singular vectors and singular values in a reduced rank base built from the descent vectors of the conjugate gradient. Trémolet (2008) proposed a method to compute the exact adjoint of **K** in the reduced rank base. In this article, we use a different method to compute the full ranked approximation of **A**. The subspace spanned by the descent vectors is completed with the background-error covariance matrix **B** (Fisher and Courtier, 1995).

## 3. Framework of the simulations

The KFS combines a dynamical model (here a two-layer QG model) with a VDAS (4D-Inc) to compute the sensitivity field. The QG model provides realistic nonlinear dynamical behaviour, and the strong initial error growth needed for an ensemble approach (as applied later in the article). The two-layer QG model allows such simulations with modest computation cost; it is described in subsection 3.1. Subsection 3.2 describes the components used for the 4D-Inc VDAS, **H**, **R** and **B** and its configuration. The experimental system is simple but designed to mimic suboptimal features present in operational NWP systems.

### 3.1. Dynamical model

The simulation is carried out using the numerical QG baroclinic two-layer model (Phillips, 1951). The equations are derived from the two-layer shallow-water equations, assuming that the ground and the tropopause are rigid. The development of the equations can be found in Pedlosky (1982). The filtering of the gravity waves leads to the potential vorticity $q$ coupled system:

$$q_1 = f_0 + \beta y + \Delta\psi_1 - \lambda^{-2}(\psi_1 - \psi_2), \quad (8)$$

$$q_2 = f_0 + \beta y + \Delta\psi_2 + \lambda^{-2}(\psi_1 - \psi_2), \quad (9)$$

$$\frac{\partial q_i}{\partial t} = -J(\psi_i, q_i), \text{ for } i \in \{1, 2\}, \quad (10)$$

where the upper/lower layer is denoted by subscript 1/2, $\psi_i$ is the stream function for the layer $i$, i.e. $\mathbf{u}_i = \left(-\partial_y\psi_i, \partial_x\psi_i\right)$. $J(\psi, q)$ is the Jacobian of $\psi$ and $q$ (with $J(A, B) = \partial_x A \partial_y B - \partial_y A \partial_x B$), $\lambda = \sqrt{gh}/f$ is the internal Rossby deformation radius (where $h$ is the height of the tropopause), and $f_0 + \beta y$ is the Coriolis parameter. The total flow is divided into a stationary basic state $\overline{q}$ and an anomaly $q'$, without making any hypothesis on their amplitudes. The basic state time evolution equation can be written $\partial_t \overline{q} + J(\overline{\psi}, \overline{q}) = F$, with $F$ a forcing term taken as $F = J(\overline{\psi}, \overline{q})$ in order to have the basic state stationary ($\partial_t \overline{q} = 0$). A dissipating term $D(q) = \nu_x \partial_x^6 q + \nu_y \partial_y^6 q$ is also added to ensure numerical stability.

The domain is rectangular in a $\beta$-plane and bi-periodic in $x$ and $y$. The model is quasi-spectral (Canuto *et al.*, 1988), with the nonlinear terms computed on a regular grid, using an elliptical truncation, $(k_x/T_x)^2 + (k_y/T_y)^2 \leqslant 1$, $k_x$ and $k_y$ being the wavenumbers. The time integration scheme is *leap-frog* with an Asselin filter (Asselin, 1972) modified following Williams (2009).

The basic state $\overline{q}$ represents a synoptic situation with baroclinicity associated with an eastward jet stream, assumed to be perfectly simulated. Therefore, no uncertainty is associated with the basic flow, i.e. **B** only takes into account the anomaly. The basic state is composed of a uniform zonal wind and a Gaussian profile jet stream at the centre of the domain. To have a large variety of scenarios in the growth of the anomaly, the jet is oscillating. The analytical form of the basic state is

$$U(x, y) = \overline{U}\left\{1 + \exp\left(\frac{y - \frac{L_y}{2} - \delta_y \cos\left(\frac{4\pi}{L_x}x\right)}{\Delta_y^2}\right)\right\},$$

with $\Delta_y = L_y/12$ the amplitude factor of the oscillations.

The anomalies are generated at each level with an eastward tilt for the lower one. They have the same Gaussian shape and amplitude, following the expression

$$q'_k = \mathcal{A} \exp\left\{\frac{(x - x_0 - \delta_{k2} d)^2}{\gamma^2} + \frac{(y - y_0)^2}{\gamma^2}\right\},$$

Copyright © 2011 Royal Meteorological Society

Table I. Parameters used in the model.

| Parameter | Symbol | Value | Units |
|---|---|---|---|
| Longitude dimension | $L_x$ | $16\,000 \times 10^3$ | m |
| Latitude dimension | $L_y$ | $8\,000 \times 10^3$ | m |
| Longitude truncation | $T_x$ | 64 | – |
| Latitude truncation | $T_y$ | 32 | – |
| Time step | d$t$ | 450 | s |
| Internal Rossby deformation radius | $\lambda$ | $300 \times 10^3$ | m |
| $\beta$ at 45 degrees | $\beta$ | $1.6 \times 10^{-11}$ | $\mathrm{m^{-1}s^{-1}}$ |
| Tri-Laplacian ($\nabla^6$) viscosity on $x$ | $\nu_x$ | $6 \times 10^{22}$ | $\mathrm{m^6s^{-1}}$ |
| Tri-Laplacian ($\nabla^6$) viscosity on $y$ | $\nu_y$ | $6 \times 10^{20}$ | $\mathrm{m^6s^{-1}}$ |
| Basic state wind amplitude (upper layer) | $\overline{U}_1$ | 17 | $\mathrm{m\,s^{-1}}$ |
| Basic state wind amplitude (lower layer) | $\overline{U}_2$ | 8.5 | $\mathrm{m\,s^{-1}}$ |
| Asselin filter parameter | $\varepsilon$ | 0.01 | – |
| Asselin filter parameter | $\alpha$ | 0.53 | – |

$(x_0, y_0)$ being the centre of the Gaussian, $\gamma$ the spread of the Gaussian and $d$ the tilt between the two layers. The values used are $\mathcal{A} = 2 \times 10^{-5}\mathrm{s}^{-1}$, $\gamma = 700$ km and $d = 500$ km. $\delta_{ij}$ is the Kronecker symbol. The values of the parameters used for the simulations are displayed in Table I.

Figure 4 shows a simulation of a deepening cyclone associated with a baroclinic interaction between a potential vorticity anomaly and a baroclinic basic state. The Gaussian anomaly is transported by the jet stream, leading to the appearance of a downstream development after 24 h. Once the dipole is formed, the baroclinic interaction feeds the anomaly (48 h) and a lambda-shaped feature (i.e. a weather perturbation) can be seen. With the growth of the anomaly, the nonlinearities elongate the structure (72 h and 96 h). The growth of the anomaly is dependent on its initial location and amplitude, making allowances for strong initial error growth.

### 3.2. *Components of the OSSE*

The study aims to evaluate the influence of the trajectory on the sensitivity area computed by the KFS. The theoretical background has been described in sections 2.3 and 3.1, and we now describe the numerical experiments. To explicitly measure the variances of the forecast score, an OSSE is set up. The background-error covariance matrix is modelled, the observation network is spatially homogeneous and the observation values are obtained from a reference nature run. As the observed parameter equals the potential vorticity, the observation error $\varepsilon^{\mathrm{o}}$ is framed so as to be compatible with $\mathbf{R}$; $\sigma^{\mathrm{o}2}$ is set at $1 \times 10^{-11}\mathrm{s}^{-2}$.

The reference run, i.e. the truth for the observations, is chosen close to but slightly shifted from the ensemble mean anomaly position and it is supposed that the potential vorticity can be directly observed. The ensemble of initial conditions is generated randomly depending on the experiment. For each member, a two-day forecast ($t_{\mathrm{verif}}$) is computed to get $\nabla S$ and the initial condition is assimilated using the observations in order to have the approximation of $\mathbf{A}_{\mathrm{c}}$ by the Lanczos algorithm. Once the sensitivity field and $\sigma_S^{\mathrm{c}2}$ are computed, the targeted observations are deployed and a second assimilation is carried out to get $\mathbf{A}_{\mathrm{c+t}}$. $\sigma_S^{(\mathrm{c+t})^2}$ can then be computed and the effectiveness of the target observations assessed.

The assimilation window is 3 h long with two series of simultaneous observations at 0 h and 3 h, and the network is supposed unchanged. During the minimization, only one outer loop is made with a convergence criterion for the norm of the gradient $\nabla J$ of $10^{-9}$; a maximum of 15 iterations is imposed for the inner loop. The enstrophy on the bottom layer is chosen as the forecast score $S$; the gradient of $S$ is $\nabla_{\mathbf{x}^{\mathrm{f}}} S = \xi_2$. The verifying area is chosen wide enough to enclose the weather feature of interest with all members; it is shown on Figure 5. The strategy for the targeted observation is to add a single observation, of the same quality as the routine network, at the maximum of $\mathbf{A}_{\mathrm{c}} \nabla S$ for each layer.

The parameters presented in Table II are common to all the experiments. To evaluate the influence of the trajectory, several components of the data assimilation systems have been used. The map of the background-error variance $\sigma^{\mathrm{b}2}$ is related to the kind of perturbation that is sought (anomaly only or anomaly and jet stream). The background-error covariance matrix $\mathbf{B}$ is modelled following Fisher (2003) as $\mathbf{B} = \Sigma \mathbf{C} \Sigma^{\mathrm{T}}$, with $\Sigma$ mapping the root of $\sigma^{\mathrm{b}2}$ in grid points and $\mathbf{C}$ the spectrum of the correlation function in Fourier space; both matrices are diagonal in their respective spaces. This leads to a homogeneous correlation function but heterogeneous covariances. The correlation function is taken as Gaussian in the two directions with length-scales $L_{px} = L_x/20$ and $L_{py} = L_y/15$. The $\sigma^{\mathrm{b}2}$ map is built from the diagnosis of the ensemble initial variance. Two routine networks were designed, a homogeneous one *ResH* which did not put any emphasis on any location of the domain, and a second one *ResD* receiving more observations over a band at the level of the jet (Figure 5).

## 4. Results and comments

In order to assess the influence of the trajectory, which depends on the components of the VDAS, i.e. the routine network and $\mathbf{B}$, a first experiment including three ensemble simulations is described. A targeting strategy that considers the ensemble of sensitivity fields is then proposed.

Copyright © 2011 Royal Meteorological Society

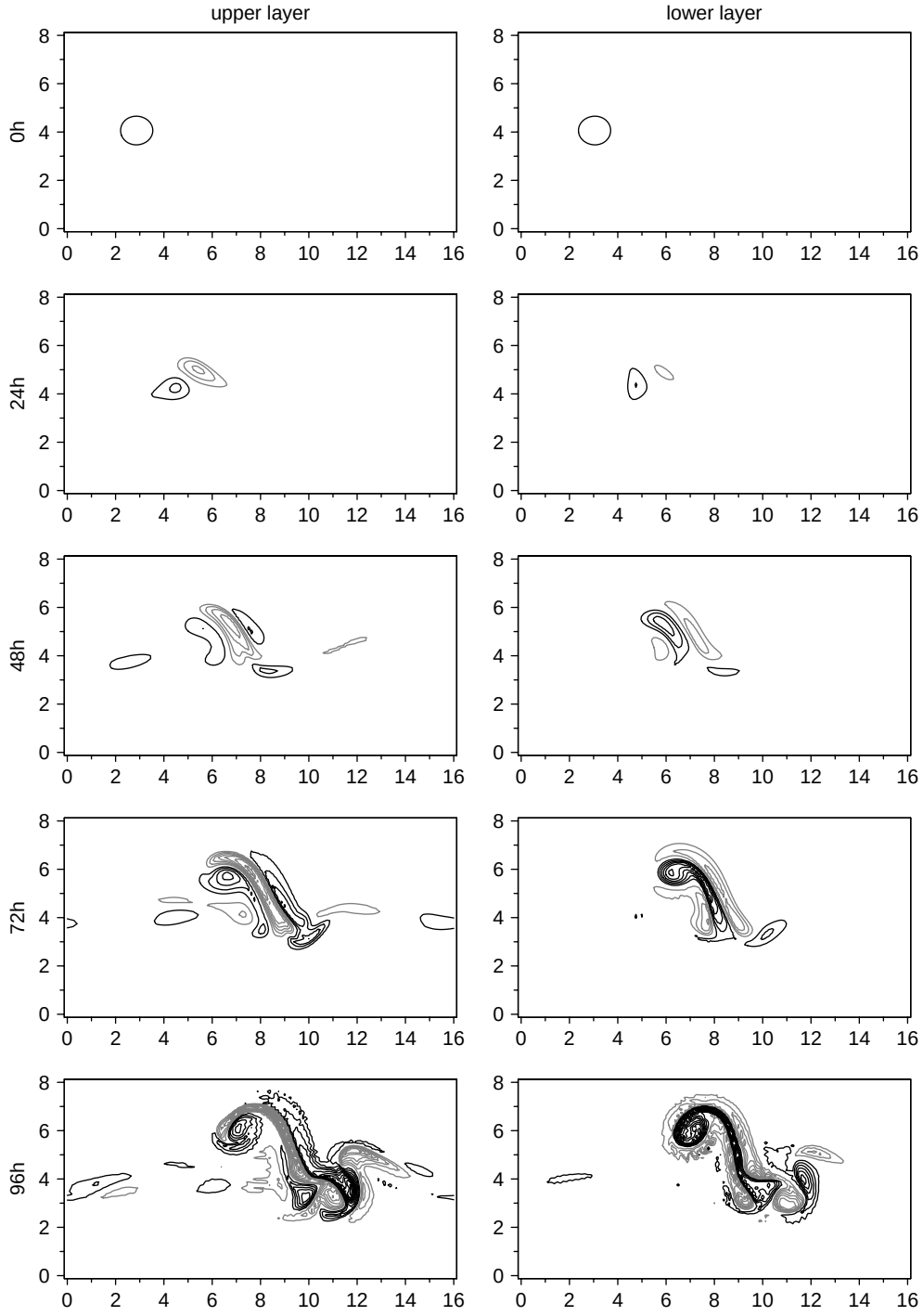



**Figure 4.** Time evolution of the vorticity anomaly over a 4 day integration. The positive values are in black and the negative ones in grey, with contour interval $2 \times 10^{-5} s^{-1}$. The unit of the $x$ and $y$ axes is $10^6$ km.

### 4.1. *Influence of trajectory and VDAS components (Exp I)*

Three experiments have been investigated to compare the influence on the KFS results of the trajectory and of the components of the VDAS. The aim is to relax the spatial constraints on the components for each experiments, i.e. the routine observations locations and the background-error variance map. The experiment configurations are summarized in Table II. The localized map of variance is shown in Figure 5, with the uniform variance given the value $1 \times 10^{-11}$. Figure 5 also shows the routine network.

In these experiments the perturbation for each member is generated by randomly shifting the anomaly in $x$ and $y$.

Table II. Experiment configurations: observation network and background-error variance pattern.

| | Observation network | $\sigma^{b2}$ |
|---|---|---|
| Exp I.a | *ResD* | localized |
| Exp I.b | *ResH* | localized |
| Exp I.c | *ResH* | uniform |

The amplitude of the anomaly is randomly perturbed by a factor following a normal distribution. This emphasizes

Copyright © 2011 Royal Meteorological Society

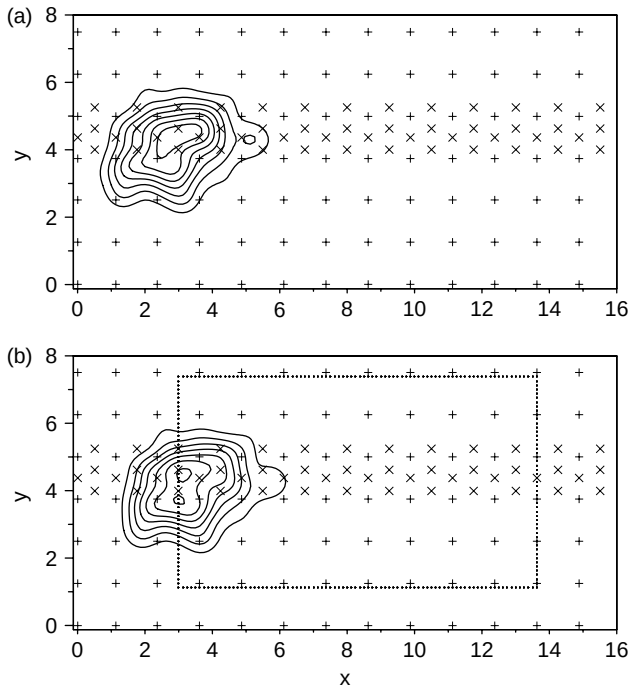



**Figure 5.** Position of the observations for the routine networks *ResD* (for the denser network) and *ResH* (for the homogeneous network), the 'X' points are the additional observations of *ResD* with respect to *ResH*. The $x$ and $y$ scales are expressed in $10^3$km. The field is the background-error variance map (a) for the upper layer and (b) the lower layer. The contour interval is $3\times10^{-11}\mathrm{s}^{-2}$. The dashed line shows the limit of the verification area.

the growth of the nonlinearities. This kind of perturbation allows many different trajectories in the ensemble. Each ensemble includes 50 members.

#### 4.1.1. *Ensemble averaged fields*

The first results shown are calculated from the ensemble averaged fields. Figures 6, 7 and 8 show the composite mean gradient of forecast score $\widetilde{\nabla S}$, the composite mean gradient weighted by the background-error covariance matrix $\mathbf{B}\widetilde{\nabla S}$, the composite mean sensitivity field $\widetilde{\mathbf{A}_c\nabla S}$ and the distribution of the targeted observation from each member for the upper level. The intensity of the grey scale is related to the value of the sensitivity field. The composite fields $\widetilde{X}$, that mix the fields at both levels, are computed from $\widetilde{f}(x,y)=\sqrt{f_1(x,y)^2+f_2(x,y)^2}$.

$\widetilde{\nabla S}$ gives information on the dynamics of the trajectories. The field $\widetilde{\nabla S}$ is focused on the area of the initial anomalies and is consistent between the three experiments (Figures 6(a), 7(a) and 8(a)). When looking at some members in particular, it is striking to see that the highest values are associated with the jet, even for initial conditions that are far away from it (not shown here). When the localized $\sigma^{b^2}$ field is used (Figures 6(b) and 7(b)) the strongest value of $\mathbf{B}\widetilde{\nabla S}$ is found at the maximum of the variance. In such a case, before any observation is assimilated, the sensitivity field is determined by the background-variance error map. The assimilation of observations induces a modification in the map of the error variance, Figure 6(c). The information from the observations reduces the uncertainties and thus the analysis-error variance. The area of high sensitivity is displaced to the edge of the band of relatively dense observations (out of $40$–$55\times10^3$ km in $y$). The southern border is emphasized by the higher values of $\nabla S$ due to the presence of the jet.

When the *ResH* is used (Figure 7(c)), the sensitive area is found at the maximum of $\sigma^{b^2}$ in the valley between the lines of observations (at $38\times10^3$ km and $50\times10^3$ km in $y$). In the final experiment I.c (Figure 8), the field of $\sigma^{b^2}$ is assumed uniform, and the sensitive area location is then driven only by the dynamical aspects and the routine network positions. As the values of $\sigma^{b^2}$ are lower in this case than the other ones in the area of interest, the values of $\widetilde{\mathbf{A}_c\nabla S}$ follow the same pattern. The impact of the routine network is less obvious than in Exp I.b, but the sensitivity gradient is lower. Further confirmation is given by the position of the targeted observations.

#### 4.1.2. *Distribution of targeted observations*

The distribution of targeted observations is shown in Figures 6(d), 7(d) and 8(d). They illustrate the variability between the members and thus the influence of the trajectory. When the variability of the maximum values of $\mathbf{A}\nabla S$ is small, the trajectory has little influence on the sensitivity field. The values of Exp I.a and I.b are similar and distributed over the sparsely observed areas; this is consistent with the discussion in 4.1.1. The largest values of $\mathbf{A}\nabla S$ are located at the maximum of $\widetilde{\mathbf{A}_c\nabla S}$. As expected, for Exp I.c the spread of values is larger and no specific area of large values is associated with the maximum of the mean sensitivity field. Several larger values are found near the point $(50, 30)$ where the jet stream is larger (not shown). The interaction between the anomaly and the jet stream is increased when the anomaly is located south of the jet stream. The area of high sensitivity is also located there. The distribution of the initial conditions enables some anomalies to get an advantageous position, leading to these strong values of sensitivity.

Figure 9 shows the correlation between the maximum of the sensitivity fields for the both levels together and the estimated variance $\sigma^2_{S_c}=\nabla S^{\mathrm{T}}\mathbf{A}_c\nabla S$, and between the maximum and the estimated variance reduction $\delta\sigma^2_{S_t}=\sigma^2_{S_{c+t}}-\sigma^2_{S_c}$. The plotted values are those from Exp I.c, but similar graphs result from the other experiments (not shown). As the expression of the variance of the score function and the sensitivity field are similar, the values appear correlated. According to the graph, an observation deployed in an area of strong sensitivity will lead to a more effective reduction of $\sigma_S$. This is in agreement with the deployment of the targeted observation in regions of high values of sensitivity field $\mathbf{A}_c\nabla S$. But, as can be seen in Figures 6(d), 7(d) and 8(d), such areas can be very trajectory dependent. According to the selected member, the optimal location can change as well as its expected efficiency.

#### 4.1.3. *Discussion*

The analysis of the average fields allows the discrimination of the different contributions to the sensitivity field. The routine network and the modelling of $\mathbf{B}$ are strong constraints on it. The action of $\mathbf{A}_c$ on the gradient $\nabla S$ emphasizes the areas of strong analysis-error variance values. In these experiments the strong values of $\nabla S$ and $\sigma^{b^2}$ are roughly collocated, thus the contribution of each cannot be

Copyright © 2011 Royal Meteorological Society

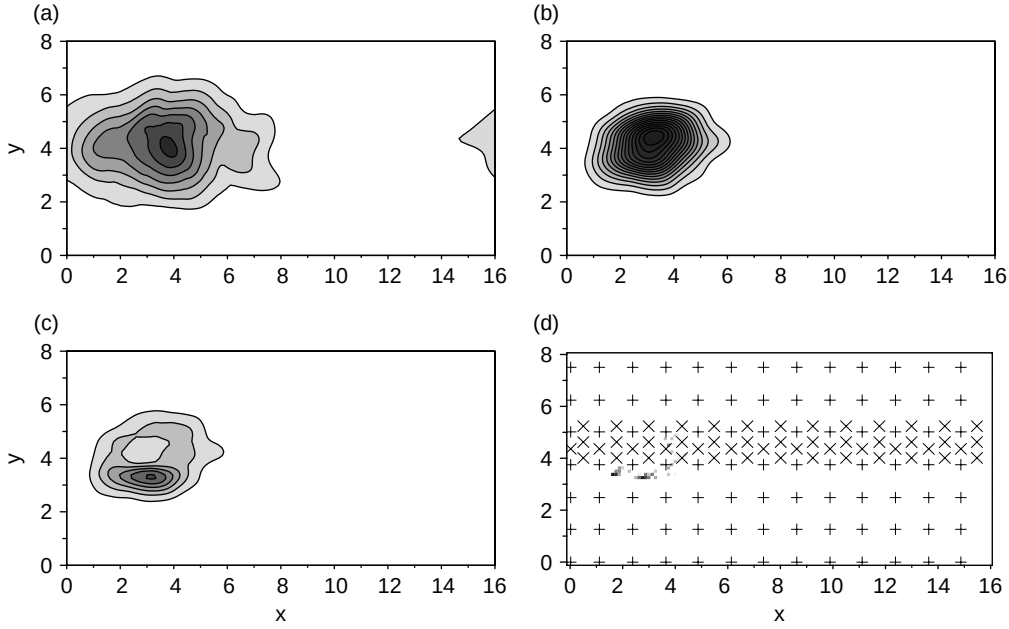



**Figure 6.** Composite field for Exp I.a. (a) forecast score gradient ensemble mean of $\nabla S$ at $t_{obs}$, (b) ensemble mean of $\mathbf{B}\nabla S$, (c) ensemble mean of $\mathbf{A}_c\nabla S$, and (d) position of the targeted observations for each member at the upper level. The grey colour scale is proportional to the intensity of the sensitivity field. The routine network is plotted as in Figure 5. The contour interval is $1\times10^{-8}$ for (a) and $5\times10^{-13}$ for the other fields.

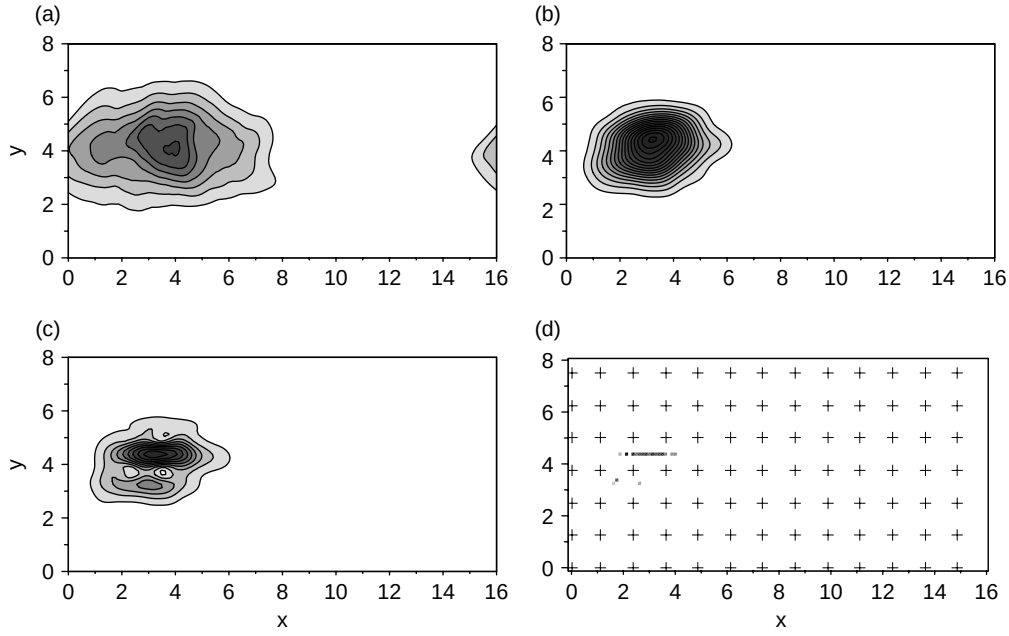



**Figure 7.** As Figure 6, but for Exp I.b.

exactly estimated. Other tests have shown that errors in the $\sigma^{b^2}$ field can lead to sensitivity fields located far away from any verifying area.

These experiments show that the sensitivity fields derived from the KFS are strongly constrained by both the mapping of the initial uncertainty $\sigma^{b^2}$ and the conventional observing network $\mathbf{H}_c$. The spread of the targeted observations locations is bigger when the number of conventional observations is reduced or when the initial uncertainty on the trajectories is more uniform. In such cases, the shape of $\mathbf{A}_c$ is less constraining on the KFS and the trajectory influences more strongly the location of the sensitivity field estimated by the KFS. Yet the distribution targeted observation shows that when of the constraints are applied on the VDAS; this leads to the values of sensitivity being mainly determined by the trajectory. Thus, depending on constraints on the VDAS, the trajectory will influence the location of the targeted observation or the KFS values.

As sensitivity values are strongly correlated to the expected reductions, the choice of a particular trajectory can lead to different deployment strategies. With regards to the selected member, a small displacement of the additional observation may lead to strong variance reduction or may have no impact. This may also be valuable for the observation impact computation, even if the use of a second-order development using the forecast and the analysis tends to balance the trajectory influence.

Copyright © 2011 Royal Meteorological Society 

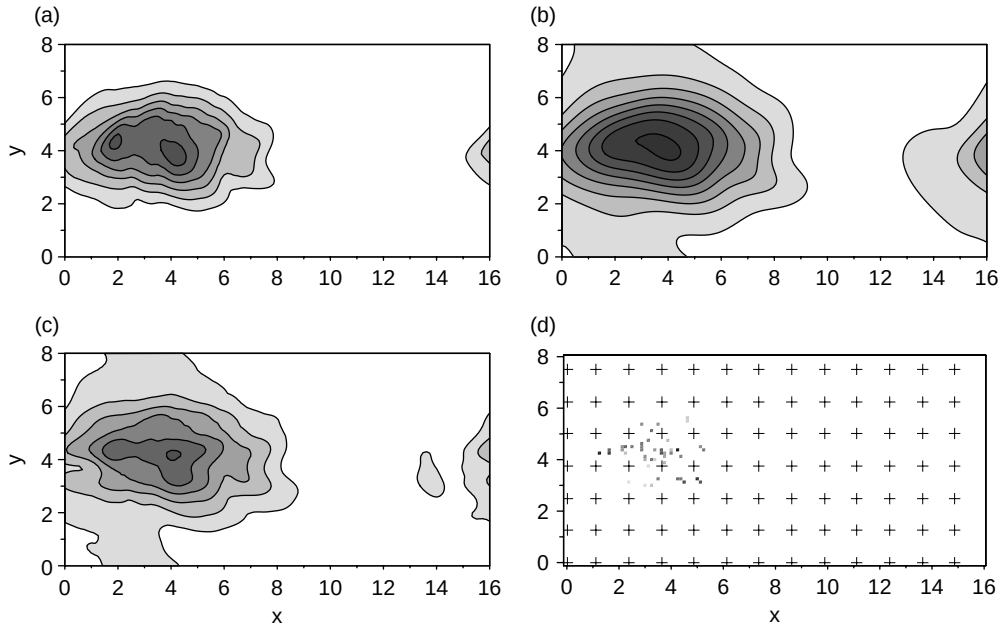



**Figure 8.** As Figure 6, but for Exp I.c. The contour interval is $5 \times 10^{-14}$ for (b), (c) and (d).

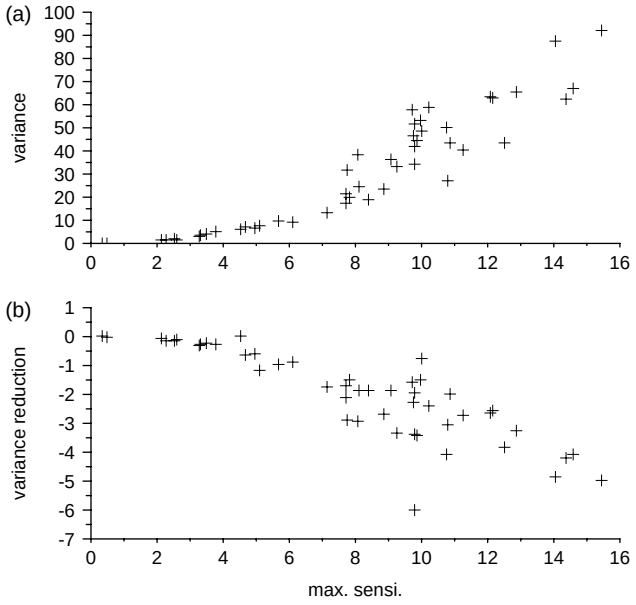



**Figure 9.** Correlation (a) between the maximum value and the estimated forecast score variance, and (b) between the maximum value and the estimated forecast score variance reduction. Each cross represents the value of a member. Sensitivity values are multiplied by $10^{13}$ and the variance by $10^{18}$.

The purpose of Exp I is primarily to assess the influence of the trajectory on the sensitivity field. The information retrieved from the ensemble was not used to this end. Results of Exp I showed that the dynamics of the simulations and the forecast score gradient are driven by the jet stream. However, in the DAS configuration for Exp I, the jet stream is not part of the state variable (neither the observed nor the assimilated variables). As a consequence, the location of the targeted observation obtained with some members of the ensemble lies in regions where the observation values are very close to zero. Such very low value observations are not realistic and are useless for assessing the efficiency of the targeted observations. To alleviate this shortcoming in the second experiment, a part of the jet stream is included in the state variable.

### 4.2. *Ensemble-based targeting strategy (Exp II)*

For this experiment the anomaly and the included part of the jet stream are shifted as in Exp I. A new **B** matrix is computed to include the uncertainty of the jet stream. The routine network *ResH* is used in order to keep some spread after the assimilation. One hundred members are used here to reduce the sampling noise. For this experiment, an ensemble assimilation system is built by adding Gaussian noise to the observations for each member, proportional to the **R** matrix ($\varepsilon^{\mathrm{o}} = \mathbf{R}\delta$ with $\delta \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$) (Burgers, 1998). This ensemble data assimilation system is used to assess the effectiveness of various targeting strategies based on the sensitivity field ensemble.

In Eq. (5), the uncertainty of the trajectory is neglected, i.e. $\mathbb{E}(\nabla_{\mathbf{x}^{\mathrm{a}}+\varepsilon^{\mathrm{a}}} S) = \nabla_{\mathbf{x}^{\mathrm{a}}} S$. The forecast score variance can be rewritten as a law of conditional probability to take into account the dependence on the trajectory:

$$\begin{aligned} \sigma_S^2 &= \mathbb{E}\left\{\mathbb{E}(\delta S\, \delta S^{\mathrm{T}} | \mathbf{x}^{\mathrm{a}})\right\} \\ &= \mathbb{E}\left\{[\nabla_{\mathbf{x}^{\mathrm{a}}} S]^{\mathrm{T}} \mathbb{E}(\varepsilon^{\mathrm{a}}\, \varepsilon^{\mathrm{aT}})[\nabla_{\mathbf{x}^{\mathrm{a}}} S]\right\} \\ &= \mathbb{E}\left\{[\nabla_{\mathbf{x}^{\mathrm{a}}} S]^{\mathrm{T}} \mathbf{A} [\nabla_{\mathbf{x}^{\mathrm{a}}} S]\right\}. \end{aligned} \tag{11}$$

$[\nabla_{\mathbf{x}^{\mathrm{a}}} S]^{\mathrm{T}} \mathbf{A} [\nabla_{\mathbf{x}^{\mathrm{a}}} S]$ is now considered as a stochastic variable. By using the ensemble in a Monte-Carlo method, the expectation is computed as the ensemble mean. The associated sensitivity field now reads $\mathbb{E}(\mathbf{A}[\nabla_{\mathbf{x}^{\mathrm{a}}} S])$, being the mean of all sensitivity fields. The optimal supplementary observation is given by the maximum value of $\mathbb{E}(\mathbf{A}[\nabla_{\mathbf{x}^{\mathrm{a}}} S])$, which is called the 'consensus observation'. This method is really easy to implement as soon as a single sensitivity can be computed and is naturally parallelizable. The interpretation of the mean sensitivity field is also straightforward. The high values can be associated with a very strong sensitivity for one member or with an area of high values present in several members. Both cases are worth investigating; the first one

Copyright © 2011 Royal Meteorological Society 

Table III. Deployment strategy comparison. The values are the fraction of the total variance without supplementary observations.

| | $\sigma_c^2$ | $\sigma_{c+t}^2$ | $\sigma_{c+m}^2$ | $\sigma_{c+r}^2$ |
|---|---|---|---|---|
| Initial time | 1 | 0.95 | 0.90 | 1.01 |
| Two-day forecast | 1 | 0.97 | 0.94 | 1.009 |

Subscript t denotes the targeted observation of each member, m the consensus and r the random observation.
The reference value at initial time $t_0 = 26 \times 10^{-8} s^{-2}$ and at $t_{verif} = 374 \times 10^{-8} s^{-2}$.

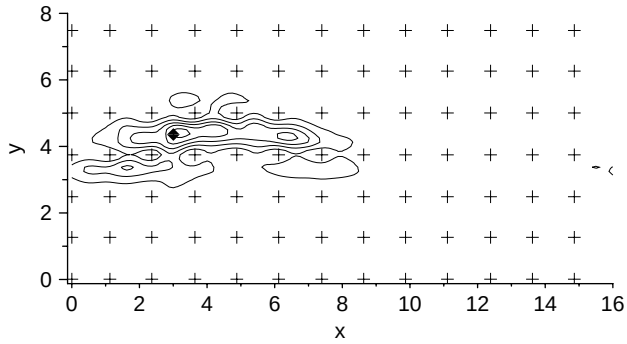



**Figure 10.** The location of the consensus observation at the upper level (diamond). The crosses are the routine observations and the contours show the analysis-error variance with interval $5 \times 10^{-11} s^{-2}$.

can be associated with a severe event scenario and the second with a more likely event.

To assess the effectiveness of the targeting strategy, one observation per level is added and the total potential vorticity variance serves as the selected measure, i.e. the sum variances for each grid point. Three different strategies are compared; in the first one, each member uses its own targeted observations (c+t). This case cannot be deployed operationally and is used as reference. In the second case, the consensus observation (c+m) is taken for each level. The third strategy uses one random supplementary observation (c+r) per level. The results are shown in Table III. In both cases of targeted observation c+t and c+m, the variance of the ensemble is reduced at the initial time and at $t_{verif}$. The randomly added observations have no impact on the ensemble spread. Out of the three strategies, the consensus is the most efficient for reducing the ensemble variance. The comparison of two computational moments shows that the impact of the targeted observation is reduced with increasing integration time.

Figure 10 shows the location of the consensus observation at the upper level; the results are very similar for the lower level. In these experiments the observations are located very close to the maximum of the analysis-error variance. The dynamics do not have the main influence on the sensitivity field here and only the uncertainty of the analysis drives the sensitivity field. By adding an observation where the members are most different, it is clear that this is the most effective strategy to reduce the variance shown in Table III.

## 5. Conclusion

The Kalman Filter Sensitivity (KFS) method combines information about the atmospheric physics and dynamics from the adjoint model and from the trajectory, together with information from the Variational Data Assimilation System (VDAS) and observation networks, to predict sensitive areas. When considering a quasi-operational NWP system (or at least high-dimensional ones), this kind of computation needs many assumptions, which cannot always be thoroughly verified. The prediction of the routine network and the accurate definition of the background-error covariance matrix are still difficult to make and the influence of the trajectory remains unknown.

The influence of the trajectory on the fields given by the KFS was studied in a rather small system with which several numerical experiments have been carried out. The results of the first experiment presented here (Exp I) show the leading role of the VDAS components in controlling the spread of the predicted location for additional targeted observations among the ensemble members. When the highest values of the dynamical sensitivity field $\nabla_{\mathbf{x}^a} S$ are already well observed, i.e. we have a good sampling of the region with routine observations, the sensitivity signal is driven by the VDAS components. Thus, there is a good agreement on the location of the supplementary observations. When the sensitivity field has fewer routine observations, it is the background-error variance field which drives the location of the additional observations. The less constraints there are on the VDAS components, the stronger is the influence of the trajectory on the behaviour of the targeted observations. Thus, the forecast of the routine observing network and the background-error covariance matrix are key elements of the quality of the sensitivity fields.

The experiments also showed that the maximum value of the sensitivity field was correlated with the estimation of the forecast score variance and its expected reduction. As the maximum value of the sensitivity field is strongly member-dependent, the choice of a deployment strategy according to a single trajectory is problematic.

The second experiment (Exp II) proposed a method to extract the information given by the ensemble of sensitivity fields. Results showed that the KFS was able to compute useful targeting strategies such as the 'consensus observation'. This strategy is easy to implement and leads to better results than the two other strategies investigated in this work. Another benefit from the 'consensus observation' method is that it can be repeated once the first targeted observation has been chosen. An optimal observation network can be computed iteratively to account for all the observations available.

The accuracy of the method should be further investigated. As has been mentioned, several assumptions are needed to compute the KFS. Among them are the first-order forecast score Taylor expansion ($dS = \nabla_{\mathbf{x}^f} S^T \varepsilon^f$) or the linear transport of the perturbations ($\|\varepsilon^f\| << \|\mathbf{x}^f\|$). The use of a second-order expansion proved to be more accurate in observation impact studies (Errico, 2007), but is harder to achieve in a prognostic framework.

Accounting for the trajectory variability is even more important when the adaptive observing platform is not fixed or only available for a very short time, but when it follows the flow for several days, as with drifting sondes. The targeted deployment must be evaluated considering all the observations which would be made by the platform along its own trajectory. Such tools dealing with the time-integrated sensitivity field have to be investigated. Measurement campaigns, such as

Copyright © 2011 Royal Meteorological Society 

Concordiasi (http://www.cnrm.meteo.fr/concordiasi/) or the Hydrological Cycle in the Mediterranean Experiment (HyMeX; http://www.hymex.org) involving, respectively, drifting sondes and boundary-layer super-pressurized balloons (BLPBs), could benefit greatly from them.

During the DTS-MEDEX 2009 field experiments, several targeting methods were used at Météo-France, such as a modified ETKF (Bishop *et al.*, 2001) and KFS. The ensemble-based and adjoint-based techniques lead to quite different sensitive areas. Another ensemble-based method, called Ensemble Sensitivity (ES; Ancell and Hakim, 2007), aims to compute the same field as the KFS. The gradient $\nabla_{\mathbf{x}^a} S$ is computed by linear regression and without an adjoint model. It would be of interest to investigate the mean of the KFS ensemble, the ETKF summary maps and the ES fields.

A recent study (Gelaro *et al.*, 2010) showed that relatively few observations (around 53%) bring useful information to the assimilation process. Such results are taken from experiments considering the whole observation network and no specific meteorological phenomena, without any data targeting concern. In the case of targeting particular weather aspects, Cardinali *et al.*(2007) showed that the optimally located observations improve the forecast. To progress further we expect any targeted observation to bring significant improvement. This also applies to limited-area models and mesoscale predictions where fewer conventional observations are available.

**Acknowledgements**

The authors would like to thank the HYMEX/BAMED project and the French Agence Nationale de la Recherche (ANR), under grant Geo-FLUIDS (ANR-09-SYSC-005-01) 'Analyse et simulation découlements fluides à partir de séquences d'images: application à latitude découlements géophysiques'. The authors acknowledge Gwendal Rivière for his contributions on interpreting the two-layer QG model behaviour, and also Nadia Fourrie, Pierrick Cébron, Jean-Antoine Maziejewski and William Lahoz for their comments on an earlier version of the article. Finally, the anonymous reviewers are thanked for their helpful comments.

## Appendix

*Estimation of analysis-error variance by the Lanczos algorithm.*

The analysis-error variance field computed by the Lanczos algorithm is compared to the variance computed by an assimilation ensemble. The ensemble is here built according to the background-error covariance matrix used in the VDAS, with Gaussian perturbation of the form $\varepsilon^{\mathrm{b}} = \mathbf{B}^{1/2}\eta$ (Fisher and Courtier, 1995) with $\eta \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{I}$ the identity matrix. The two analysis-error variance fields are shown in Figure A.1. The structure of the *ResH* can be recognized in the two fields. Even if the values are not exactly the same, the shapes of the fields are quite similar and would lead to the same sensitivity field. The comparison between variance fields issued from different trajectories shows good robustness of the method regarding the values of the increment $\mathbf{d}$ (not shown here). This fits well with the prognostic framework of the adaptive observations where these values are not known.

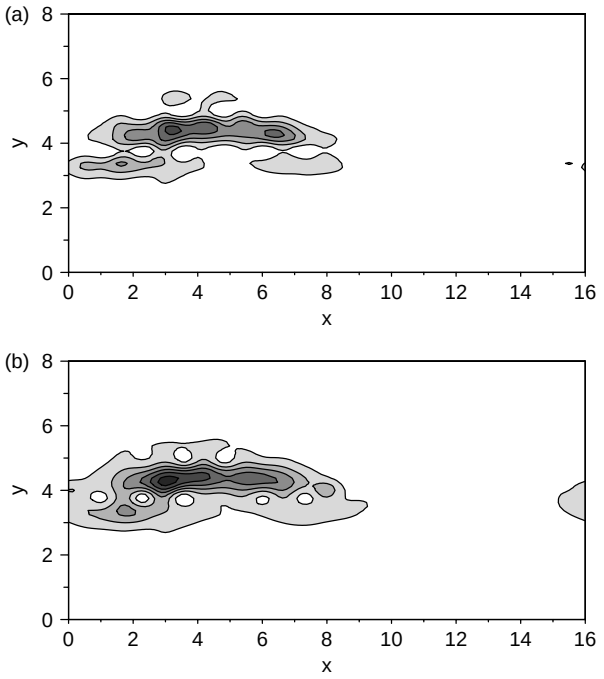



**Figure A.1.** Upper-layer analysis-error variance computed (a) from the analysis ensemble and (b) by the Lanczos algorithm. The holes are due to the observations. The contour interval is $5 \times 10^{-11}\mathrm{s}^{-2}$.

## References

Ancell B, Hakim GJ. 2007. Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Weather Rev.* **135**: 4117–4134.

Asselin R. 1972. Frequency filter for time integrations. *Mon. Weather Rev.* **100**: 487–490.

Baker NL, Daley R. 2000. Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Meteorol. Soc.* **126**: 1431–1454.

Bergot T, Doerenbecher A. 2002. A study on the optimization of the deployment of targeted observations using adjoint-based methods. *Q. J. R. Meteorol. Soc.* **128**: 1689–1712.

Bishop CH, Etherton BJ, Majumdar SJ. 2001. Adaptive sampling with the Ensemble Transform Kalman Filter. Part I: Theoretical aspects. *Mon. Weather Rev.* **129**: 420–436.

Burgers G. 1998. On the analysis scheme in the Ensemble Kalman Filter. *Mon. Weather Rev.* **126**: 1719–1724.

Canuto C, Hussaini M, Quarteroni A, Zang T. 1988. *Spectral Methods in Fluid Dynamics*. Springer Verlag: Berlin and New York.

Cardinali C. 2009. Monitoring the observation impact on the short-range forecast. *Q. J. R. Meteorol. Soc.* **135**: 239–250.

Cardinali C, Buizza R, Kelly G, Shapiro M, Thépaut J-N. 2007. The value of observations. III: Influence of weather regimes on targeting. *Q. J. R. Meteorol. Soc.* **133**: 1833–1842.

Courtier P, Thépaut J-N, Hollingsworth A. 1994. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.* **120**: 1367–1387.

Daescu DN. 2008. On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation. *Mon. Weather Rev.* **136**: 3050–3065.

Errico RM. 2007. Interpretations of an adjoint-derived observational impact measure. *Tellus A* **59**: 273–276.

Fisher M. 2003. 'Background-error covariance modelling'. In *Proceedings of the Seminar on Recent developments in data assimilation for atmosphere and ocean*, 8–12 September 2003. ECMWF: Reading, UK. 45–63.

Fisher M, Courtier P. 1995. 'Estimating the covariance matrices of analysis and forecast error in variational data assimilation'. Tech. Memo. 220. ECMWF: Reading, UK.

Gelaro R, Zhu Y. 2009. Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus A*, **61**: 179–193.

Gelaro R, Langland RH, Pellerin S, Todling R. 2010. The THORPEX observation impact intercomparison experiment. *Mon. Weather Rev.* **138**: 4009–4025.

Copyright © 2011 Royal Meteorological Society

Gilbert J-C, Lemaréchal C. 1989. Some numerical experiments with variable storage quasi-Newton algorithms. *Mathematical Programming*.
Ide K, Courtier P, Ghil M, Lorenc A. 1997. Unified notation for data assimilation: Operational, sequential and variational. *J. Meteorol. Soc. Japan* **75**: 181–189.
Palmer TN, Gelaro R, Barkmeijer J, Buizza R. 1998. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* **55**: 633–653.
Pedlosky J. 1982. *Geophysical fluid dynamics*. Springer-Verlag: New York and Berlin.
Phillips NA. 1951. A simple three-dimensional model for the study of large-scale extratropical flow patterns. *J. Atmos. Sci.* **8**: 381–394.
Trémolet Y. 2008. Computation of observation sensitivity and observation impact in incremental variational data assimilation. *Tellus A* **60**: 964–978.
Williams PD. 2009. A proposed modification to the Robert–Asselin time filter. *Mon. Weather Rev.* **137**: 2538–2546.

Copyright © 2011 Royal Meteorological Society 

# Annexe I

# Code C++ pour le calcul de la VRF avec un ensemble

Cette annexe esquisse une implémentation du calcul de la VRF, codée en C++ et utilisant un ensemble dans l'environnement OOPS (algorithme décrit au chapitre 8). Les deux équations principales de l'algorithme sont le calcul du vecteur de sensibilité

$$\begin{aligned} \mathbb{E}(\boldsymbol{\varepsilon}^a \, \delta S) &= \mathbb{E}(\boldsymbol{\varepsilon}^a \boldsymbol{\varepsilon}^{aT} [\nabla_{\boldsymbol{x}^a} S]) \\ &= \boldsymbol{A}_c [\nabla_{\boldsymbol{x}^a} S]. \end{aligned} \tag{I.1}$$

et le calcul du champ de VRF en lui même,

$$\delta \boldsymbol{\sigma}^2_{S_t} = \boldsymbol{s} \odot \boldsymbol{s} \odot (\sigma_t^{o2} \mathbf{1} + \boldsymbol{\sigma}^{\boldsymbol{a2}})^{-1}. \tag{I.2}$$

Ce code reste partiel car certaines fonctionnalités sont encore à développer dans OOPS (gestion des ensembles hors perturbations, la classe "fonction score" et le calcul de l'inverse de chaque membre d'un *vecteur*). Cette proposition de code n'a été ni testée, ni compilée.

```
int execute(const util : :Config & fullConfig) {
// ————————- VRF ENSEMBLE ————————-

// Setup time loop
const util : :Config fcConfig(fullConfig, "forecast") ;
const util : :DateTime obsDate(fcConfig.getData("begin")) ;
const util : :DateTime verifDate(fcConfig.getData("end")) ;
// Setup resolution
const util : :Config resolConfig(fullConfig, "resolution") ;
const Geometry resol(resolConfig) ;

const util : :Config xaConfig(fullConfig, "VRFAnalysisEnsemble") ;
const util : :Config xfConfig(fullConfig, "VRFForecastEnsemble") ;

// lecture de l'ensemble d'analyse
Ensemble<MODEL> Xa_(obsDate,xaConfig) ;
Ensemble<MODEL> Xf_(verifDate,xfConfig) ;
// le constructeur va chercher l'ensemble des valeurs filename dans l'objet Config passé en argument
// on extrait un memnre (de type Increment) par Xa_[i]

// Calcul des moyennes de l'ensemble (à implementer)
xaMean_ = Xa_->mean() ;
```

```
xfMean_ = Xf_->mean();

// Calcul des dS et du vecteur sensibilité. La classe ScoreFunction est à créer dans OOPS.
const util : :Config ScoreConfig(fullConfig, "ScoreFunction");
const ScoreFunction<MODEL> S_(ScoreConfig);

// Initialisation du vecteur de sensibilité
sensi_ = new Increment<MODEL> (resol, xaMean_->variable(), obsDate);
sensi_.zeros();

const int members = fullConfig.getInt("members");
double dS_ [members];
const double rr = 1.0/static_cast<double>(members);

// Bouche sur les membres pour le calcul de la sensibilité
for (int jm = 0; jm < members; ++jm) {
dS_[jm] = S_.getValue(Xf_[jm]);

Increment<MODEL> * dx = new Increment<MODEL>(resol, sensi_->variable(), obsDate);

// calcul de sensi=E(dS epsâ)
dx->diff(Xa_[jm]),xaMean_);
dx *= rr*dS_[jm];
sensi_ += dx;
}

// Calcul de la variance d'erreur d'analyse
Increment<MODEL> varA_= new Increment<MODEL>(resol, sensi_->variable(), obsDate);
varA_.zeros();

const double rk = 1.0/(static_cast<double>(members)-1);
// Bouche sur les membres pour le calcul de la variance d'erreur d'analyse
for (int jm = 0; jm < members; ++jm) {
Increment<MODEL> * dx = new Increment<MODEL>(resol, sensi_->variable(), obsDate);
dx->diff(Xa_[jm]),xaMean_);
dx.dot_product_with(dx);
dx *= rk;
varA += dx;
}

// Ajout de la variance d'erreur de l'observation supplémentaire
const double varR = static_cast<double>(fullConfig.get("varR"));
varA += varR;

// Inserver les valeurs de varA(à implementer)
varAinv= varA->inv();

// Calcul de la VRF
VRF_ = dot_product(sensi,sensi);
VRF_ = dot_product(VRF_,varAinv);
```

```
VRF_->write(fullConfig);

return 0;
}
```