



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *jeudi 10 décembre 2015* par :

RAPHAËL LEGRAND

**Utilisation des déformations spatiales
en assimilation de données**

JURY

FRANK ROUX

DIDIER AUROUX

EMMANUEL COSME

ÉTIENNE MÉMIN

THIBAUT MONTMERLE

YANN MICHEL

Professeur

Professeur

Maître de conférences

Directeur de Recherche

Ingénieur de Recherche

Ingénieur de Recherche

Président du Jury

Rapporteur

Rapporteur

Examineur

Directeur de thèse

Co-encadrant de thèse

École doctorale et spécialité :

SDU2E : Océan, Atmosphère et Surfaces Continentales

Unité de Recherche :

CNRM-GAME/Météo-France

Directeur(s) de Thèse :

Thibaut Montmerle et Yann Michel

Rapporteurs :

Didier Auroux et Emmanuel Cosme

À Bérangeère Legrand, ma grand-mère.

Remerciements

Si ma vie était un 4D-Var, je remercierais ma famille et mes amis pour m'avoir permis d'être une ébauche pas trop bancal. Leur soutien constant a été essentiel. Notamment cette dernière année où j'ai fait face à la thèse et au concours.

Si ma vie était un 4D-Var, je remercierais tout particulièrement cinq personnes que j'ai eu la chance de rencontrer durant mon parcours scientifique et qui ont été des observations d'une grande importance. Merci à Christine Lac, Didier Ricard, Thomas Auligné, Yann Michel et Thibaut Montmerle de m'avoir transmis leur passion pour la recherche.

Enfin, si ma vie était un 4D-Var, je remercierais ma moitié, Emilie, d'avoir été un si beau modèle. Tu as été une aide, un soutien et un réconfort de tous les instants.

Résumé

L'assimilation de données permet de construire un état initial du modèle (l'analyse) à partir de deux sources d'information : les observations disponibles et une prévision récente (l'ébauche). L'importance relative de chacune des sources dépend du modèle d'erreurs qui leur est attribué. Le modèle le plus commun en prévision numérique du temps (PNT) consiste à formuler l'erreur d'ébauche comme un incrément additif en amplitude et, avec une approche probabiliste, de spécifier sa fonction de densité de probabilité (PDF) par une Gaussienne spécifiée avec une moyenne nulle et une matrice de covariance \mathbf{B} . Le premier problème abordé durant cette thèse est le manque de dépendance au flux des modélisations de \mathbf{B} . Le deuxième est l'écartement à l'hypothèse Gaussienne notamment en présence d'erreurs de déplacement. La démarche suivie est d'utiliser des déformations spatiales dans un cadre ensembliste pour raffiner la modélisation des corrélations d'erreurs d'ébauche, et de corriger les erreurs de déplacement pour tenter de se rapprocher de l'hypothèse Gaussienne.

La première partie du travail de thèse consiste à améliorer la modélisation de \mathbf{B} , en prenant en compte objectivement l'anisotropie des corrélations d'erreurs d'ébauche à l'aide de déformations spatiales estimées statistiquement à partir d'un ensemble de prévisions générées avec un ensemble d'assimilation (EDA). Cette méthode statistique (ST) est appliquée à une simulation réelle du modèle global de Météo-France ARPEGE, et son efficacité est comparée à celle de la transformée géostrophique (GT) à l'aide de diagnostics d'anisotropie. Les résultats montrent que la ST est plus efficace que la GT avec une prise en compte plus systématique de l'anisotropie des corrélations spatiales.

Dans une deuxième partie, une documentation de la non-Gaussianité (NG) des erreurs d'ébauche d'AROME, le modèle à échelle convective de Météo-France, est proposée. La NG des distributions d'erreurs d'ébauche, diagnostiquées à partir d'un EDA, est quantifiée par un indice uniquement basé sur leur asymétrie et leur aplatissement. La NG diagnostiquée a une forte dépendance spatiale et temporelle, avec des structures spatiales qui suivent les structures météorologiques. Le lien avec certains processus non-linéaires, liés par exemple à la présence de nuages, est notamment observé. Les diagnostics montrent également que le processus d'analyse réduit fortement la NG observée dans l'ébauche, et que la vorticit  et la divergence sont les variables de contr le de l'assimilation les moins Gaussiennes.

Une des causes possibles de la NG  tant la pr sence d'erreurs de d placement, la troisi me partie de ce travail de th se s'attache   documenter les apports de ce mod le d'erreurs alternatif. Un algorithme d'estimation des erreurs de d placements a tout d'abord  t  mis en place. Cet algorithme, qui s'inspire des m thodes d'estimation du flot optique, se base sur une approche variationnelle quasi-lin aire, quasi Gaussienne et multi- chelles. Le comportement de la m thode a  t  test  sur des cas id alis s, puis sur des cas r els bas s sur des pr visions AROME. Ceci permet d'envisager   l'avenir l'estimation statistique des principaux param tres caract risants ce type d'erreurs, ainsi que leur exploitation dans des algorithmes visant   corriger ces erreurs en amont de la correction d'amplitude.

Abstract

Data assimilation provides initial conditions (the analysis) for the model from two sources of information : the available observations and a short term forecast (the background). The relative weights of both sources depend on the chosen error model. Usually, background errors are modeled as additive amplitude errors and, following a probabilistic approach, their probability density function (PDF) are modeled as Gaussians, specified with a zero mean and a covariance matrix \mathbf{B} . For this PhD, the tackled issue are the lack of flow dependency of \mathbf{B} and the non-Gaussianity (NG) frequently arising in numerical weather prediction (NWP), especially in case of displacement error. Using spatial deformation methods and ensemble of forecasts generated from an ensemble of data (EDA), the aim of this work is to refine the model of background error correlations, and to correct displacement errors as a possible way to get closer to the Gaussian hypothesis.

A first part of this work is to improve the \mathbf{B} modeling, by accounting objectively for the anisotropy of the background error correlations thanks to spatial deformation, which is statistically estimated from an EDA. This statistical method (ST) is applied to a real dataset of ARPEGE, the global model at Météo-France. The ST is compared to an other spatial warping method, the geostrophic transform (GT), and their efficiency are quantified by an anisotropy index. Results show a better efficiency of the ST than the GT, with a more systematic accounting of the anisotropy of the spatial background correlations.

The second part of this work is a documentation of the NG for the background errors of AROME, the convective scale model at Météo-France. Distributions are diagnosed from an EDA, and their Gaussianity is quantified with a statistical index only based on two shape parameters, skewness and kurtosis. The diagnosed NG is spatially and temporally dependent, with large similarities between meteorological and NG structures. Links with nonlinear processes, linked for example to cloud covering, are observed. Diagnostics also show that the analysis step is largely decreasing the observed NG of the background, and that the most non-Gaussian variables in control space are vorticity and divergence.

Since, a possible reason of the NG is the displacement errors, the third part of this PhD work aims at getting insight in the documentation of this alternative error model. First, a quasi-linear, quasi-Gaussian and multi-scale method, inspired from optical flow estimation methods, is derived. It has been tested on ideal case and on real cases based on AROME forecast fields. Future works may be, first, a statistical estimation of the main parameters characterizing this kind of error, and second, the use of this algorithm to correct displacement error upstream of the additive error correction.

Table des matières

Introduction	12
1 Fondamentaux de l'assimilation de données variationnelle	19
1.1 Les systèmes de prévision de Météo-France	19
1.2 Formulation Bayésienne	24
1.3 Cas particulier des erreurs additives et Gaussiennes	27
1.4 Modélisation de la matrice \mathbf{B}	33
1.5 Conclusion	40
2 Représentation anisotrope des corrélations d'erreurs d'ébauche	42
2.1 Contexte de l'étude et résumé	42
2.2 Version originale de l'article publié	44
3 Diagnostics de non-Gaussianité des erreurs de prévision et d'analyse aux échelles convectives	61
3.1 Contexte de l'étude et résumé de l'article	61
3.2 Version originale de l'article publié	63
4 Correction des erreurs de déplacement en PNT	77
4.1 Introduction	77
4.2 Méthodes d'estimation existantes en PNT	80
4.3 Approche Bayésienne	83
4.4 Estimation quasi-linéaire quasi-Gaussienne	85
4.5 Description technique de la méthode implémentée	87
4.6 Test sur des cas idéalisés	95
4.7 Application à un cas réel AROME-WMED	101
4.8 Conclusions	109
Conclusions et perspectives	113
A Annexes	115
A.1 Évolutions des systèmes opérationnels de Météo-France en 2015	115

Introduction

A FIN de prévoir le temps qu'il fera dans les prochains jours, le travail du prévisionniste peut se décomposer en au moins deux parties : l'analyse et la prévision. L'analyse consiste à vérifier, dans le temps présent, si l'état atmosphérique qui a été prévu par un modèle de prévision numérique du temps (PNT) est en adéquation avec les observations météorologiques satellitaires ou autres (dites conventionnelles : stations terrestres, radar météorologiques, radiosondages, etc.). Parmi tous les systèmes de prévisions maintenus par les centres nationaux dans le monde entier, le prévisionniste choisit l'un de ceux qu'il considère comme le plus fiable et qu'il utilisera pour la prévision du temps futur. Cette étape d'analyse est également l'occasion de comprendre la situation météorologique en cours à au moins deux niveaux de compréhension : la situation synoptique, caractérisée par les structures météorologiques d'échelle spatiale de plus de 500km (telles que les fronts, ou les dépressions), et les multiples phénomènes régionaux ou locaux ayant des échelles spatiales inférieures à quelques centaines de kilomètre (tels que les cellules orageuses). La méthodologie conseillée est d'effectuer une descente d'échelle en commençant par s'intéresser à la situation synoptique, pour ensuite en déduire les liens avec la méso-échelle, et le temps ressenti localement. Suite à cette première phase d'analyse intervient la phase de prévision qui vise à estimer quel sera le temps ressenti dans le futur. Le mode opératoire suivi pour une prévision à courte échéance (inférieure à un ou deux jours) est différent de celui qui est suivi pour une prévision plus longue allant de quelques jours à plus d'une semaine. Pour la courte échéance, la prévision est détaillée avec une description assez précise de l'évolution de structures de petites échelles. Ce type de prévision repose principalement sur des modèles météorologiques de fines échelles tels que AROME pour Météo-France. Du fait de leur échelle temporelle réduite, il est difficile de prévoir correctement les phénomènes de petites échelles après un ou deux jours. Pour la prévision à plus longue échéance, le bulletin météorologique se base sur la description des structures de plus grandes échelles à l'aide de modèles ayant une résolution spatiale moindre (ARPEGE à Météo-France), pour en déduire au mieux le temps ressenti. Cette limite temporelle dans la prévision d'un phénomène météorologique est liée aux incertitudes contenues dans les prévisions météorologiques. C'est ce qu'on appelle l'horizon de prévisibilité : c'est la durée maximale pour laquelle il est possible de prévoir de façon unique l'évolution d'un phénomène (Joly & Descamps, 2009).

Les erreurs de prévision mentionnées au paragraphe précédent peuvent être rassemblées en deux classes : les erreurs de modélisation et les erreurs dans les conditions initiales de la prévision. Les erreurs de modélisation regroupent une large gamme d'inexactitudes liées à la paramétrisation physique et à la dynamique du modèle, auxquelles se rajoutent, pour un modèle à aire limitée comme AROME, les erreurs commises dans la paramétrisation des conditions aux limites latérales de la grille du modèle. Afin de fournir des prévisions de la meilleure qualité possible, ces erreurs sont régulièrement corrigées : le système de PNT est organisé en cycles quotidiens qui se composent d'une phase de propagation et d'une phase d'assimilation. Cette dernière vise à corriger les erreurs de prévisions issues de l'étape de propagation précédente pour fournir un nouvel état initial plus proche d'un état observé de l'atmosphère, pour l'étape de propagation suivante. Habituellement, l'état de l'atmosphère prévu à la fin de l'étape de propagation est appelé ébauche, et celui modélisé à la fin de l'étape d'assimilation est appelé analyse. L'analyse est calculée à partir des informations disponibles au moment de l'étape d'assimilation : l'ébauche issue du cycle précédent et toutes les observations disponibles

effectuées dans un laps de temps, appelé fenêtre d’assimilation, située de part et d’autre de l’instant où est effectuée l’analyse. On parle d’assimilation de données. L’apprentissage, par le système de prévision, des informations contenues dans les observations et l’ébauche, doit suivre au moins trois caractéristiques. Tout d’abord, il n’est pas possible de reconstituer un état complet de l’atmosphère uniquement à partir des observations. En effet, le nombre d’observations disponibles est bien inférieur au nombre de degrés de liberté du modèle. Ensuite, l’assimilation doit prendre en compte les erreurs contenues dans les observations et dans l’ébauche. Ceci est notamment possible avec un formalisme probabiliste, où l’état de l’atmosphère et où les observations sont considérés comme des réalisations de processus aléatoires issus de fonctions de densité de probabilité (PDF). Ces PDFs d’erreurs d’observations et d’ébauche permettent d’attribuer un poids relatif à l’information venant de chacune de ces deux sources d’information. Enfin, l’état d’analyse doit respecter les grands équilibres qui sont décrits par les équations du modèle, afin de limiter la propagation d’ondes indésirables dans les premiers instants de la phase de propagation. Ce phénomène est habituellement désigné sous le terme de « spin-up ».

Les développements exposés dans cette thèse se focalisent sur la spécification des erreurs d’ébauche et de leurs PDFs déduites d’ensemble de prévisions. Le cadre le plus courant en assimilation de données pour la PNT est de considérer que l’erreur d’ébauche est la différence entre l’état vrai de l’atmosphère et l’ébauche, i.e une erreur additive d’amplitude. De plus, la PDF qui lui est associée est modélisée, la plupart du temps, par une Gaussienne. Avec ces deux hypothèses, les statistiques d’erreurs d’ébauche sont complètement décrites par leur moyenne (supposée nulle) et par leur matrice de covariances, nommée communément \mathbf{B} . Comme mentionné dans Bocquet et al. (2010), le choix du modèle Gaussien est fait notamment pour sa fréquence d’occurrence dans les processus naturels (conséquence du théorème central limite), ainsi que pour le nombre limité de paramètres nécessaire à sa caractérisation (i.e ses deux premiers moments). Le deuxième point est essentiel pour les problèmes de grande taille comme la PNT qui possède un nombre de degrés de liberté de l’ordre de $\mathcal{O}(10^7)$.

L’information contenue dans la matrice \mathbf{B} est essentielle. En effet, pour une PDF Gaussienne et comme rappelé dans Rasmussen & Williams (2006), c’est la matrice de covariance qui définit la notion de *similarité* entre les points de grilles du modèle, c’est à dire la ressemblance du comportement de l’erreur entre deux points selon leurs positions respectives. Cette information est notamment nécessaire à la propagation de l’information contenue dans une observation ponctuelle vers les points de grille environnant. Dans la pratique, la spécification de la matrice \mathbf{B} se heurte à deux écueils (Bannister, 2008a). Le premier est que l’état « vrai » permettant de définir l’erreur d’ébauche ne peut pas être connu. Le deuxième est que la taille de la matrice \mathbf{B} rend impossible son stockage complet. Pour ces deux raisons, il est nécessaire de modéliser \mathbf{B} sous la forme d’un opérateur de covariances simplifié, permettant d’effectuer efficacement son produit avec un vecteur. En PNT, la matrice \mathbf{B} est habituellement séparée en trois sous-opérateurs (Derber & Bouttier, 1999) : l’opérateur de covariances inter-variables, l’opérateur d’auto-covariances verticales, et l’opérateur d’auto-covariances horizontales. Ce dernier terme inclut les variances et les auto-corrélations horizontales d’erreurs d’ébauche. À Météo-France, le modèle opérationnel AROME utilise des auto-covariances (verticales et horizontales) de forme Gaussienne, qui sont statiques (ne varient pas avec le temps : on parle de covariances climatologiques), homogènes (identiques pour tous les points de grille d’un même niveau modèle), isotropes (seule la distance entre deux points de grille d’un même niveau modèle détermine leur *similarité*), et non-séparables (les covariances horizontales varient avec les niveaux verticaux, et les covariances verticales varient selon l’échelle horizontale du phénomène modélisé; Courtier et al. (1998)). Pour le modèle global ARPEGE, les variances d’erreurs d’ébauche sont spécifiées de façon non statique et non homogène (Raynaud et al., 2011), et les corrélations

horizontales ont également une dépendance au flux grâce à la formulation en ondelettes présentée par (Berre et al., 2015).

Plusieurs études ont montrés que les statistiques d'erreurs d'ébauche dépendent notamment a) du modèle et de sa résolution (Ménétrier et al., 2014), b) des phénomènes météorologiques d'intérêt (par exemple zones précipitantes ou non précipitantes, Montmerle & Berre (2010), ou bien dans les zones de brouillard, Ménétrier & Montmerle (2011)), c) du régime de temps modélisé (par exemple situations anticycloniques ou dépressionnaires, Brousseau et al. (2011)), et enfin d) de la densité du réseau d'observations (ces quatre points seront regroupés par la suite sous le terme de "dépendance au flux météorologique"). Une modélisation statique, homogène, et isotrope des corrélations d'erreurs d'ébauche rend malheureusement impossible la prise en compte de ces caractéristiques dans le système d'assimilation. Le premier axe de ce travail de thèse a été d'associer un opérateur de déformation spatiale à la modélisation actuelle des auto-corrélations horizontales d'erreurs d'ébauche, de manière à prendre en compte une partie de l'anisotropie de ces corrélations due à leur dépendance au flux météorologique. L'étude proposée, qui a fait l'objet d'une première publication (Legrand & Michel, 2014), s'est attachée à tester sur un cas réel, prévu par le modèle ARPEGE, un opérateur de déformation spatiale estimé statistiquement selon Michel (2013a) et Michel (2013b). Cet opérateur est comparé à un autre type de déformation spatiale : la transformée géostrophique (Desroziers, 1997). La question de l'objectivité de l'estimation des opérateurs de déformations a été prise en compte. En effet, contrairement à la transformée géostrophique, aucune hypothèse *a priori* n'est faite sur la forme des déformations avec la méthode statistique, l'information étant uniquement extraite à partir d'un ensemble de prévision.

Comme Lawson & Hansen (2005) le remarquent, modéliser l'erreur d'ébauche uniquement avec une erreur additive n'est pas un choix restrictif. Par contre modéliser cette erreur additive comme une réalisation issue d'une PDF Gaussienne est restrictif. Malgré le constat mentionné précédemment que les distributions Gaussiennes se rencontrent fréquemment dans les phénomènes naturels, la non-Gaussianité se rencontre couramment en PNT. En effet, propager un état de l'atmosphère et le champ d'erreur qui lui est associé avec un modèle non-linéaire entraîne nécessairement le développement de distributions d'erreurs non-Gaussiennes (Lawson & Hansen, 2005; Bocquet et al., 2010). Ceci est d'autant plus vrai pour les modèles à mailles fines tels qu'AROME, qui permettent la résolution explicite de processus dont la croissance des erreurs est très rapidement fortement non-linéaire (tels que ceux liés à la couverture nuageuse, aux précipitations, ou à la convection). Une conséquence directe est leur faible horizon de prévisibilité : quelques heures pour des phénomènes dont la taille caractéristique est de l'ordre de 20km. Discutant les améliorations nécessaires à la mise en place d'un système d'assimilation capable de produire des analyses de couvertures nuageuses, Auligné et al. (2011) soulignent que les distributions d'erreurs en situation nuageuse sont clairement non-Gaussiennes, et ils attribuent la non-Gaussianité des erreurs pour les variables nuageuses à au moins deux causes : les erreurs de positionnement des structures nuageuses prévues par rapport aux structures observées et les limites intrinsèques à la physique des nuages qui se basent sur de nombreux processus à seuil (niveau de saturation, changement de phase, etc.). Dans le cas où l'erreur d'ébauche est non-Gaussienne mais où l'on utilise un algorithme d'assimilation qui repose sur l'utilisation d'un modèle Gaussien des erreurs, l'étape d'assimilation peut produire une analyse mathématiquement acceptable (qui respecte un certain critère d'optimalité) mais dont la cohérence physique est discutable. Un exemple est donné par Ravela et al. (2007) dans un cas idéalisé où les observations sont échantillonnées à partir d'un état vrai connu, alors que l'ébauche étant identique en amplitude à cet état, mais est translatée. L'état d'analyse produit est distordu de manière irréaliste, et ne ressemble ni à la structure représentée par

les observations, ni à celle représentée par l'ébauche. Étant donné que le système de prévision opérationnelle de Météo-France utilise des algorithmes d'assimilation se basant sur une modélisation Gaussienne des erreurs d'ébauche, une deuxième partie de cette thèse s'est attachée à documenter et à quantifier les écarts à la Gaussianité des erreurs de prévision et d'analyse, et ainsi identifier les zones susceptibles de causer des problèmes lors de l'étape d'assimilation. Les résultats sont présentés pour le modèle AROME, en utilisant une modélisation des PDFs d'erreurs d'ébauche par des méthodes ensemblistes. Ce travail a fait l'objet de la publication d'un deuxième article (Legrand et al., 2015).

Comme cela est mentionné par Bocquet et al. (2010), effectuer l'étape d'analyse sans supposer la Gaussianité des erreurs (d'ébauche et d'observations) est possible, avec par exemple l'utilisation de filtres à particules (Van Leeuwen, 2009). Néanmoins ce type de méthode n'est pas utilisable pour les problèmes de grande taille puisqu'elles requièrent un nombre de particules qui varie de façon exponentielle par rapport à la taille du problème (Snyder et al., 2008). Pour la PNT il semble y avoir un consensus (Bocquet et al., 2010; Auligné et al., 2011) sur la voie à suivre, au moins comme point de départ, pour l'amélioration des systèmes d'assimilation de données pour la prise en compte des erreurs de prévisions non-Gaussiennes (associées aux processus nuageux par exemple) : conserver le formalisme actuel quasi-linéaire (e.g 4D-Var, 3D-Var), et considérer les effets non-Gaussiens uniquement comme une déviation du modèle Gaussien. Afin de prendre en compte une partie de la non-Gaussianité, Lawson & Hansen (2005) et Ravela et al. (2007) proposent par exemple l'ajout d'une erreur de déplacement dans la modélisation de l'erreur d'ébauche en plus de l'erreur additive d'amplitude. En effet, dans un cadre théorique, il a été prouvé que des erreurs de déplacement peuvent entraîner la non-Gaussianité des erreurs d'ébauche (Lawson & Hansen, 2005).

De fait, la nécessité de prendre en compte le mauvais positionnement des structures modélisées est une partie importante du métier de prévisionniste. Une première approche consiste à estimer visuellement les erreurs de déplacement par rapport aux observations ou à d'autres modèles. On parle de « calage modèle ». Cette méthode est passive puisqu'elle ne corrige pas les erreurs de déplacement, mais elle permet d'anticiper pour les échéances suivantes les erreurs commises par le modèle. Ceci peut être crucial dans les cas de fortes tempêtes ou de cyclones pour l'anticipation de l'heure d'arrivée de l'évènement dangereux sur les côtes habitées. Depuis une dizaine d'années, avec l'outil CTPIni (*Correction du Tourbillon Potentiel Initial*), les prévisionnistes peuvent corriger manuellement les erreurs de déplacement et recalculer de nouvelles prévisions à partir du nouvel état initial "aligné". Néanmoins, même si certaines erreurs de déplacement sont assez évidentes pour être estimées visuellement par le prévisionniste, l'utilisation d'algorithme automatique peut permettre un alignement plus systématique et plus précis de l'ébauche.

Les erreurs de déplacement sont connues dans le domaine de traitement d'image sous le nom de flot optique. Toutes les méthodes d'estimation du flot optique sont basées sur la conservation temporelle d'un paramètre du problème (Mémin, 2003) : soit un paramètre géométrique (comme le contour d'une structure à déplacer) soit un paramètre photométrique comme l'intensité lumineuse (i.e lors de la correction de l'erreur de déplacement, la valeur de chaque point de grille est conservée ; seule la position du point de grille est modifiée, Horn & Schunck, 1981). Une grande partie des méthodes d'estimation d'erreurs de déplacement en PNT font l'hypothèse de conservation de l'intensité lumineuse, qui permet une estimation dense du flot optique. Néanmoins cette hypothèse ne suffit pas pour s'assurer que le problème est bien posé selon Hadamard (1902) i.e que la solution existe, qu'elle soit unique, et qu'elle soit stable (de petites variations dans les conditions initiales ou dans les termes de sources n'entraîne que de petites

variations de la solution, Maz'ya & Shaposhnikova, 2012). Afin d'assurer ces trois critères, une possibilité est de contraindre le champ d'erreurs de déplacement selon une information *a priori* que l'on aurait sur sa forme. Le choix de cette contrainte est non seulement très vaste, mais surtout dépend du type de problème à résoudre (Souopgui et al., 2012). Actuellement, la plupart des études effectuées en PNT contraignent soit la norme du déplacement lui-même, soit celle de son gradient, soit celle de sa divergence, ou encore celle de son rotationnel. Les normes utilisées pour ces termes de contraintes peuvent être caractérisées, dans le cas général, selon une matrice de covariances (Tarantola, 2005). Par exemple, pour la norme euclidienne, cette matrice est diagonale et les termes diagonaux sont tous égaux à 1. Avec une approche Bayésienne et Gaussienne du problème d'estimation des erreurs de déplacement (Ravela et al., 2007; Aonashi & Eito, 2011), la matrice associée à la norme du déplacement lui-même (régularisation de Tikhonov, Tikhonov, 1963) contient les covariances d'erreurs de déplacement, et est notée \mathbf{D} par la suite.

Enfin, de façon pratique, avec l'utilisation d'une modélisation des erreurs d'ébauche "mixte" rassemblant les erreurs additives (d'amplitude) et les erreurs de déplacement, l'étape d'assimilation est habituellement décomposée en deux sous-étapes effectuées séquentiellement (Ravela et al., 2007) : une première étape de correction des erreurs de déplacements, puis une deuxième étape visant à corriger les erreurs additives à partir d'une ébauche "alignée". De manière analogue à ce qui est déjà classiquement utilisé, l'étape "d'alignement" revient à minimiser une fonction de coût non-linéaire et non-quadratique. Pour cette thèse, un algorithme quasi-linéaire de type Gauss-Newton est testé. Celui-ci a l'avantage de ne pas nécessiter l'estimation de la dérivée d'ordre deux de la fonction de coût (Gratton et al., 2007), et d'avoir un formalisme proche de celui utilisé pour la correction des erreurs d'amplitude (e.g) à Météo-France.

Plan de la thèse

Le chapitre 1 introduit les bases de l'assimilation de données nécessaires à la compréhension de la suite du manuscrit. Les trois concepts de modélisation des covariances d'ébauche, de non-Gaussianité, et d'erreur de déplacement seront mis en équations à cette occasion.

Le chapitre 2 expose un travail effectué sur la possibilité de raffiner la modélisation de la matrice \mathbf{B} avec l'utilisation de déformations spatiales pour modéliser de l'anisotropie, dans un contexte où les corrélations spatiales d'erreur additive d'ébauche sont modélisées de façon statiques, homogènes et isotropes. Les modifications apportées par un opérateur de déformation spatiale estimée statistiquement (Michel, 2013b,a) sont comparées objectivement à celles apportées par la transformée géostrophique (Desroziers, 1997).

Le chapitre 3 présente des résultats de diagnostics de non-Gaussianité pour le modèle AROME. Ceci constitue une documentation de la non-Gaussianité des erreurs additive d'amplitude de prévision et d'analyse pour ce modèle à échelle convective. L'accent est mis notamment sur le lien entre la non-Gaussianité et les processus nuageux ainsi que sur le comportement de la non-Gaussianité durant l'étape d'assimilation.

Le chapitre 4 présente le formalisme utilisé dans cette thèse pour l'estimation et la correction des erreurs de déplacement préalablement à l'étape de correction habituelle des erreurs additives d'amplitude. Une revue des méthodes déjà utilisées pour la PNT est ensuite proposée. Enfin une méthode originale quasi-linéaire, quasi-Gaussienne d'estimation des erreurs de déplacement est développée et testée.

1 | Fondamentaux de l'assimilation de données variationnelle

En plus d'une présentation des systèmes opérationnels ARPEGE et AROME de Météo-France, l'objet de ce chapitre est de fournir quelques éléments de compréhension de l'assimilation de données variationnelle. La présentation de ces concepts est l'occasion d'introduire les thématiques développées durant cette thèse : les erreurs de déplacement, la Gaussianité, ainsi que la modélisation de la matrice \mathbf{B} .

Sommaire

1.1	Les systèmes de prévision de Météo-France	19
1.1.1	ARPEGE	20
1.1.2	AROME	22
1.2	Formulation Bayésienne	24
1.2.1	Notations	24
1.2.2	Approche Bayésienne	25
1.2.3	Les modèles d'erreurs	26
1.2.4	Modélisation Gaussienne des PDFs d'erreurs	27
1.3	Cas particulier des erreurs additives et Gaussiennes	27
1.3.1	Hypothèse Gaussienne	28
1.3.2	Définition du problème d'optimisation	28
1.3.3	Résolution du problème d'optimisation	30
1.3.4	Configurations opérationnelles	32
1.3.5	Cas linéaire Gaussien et filtres de Kalman	33
1.4	Modélisation de la matrice \mathbf{B}	33
1.4.1	Rôle de \mathbf{B}	34
1.4.2	Modélisation de \mathbf{B} en sous-opérateurs	35
1.4.3	Méthodes d'estimation de \mathbf{B}	35
1.4.4	L'assimilation d'ensemble	37
1.4.5	Configurations opérationnelles	39
1.5	Conclusion	40

1.1. Les systèmes de prévision de Météo-France

Pour la prévision du temps sur la métropole, le système de PNT opérationnelle de Météo-France se compose d'un modèle global ARPEGE et d'un modèle à aire limitée AROME. Depuis le début du travail de thèse, des modifications importantes des systèmes ARPEGE et AROME ont été effectuées, notamment en ce qui concerne leurs résolutions. Les résultats exposés ici se basent sur les configurations opérationnelles utilisées en 2012. Les évolutions de la version de 2015 sont présentées pour information en annexes A.1.1 et A.1.2. Dans un premier temps, seuls des aspects généraux seront présentés pour les deux modèles. Des détails plus techniques sur les systèmes d'assimilation seront donnés aux sections 1.3.4 et 1.4.5.

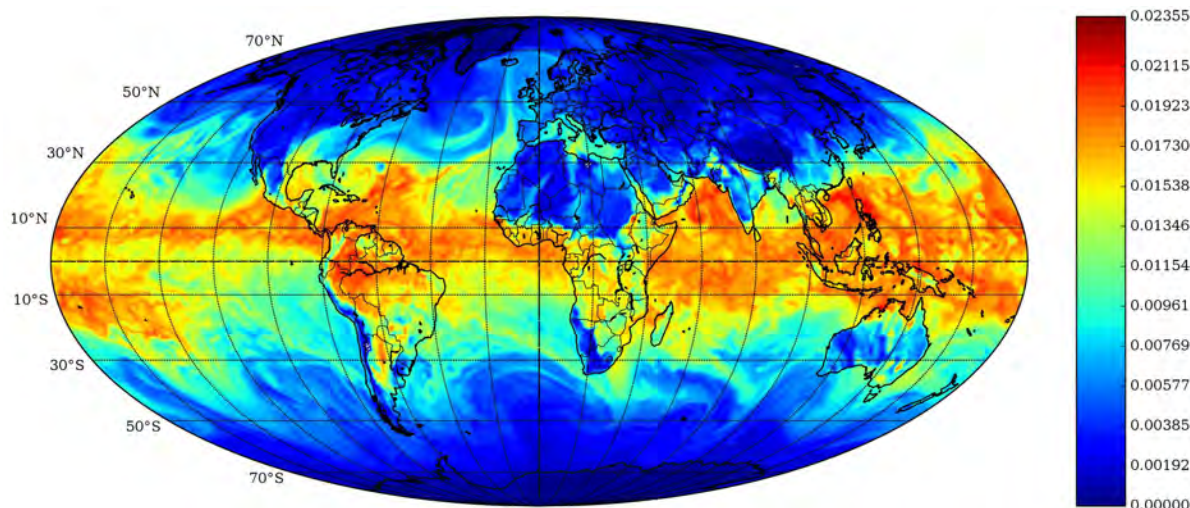


Figure 1.1. – Prévion ARPEGE à 6 heures d'échéance du champ d'humidité spécifique ($kg.kg^{-1}$) à son premier niveau près de la surface, valide le 08/11/2011 à 06 UTC (réseau de 00 UTC).

1.1.1. ARPEGE

Modèle

Le modèle ARPEGE¹ est un modèle global qui a la particularité d'avoir une grille à maille variable : la résolution sur la France métropolitaine est plus élevée qu'aux antipodes. Un exemple de prévision sur le globe est donné par la Fig.1.1. La taille horizontale de la maille est d'environ 10km pour la métropole et de 60km aux antipodes. La grille est découpée sur la verticale selon des coordonnées hybrides dépendant de la pression de surface. Les niveaux modèles suivent le relief au sol et sont relaxées en altitude vers les surfaces isobares (Sangster, 1960; Simmons & Strüfing, 1981; Laprise, 1992). L'atmosphère est décrite sur 70 niveaux verticaux, avec le niveau le plus bas à 17 mètres au dessus de la surface et un niveau supérieur à 0.1hPa ($\approx 70km$, dans la mésosphère). Un resserrement des niveaux les plus bas est utilisé pour permettre de privilégier la description des phénomènes de basses couches. La surface est décrite par le schéma ISBA². Les variables pronostiques du modèle sont le vent zonal (U) et méridien (V), la température (T), l'humidité spécifique (q), la pression de surface, l'eau nuageuse (liquide et solide selon la température), les hydrométéores précipitants (liquide et solide selon la température) et enfin l'énergie cinétique turbulente.

L'état initial d'une prévision est propagé dans le temps selon les équations primitives en utilisant un schéma dynamique de type semi-implicite, semi-lagrangien. La résolution horizontale supérieure à 10 km nécessite de représenter certains processus sous-maille sous forme de paramétrisation. C'est le cas pour le rayonnement, des flux de surface, du couplage océan-atmosphère, des changements d'état associés aux différents hydrométéores, et de la représentation de la convection. Une fois calculée, la contribution de chacune de ces paramétrisations est ajoutée aux équations du système.

La production journalière des prévisions est organisée cycliquement en 4 réseaux dits synoptiques : 00 UTC, 06 UTC, 12 UTC, et 18 UTC. À chaque réseau, la prévision du réseau

1. Action de Recherche Petite Échelles Grande Échelles, Pailleux et al. (2000)

2. Interaction Sol-Biosphère-Atmosphère, Noilhan & Mahfouf (1996)

antérieur appelée ébauche, est recalée vers les observations grâce à un schéma d'assimilation de données de type 4D-Var incrémental (Le Dimet & Talagrand, 1986; Courtier et al., 1994). Ce schéma permet d'assimiler les observations contenues dans une fenêtre d'assimilation de 6 heures centrée sur chacun des quatre réseaux synoptiques. À partir de cet état initial plus récent, appelé analyse, de nouvelles prévisions sont ensuite calculées avec une échéance maximum de 102 heures pour ARPEGE.

Système d'assimilation

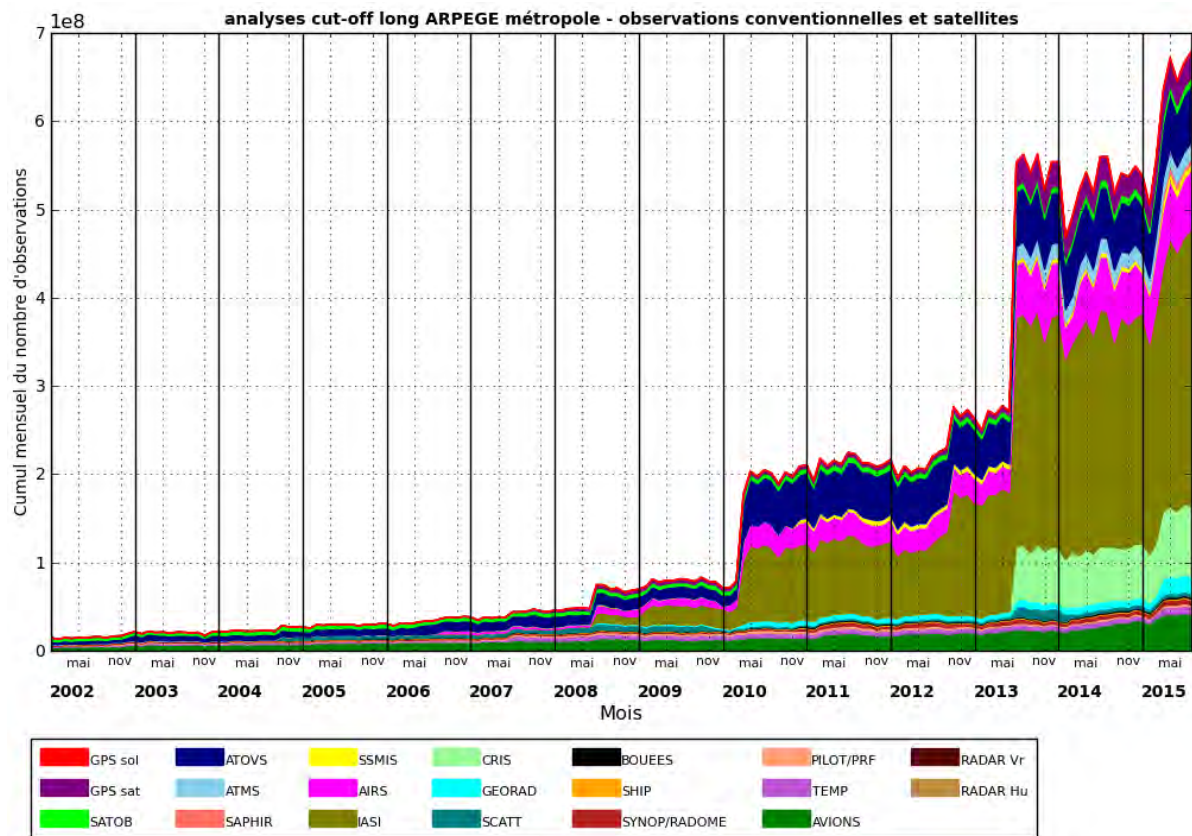


Figure 1.2. – Évolution temporelle du nombre d'observations satellites et conventionnelles utilisées pour le système d'assimilation d'ARPEGE.

Lors de l'étape d'assimilation sont prises en compte des observations satellites et des observations dites conventionnelles (autres que satellites : mesures issues de stations météorologiques terrestres, de capteurs embarqués sur des bouées, bateaux, avions, ballon-sondes, etc...). Comme cela est représenté par la Fig.1.2, le nombre d'observations utilisées par le système d'assimilation d'ARPEGE augmente avec les années. On remarque notamment, à partir de 2008, l'utilisation croissante des observations de radiances infrarouges du satellite IASI. Finalement en 2015, environ 10^6 observations sont prises en compte à chaque réseau (non montré).

Toutes les observations disponibles ne sont pas nécessairement utilisées par le système d'assimilation. D'après la Fig.1.3 synthétisant les pourcentages d'observations disponibles et utilisées, on remarque pour le modèle ARPEGE que :

- les observations satellites représentent pratiquement la totalité du volume de données entrantes. Ceci est notamment dû à l'augmentation de la résolution spatiale des capteurs, ainsi qu'à l'apparition d'une nouvelle génération de sondeurs hyperspectraux de

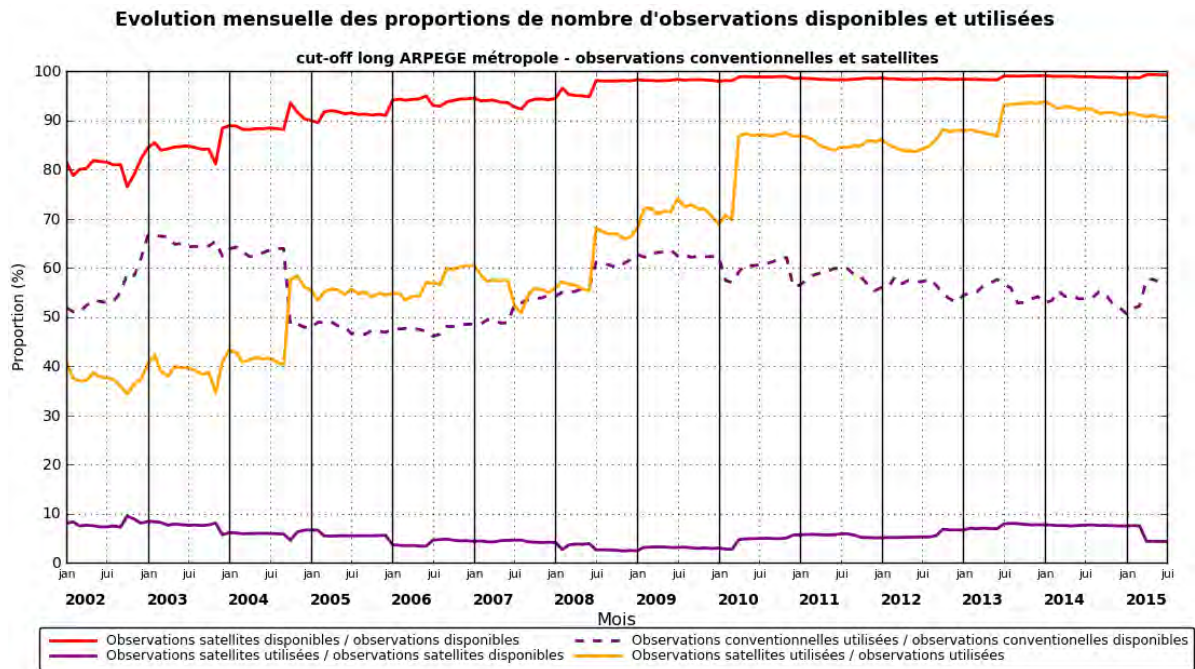


Figure 1.3. – Évolution temporelle de l’utilisation des observations satellites et conventionnelles pour le système d’assimilation d’ARPEGE.

type AIRS ou IASI possédant des centaines de canaux.

— seuls 5% des observations satellites sont utilisées par le processus d’assimilation (contre $\approx 58\%$ pour les observations conventionnelles). Les raisons du rejet de 95% des données satellitaires sont les suivantes (cf Martinet, 2013) :

1. la redondance spectrale entre les différents canaux sensibles aux mêmes couches atmosphériques ;
2. la forte contamination par les nuages des données infrarouges, et par la pluie pour les données micro-ondes ;
3. non-Gaussianité des innovations nécessitant l’utilisation de contrôles de qualité (une discussion sur la non-Gaussianité des erreurs d’ébauche est proposée au chapitre 3) ;
4. la méconnaissance des corrélations d’erreurs d’observations. C’est une des raisons de l’utilisation d’un modèle simplifié où la matrice de covariance des erreurs d’observations (\mathbf{R}) est diagonale. Un écrémage est alors nécessaire pour ne pas considérer des observations corrélées spatialement ;

— malgré la proportion importante d’observations satellites rejetées, elles restent majoritaires ($\approx 90\%$) parmi les observations réellement utilisées.

1.1.2. AROME

Modèle

Le modèle AROME¹ est conçu à partir de deux modèles déjà existants dans les années 2000 : la paramétrisation physique est reprise du modèle de recherche Méso-NH (Lafore et al., 1998) et le cœur dynamique reprend celui du modèle à aire limitée ALADIN-NH (Bubnová et al.,

1. Applications de la Recherche à l’Opérationnel à Méso-Échelle, Seity et al. (2011)

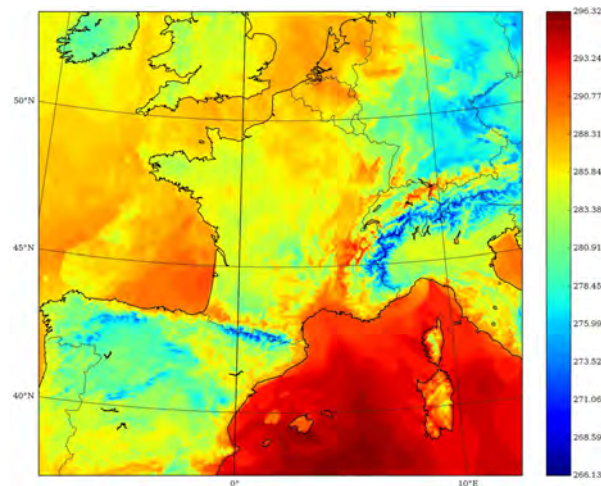


Figure 1.4. – Prévision AROME-France à 3 heures du champ de température (K) au premier niveau de surface, valide le 04/11/2011 à 03 UTC.

1995). En opérationnel depuis décembre 2008, la version AROME-France fournit des prévisions à des échéances maximales de 30h. Sa résolution horizontale est de 2.5km sur un domaine centré sur la France métropolitaine (cf figure 1.4). Par rapport à ARPEGE, sa haute résolution lui permet de modéliser explicitement l'évolution de processus de fine échelle comme par exemple la convection, les phénomènes de brises en basses couches, ainsi que les effets radiatifs de l'orographie (Senkova et al., 2007). Les conditions aux limites latérales du domaine sont fournies par le modèle ARPEGE, et les champs de surface (humidité et température du sol, etc...) sont décrits par le modèle SURFEX¹. AROME est un modèle spectral qui nécessite la bipériodisation des champs. L'atmosphère est découpée sur 60 niveaux verticaux utilisant des coordonnées hybrides. Le premier niveau est situé à 10m avec une augmentation du nombre de niveaux dans les basses couches (ceci permet une meilleure représentation des phénomènes de basse couche et des interactions avec la surface) et le dernier à 1hPa. La dynamique du modèle est contrôlée par un schéma semi-lagrangien, semi-implicite. Les variables pronostiques d'AROME sont au nombre de 12 : le vent (zonal et méridien), la température, l'humidité spécifique, la pluie, la neige, la neige roulée (graupel), l'eau nuageuse liquide (goutte d'eau) et solide (cristaux de glace), ainsi que l'énergie cinétique turbulente et deux autres variables non-hydrostatiques (Bénard et al., 2010; Seity et al., 2011).

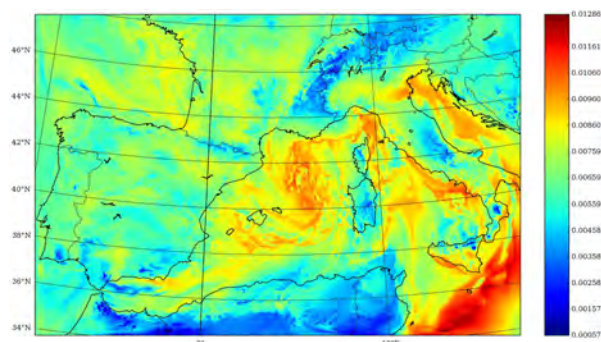


Figure 1.5. – Prévision AROME-WMED à 3 heures du champ d'humidité spécifique ($kg.kg^{-1}$) au niveau modèle 52 (≈ 950 hPa), valide le 08/11/2011 à 03 UTC.

1. Surface Externalisée, Le Moigne et al. (2009)

Pour les besoins de la campagne d'étude HyMeX¹, une autre version opérationnelle du modèle AROME, AROME-WMED², a produit des prévisions à une résolution horizontale de 2.5 km sur un domaine couvrant l'Ouest de la Méditerranée (cf Fig.1.5). Pour cette version, l'échéance maximale de prévision a été prolongée de 18 heures par rapport à AROME-France (soit 48h). Tout le système de prévision/assimilation d'AROME-WMED a été conçu à partir d'AROME-France avec néanmoins quelques différences liées soit au changement de domaine, soit à la volonté de se focaliser sur les épisodes fortement précipitants e.g recalibrage de la matrice covariance d'erreur d'ébauche sur la période automnale d'octobre 2011, caractérisée par des phénomènes convectifs dans le sud de la France et en Catalogne (Fourrié et al., 2015).

Système d'assimilation

Le cycle d'assimilation de 3h utilise un algorithme d'assimilation de type 3D-Var avec des fenêtres d'assimilation de plus ou moins 1h30 centrées sur les réseaux. Toutes les observations disponibles sont donc considérées comme effectuées au milieu de la fenêtre. Les prévisions à longue échéance (30h) ne sont produites que pour les quatre réseaux synoptiques quotidiens : 00 UTC, 06 UTC, 12 UTC, et 18 UTC. Les réseaux intermédiaires ne servent qu'à produire des ébauches (correspondant à des prévisions 3h) pour le réseau suivant.

En ce qui concerne les observations, l'une des différences avec ARPEGE est l'utilisation des données issues des radars météorologiques. L'assimilation des vitesses radiales (Montmerle & Faccani, 2009) et des réflectivités (Wattrelot et al., 2014) permet ainsi d'analyser la structure tridimensionnelle du vent et de l'humidité relative dans les zones de précipitations. Selon Brousseau et al. (2014), ce type d'observations est particulièrement efficace en période pluvieuse pour contraindre les champs de vents et d'humidité spécifique en moyenne et haute troposphère pour des échelles caractéristiques inférieures à 200km.

1.2. Formulation Bayésienne

Selon une approche classique en PNT, l'analyse est définie comme l'état le plus probable associé à la PDF définissant la probabilité d'occurrence des vecteurs d'état dans l'espace modèle en ayant connaissance des observations disponibles, de l'ébauche, et de leurs statistiques d'erreurs associées. Habituellement, cette approche probabiliste est associée notamment à l'utilisation de la règle de Bayes et à l'utilisation de l'hypothèse Gaussienne. Dans cette section, le parti-pris est de pousser le plus loin possible la dérivation du processus d'assimilation sans faire de différence entre les modèles d'erreurs d'ébauche choisis (e.g erreurs d'amplitude ou de déplacement).

1.2.1. Notations

La représentation par le modèle de PNT de l'état de l'atmosphère à un instant t_l est contenue dans un vecteur noté \mathbf{x}^l appartenant à l'espace modèle. La prévision de l'état de l'atmosphère à l'instant suivant t_{l+1} , notée \mathbf{x}^{l+1} , est obtenue à partir de \mathbf{x}^l tel que

$$\mathbf{x}^{l+1} = \mathcal{M}_{l \rightarrow l+1}(\mathbf{x}^l)$$

1. HYdrological cycle in the Mediterranean EXperiment, Drobinski et al. (2014). Campagne d'étude des cycles hydrologiques du bassin méditerranéen, visant notamment à mieux connaître leurs variabilités climatiques, ainsi qu'à une meilleure compréhension des événements extrêmes qui impactent cette zone (Fourrié et al., 2015).

2. Application of Research to Operations at MEsoscale West MEDiterranean Sea, Fourrié et al. (2015)

avec \mathcal{M} l'opérateur non-linéaire de propagation. La qualité de \mathbf{x}^l peut être quantifiée en le comparant aux observations \mathbf{y}_o^l faites de l'atmosphère à l'instant t_l (espace des observations). Pour permettre cette comparaison, l'opérateur d'observations \mathcal{H} (généralement non-linéaire) forme une image \mathbf{y}^l dans l'espace des observations d'un vecteur d'état \mathbf{x}^l de l'espace modèle :

$$\mathbf{y}^l = \mathcal{H}(\mathbf{x}^l).$$

L'opérateur d'observations permet ainsi de générer un équivalent modèle des observations en chaque points du réseau d'observations. Il se constitue donc classiquement d'un modèle permettant de calculer une variable observée à partir des variables simulées (modèle de transfert radiatif pour les radiances satellites par exemple), et d'un opérateur d'interpolation spatiale.

Par la suite l'ensemble des observations effectuées entre les instants t_0 et $t_{\mathcal{T}}$ seront regroupées dans $\mathbf{Y}_o^{\mathcal{T}}$ tel que

$$\mathbf{Y}_o^{\mathcal{T}} = (\mathbf{y}_o^1, \dots, \mathbf{y}_o^{\mathcal{T}})$$

De même, l'ensemble des prévisions, \mathbf{x}^l , valides durant la fenêtre temporelle $[t_0, t_{\mathcal{T}}]$ seront regroupées dans $\mathbf{X}^{\mathcal{T}}$, nommée trajectoire, et définie telle que

$$\mathbf{X}^{\mathcal{T}} = (\mathbf{x}^0, \dots, \mathbf{x}^{\mathcal{T}}).$$

Une étape d'assimilation de données variationnelle consiste à obtenir la meilleure trajectoire $\mathbf{X}_a^{\mathcal{T}}$ (nommée trajectoire d'analyse) durant une fenêtre temporelle $[t_0; t_{\mathcal{T}}]$ (nommée fenêtre d'assimilation), à partir des informations contenues dans les observations $\mathbf{Y}_o^{\mathcal{T}}$ et dans l'ébauche \mathbf{x}_b (état prévu de l'atmosphère en début de fenêtre d'assimilation). L'ébauche est habituellement une prévision à courte échéance valide à t_0 et initialisée à $t_{-\mathcal{T}}$ au cycle d'assimilation précédent. Le processus d'assimilation tentera de corriger

- les erreurs d'ébauche incluant toutes les erreurs dans les conditions initiales de la prévision
- les erreurs modèles issues des inexactitudes de \mathcal{M}
- les erreurs d'observations comprenant les erreurs de mesure, de représentativité, mais également les inexactitudes dans la formulation de \mathcal{H} (Lorenç, 1986).

La connaissance des statistiques de ces trois types d'erreurs permet d'attribuer une confiance relative plus grande à l'une ou à l'autre source d'information (observations ou ébauche).

1.2.2. Approche Bayésienne

La prise en compte des statistiques d'erreurs est permise grâce à l'utilisation d'un formalisme probabiliste où ce sont les PDFs d'erreurs qui sont manipulées plutôt que les vecteurs d'état (qui n'en sont qu'une réalisation). La probabilité conditionnelle associée à toute trajectoire $\mathbf{X}^{\mathcal{T}}$, issue du processus d'assimilation, est notée $\mathcal{P}(\mathbf{X}^{\mathcal{T}}|\mathbf{Y}^{\mathcal{T}})$ et peut être définie en suivant la règle de Bayes telle que

$$\mathcal{P}(\mathbf{X}^{\mathcal{T}}|\mathbf{Y}^{\mathcal{T}}) \propto \mathcal{P}(\mathbf{x}^0) \prod_{l=1}^{\mathcal{T}} \mathcal{P}(\mathbf{y}_o^l|\mathbf{x}^l) \mathcal{P}(\mathbf{x}^l|\mathbf{x}^{l-1}) \quad (1.1)$$

$\mathcal{P}(\mathbf{x}^0)$ est la probabilité *a priori* associée aux erreurs d'ébauche, $\mathcal{P}(\mathbf{y}_o^l|\mathbf{x}^l)$ contient les statistiques d'erreurs d'observations, et $\mathcal{P}(\mathbf{x}^l|\mathbf{x}^{l-1})$ regroupe les statistiques d'erreurs modèle formées entre l'instant t_l et t_{l+1} . La trajectoire analysée est calculée selon un critère de maximum de vraisemblance i.e \mathbf{X}_a^T est estimé comme étant le mode de $\mathcal{P}(\mathbf{X}^T|\mathbf{Y}^T)$. L'étape d'assimilation est alors interprétée de manière équivalente comme un problème de minimisation de la fonction coût $\mathcal{J} = -\ln(\mathcal{P}(\mathbf{X}^T|\mathbf{Y}^T))$.

À partir de l'Eq.1.1, il est nécessaire de faire au moins deux choix : le choix d'un modèle d'erreurs et le choix d'une modélisation pour les PDFs d'erreurs. Les PDFs des erreurs d'ébauche, des erreurs d'observations et des erreurs modèle sont notées par la suite \mathcal{P}_b , \mathcal{P}_o , et \mathcal{P}_q .

1.2.3. Les modèles d'erreurs

Le choix du modèle d'erreur définit la manière avec laquelle \mathbf{X}^T va être modifiée pour obtenir \mathbf{X}_a^T . Le modèle le plus courant est le modèle d'erreurs additives, mais il en existe au moins deux autres : les erreurs multiplicatives, et les erreurs de déplacement. Pour l'erreur d'ébauche, définie ici comme l'écart entre l'ébauche \mathbf{x}_b et un état de l'atmosphère \mathbf{x}^0 , ces modèles peuvent respectivement s'écrire (Lawson & Hansen, 2005) :

$$\begin{cases} \mathbf{x}^0 = \mathbf{x}_b + \mathbf{a} \\ \mathbf{x}^0 = (1 + \mathbf{m}) \mathbf{x}_b \\ \mathbf{x}^0 = \mathbf{x}_{b,[\mathbf{s}+\mathbf{d}]} \end{cases}$$

avec \mathbf{a} une erreur additive, \mathbf{m} une erreur multiplicative, et \mathbf{d} une erreur de déplacement. Pour définir le cadre d'utilisation des erreurs de déplacement, les notations suivantes sont introduites. Elles seront réutilisées dans les chapitres 4. Si le vecteur d'état \mathbf{x} contient l'état de l'atmosphère prévu en chacun des points de la grille du modèle, le vecteur d'état $\mathbf{x}_{[\mathbf{s}+\mathbf{d}]}$ contient en chaque point de cette même grille l'état de l'atmosphère qui aura été déplacé par le vecteur de déplacement \mathbf{d} . \mathbf{s} représente les coordonnées de chacun des points de la grille du modèle.

Les coordonnées de chacun de ces points de grilles sont contenues dans le vecteur \mathbf{s} . Pour une grille tridimensionnelle contenant n points de grille, \mathbf{s} est de taille $n \times 3$. Un vecteur d'état \mathbf{x} déformé par un déplacement \mathbf{d} sera alors noté $\mathbf{x}_{[\mathbf{s}+\mathbf{d}]}$, où $\mathbf{s} + \mathbf{d}$ définit pour chacun des points de grille leur nouvelle position.

Sans hypothèse supplémentaire, les modèles d'erreurs additives ou multiplicatives ne sont pas restrictifs dans le sens où il est toujours possible de trouver \mathbf{a} ou \mathbf{m} permettant de modéliser les écarts entre l'ébauche et n'importe quel vecteur d'état \mathbf{x}^0 . Pour l'erreur multiplicative, cette généralité n'est cependant valable seulement si aucun point de grille de l'ébauche (\mathbf{x}_b) n'est nul. Pour l'erreur de déplacement, le nombre de degrés de liberté n'est pas aussi grand que pour les deux autres modèles lorsque l'hypothèse de conservation de la luminance est faite (Horn & Schunck, 1981). En effet cette dernière impose à la transformation $\mathbf{x} \rightarrow \mathbf{x}_{[\mathbf{s}+\mathbf{d}]}$ de conserver la valeur de chaque point de grille, en ne modifiant que leur position. L'utilisation de modèles d'erreurs non-additives est habituellement faite en complément des erreurs additives avec des approches mixtes. Par exemple Fletcher (2007) propose l'utilisation d'erreurs additives et multiplicatives pour autoriser l'emploi de PDFs de type log-normale dans le processus d'assimilation. Lawson & Hansen (2005) ainsi que Ravela et al. (2007) choisissent de corriger les erreurs de déplacement en plus des erreurs additives, ce qui comme ils le montrent dans un

cadre simplifié, permet de rendre les erreurs de prévision bien plus Gaussiennes en amont de la correction d'erreurs additives.

La prise en compte des erreurs de déplacement dans le processus d'assimilation de données en PNT est l'objet de la troisième partie du travail de cette thèse, et est couvert par les chapitres 4.

1.2.4. Modélisation Gaussienne des PDFs d'erreurs

Une estimation objective de la PDF des erreurs d'un processus est possible grâce à des méthodes d'exploration aléatoires dites de Monte-Carlo (Leith, 1974; Tarantola, 2005). Leur efficacité est néanmoins fortement dépendante de la taille de l'échantillon utilisé. Par exemple, les filtres à particules nécessitent l'emploi d'un échantillon dont la taille varie exponentiellement avec les dimensions du problème (Snyder et al., 2008; Van Leeuwen, 2009; Bocquet et al., 2010). Pour les problèmes de grande taille tels que la PNT, une hypothèse simplificatrice est nécessaire pour pouvoir résoudre le problème à moindre coût. L'hypothèse couramment utilisée est que les PDFs d'erreurs sont des Gaussiennes multidimensionnelles définies pour une erreur \mathbf{r} quelconque par

$$\mathcal{P}(\mathbf{r}) \propto \exp \left[-\frac{1}{2} (\mathbf{r} - E[\mathbf{r}])^T \mathbf{C}^{-1} (\mathbf{r} - E[\mathbf{r}]) \right] \quad (1.2)$$

avec \mathbf{C} la matrice de covariances d'erreurs de \mathbf{r} , et $E[\mathbf{r}]$ son espérance mathématique. Les méthodes de type Monte-Carlo peuvent alors être utilisées pour fixer la paramétrisation du modèle Gaussien.

1.3. Cas particulier des erreurs additives et Gaussiennes

Si les erreurs sont supposées être additives, les erreurs d'ébauche, d'observations, et du modèle sont définies par rapport à une trajectoire \mathbf{X}^T , telles que :

$$\begin{cases} \delta \mathbf{x}_b^0 = \delta \mathbf{x} = \mathbf{x}_b - \mathbf{x}^0 & \text{(erreurs d'ébauche)} \\ \delta \mathbf{x}_q^l = \mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1}) & \text{(erreurs modèle)} \\ \delta \mathbf{y}^l = \mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l) & \text{(erreurs d'observations).} \end{cases}$$

De même, en prenant comme référence la "vraie" trajectoire \mathbf{X}_t^T , les vraies erreurs d'ébauche, d'observations et du modèle sont exprimées telles que

$$\begin{cases} \epsilon_b = \mathbf{x}_b - \mathbf{x}_t^0 & \text{(les vraies erreurs d'ébauche)} \\ \epsilon_q^l = \mathbf{x}_t^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}_t^{l-1}) & \text{(les vraies erreurs modèle)} \\ \epsilon_o^l = \mathbf{y}_o^l - \mathcal{H}(\mathbf{x}_t^l) & \text{(les vraies erreurs d'observations)} \end{cases}$$

Les matrices de covariances de ces trois types d'erreurs sont notées respectivement \mathbf{B} , \mathbf{Q} , et \mathbf{R} . Elles sont définies telles que

$$\mathbf{B} = E \left[\epsilon_b \epsilon_b^T \right], \quad \mathbf{Q}^l = E \left[\epsilon_q^l (\epsilon_q^l)^T \right], \quad \mathbf{R}^l = E \left[\epsilon_o^l (\epsilon_o^l)^T \right] \quad (1.3)$$

Avec ce modèle d'erreurs additives, et en supposant l'indépendance deux à deux entre les trois

types d'erreurs, l'Eq.1.1 devient (Bocquet et al., 2010)

$$\mathcal{P}_a(\mathbf{X}^T) \propto \mathcal{P}_b(\mathbf{x}^0 - \mathbf{x}_b) \prod_{l=1}^T \mathcal{P}_o(\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l)) \mathcal{P}_q(\mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1})). \quad (1.4)$$

1.3.1. Hypothèse Gaussienne

Avec une modélisation Gaussienne de \mathcal{P}_b , \mathcal{P}_o , et \mathcal{P}_q , les PDFs d'erreurs sont entièrement spécifiées par leur moyenne et leur matrice de covariances d'erreurs :

- les erreurs d'ébauche $(\mathbf{x}^0 - \mathbf{x}_b)$ sont distribuées selon \mathcal{P}_b qui est modélisée telle que

$$\mathcal{P}_b(\mathbf{x}^0 - \mathbf{x}_b) \propto \exp \left[-\frac{1}{2} (\mathbf{x}^0 - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x}^0 - \mathbf{x}_b) \right]. \quad (1.5)$$

Cette expression de \mathcal{P}_b implique que les erreurs d'ébauche sont non biaisées i.e $E[\mathbf{x}_b - \mathbf{x}_t^0] = 0$.

- pour $l = 1..T$, les innovations définies par $(\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l))$ sont distribuées selon \mathcal{P}_o qui est modélisée telle que

$$\mathcal{P}_o(\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l)) \propto \exp \left[-\frac{1}{2} (\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l))^T (\mathbf{R}^l)^{-1} (\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l)) \right]. \quad (1.6)$$

Cette expression de \mathcal{P}_o implique que les erreurs d'observations sont non biaisées i.e $E[\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}_t^l)] = 0$.

- pour $l = 1..T$, les erreurs modèle $(\mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1}))$ sont supposées être distribuées selon \mathcal{P}_q modélisée telle que

$$\mathcal{P}_q(\mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1})) \propto \exp \left[-\frac{1}{2} (\mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1}))^T (\mathbf{Q}^l)^{-1} (\mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1})) \right]. \quad (1.7)$$

Cette expression de \mathcal{P}_q implique que les erreurs du modèle sont supposées être non biaisées avec $E[\mathbf{x}_t^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}_t^{l-1})] = 0$.

1.3.2. Définition du problème d'optimisation

De la même manière que Trémolet (2006), il a été choisi dans cette partie de présenter le 4D-Var et le 3D-Var comme des approximations du 4D-Var à contrainte faible.

4D-Var à contrainte faible

Avec un approche de type maximum de vraisemblance, la trajectoire analysée \mathbf{X}_a est le mode de \mathcal{P}_a i.e le minimum de $\mathcal{J} = -\ln(\mathcal{P}_a)$. Avec le modèle additif et Gaussien des erreurs (Eq.1.4,

1.5, 1.6, et 1.7), \mathcal{J} est donnée telle que¹ :

$$\mathcal{J}(\mathbf{X}^T) = \frac{1}{2} \|\mathbf{x}^0 - \mathbf{x}_b\|_{\mathbf{B}^{-1}}^2 + \sum_{l=1}^T \left[\frac{1}{2} \|\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^l)\|_{(\mathbf{R}^l)^{-1}}^2 + \frac{1}{2} \|\mathbf{x}^l - \mathcal{M}_{l-1 \rightarrow l}(\mathbf{x}^{l-1})\|_{(\mathbf{Q}^l)^{-1}}^2 \right] \quad (1.8)$$

Un système d'assimilation variationnelle basé sur la minimisation de cette fonction coût, est nommé 4D-Var à contrainte faible : "Var" pour variationnelle, "4D" pour signifier que les observations contraignent une trajectoire qui est l'évolution temporelle d'un champ tridimensionnel, et "contrainte faible" signifie que l'étape de propagation dans la fenêtre d'assimilation ne suit pas exactement la trajectoire déterministe du modèle du fait de la correction des corrections additives des erreurs du modèle. Une description plus complète du 4D-Var à contrainte faible peut être trouvée dans Zupanski (1997) et Trémolet (2006). L'utilisation de cette approche nécessite de spécifier les matrices de covariances d'erreurs \mathbf{Q}^l , \mathbf{B} , et \mathbf{R}^l , ainsi que de manipuler un vecteur trajectoire de très grande taille (vecteur de dimension égale au produit du nombre de sous-créneaux temporels qui composent une fenêtre d'assimilation avec la taille d'un vecteur d'état du modèle).

4D-Var

Une première simplification du 4D-Var à contrainte faible est de négliger les erreurs modèle. Ainsi toute la trajectoire est entièrement spécifiée à partir de l'état initial de la fenêtre d'assimilation : $\mathbf{x}^l = \mathcal{M}_{0 \rightarrow l}(\mathbf{x}^0)$. La fonction coût associée à cette simplification de l'Eq.1.8 est donnée par

$$\mathcal{J}(\mathbf{x}^0) = \frac{1}{2} \|\mathbf{x}^0 - \mathbf{x}_b\|_{\mathbf{B}^{-1}}^2 + \sum_{l=1}^T \frac{1}{2} \|\mathbf{y}_o^l - \mathcal{H}(\mathcal{M}_{0 \rightarrow l}(\mathbf{x}^0))\|_{(\mathbf{R}^l)^{-1}}^2 \quad (1.9)$$

avec comme argument de \mathcal{J} , le vecteur d'état du modèle en début de fenêtre d'assimilation, plutôt que la trajectoire entière. Cette méthode est nommée 4D-Var à contrainte forte, ou plus simplement 4D-Var (Le Dimet & Talagrand, 1986). Habituellement le premier terme du membre de droite est noté \mathcal{J}_b . Il impose une contrainte de rappel vers l'ébauche, alors que le deuxième terme qui est un rappel vers les observations, est noté \mathcal{J}_o .

3D-Var

Une simplification supplémentaire de l'Eq.1.9 est de négliger la propagation par le modèle durant la fenêtre d'assimilation en utilisant l'identité comme opérateur modèle : $\mathcal{M}_{0 \rightarrow l}(\mathbf{x}^0) = \mathbf{x}^0$. Ceci implique que toutes les observations comprises dans la fenêtre d'assimilation sont considérées comme ayant été effectuées au même moment que l'ébauche. Cette approche est noté 3D-Var et la fonction coût associée est donnée par

$$\mathcal{J}(\mathbf{x}^0) = \frac{1}{2} \|\mathbf{x}^0 - \mathbf{x}_b\|_{\mathbf{B}^{-1}}^2 + \sum_{l=1}^T \frac{1}{2} \|\mathbf{y}_o^l - \mathcal{H}(\mathbf{x}^0)\|_{(\mathbf{R}^l)^{-1}}^2 \quad (1.10)$$

1. Pour la concision des notations, la norme d'un vecteur \mathbf{r} , caractérisée par l'inverse d'une matrice de covariance \mathbf{C} , est introduite telle que (Tarantola, 2005)

$$\|\mathbf{r}\|_{\mathbf{C}^{-1}}^2 = \mathbf{r}^T \mathbf{C}^{-1} \mathbf{r}$$

La norme euclidienne est retrouvée en prenant $\mathbf{C} = \mathbf{I}$.

Cette hypothèse implique que, pour les longues fenêtres d'assimilation, le 3D-Var peut être moins efficace que le 4D-Var (Lorenç & Rawlins, 2005). Pour palier en partie ce défaut, des fenêtres d'assimilation plus courtes sont généralement utilisées. Comme nous l'avons vu à la section 1.1.2, c'est notamment le cas pour le système d'assimilation d'AROME basé sur des fenêtres d'assimilation de trois heures, avant de passer récemment sur des fenêtres d'une heure.

3D-FGAT

On peut citer une troisième variante du 4D-Var à contrainte faible : le 3D-FGAT (Fisher & Anderson, 2001; Lawless, 2010). C'est une méthode variationnelle intermédiaire entre le 4D-Var et le 3D-Var, où la correction apportée à l'ébauche $\delta \mathbf{x} = \mathbf{x}^0 - \mathbf{x}_b$ est stationnaire sur toute la fenêtre d'assimilation comme pour le 3D-Var, mais où chaque observation est, comme pour le 4D-Var, comparée à l'ébauche propagée par le modèle complet jusqu'au moment de l'observation. La fonction coût associée est définie telle que

$$\mathcal{J}(\delta \mathbf{x}) = \frac{1}{2} \|\delta \mathbf{x}\|_{\mathbf{B}^{-1}}^2 + \sum_{l=1}^T \frac{1}{2} \|\mathbf{y}_o^l - \mathcal{H}(\mathcal{M}_{0 \rightarrow l}(\mathbf{x}^b) + \delta \mathbf{x})\|_{(\mathbf{R}^l)^{-1}}^2. \quad (1.11)$$

Cette méthode nécessite une formulation dite incrémentale (Courtier et al., 1994), où la solution recherchée est un incrément d'analyse $\delta \mathbf{x}_a = \mathbf{x}_a - \mathbf{x}_b$, plutôt que l'analyse elle-même.

1.3.3. Résolution du problème d'optimisation

Les systèmes d'assimilation des modèles ARPEGE et AROME utilisent respectivement des schémas de types 4D-Var et 3D-Var basés sur des formulations incrémentales. Pour rester dans un cadre plus simplifié, seule la résolution du 3D-Var incrémental sera abordée dans cette section. Ce choix est suffisant pour introduire le concept de modélisation de la matrice \mathbf{B} , ainsi que pour servir de base à la mise en place de l'algorithme variationnel d'estimation des erreurs de déplacement présenté au chapitre 4.

L'opérateur d'observations \mathcal{H} étant généralement non-linéaire, la minimisation de \mathcal{J} est un problème non-linéaire non-quadratique de type moindres carrés. Il est donc nécessaire d'utiliser des méthodes de minimisation non linéaires et qui soient adaptées aux problèmes de grande taille tels que la PNT.

Méthode quasi-linéaire de type Gauss-Newton

En suivant Lawless et al. (2005) et Gratton et al. (2007), la résolution quasi-linéaire du 3D-Var incrémental (ou du 4D-Var incrémental) est assimilée à une résolution de type Gauss-Newton. La minimisation est ici un problème des moindres carrés où \mathcal{J} est alors factorisable par

$$2\mathcal{J}(\mathbf{x}) = f(\mathbf{x})^T f(\mathbf{x})$$

avec $f(\mathbf{x})$ définie par identification avec l'Eq.1.10 telle que

$$f(\mathbf{x}) = \begin{pmatrix} \mathbf{B}^{-\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \mathbf{x} - \mathbf{x}_b \\ \mathbf{y}^0 - \mathcal{H}(\mathbf{x}) \end{pmatrix}.$$

Pour simplifier les notations, dans cette formulation contrairement à l'Eq.1.10, la dépendance temporelle des observations et de leurs statistiques d'erreurs n'est plus explicitée. \mathbf{F} , le jacobien

de $f(\mathbf{x})$ est alors définie par

$$\mathbf{F} = \begin{pmatrix} \mathbf{B}^{-\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \mathbf{I} \\ \mathbf{H} \end{pmatrix}$$

avec \mathbf{I} la matrice identité, et \mathbf{H} l'opérateur tangent linéaire de \mathcal{H} . L'algorithme de minimisation de la fonction coût \mathcal{J} par une méthode de type Gauss-Newton est alors définie telle que (Gratton et al., 2007) :

Algorithme de Gauss-Newton

- | | |
|--|---------------------------|
| 1. $\mathbf{x}_0 = \mathbf{x}_b$ | (initialisation) |
| 2. à faire jusqu'à la convergence | (boucle externe sur k) |
| — résoudre $\mathbf{F}(\mathbf{x}_k)^T \mathbf{F}(\mathbf{x}_k) \widehat{\delta \mathbf{x}}_k = -\mathbf{F}(\mathbf{x}_k)^T f(\mathbf{x}_k)$ | (boucle interne) |
| — $\mathbf{x}_{k+1} = \mathbf{x}_k + \widehat{\delta \mathbf{x}}_k$ | (actualisation) |
| 3. $\mathbf{x}_a = \mathbf{x}_{k+1}$ | (état d'analyse) |
-

Chaque boucle interne (cf algorithme de Gauss-Newton) est en fait la minimisation par rapport à un incrément d'amplitude $\delta \mathbf{x}_k$, de la fonction coût quadratique, J , définie comme la linéarisation de \mathcal{J} autour de \mathbf{x}_k telle que

$$J(\delta \mathbf{x}_k) = \frac{1}{2} \|\mathbf{x}_k + \delta \mathbf{x}_k - \mathbf{x}_b\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_k) - \mathbf{H} \delta \mathbf{x}_k\|_{\mathbf{R}^{-1}}^2. \quad (1.12)$$

La forme analytique de $\widehat{\delta \mathbf{x}}_k$, défini comme l'argument de J à son unique minimum, est donnée par (deux formulations équivalentes sont proposées)

$$\widehat{\delta \mathbf{x}}_k = \left(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right)^{-1} \left[\mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}_o - \mathcal{H}(\mathbf{x}_k)) - \mathbf{B}^{-1} (\mathbf{x}_k - \mathbf{x}_b) \right] \quad (1.13)$$

$$= \mathbf{B} \mathbf{H}^T \left(\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R} \right)^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_k) - \mathbf{R} \mathbf{H} \mathbf{B}^{-1} (\mathbf{x}_k - \mathbf{x}_b) \right) \quad (1.14)$$

avec l'utilisation de la formule de Sherman-Morrison-Woodbury pour passer de l'Eq.1.13 à l'Eq.1.14 (Sherman & Morrison, 1949, 1950).

L'analogie entre la résolution du 3D-Var (ou 4D-Var) incrémental et l'algorithme de Gauss-Newton n'est valide que si \mathbf{H} est exactement le tangent linéaire de \mathcal{H} . La convergence du système entier dans le cas où des simplifications sont apportés à \mathbf{H} et où les boucles internes sont stoppées avant la convergence exacte est étudiée par Gratton et al. (2007).

Préconditionnement du problème linéarisé

Étant donné que les matrices \mathbf{B} , \mathbf{R} , et \mathbf{H} ne sont disponibles que sous forme d'opérateurs, l'incrément $\delta \mathbf{x}$, solution de chaque boucle interne, est estimé par l'inversion du système linéaire donné par la boucle interne de l'algorithme de Gauss-Newton tel que

$$\left(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \right) \widehat{\delta \mathbf{x}} = \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y}_o - \mathcal{H}(\mathbf{x})) - \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) \quad (1.15)$$

(Les indices k ont été retirés pour simplifier les équations). Le système 1.15 est mal conditionné et l'opérateur \mathbf{B}^{-1} est difficilement implémentable. \mathbf{R} est, lui, habituellement modélisé par

une matrice diagonale ce qui permet d'obtenir facilement l'opérateur inverse \mathbf{R}^{-1} . Il est alors nécessaire de recourir à une formulation préconditionnée (Haben et al., 2011; Gürol et al., 2013; Nocedal & Wright, 2006) obtenue par un changement de la variable de contrôle (CVT). En PNT, au moins deux choix sont possibles pour la nouvelle variable de contrôle $\boldsymbol{\lambda}$ du système préconditionné :

- avec un préconditionnement par $\mathbf{B}^{\frac{1}{2}}$, $\boldsymbol{\lambda} = \mathbf{U}^{-1}\boldsymbol{\delta x}$, avec \mathbf{U} vérifiant $\mathbf{B} = \mathbf{U}\mathbf{U}^T$.
- avec un préconditionnement par \mathbf{B} , $\boldsymbol{\lambda} = \mathbf{B}^{-1}\boldsymbol{\delta x}$.

Même si c'est le préconditionnement par $\mathbf{B}^{\frac{1}{2}}$ qui est utilisé en opérationnel pour les modèles ARPEGE et AROME, seul le préconditionnement par \mathbf{B} est présenté, et utilisé durant cette thèse. Il a l'avantage de s'affranchir de la factorisation de \mathbf{B} tout en apportant des performances équivalentes au préconditionnement par $\mathbf{B}^{\frac{1}{2}}$. Avec un système préconditionné par \mathbf{B} , l'Eq.1.15 devient

$$\left(\mathbf{I} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B}\right) \boldsymbol{\lambda} = \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x})\right) - \mathbf{B}^{-1} \left(\mathbf{x} - \mathbf{x}_b\right) \quad (1.16)$$

La matrice Hessienne $\left(\mathbf{I} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B}\right)$ a notamment la particularité d'être symétrique (auto-adjointe) par rapport au produit scalaire défini par $\mathbf{x}_1^T \mathbf{B} \mathbf{x}_2$ pour tous vecteurs d'état $\mathbf{x}_1, \mathbf{x}_2$. Le système linéaire 1.16 peut alors être inversé avec une méthode itérative telle que le gradient conjugué (méthode BCG, Gürol et al., 2013).

Schéma multi-incrémental

Afin de réduire les coûts de calcul de l'étape de minimisation de \mathcal{J} , mais également pour se placer à des résolutions où la linéarisation de certains opérateurs est plus aisée du fait de l'utilisation de physique simplifiée pour le modèle (Veersé & Thépaut, 1998), les boucles internes peuvent être effectuées à des résolutions réduites par rapport à la pleine résolution du modèle (Trémolet, 2007). Pour un système ayant au moins deux itérations dans la boucle externe, les résolutions spatiales de calcul sont choisies de façon croissante (mailles de calcul de plus en plus fines).

1.3.4. Configurations opérationnelles

Pour le 4D-Var incrémental d'ARPEGE, afin de minimiser la fonction coût \mathcal{J} , la méthode quasi-linéaire exposée à la section 1.3.3 est paramétrisée telle que :

- la boucle externe se compose de deux itérations i.e d'une seule re-linéarisation du problème.
- chacun des deux problèmes linéarisés est résolu par la méthode itérative du gradient conjugué, dont le nombre d'itérations dans la boucle interne est fixé respectivement à 25 puis 30 itérations.
- le coût de calcul engendré par la minimisation est diminué grâce à la dégradation des résolutions de calcul de chacune des boucles internes. Alors que le modèle déterministe est en pleine résolution spectrale à T798, la première boucle interne est effectuée à T107, et la seconde à T323
- chaque boucle interne utilise un préconditionnement par $\mathbf{B}^{\frac{1}{2}}$.

Le 3D-Var incrémental d'AROME minimise \mathcal{J} lors d'une unique minimisation de 50 itérations, à pleine résolution (Seity et al., 2011). La minimisation du problème linéarisé quadratique utilise également un préconditionnement par $\mathbf{B}^{\frac{1}{2}}$.

1.3.5. Cas linéaire Gaussien et filtres de Kalman

Avec un modèle d'erreurs additif et Gaussien, ainsi qu'un cadre totalement linéaire (avec notamment un opérateur d'observation linéaire), la fonction coût \mathcal{J} définie par l'Eq.1.12 est quadratique et l'estimation de \mathbf{x}_a ne nécessite qu'une unique itération dans la boucle externe. En reprenant l'Eq.1.14 avec $\mathbf{x}_k = \mathbf{x}_b$, l'état d'analyse s'exprime alors par

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{B}\mathbf{H}^T \left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R} \right)^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b) \right) \quad (1.17)$$

Cette solution est appelée le *Best Linear Unbiased Operator* (BLUE, Kalnay (2003)). La suite d'opérateurs $\mathbf{B}\mathbf{H}^T \left(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R} \right)^{-1}$ est habituellement nommée "gain de Kalman", et est notée \mathbf{K}_{BLUE} .

Un des avantages du cadre linéaire Gaussien est que les matrices de covariances d'erreurs de l'état du modèle (ébauche et analyse) peuvent être explicitement calculées durant tout le cycle d'assimilation (étape de propagation et étape d'assimilation). Un système d'assimilation complet dans ce cadre linéaire et Gaussien est appelé filtre de Kalman et peut être exprimé tel que

$$\left\{ \begin{array}{ll} \mathbf{x}_b = \mathbf{M}\mathbf{x}_a^- & \text{(propagation du vecteur d'état)} \\ \mathbf{B} = \mathbf{M}\mathbf{A}^- \mathbf{M}^T + \mathbf{Q} & \text{(propagation des covariances d'erreurs)} \\ \mathbf{x}_a = \mathbf{x}_b + \mathbf{K}_{BLUE} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b) \right) & \text{(actualisation du vecteur d'état)} \\ \mathbf{A} = \left(\mathbf{I} - \mathbf{K}_{BLUE} \mathbf{H} \right) \mathbf{B} & \text{(actualisation des covariances d'erreurs)} \end{array} \right.$$

avec \mathbf{M} le modèle linéaire, \mathbf{Q} la matrice de covariances d'erreurs modèle, et l'exposant "-" indiquant l'instant de l'analyse du cycle précédent.

Les défauts du filtre de Kalman sont la non prise en compte des non-linéarités, le coût prohibitif du stockage des matrices de covariances d'erreurs lorsqu'il est appliqué à des problèmes de grande taille, et la dégénérescence du filtre due à la sous-dispersivité de l'ensemble. De nombreuses variantes du filtre de Kalman sont disponibles pour alléger ces contraintes. Par exemple le filtre de Kalman étendu (EKF, Gelb, 1974), permet de prendre en compte une partie des non-linéarités du modèle et de l'opérateur d'observations. Une autre variante largement utilisée est le filtre de Kalman d'ensemble (ENKF, Evensen, 1994, 2003). L'ENKF permet de propager, avec le modèle non-linéaire complet, les statistiques d'erreurs d'ébauche, sans avoir à les stocker explicitement grâce à l'utilisation d'un ensemble (voir section 1.4.3).

L'approche choisie dans cette section est de dériver la formulation du filtre de Kalman comme une approche linéaire Gaussienne du problème d'estimation Bayésienne des erreurs additives d'amplitude. L'approche inverse est également possible, i.e de retrouver les équations du filtre de Kalman (notamment le BLUE Eq.1.17), uniquement avec les hypothèses de linéarité et de nullité des biais d'erreurs additives d'amplitudes, puis de faire l'hypothèse de Gaussianité des PDFs d'erreurs pour faire le lien avec le problème d'estimation Bayésienne.

1.4. Modélisation de la matrice B

Dans un cadre Gaussien, les matrices de covariances d'erreurs \mathbf{B} , \mathbf{R} ou \mathbf{Q} définissent à elles-seules les PDFs d'erreurs non biaisées. Leur spécification exacte n'est pas possible pour deux raisons. La première est que leurs définitions, regroupées dans l'Eq.1.3, utilisent l'état vrai,

qui est inconnu. La deuxième raison, valable pour \mathbf{B} et \mathbf{Q} est que leur grande taille rend impossible leur stockage et leur manipulation, même s'il était possible de les calculer. Pour \mathbf{B} , par exemple, la taille d'un vecteur d'état dans l'espace modèle a pour ordre de grandeur $\mathcal{O}(10^9)$. \mathbf{B} est donc de taille $\mathcal{O}(10^{18})$ (soit 8 exaotets pour une représentation en réel à double précision) ce qui est bien au dessus des moyens informatiques actuels. Il est donc nécessaire d'exprimer ces matrices comme une suite d'opérateurs simplifiés. Ces opérateurs doivent être optimisés pour que le produit entre une matrice de covariances ainsi constituée et un vecteur d'état soit le moins coûteux possible. C'est en effet une opération très fréquente durant le processus d'assimilation étant donné que \mathbf{B} est appliquée à chaque itération de la minimisation. Dans cette partie, seule la modélisation de l'opérateur de covariances d'erreurs d'ébauche (noté également \mathbf{B}) est abordée.

1.4.1. Rôle de \mathbf{B}

Un cadre idéalisé du BLUE est proposé pour comprendre le rôle joué par chacun des coefficients de la matrice \mathbf{B} (inspiré de Bannister (2008a) et Ménétrier (2014)). Supposons que l'on s'intéresse à un modèle ayant deux variables pronostiques U et V , et utilisant une grille comportant N points de grille numérotés par $k = 1..N$. On veut assimiler une unique observation $y_{o,U}^k$ de la variable U située exactement sur un point de grille noté k . Dans ce cas, \mathbf{R} est un scalaire fixant la variance d'erreur d'observations $\sigma_{o,U}^2$, et \mathcal{H} est un vecteur nul sauf pour U au niveau du coefficient k où il est égal à 1. Tout vecteur d'état \mathbf{x} peut être noté $\mathbf{x}^T = (\mathbf{x}_U^T, \mathbf{x}_V^T)$, comme la concaténation de \mathbf{x}_U , et \mathbf{x}_V , les champs de U et de V . La matrice \mathbf{B} définie selon l'Eq.1.3 prend la forme

$$\mathbf{B} = \left(\begin{array}{c|c} \mathbf{B}_U & \mathbf{B}_{UV} \\ \hline \mathbf{B}_{UV} & \mathbf{B}_V \end{array} \right)$$

avec \mathbf{B}_U , \mathbf{B}_V les matrices d'auto-covariances de U et de V , et \mathbf{B}_{UV} la matrice de covariances inter-variables entre U et V . Par exemple, le coefficient (i,j) de la matrice \mathbf{B}_U , noté $\mathbf{B}_U(i,j)$, est défini par

$$\mathbf{B}_U(i,j) = E \left[(\mathbf{x}_{b,U}(i) - \mathbf{x}_{t,U}(i)) (\mathbf{x}_{b,U}(j) - \mathbf{x}_{t,U}(j)) \right]$$

Les coefficients diagonaux ($i = j$) de \mathbf{B}_U et \mathbf{B}_V sont les variances d'erreurs pour U et V . De même par définition $\mathbf{B}_{UV}(i,j) = E \left[(\mathbf{x}_{b,U}(i) - \mathbf{x}_{t,U}(i)) (\mathbf{x}_{b,V}(j) - \mathbf{x}_{t,V}(j)) \right]$.

Selon l'expression du BLUE (Eq.1.17), l'incrément d'analyse pour les variables U et V , en un point de grille l , est donné tel que

$$\begin{cases} \mathbf{x}_{a,U}(l) - \mathbf{x}_{b,U}(l) = \mathbf{B}_U(l,k) \times \frac{\mathbf{y}_{o,U}(k) - \mathbf{x}_{b,U}(k)}{\mathbf{B}_U(k,k) + \sigma_{o,U}^2} \\ \mathbf{x}_{a,V}(l) - \mathbf{x}_{b,V}(l) = \mathbf{B}_{UV}(l,k) \times \frac{\mathbf{y}_{o,U}(k) - \mathbf{x}_{b,U}(k)}{\mathbf{B}_U(k,k) + \sigma_{o,U}^2} \end{cases} \quad (1.18)$$

Comme cela est mentionné dans l'introduction, \mathbf{B} quantifie la similarité qu'il y a entre deux points de grille. Cette similarité dépend de la position respective des deux points ainsi que des variables représentées. De manière plus détaillée, l'Eq.1.18 indique que

- au point d'observation k , l'incrément d'analyse pour la variable observée U est fixé à $\frac{\mathbf{B}_U(k,k)}{\mathbf{B}_U(k,k) + \sigma_{o,U}^2} (\mathbf{y}_{o,U}(k) - \mathbf{x}_{b,U}(k))$. Si la variance d'erreurs d'ébauche $\mathbf{B}_U(k,k)$ est d'un ordre de grandeur supérieur à la variance d'erreur d'observations $\sigma_{o,U}^2$, l'observation est beaucoup plus fiable que l'ébauche et donc $\mathbf{x}_{a,U}(k) \approx \mathbf{y}_{o,U}(k)$

- la valeur de l'incrément pour un point, $l \neq k$, est pondérée par $\mathbf{B}_U(l, k)$. La diffusion de l'information dans les trois dimensions de l'espace, du point d'observation vers les autres points de grille, est modélisée par les termes non-diagonaux de \mathbf{B}_U : les auto-covariances horizontales et verticales
- la valeur de l'incrément pour la variable non-observée V , est pondérée par les coefficients $\mathbf{B}_{UV}(l, k)$ des blocs non-diagonaux de \mathbf{B} . L'opérateur \mathbf{B}_{UV} permet de retranscrire les corrélations qui existe entre les variables U et V : les relations de balance.

La propagation de l'information du point d'observation vers les autres points de grille est notamment importante pour les zones peu ou pas observées qui, grâce aux covariances spatiales contenues dans \mathbf{B} , vont elles aussi bénéficier de l'information apportée par les observations.

1.4.2. Modélisation de \mathbf{B} en sous-opérateurs

Derber & Bouttier (1999) proposent une modélisation de \mathbf{B} telle que $\mathbf{B} = \mathbf{U}\mathbf{U}^T$, la matrice \mathbf{U} étant elle-même est découpée en sous-opérateurs permettant de prendre en compte les relations de balances, les variances et les corrélations spatiales verticales et horizontales. \mathbf{U} peut alors s'écrire :

$$\mathbf{U} = \mathbf{K}\mathbf{B}_u^{\frac{1}{2}}$$

avec

- \mathbf{K} est un opérateur de balance permettant de propager l'information contenue dans l'incrément d'analyse des variables observées vers les variables non-observées. Inversement, \mathbf{K}^{-1} permet de retirer au moins une partie des covariances inter-variables ce qui permet de manipuler des nouvelles variables qui ne sont pas corrélées entre elles. Chacune des variables est alors séparée en deux parties : une partie débalancée (supposée décorrélée des autres variables débalancées) et une partie balancée retirée par \mathbf{K}^{-1} . \mathbf{K} est habituellement un opérateur multi-linéaire de régression.
- $\mathbf{B}_u = \mathbf{B}_u^{\frac{1}{2}}\mathbf{B}_u^{\frac{T}{2}}$ est l'opérateur de covariances des erreurs d'ébauche pour les variables non-balancées. Sa représentation matricielle est diagonale par bloc, chaque bloc contenant les auto-covariances spatiales pour une variable non-balancée. $\mathbf{B}_u^{\frac{1}{2}}$ est appelée transformée spatiale, et est construite telle que

$$\mathbf{B}_u^{\frac{1}{2}} = \mathbf{V}^{\frac{1}{2}}\mathbf{C}_v^{\frac{1}{2}}\mathbf{C}_h^{\frac{1}{2}}$$

où $\mathbf{V}^{\frac{1}{2}}$ contient les écarts-types d'erreurs d'ébauche, $\mathbf{C}_v^{\frac{1}{2}}$ contient les corrélations entre les niveaux verticaux du modèle, et $\mathbf{C}_h^{\frac{1}{2}}$ contient les corrélations horizontales entre les points de grille.

Plus de détails peuvent être trouvés, par exemple, dans Bannister (2008b). Cette formulation a au moins deux avantages. Le premier est que, par construction, la matrice \mathbf{B}_u modélisée vérifie bien deux propriétés nécessaires d'une matrice de covariance d'erreurs (Rasmussen & Williams, 2006) : la symétrie et la semi-définie positivité ($\mathbf{C}_v^{\frac{1}{2}}$, $\mathbf{C}_h^{\frac{1}{2}}$, et $\mathbf{V}^{\frac{1}{2}}$ étant des racines carrées de matrices de covariances). Le deuxième avantage est que la factorisation $\mathbf{B} = \mathbf{K}\mathbf{B}_u^{\frac{1}{2}}\mathbf{B}_u^{\frac{T}{2}}\mathbf{K}^T$ permet d'utiliser le préconditionnement par $\mathbf{B}_u^{\frac{1}{2}}$ sans développement supplémentaire.

1.4.3. Méthodes d'estimation de \mathbf{B}

Dans cette section, sont présentés trois algorithmes d'approximation des erreurs d'ébauche à partir desquelles il est possible de calibrer l'opérateur \mathbf{B} (selon l'Eq.1.3). Ils sont introduits

selon un coût de calcul croissant : méthode des innovations, méthode NMC, et méthode d'ensemble.

Méthode des innovations

Dans le cas où l'on ne possède pas d'ensemble de prévisions, il est possible d'estimer \mathbf{B} à partir des covariances d'innovations $(\mathbf{y} - \mathcal{H}(\mathbf{x}_b))$ (Hollingsworth & Lönnberg, 1986). En effet, en supposant que les erreurs d'observations et d'ébauches sont décorréelées, et que les erreurs d'observations ne sont pas corrélées spatialement (\mathbf{R} est diagonale), les termes non-diagonaux de \mathbf{B} (pas les variances) peuvent être approximés dans l'espace des observations tels que

$$\mathbf{HBH}^T = E[(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b))(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_b))^T]$$

L'espérance est approximée par une moyenne spatiale et/ou temporelle. Les variances sont ensuite extrapolées à partir des covariances à courtes distances.

La limitation de cette méthode est qu'elle nécessite le besoin d'avoir des observations décorréelées en nombre suffisant, ainsi que des erreurs d'ébauche et d'observations non-biaisées. De plus, l'estimation des statistiques d'erreurs d'ébauche est faite dans l'espace des observations (\mathbf{B} elle-même n'est jamais calculée Bannister, 2008a).

Méthode NMC

Par rapport à la méthode des innovations, la méthode NMC (qui ne nécessite pas non plus d'ensemble) permet de s'affranchir de l'utilisation explicite des observations dans l'estimation de \mathbf{B} (Parrish & Derber, 1992). En effet l'erreur d'ébauche à un instant k est approximée par la différence entre deux prévisions d'échéances différentes mais valide à l'instant k . Pour éviter les effets diurnes, des prévisions décalées de exactement 24h sont habituellement choisies. Par exemple Pereira & Berre (2006) utilisent des prévisions à 12 et 36 heures d'échéance (respectivement \mathbf{x}_{12} et \mathbf{x}_{36}) pour définir \mathbf{B} telle que

$$2\mathbf{B} = E[(\mathbf{x}_{36} - \mathbf{x}_{12})(\mathbf{x}_{36} - \mathbf{x}_{12})^T]$$

D'après Bannister (2008a), les limitations inhérentes à la méthode NMC sont tout d'abord le manque de justification du lien existant entre l'erreur d'ébauche (habituellement une prévision à courte échéance) et une différence de prévision à plus longue échéance. De plus, le manque d'information dans les zones peu observées conduisent à une sous-dispersion dans ces dernières (Pereira & Berre, 2006). Enfin, la méthode NMC nécessite une moyenne temporelle sur une longue période, ce qui limite la dépendance au flux météorologique des covariances estimées.

Méthode d'ensemble

Pour faire un "sondage" dans les statistiques d'erreurs d'ébauche (Joly & Descamps, 2009), les méthodes d'ensemble utilisent un échantillon de plusieurs répliques, perturbées les unes par rapport aux autres, de prévisions que l'on fait évoluer en parallèle. Si l'on note $\mathbf{x}_{p=1..N_s}$ chacun des membres d'un ensemble de taille N_s , l'espérance mathématique de l'Eq.1.3 est alors approximée par une moyenne sur les membres de l'ensemble. \mathbf{B} peut être estimée telle que (formulation non-biaisée)

$$\mathbf{B} = \frac{1}{N_s - 1} \sum_{p=1}^{N_s} (\mathbf{x}_p - \bar{\mathbf{x}})(\mathbf{x}_p - \bar{\mathbf{x}})^T$$

avec $\bar{\mathbf{x}}$ la moyenne sur l'ensemble

$$\bar{\mathbf{x}} = \frac{1}{N_s} \sum_{p=1}^{N_s} \mathbf{x}_p$$

De même que pour les autres méthodes de type Monte-Carlo, un nombre élevé de membres doit apporter une meilleure estimation des vraies covariances. En PNT, les tailles d'ensemble reste cependant faible avec $N_s = o(10^2)$, nécessitant l'utilisation de méthodes de filtrage du bruit d'échantillonnage (e.g, Pannekoucke et al., 2007; Raynaud, 2010; Ménétrier, 2014).

Il existe de nombreuses configurations pour générer un ensemble de prévisions, par exemple à partir d'une des variantes du filtre de Kalman (ENKF), ou, dans un cadre variationnel, d'une assimilation d'ensemble (EDA, cf section 1.4.4). Ménétrier (2014) propose trois critères pour les différencier les uns des autres :

1. l'utilisation ou non d'une prévision non perturbée (dite déterministe), et des différences de configuration entre la déterministe et celle de chacun des membres perturbés de l'ensemble (différences de modèle ou de résolution)
2. les degrés de complexité du modèle utilisé pour \mathbf{B} . Par exemple, si on utilise un modèle Gaussien pour les covariances spatiales, l'ensemble les spécifie entièrement uniquement avec deux paramètres : des variances et des longueurs de portées d'erreurs d'ébauche.
3. le mode de perturbations des membres, les uns par rapport aux autres, notamment comment ces perturbations sont remaniées d'un cycle à l'autre.

Une comparaison de l'estimation faite avec une méthode d'ensemble et une méthode NMC est faite par Pereira & Berre (2006). Dans ce travail de thèse, les statistiques des erreurs d'ébauche seront uniquement estimées à l'aide d'EDAs.

1.4.4. L'assimilation d'ensemble

Météo-France utilise en opérationnel des méthodes d'ensemble pour caractériser \mathbf{B} du modèle ARPEGE grâce à une EDA : l'AEARP¹. Avec une EDA, la distribution des erreurs d'ébauche est modélisée par un ensemble de répliques perturbées de la prévision déterministe. Une représentation simplifiée de l'évolution d'un ensemble d'assimilation est donné par la Fig.1.6. Le système d'assimilation/prévision utilisé pour faire évoluer chacune des répliques reprend la plupart des caractéristiques du système déterministe avec notamment le même modèle de propagation non-linéaire \mathcal{M} , et le même algorithme d'assimilation (4D-Var incrémental pour l'AEARP). Néanmoins, des économies sont nécessaires pour pouvoir respecter les contraintes imposées par les ressources de calcul disponibles tout en conservant une taille d'ensemble assez grande. Par exemple, cela est fait grâce à l'utilisation de résolutions de travail dégradées par rapport à celle du déterministe. Pour l'AEARP, une autre simplification est ajoutée lors de l'analyse puisque la minimisation ne s'effectue qu'avec une unique itération dans la boucle externe.

Plusieurs modes de perturbations des membres d'un ensemble sont possibles. Pour l'AEARP, les perturbations sont imposées de deux façons :

- chaque membre assimile un jeu d'observations différent obtenu en perturbant selon \mathbf{R} les observations non perturbées, \mathbf{y}_o .
- les perturbations de prévisions (écarts entre l'ébauche déterministe et chacune des ébauches perturbées) sont inflatées par un coefficient multiplicateur définie selon Raynaud et al. (2012) comme paramétrisation simplifiée des erreurs modèle.

1. Assimilation d'Ensemble d'ARPEGE, Berre et al. (2006); Berre & Desroziers (2010)

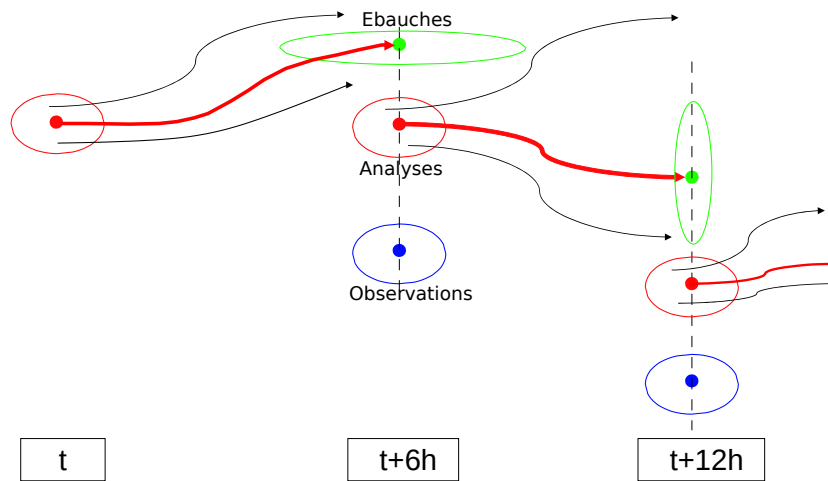


Figure 1.6. – D’après Ehrendorfer (2008), schéma expliquant le principe du déroulement d’un ensemble d’assimilation durant un cycle d’assimilation de données. Chaque flèche entre les réseaux d’assimilation (t , $t+6h$, et $t+12h$) représente la propagation d’une analyse perturbée par le modèle non linéaire \mathcal{M} pour donner une ébauche perturbée au réseau suivant. Pour chacun des réseaux, le point vert représente la prévision déterministe initialisée par l’analyse non-perturbée du réseau précédent (point rouge). Le point bleu représente les observations \mathbf{y}_o non perturbées. Les distributions d’erreurs d’ébauche (caractérisées par \mathbf{B}), d’analyse (caractérisées par la matrice de covariance d’erreurs d’analyse \mathbf{A}), et d’observations (caractérisées par \mathbf{R}) sont modélisées respectivement par l’ensemble des ébauches perturbées (ellipses vertes), l’ensemble des analyses perturbées (ellipses rouges), et l’ensemble des observations perturbées (ellipses bleues). Chaque analyse perturbée est la combinaison d’une ébauche perturbée et d’un jeu d’observations perturbées.

L’AEARP utilisée pour cette thèse a été mise en place par Yann Michel, et se compose de 90 membres.

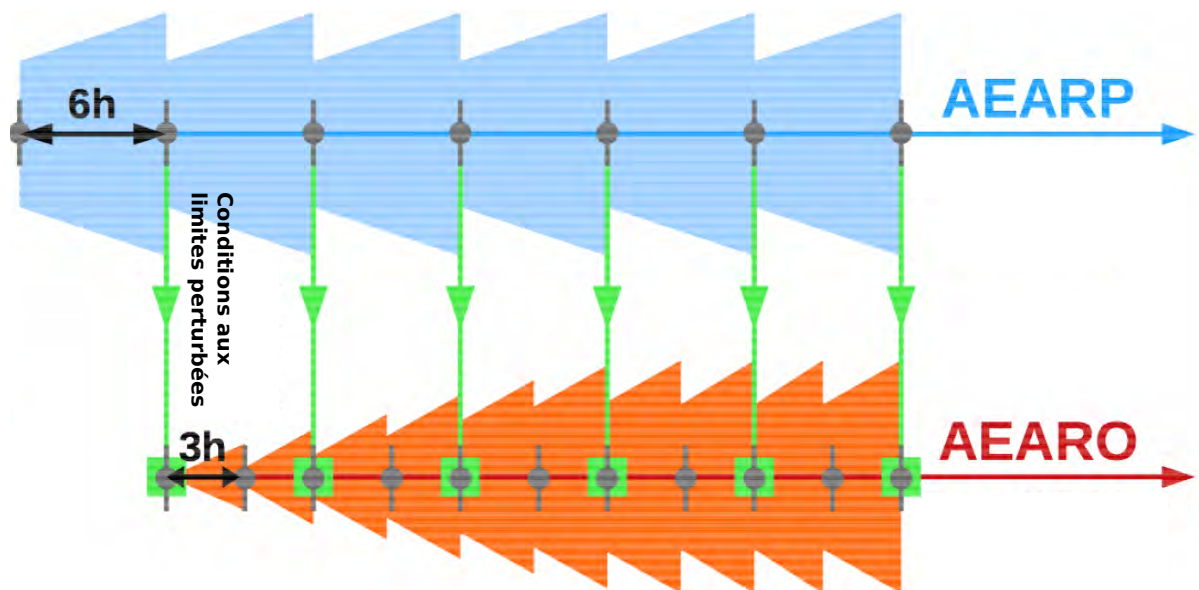


Figure 1.7. – Schéma d’évolution du système ensembliste AEARO+AEARP (figure largement inspirée d’une illustration produite par Benjamin Ménétrier).

Les expériences présentées dans cette thèse utilisent également une EDA pour AROME (nommée AEARO) qui n'est pas encore utilisée de manière opérationnelle (Fisher, 2003; Kucukkaraca & Fisher, 2006; Berre et al., 2006). La version expérimentale utilisée est construite selon Ménétrier et al. (2014) et se compose également de 90 membres. Comme pour le déterministe, le schéma d'assimilation utilisé est de type 3D-Var. Les perturbations d'ensemble sont obtenues par perturbation des observations (de même que l'AEARP) et par perturbation des conditions aux limites latérales (chaque membre de l'AEARO utilise un membre différent de l'AEARP comme champ coupleur). Aucune inflation des perturbations d'ébauche n'est utilisée. L'organisation du système AEARO+AEARP est schématisée par la Fig.1.7. Ménétrier (2014) note qu'une période d'initialisation assez courte de 3 réseaux est nécessaire pour déclarer l'ensemble "assez dispersé".

Deux AEAROs ont été utilisés durant cette thèse :

- le chapitre 3 et 4 utilise l'AEARO d'AROME-France conçu et utilisé par Ménétrier et al. (2014). Il est noté AEARO-France par la suite, et sauf si cela est spécifié autrement ce sont les prévisions à 3h valides le 4 Novembre 2011 à 03 UTC qui sont utilisées.
- le chapitre 4 utilise également un ensemble d'assimilation pour le modèle AROME-WMED. Il a été construit par Yann Michel selon la même conception qu'AEARO-France. Il est noté AEARO-WMED par la suite, et sauf si cela est spécifié autrement ce sont les prévisions à 3h valides le 8 Novembre 2011 à 03 UTC qui sont utilisées. Cette date a été choisie pour le cas de Médicane en cours.

1.4.5. Configurations opérationnelles

Dans cette section sont présentés les éléments techniques de la modélisation de \mathbf{B} pour les configurations opérationnelles des systèmes d'assimilation d'ARPEGE et d'AROME.

ARPEGE déterministe

Pour ARPEGE, les principales caractéristiques de la matrice \mathbf{B} utilisée dans le 4D-Var incrémental sont

- l'opérateur de balance est défini selon Derber & Bouttier (1999) et Fisher (2003).
- les variances d'erreurs sont estimées à partir de l'AEARP (cf section 1.4.4) composée de 6 membres à une résolution horizontale dégradée avoisinant les 50km.
- les corrélations horizontales sont modélisées de manière homogène et isotrope par une formulation diagonale dans l'espace de Fourier.
- la direction verticale est dite non-séparable de l'horizontale i.e les corrélations verticales dépendent des échelles horizontales représentées ainsi que de la position géographique (Ingleby, 2001). Les corrélations verticales sont modélisées à l'aide d'une projection sur une base de modes propres verticaux (Bannister, 2008b).

AROME déterministe

Les principales différences de la matrice \mathbf{B} utilisée par le 3D-Var incrémental d'AROME par rapport à celle d'ARPEGE sont

- une approche multivariée également pour l'humidité spécifique (Berre, 2000), qui permet notamment de prendre en compte le fort couplage entre cette variable et la divergence dans les zones précipitantes (Montmerle & Berre, 2010).

- l'utilisation de variances d'erreurs climatologiques, calibrées par Brousseau et al. (2011) à partir d'EDAs de 6 membres lancées sur des périodes hivernales et estivales. La possibilité d'estimer les covariances d'erreurs d'ébauche du jour à partir d'un EDA pour AROME a été investiguée par Ménétrier et al. (2014). Malgré le développement de techniques efficaces de filtrages objectifs du bruit d'échantillonnage, la prise en compte des variances filtrées dans \mathbf{B} montre, dans la configuration actuelle, un impact neutre ou négatif sur les scores de prévisions (Ménétrier, 2014).
- les corrélations horizontales sont modélisées de manière homogène et isotrope par une formulation diagonale dans l'espace bi-Fourier.

1.5. Conclusion

Les concepts de base associés aux systèmes d'assimilation variationnelle les plus courants, basés sur une modélisation additive Gaussienne des erreurs, ont été introduits dans ce chapitre. Par la même occasion, les configurations des systèmes de PNT ARPEGE et AROME, utilisées en opérationnel à Météo-France, ont été présentées. L'accent a été mis sur la caractérisation des erreurs d'ébauche, en laissant de côté les enjeux liés aux observations.

Trois problématiques peuvent être mises en avant. La première est la nécessité dans la pratique de faire des simplifications importantes pour la modélisation des opérateurs de covariances tels que \mathbf{B} , alors que cette dernière a une influence cruciale sur la forme des incréments d'analyse. Les corrélations spatiales d'erreurs d'ébauche sont notamment impactées. La deuxième problématique est l'importance de l'hypothèse Gaussienne dans la dérivation des algorithmes d'assimilation, alors que sa validité ne peut pas être totalement assurée notamment à cause de l'utilisation de modèles de propagation non-linéaires. Le troisième point soulevé est la possibilité d'utiliser des modèles d'erreurs alternatifs tels que les erreurs de déplacement, tout en gardant une approche Bayésienne de type maximum de vraisemblance.

2 | Représentation anisotrope des corrélations d'erreurs d'ébauche

Dans ce chapitre, une méthode originale permettant d'appliquer une certaine dépendance au flux météorologique à la modélisation des corrélations horizontales d'erreurs d'ébauche est présentée. L'étude est effectuée dans le cadre courant où l'erreur d'ébauche est constituée uniquement des erreurs additives d'amplitude dont la PDF est spécifiée par un modèle Gaussien. Les erreurs de déplacement ne sont donc pas prises en compte ici.

Les développements et les résultats présentés ont été publiés dans la revue Tellus A dans un article intitulé « Modelling background error correlations with spatial deformations : a case study » (Legrand & Michel, 2014).

Sommaire

2.1	Contexte de l'étude et résumé	42
2.1.1	Contexte	42
2.1.2	Déformations des corrélations d'erreurs d'ébauche	42
2.1.3	Problématique et méthodologie	43
2.1.4	Résultats	43
2.1.5	Perspectives	44
2.2	Version originale de l'article publié	44

2.1. Contexte de l'étude et résumé

2.1.1. Contexte

Actuellement, le système d'assimilation du modèle ARPEGE utilise en opérationnel des corrélations horizontales d'erreurs d'ébauche basées sur une formulation en ondelettes dont les coefficients sont calculés à partir d'une ensemble d'assimilation du jour (Fisher, 2003; Varella et al., 2011; Berre et al., 2015). Au moments de l'étude toutefois, ces corrélations étaient représentées par une matrice diagonale spectrale, impliquant dans l'espace points de grille des corrélations homogènes et isotropes. De manière à rajouter de la dépendance au flux météorologique dans la représentation de ces corrélations (qui peut apparaitre le long des zones baroclines par exemple), une approche alternative aux ondelettes, basée sur l'utilisation des déformations spatiales, a été testée.

2.1.2. Déformations des corrélations d'erreurs d'ébauche

La méthode présentée ici consiste à ajouter à la matrice de corrélations horizontales des erreurs d'ébauche un opérateur de déformation spatiale. L'objectif est d'estimer un opérateur capable de prendre en compte la dépendance au flux météorologique des corrélations d'erreurs. Notamment, appliquer l'inverse de cette déformation aux distributions d'erreurs d'ébauche doit amener à une meilleur adéquation avec la modélisation « diagonale spectrale » (i.e corrélations homogènes isotropes). Afin de déterminer quel opérateur de déformation il faut utiliser, une première approche consiste à se concentrer sur une caractéristique de la dépendance au flux

météorologique et de formuler *a priori* l'opérateur de déformation à utiliser pour que les corrélations associées suivent cette même caractéristique. Par exemple Rossby (1937) constate que le long des surfaces isentropes (surface d'égale température potentielle), les discontinuités des fronts sont atténuées, ce qui implique une meilleure validité de l'hypothèse d'isotropie des corrélations d'erreurs. Benjamin (1989) propose alors d'appliquer un changement de coordonnées afin d'effectuer l'étape d'assimilation selon les isentropes (« espace inverse ») plutôt que selon les isobares (« espace déformé »). De même, en se basant sur la théorie du semi-géostrophisme, Desroziers (1997) propose d'utiliser comme opérateur de déformation inverse la transformée géostrophique (GT, Hoskins (1975)) qui permet de s'affranchir, dans l'espace inverse, d'une partie des effets agéostrophiques notamment responsables du « pincement » des fronts. Ce type d'approche où la formulation de la déformation est connue *a priori* sera qualifiée « d'approche physique ».

Un deuxième type d'approche, plus général, consiste à calculer la déformation sans faire d'hypothèse *a priori* sur sa forme. Michel (2013a) et Michel (2013b) proposent l'estimation statistique de l'opérateur de déformation spatiale à partir d'une méthode ensembliste. Cette méthode est nommée par la suite transformée statistique (ST). Cette estimation objective repose sur deux points importants. Premièrement, l'hypothèse est faite que les propriétés de non-stationnarité et d'anisotropie des corrélations issues d'un processus F peuvent être correctement représentées par une déformation spatiale \mathbf{d} d'un processus R spatialement stationnaire : $F(\mathbf{x}) = R(\mathbf{d}(\mathbf{x}))$. Le deuxième point est que selon Clerc & Mallat (2002), si cette hypothèse est vérifiée, il est possible de retrouver cette déformation à partir du champ de variance de la dérivée spatiale du processus déformé F .

2.1.3. Problématique et méthodologie

Durant cette thèse, le travail effectué sur la représentation de l'anisotropie des corrélations d'erreurs horizontales d'ébauche a consisté à tester sur un cas réel les améliorations apportées par la ST. La méthodologie utilisée est la suivante. Tout d'abord l'opérateur de déformation spatiale est estimé à partir d'un ensemble d'assimilation. Ensuite, la déformation inverse est appliquée à chacun des membres de l'ensemble. Enfin l'isotropisation et l'homogénéisation induites par la déformation inverse sont quantifiées à partir de diagnostics calculés sur l'ensemble dans l'espace déformé et dans l'espace inverse. Les résultats obtenus avec la ST sont comparés à ceux obtenus avec une autre déformation spatiale déjà existante : la GT.

Pour quantifier les effets de la ST et de la GT, un indice d'anisotropie et une longueur de corrélation sont calculées pour chacun des points de grille du domaine. Ces deux diagnostics sont déduits du tenseur de corrélations local en chaque point de grille (Michel, 2013b; Ménétrier et al., 2014).

Les tests sont effectués sur un ensemble d'assimilation de 90 membres du modèle ARPEGE pour la situation du 7 novembre 2011 à 12 UTC. Cette situation est notamment caractérisée par un cas de fort cyclonisme en méditerranée, proche de la dynamique des cyclones tropicaux (Médicane; Chaboureau et al., 2012), et par un front froid actif dans le proche Atlantique. Le domaine d'étude s'étend de la Corne de l'Afrique, au sud-est, jusqu'à Terre-Neuve, au nord-ouest.

2.1.4. Résultats

Dans la troposphère, la ST et la GT ont toutes les deux un impact positif sur l'anisotropie avec un indice d'anisotropie plus faible dans l'espace inverse que dans l'espace déformé. La

plus grande amélioration est située dans la haute troposphère. D'après notre diagnostic d'anisotropie, l'effet de la ST est systématiquement plus important que celui de la GT. Ceci est vrai même dans la zone frontale où les effets de l'agéostrophisme pourraient favoriser la GT (Semple, 2001). De plus, dans la troposphère libre (au dessus de la couche limite), l'isotropisation moyenne est associée à une homogénéisation dans le sens où l'écart-type sur le domaine de l'indice d'anisotropie est également diminué. Ceci est observé pour la ST et la GT, mais dans une bien moindre mesure pour cette dernière.

Pour les longueurs de corrélation, la ST semble également diminuer leur écart-type pour la partie basse de la troposphère (approximativement jusqu'au niveau pression 700hPa). Néanmoins, pour les niveaux supérieurs de la troposphère, la baisse moyenne de l'isotropie et de ses variations s'accompagnent d'une augmentation de l'hétérogénéité des longueurs de corrélation. En plus d'une faible efficacité à diminuer l'hétérogénéité de l'isotropie, la GT augmente l'hétérogénéité des longueurs de corrélation de façon systématique (à part pour l'humidité spécifique au dessus de 300hPa dans la troposphère et stratosphère, où la quantité d'humidité spécifique est négligeable). Pour les niveaux stratosphériques, l'anisotropie moyenne est diminuée par la ST. Les zones de très fortes anisotropies sont également moins prononcées dans l'espace inverse.

Le dernier résultat de cette étude est la ressemblance entre les déformations estimées dans la troposphère libre pour les quatre variables modèles testées, à savoir la vorticit , la divergence, la temp rature, et l'humidit  sp cifique. Ceci est vrai malgr  le retrait au pr alable de corr lations inter-variables. L'utilisation d'une unique d formation pour toutes les variables est donc envisageable.

2.1.5. Perspectives

Les r sultats de la ST sont encourageants, mais il est clair que toute l'anisotropie ne peut pas  tre mod lis e uniquement par les d formations spatiales d'un mod le isotrope et stationnaire. Une premi re possibilit  est d' tendre la m thode ST   l'estimation de d formations dans les trois dimensions de l'espace. De plus, l'utilisation suppl mentaire d'autres m thodes de mod lisation des corr lations pourrait  galement ajouter une plus-value. Par exemple l'apport de la ST pourrait  tre test  en compl ment de la mod lisation en ondelettes utilis es actuellement en op rationnel. Enfin, les d formations spatiales pourraient  galement  tre utilis es dans un cadre ensembliste. En effet, l' chantillonnage direct des corr lations d'erreurs   partir d'un ensemble d'assimilation semble   l'heure actuelle la m thode donnant les meilleurs r sultats pour la prise en compte de la d pendance au flux m t orologique. Cette technique est cependant sujette   des probl matiques de bruits d' chantillonnage, ce qui n cessite l'application de fonctions de localisations aux corr lations  chantillonn es (Houtekamer & Mitchell, 2001). De telles fonctions, permettant notamment d'att nuer le bruit d' chantillonnage avec la distance, sont couramment homog nes et isotropes. Or comme l'a montr  M n trier et al. (2014), les longueurs caract ristiques (ou longueurs de localisations) d pendent fortement de la variable et du flux m t orologique. Dans ce cadre, les d formations spatiales pourraient  tre utilis es pour former des fonctions de localisations anisotropes et h t rog nes (Michel, 2013a).

2.2. Version originale de l'article publi 

Modelling background error correlations with spatial deformations: a case study

By RAPHAËL LEGRAND* and YANN MICHEL, *Météo-France/CNRS CNRM/GAME UMR 3589, 42 avenue Coriolis, FR-31057 Toulouse, France*

(Manuscript received 3 February 2014; in final form 22 September 2014)

ABSTRACT

A long-term goal in variational data assimilation is to improve the anisotropy of background error correlations. One way to achieve anisotropic correlations is to introduce spatial deformations. This deformation can be specified a priori for instance by using the geostrophic transform (GT) as introduced by Desroziers (1997). The deformation can also be estimated from a purely statistical point of view (Michel, 2013a). The aim of this study is to evaluate the performance of such spatial deformation techniques for the use of background error modelling. A large ensemble of variational assimilations with perturbed observations is set up on a case study with the global ARPEGE model. An anisotropy index and a length scale diagnostic are defined to compare objectively the effectiveness of the deformations. This effectiveness is measured as the ability of the inverse spatial deformations to make the correlations more isotropic or more homogeneous. The results are shown to depend on the vertical level and on the variable. Generally, the statistical deformation is able to reduce the anisotropy while the GT is giving much smaller improvements that are, in this case study, confined to the frontal area of an extratropical cyclone.

Keywords: data assimilation, anisotropy, background error, ensemble, spatial deformation

1. Introduction

In Numerical Weather Prediction (NWP), the analysis step aims to estimate the best initial conditions for the next forecast by assimilating observations. Due to the lower number of observations compared to the model grid size, one usually resolves a full rank problem by using a short-range forecast called the background. The uncertainty of the background is taken into account by the background error covariance matrix \mathbf{B} . As reviewed by Bannister (2008), most operational centres use a control variable transform (CVT). A square root of \mathbf{B} is conveniently modelled as a sequence of transforms: the balance operator accounting for the cross-covariances between variables and the spatial transform accounting for the covariances for every unbalanced control variable. The spatial transform is also usually split in three components: the variances, the vertical and the horizontal correlations. This paper is focusing on horizontal correlations.

The variances of the background errors are the diagonal terms of \mathbf{B} . In the ARPEGE¹ assimilation scheme, they are estimated from a small ensemble of forecasts (currently six members in AEARP²) (Raynaud et al., 2011) and thus they are flow-dependent. The correlations are still modelled in a static way with a diagonal matrix in spectral space. This modelling yields homogeneous and isotropic correlations on the sphere (Courtier et al., 1998). To increase the heterogeneity in time and space, Varella et al. (2011) model the correlations with a diagonal matrix now in wavelet space computed over a 3-week calibration period. It was shown to improve the heterogeneity of the modelled correlations, in particular by modelling larger length scales in tropical areas and smaller length scales in mid-latitude areas. The correlation modelled with the isotropic wavelets (Fisher, 2004) do allow for heterogeneity, but only weak anisotropy.

Alternatively (or in a complementary way), spatial deformations have been proposed to model the heterogeneity and the anisotropy of the correlations. A first class of

*Corresponding author.
email: raphael.legrand@meteo.fr

¹Action de Recherche Petite Échelle Grande Échelle (Pailleux et al., 2000).

²Assimilation d'Ensemble ARPEGE.

methods is attempting to model the flow dependency of the correlation with an *a priori* formulation of the deformation (hereafter *physical* approaches). For example, Benjamin (1989) proposes an isentropic vertical transform. Desroziers (1997) adapts a coordinate change, the geostrophic transform (hereafter GT) inspired from the semi-geostrophic formalism (Hoskins and Bretherton, 1972; Hoskins, 1975).

The estimation of the statistical deformation used in this paper is part of the second class of methods where the deformation is estimated over an ensemble. As in Michel (2013b), and Michel (2013a), the image processing *shape from texture algorithm* (Clerc and Mallat, 2002) is adapted to objectively model heterogeneity of background error correlations from the statistics of an ensemble of forecasts. This algorithm is named hereafter the statistical transform (ST).

Spatial deformations have a moderate computational cost (e.g. one interpolation) and fit well in the classical CVT formalism, such that they are good candidates for background error modelling. However, any correlation function is generally not the deformation of a stationary one, as studied by Perrin and Senoussi (1999). Thus, it is legitimate to check whether the spatial deformation introduces useful anisotropy in the structure functions. A slightly different point of view is employed in this paper: heterogeneity and anisotropy of the correlations are compared between the physical (deformed) space and the computational (inverse) space. The aim of the paper is to document the differences of anisotropies and heterogeneities between direct and inverse spaces on a case study, using objective measures.

The paper is organised as follows: in Section 2, the two methods to estimate the deformation are described. The developments implemented for this study and the setup of the experiments are explained in Section 3. Results of the ST are objectively compared with those of the GT in Section 4. Summary and discussions are given in Section 5.

2. Spatial deformations

2.1. The statistical transform

Clerc and Mallat (2002) introduce a wavelet-based algorithm capable of statistically estimating the deformation of a stationary process from a single realisation (with application to image processing). Michel (2013b) proposes to use this algorithm to model the non-stationary and anisotropy properties of the spatial correlations. This assumes that the background errors (after normalisation by their standard deviations) are well represented by spatial deformations of a spatially stationary process. Michel (2013b) also shows that the algorithm to estimate the deformation can rely on local measures of length scales (Pereira and

Berre, 2006) rather than wavelets, with very similar results. Michel (2013a) applies the algorithm to the modelling of background error correlations for a convective scale model and finds that the deformed correlations have anisotropy properties close to the raw correlations. For the sake of completeness, this section is a summary of the formalism of the ST as introduced in Clerc and Mallat (2002) and Michel (2013a).

Clerc and Mallat (2002) study deformed processes defined as $F(\mathbf{x}) = R(\mathbf{d}(\mathbf{x}))$, where $\mathbf{x} \in \mathbb{R}^n$ and \mathbf{d} is a function from \mathbb{R}^n to \mathbb{R}^n . In the rest of the article, we will work in two dimensions of space ($n=2$). Information about deformed processes is gained from a continuous wavelet analysis, which is computed as the inner product $\langle F, \psi_{\mathbf{u}, \mathbf{S}} \rangle$ where $\mathbf{u} \in \mathbb{R}^2$ is the spatial position and \mathbf{S} a warping matrix that determines the shape and orientation of the wavelet. A key result from Clerc and Mallat (2002) is that such wavelet analyses of F and R can be related through a migration property:

$$\langle F, \psi_{\mathbf{u}, \mathbf{S}} \rangle = \langle R, \psi_{\mathbf{d}(\mathbf{u}), J_d \mathbf{S}} \rangle \quad (1)$$

which is valid only for vanishing warping matrices $\det(\mathbf{S}) \rightarrow 0$ and where J_d is the Jacobian of the deformation d . Because the assumption is made that R is a spatially stationary process, its wavelet analysis does not depend on the position (in a statistical sense). More precisely, the warpogram is defined to be the variance in wavelet space:

$$\omega(\mathbf{u}, \mathbf{S}) = \mathbb{E}\{|\langle F, \psi_{\mathbf{u}, \mathbf{S}} \rangle|^2\} \quad (2)$$

Following Clerc and Mallat (2002), the \mathbf{u} -derivatives of the warpogram of R vanish, while using the migration property the warpogram of F follows the so-called texture gradient equation:

$$\partial_{u_1, u_2} \omega(\mathbf{u}, \mathbf{S}) - \sum_{i,j=1}^2 v_{i,j}^{1,2}(\mathbf{u}, \mathbf{S}) c_{i,j}(\mathbf{u}, \mathbf{S}) = 0 \quad (3)$$

and where the coefficients $c_{i,j}$ are defined by:

$$\begin{pmatrix} c_{11}(\mathbf{u}, \mathbf{S}) & c_{12}(\mathbf{u}, \mathbf{S}) \\ c_{21}(\mathbf{u}, \mathbf{S}) & c_{22}(\mathbf{u}, \mathbf{S}) \end{pmatrix} = \begin{pmatrix} \frac{\partial \omega}{\partial s_{11}} & \frac{\partial \omega}{\partial s_{12}} \\ \frac{\partial \omega}{\partial s_{21}} & \frac{\partial \omega}{\partial s_{22}} \end{pmatrix} \times \mathbf{S}^T \quad (4)$$

The ST algorithm thus follows these steps:

- given some realisations of F , compute their wavelet transforms;
- compute an estimate of the warpogram, the variance in wavelet space;
- compute spatial and scale derivatives of this warpogram (using for instance smoothed finite differences);
- estimate the slopes in wavelet space $v_{i,j}$ from a least-square solution of the texture gradient eq. (3).

The slopes in wavelet space $v_{i,j}$ are directly related to the Jacobian of the deformation:

$$\begin{pmatrix} v_{11}^{1,2}(\mathbf{u}, S) & v_{12}^{1,2}(\mathbf{u}, S) \\ v_{21}^{1,2}(\mathbf{u}, S) & v_{22}^{1,2}(\mathbf{u}, S) \end{pmatrix} \approx J_d^{-1}(\mathbf{u}) \partial_{u_1, u_2} J_d(\mathbf{u}) \quad (5)$$

The term $J_d^{-1}(\mathbf{u}) \partial_{u_1, u_2} J_d(\mathbf{u})$ that is estimated from the wavelet analysis is known as the ‘deformation gradient’. More precisely it is the relative spatial variations of the Jacobian in the two spatial directions, whereas one would like to estimate the deformation itself. The solution was introduced by Michel (2013a), who shows that it is possible to recover the deformation from the deformation gradient by solving two elliptic partial differential equations complemented with appropriate boundary conditions. In this paper, Dirichlet conditions on the perturbations are used. This means that the deformation will not affect the boundaries of the domain. After this step, it is also possible to compute an approximate inverse deformation \mathbf{d}^{-1} to be applied on F to recover the process $R = F(\mathbf{d}^{-1})$.

This method estimates the deformation \mathbf{d} relying on variances computation in wavelet space. The robustness of this estimate obviously depends on the sampling. In Clerc and Mallat (2002), the deformation is estimated from a single realisation of F and the sampling proposed uses several scales and wavelet orientations, as well as the use of spatial averaging. In the context of data assimilation, the structure of background errors is usually estimated over ensembles. In Michel (2013a) and in this paper, the sampling over scales and orientations is enriched by an ensemble of realisations (an ensemble of 90 forecasts) of F . For the same purpose of increasing the robustness of the estimated deformation, a spatial smoothing of the derivatives of the wavelet coefficient variances is applied. This is done with recursive filters (Purser et al., 2003a) that convolve fields with a Gaussian kernel defined by a chosen filter length L . In this study a fourth-order recursive filter (with a single sweep) is used. Values of L are discussed in Section 3.5. All other settings of the ST are the same as in Michel (2013a).

The model assumes that it is reasonable to represent background error correlations by the deformation of a homogeneous isotropic correlation model. Michel (2013b) argues that this is probably better than the usual assumption that background error correlations can be represented by a (quasi-) homogeneous isotropic correlation model (Derber and Bouttier, 1999; Berre, 2000). However, it was not proven whether these assumptions of homogeneity and isotropy hold better in computational (inverse) space. This is studied in Section 4, where the behaviour of the GT proposed by Desroziers (1997) is also described.

2.2. The geostrophic transform

The use of the GT in data assimilation is (loosely) justified from a physical point of view, rather than from a statistical point of view. Hoskins (1975) introduces a set of equations under the geostrophic momentum approximation that are related to the more common quasi-geostrophic equations by a transformation of the horizontal coordinates. This transformation is more capable of describing the formation of fronts than the quasi-geostrophic system, yet the solutions they produce are very similar from a conceptual point of view (e.g. they only differ by a deformation). The formulation of the coordinate change $(x, y, z, t) \rightarrow (X, Y, Z, T)$ to pass in transformed space is given as (Hoskins, 1975):

$$\begin{aligned} X &= x + \frac{1}{f} v_g \\ Y &= y - \frac{1}{f} u_g \\ Z &= z, T = t \end{aligned}$$

with f the Coriolis parameter and (u_g, v_g) are the components of the geostrophic wind. The direction of the coordinate transform is orthogonal to the geostrophic wind. One property of the coordinate change estimated with GT is to increase the horizontal resolution of frontal areas in transformed space (Nordeng, 1998), which induces a stretching of the (x, y) grid-lines [in transformed space compared to the regular grid (X, Y)] along these zones. As an illustration, Fig. 1, shows the superposition of vorticity map and a regular grid bent by the deformation \mathbf{d} estimated with the GT (here without any spatial filtering).

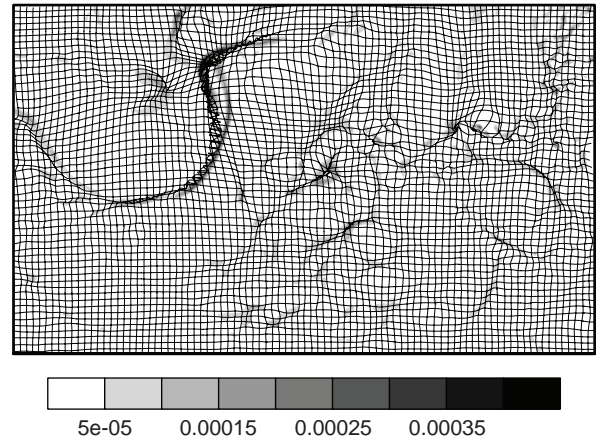


Fig. 1. Relative vorticity (shadings, s^{-1}) in the ARPEGE analysis of the 7th of November 2011 (12UTC) at model level 60 (≈ 900 hPa). Also shown is the Geostrophic Transform, represented here as grid lines bending of a regular grid induced by the GT (solid black lines). The geographical contours are not represented for clarity but the domain is the same as in Fig. 2.

Positive values of vorticity are associated with a compression of the grid (e.g. lower values of the Jacobian determinant of the deformation). The bending is pointing perpendicularly to the frontal axis.

Desroziers (1997) suggests that this transform of the horizontal coordinates could be used in data assimilation to improve the analysis in frontal areas. The coordinate change proposed by Desroziers (1997) is inspired by the semi-geostrophic theory but differs by many aspects:

- the coordinate transform is extended to the sphere by introducing an empirical relaxation of the Coriolis factor near the equator;
- the coordinate transform is applied on model level, which are based on a hybrid sigma-pressure coordinate scheme rather than height;
- an iterative spatial filtering scheme is introduced to filter out some small-scale phenomena in the transform;
- the geostrophic-wind components are approximated by the components of the non-divergent part of the wind.

The transform of the horizontal coordinates proposed by Desroziers (1997) thus differs significantly from the original geostrophic coordinates introduced by (Hoskins, 1975), yet it will still be named the ‘GT’.

The GT has been implemented in the variational assimilation scheme of the Met-Office (Semple, 2001). It was confirmed on several meteorological situations that the GT increases the flow dependency of the structure functions in baroclinic areas with especially a stretching of the correlations along fronts and cyclonic areas. The overall meteorological impact was found to be both small and neutral in terms of forecasts scores. Some cases show an improvement of the precipitation rate during the forecast. But the overall improvements do not appear clearly enough in this context to justify an operational implementation.

The computation of the GT in this paper follows the choices of Desroziers (1997). First, a relaxed Coriolis parameter $f_r(\theta) = (1 - \exp(-\frac{\theta^2}{2\theta_0^2})) \times \Omega \sin(\theta)$ has been used to avoid issues in near-equator areas. θ is the latitude, $\theta_0 = 15^\circ$ and Ω the earth spinning speed. The geostrophic wind is approximated by the non-divergent wind. The transform is applied on model levels. Some horizontal filtering is also introduced (see Section 3.5). As we are working with ensembles, one more choice has to be made whereas the GT is using the mean wind or a different wind for each member of the ensemble. In this paper, the second choice has been chosen to avoid the averaging and smoothing of the transformation in areas of interest (frontal areas for example) that could be due to displacement errors between members of the ensemble.

2.3. Link between the transport and the deformation

In the Kalman Filter equations, the background error covariance matrix is obtained from the propagation in time of the analysis error covariance matrix (plus the model error covariance matrix). The statistical framework suggested here uses a deformation of a stationary model. Therefore, there is an analogy between the deformation and the wind that deforms covariances in geophysical flows. Snyder et al. (2003) show that in a quasi-geostrophic model, the structure of forecast errors in potential vorticity is dominated by the transport (advection of reference-flow PV by the error velocity). In the cases where transport is the dominating process in the tangent-linear model, it may be particularly valid to model to background error covariances as deformation (transport) of the original background error covariance matrix. In reality however, the structure of the background error covariance and correlation also depends on other factors such as the quality and density of the observation network (Bouttier, 1994; Hamill et al., 2002) and generally on the regime of the flow.

The aim of this paper is to expand the study of Michel (2013a) by applying the ST and the GT on a new ARPEGE case study, on a wide domain for all vertical levels and variables to qualify and to quantify the possible improvements in anisotropy and heterogeneity brought by the deformation.

3. Experimental design

3.1. The meteorological situation

The meteorological situation is on the 7th of November 2011, 12UTC. As shown in Fig. 2, the main feature occurring is an active front over the Atlantic. It has a wide meridian extension. A strip of strong low-level vorticity characterises the cold front. The low pressure area associated with this frontal feature is below 970 hPa. The upper level dynamics (300 hPa) is also strong with a jet stream northward along the front with wind speeds around 70 m.s^{-1} . This situation of the 7th of November 2011 is also noticeable for a case of *Medicane* (Chaboureau et al., 2012) over the Mediterranean Sea. The signature of this cyclonic system is visible in the Western Basin of the Mediterranean Sea (40N, 6E) as a vorticity vortex at 900 hPa (Fig. 2).

The transforms have been implemented in a Cartesian system. Thus, this study is using a large size domain extending from North American coast to Middle East. The fields used in this study are bi-dimensional and are obtained from a conformal projection. This means that the notion of anisotropy of the structure functions is the same as on the sphere. The large extension of the domain enables to see the ability of the algorithm to tackle several types of

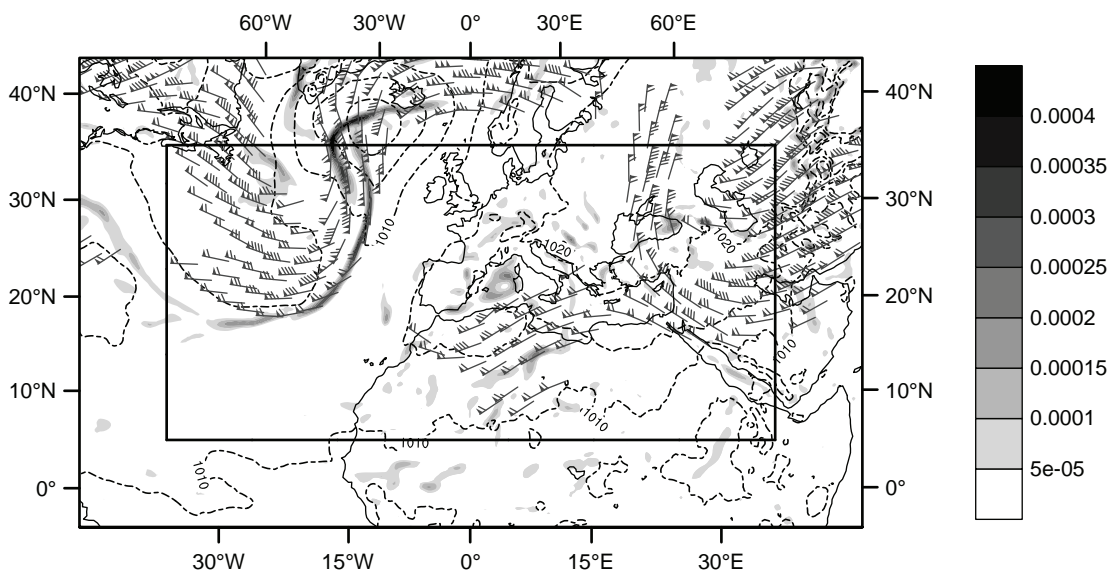


Fig. 2. Map of the meteorological situation of the 7th of November 2011 (12UTC) from the ARPEGE analysis. Vorticity (shadings, 10^{-4}s^{-1}) at model level 60 (≈ 900 hPa), sea surface pressure (dashed lines every 5 hPa) and winds at model level 35 (≈ 300 hPa) (only values over 60 kt, drawn in wind barbs). The large rectangle delimits the computational domain.

features (front, *Medicane*, etc.). Then it allows us to see if the ST is going to locally stationarise areas around each feature or if the stationarisation is going to be globally applied over the whole domain. Moreover, it allows one to see if the stationarisation induced by the ST on a domain much wider than the strongly anisotropic areas (frontal area for example) will undeform the correlation function to a globally lower anisotropy, instead of the high anisotropy of the concerning areas. The extension of the actual computational domain is represented by the large rectangle in Fig. 2.

3.2. The ARPEGE ensemble

Météo-France runs a six-member ensemble data assimilation system, which consists of four dimensional assimilation (4D-VAR) with explicitly perturbed observations, and implicitly perturbed analyses and backgrounds through the cycling (Berre and Desroziers, 2010). This study uses a 90-member ensemble of ARPEGE four-dimensional assimilations and forecasts. The ensemble has been cycled during 12 days from the 30th to the 8th of November 2011. Every day, four 4D-VAR assimilation cycles were done at 00, 06, 12, 18UTC, each time followed by a 6-hour forecast. The spectral resolution of every member is T399, which represents a spatial resolution of approximately 27 km. The model has 70 hybrid sigma-pressure coordinate vertical levels. For various reasons, including the neglect of model error, such ensembles generally underestimate the variance of background errors. In order to counterbalance the underestimation of dispersion of the ensemble, an on-line inflation ($\approx 10\%$) of the ensemble perturbations has been

applied during each cycle, as described by Raynaud et al. (2012) but including the humidity variable. This ensemble was run for the need of the HYMEX program (Ducrocq et al., 2013) that studies high impact weather events in the Mediterranean (like the *Medicane*).

3.3. Methodology

To evaluate the impact of the spatial deformations, the sample correlations of the ensemble in physical space (before the deformation) are compared with the sample correlations of the same ensemble but in transformed space (after the deformation). The improvements given by the deformations are quantified by the decrease of the anisotropy and heterogeneity between physical space and transformed space. The steps of the methodology are:

- Estimation of the geostrophic and statistical deformations from the ensemble in physical space.
- Applying the inverses of the geostrophic and statistical deformations to obtain the undeformed ensembles in transformed space.
- Quantifying and comparing anisotropies of the ensembles in physical and transformed space.

The estimation of the geostrophic and statistical deformations is covered in Sections 2.1 and 2.2. The framework used in this paper stays close to the one proposed by Desroziers (1997), where the deformation and its inverse are computed with a semi-Lagrangian advection scheme. Less than 10 iterations were necessary for convergence of

the semi-Lagrangian advection scheme. The method to measure anisotropies is described next.

3.4. Diagnostics of anisotropy and heterogeneity

Bouttier (1993) uses the inertia matrix \mathbf{M} of the correlation function to compute the main anisotropy vector. Pereira and Berre (2006) introduce a low-cost formula to compute the inertia matrix from the error variances, the derivatives of error variances, and correlation between derivatives of the errors along each spatial direction. Michel (2013b) [eq. (14) for the one-dimensional case] and Ménétrier et al. (2014) [eq. (20) for the multi-dimensional case] propose an updated formulation of the inertia matrix that guarantees its positive-definiteness. In this study, the latter formulation is chosen, with \mathbf{M} given by

$$\mathbf{M} = \begin{pmatrix} \text{Var}\left(\frac{\partial\eta}{\partial x_1}\right) & \text{Cov}\left(\frac{\partial\eta}{\partial x_1}, \frac{\partial\eta}{\partial x_2}\right) \\ \text{Cov}\left(\frac{\partial\eta}{\partial x_2}, \frac{\partial\eta}{\partial x_1}\right) & \text{Var}\left(\frac{\partial\eta}{\partial x_2}\right) \end{pmatrix} \quad (6)$$

where Var and Cov denote the variance and the covariance. $\eta = \frac{F}{\sigma(F)}$ is the error normalised by its standard deviation. x_1 and x_2 are the spatial coordinates of the projection plane. Partial derivatives are evaluated with a centred finite differences method.

The anisotropy index O is defined by

$$O = 1 - \frac{l_2}{l_1} = 1 - \sqrt{\frac{\lambda_1}{\lambda_2}}$$

where λ_1 and λ_2 are respectively the smallest and the largest eigenvalues of the inertia matrix \mathbf{M} . Their inverse square-roots $l_{i=1,2} = \frac{1}{\sqrt{\lambda_{i=1,2}}}$ are the correlation length scales in the direction of the eigenvectors. Since $\lambda_1 < \lambda_2$, $l_1 > l_2$ and $O \in [0,1]$. The higher is O , the more anisotropic are the correlations. This diagnostic is also dealing with heterogeneity of the correlation. Indeed correlations are spatially stationary if l_1 and l_2 are constant over the whole domain.

There is no need to explicitly diagonalise the inertia matrices at each grid point to compute the anisotropy index O . The determinant of \mathbf{M} equals the eigenvalues product $P = \det(\mathbf{M}) = \lambda_1 \times \lambda_2$, and the trace of \mathbf{M} equals the sum $S = \text{tr}(\mathbf{M}) = \lambda_1 + \lambda_2$. Then the eigenvalues are computed as the solution of the second order polynomial $X^2 - SX + P = 0$. The eigenvalues are expressed by

$$\lambda_1 = \frac{S - \sqrt{S^2 - 4P}}{2}$$

$$\lambda_2 = \frac{S + \sqrt{S^2 - 4P}}{2}$$

This allows fast computation of the anisotropy index.

The length scale and anisotropy diagnostics introduced by Pereira and Berre (2006), Michel (2013b), and Ménétrier et al. (2014) are local, e.g. they are related to the local behaviour of the correlation functions. Figure 3 shows a simplified representation of correlation function with elliptic shapes. Ellipses have their major axes in the direction of the main eigenvectors (associated with the smallest eigenvalues λ_1). Ratios between minor and major axe lengths equal $\frac{l_2}{l_1}$. The simplified visualisation of the correlation function with ellipses is compared with raw correlation functions at 3×4 observation points. Anisotropy, length scales and orientation of each ellipse shows good similarities with the raw correlation function. This indicates that for the global model ARPEGE, these low-cost diagnostics are actually rather good representation of the correlation functions, and that they can be used for our purpose of comparing the effect of deformations on the correlation structure. In addition, a total length scale can be computed. Similarly to Ménétrier et al. (2014), it is defined here as the geometric mean $L_t = \sqrt{l_1 \times l_2}$.

3.5. Invertibility of the deformation

The framework requires the estimated deformation to be invertible. This is not ensured by construction, e.g. both the ST and the GT can be singular. The choice made in this study is to enforce invertibility through spatial smoothing.

First estimations of the ST showed that the deformations are non-invertible only at a few grid points. This particularly affects the specific humidity error fields. To avoid this issue, the smoothing step [eq. (22) of Clerc and Mallat (2002)] initially used to increase the robustness of the statistical estimation is reinforced. Indeed large filter lengths L are chosen to smooth enough fields to avoid change of sign of $\det(J_d)$. Too large values of the filter lengths would lose signal in the estimated deformation, such that a compromise should be done. Michel (2013a) averages the spatial derivatives in wavelet space with a length of $L = 8$ grid points. In our case this was found to be insufficient to ensure invertibility.

Figure 4 shows $\det(J_d)$ values and the deformed grid associated to \mathbf{d}^{-1} for three values of filter length $L = 8$, $L = 12$, and $L = 20$ grid points. With a similar set up of $L = 8$ grid points as Michel (2013a), non-invertible areas remain. For $L = 20$ the loss of resolution may be too high. A filter length $L = 12$ grid points for every variable and every level is the lower value that avoids the null or negative values of $\det(J_d)$ in the tropospheric levels. For the rest of this paper $L = 12$ grid points will be chosen as the filter length for all four variables and all levels.

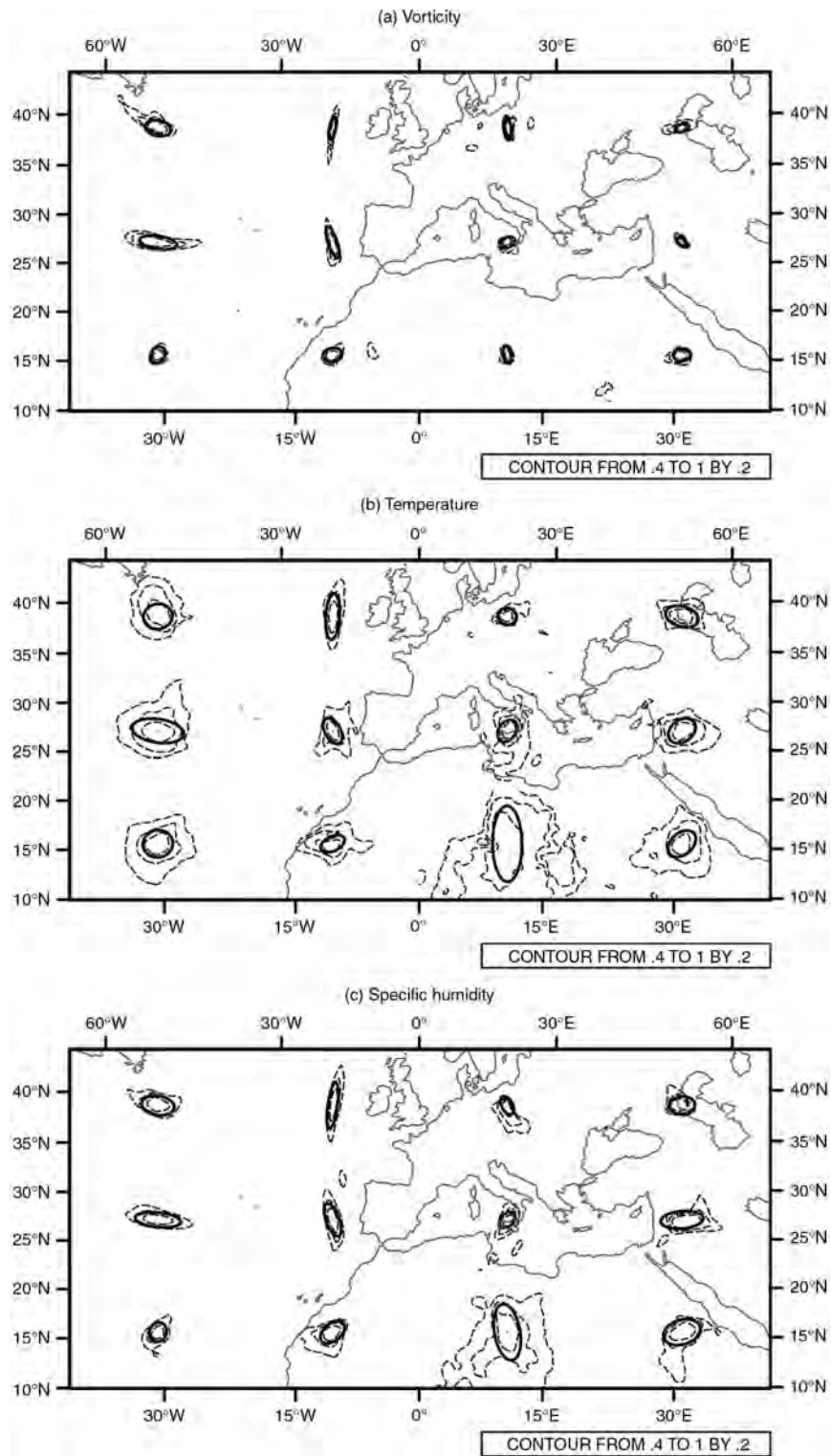


Fig. 3. Superposition of the correlation function (contoured at 0.4, 0.6, and 0.8 in dashed thin lines) with the ellipses determined from the inertia matrix (solid bold lines). Correlations are computed at level 60 (≈ 900 hPa) for (a) vorticity, (b) temperature, and (c) specific humidity. For divergence the superposition is very similar to the vorticity one but with smaller correlation lengths (not shown).

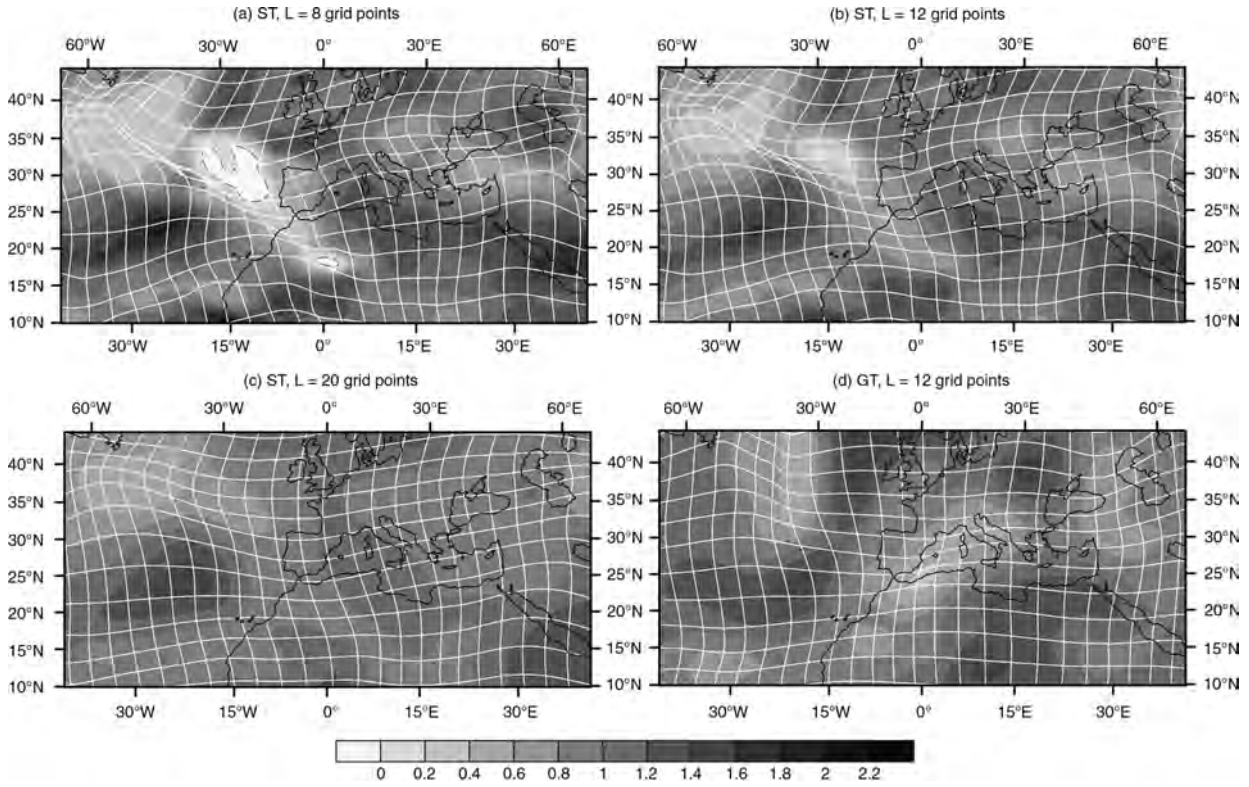


Fig. 4. Grid-lines bending of a regular grid induced by the deformation \mathbf{d} . The resulting deformed grid represents in transformed space the contour of the physical space coordinate. Deformation presented in (a), (b), and (c) are objectively estimated with the ST. Different values of L are presented: (a) $L = 8$ grid points, (b) $L = 12$ grid points, and (c) $L = 20$ grid points. Backgrounds of each panel represent values of $\det(J_d)$ the Jacobian determinant of the deformation at each grid point. Changing sign areas (noted with dashed thin lines) for $\det(J_d)$ induced a crossing of deformed grid-lines causing non-invertibility. Those results are for the specific humidity at level 35 (≈ 300 hPa). (d) The corresponding grid-lines bending and the Jacobian determinant of the deformation estimated with the GT, with $L = 12$ grid points.

In stratospheric levels, deformations have been found to remain non invertible near the domain boundaries despite this tuning of the filter length L . This was traced back to the way the wavelet coefficients are computed using fast Fourier transforms. Improvement was achieved with a better biperiodisation approach. The chosen solution was to periodise error fields with a periodic cubic splines and then to zero out the wavelet coefficients that are located in the cone of influence (Mallat, 2008) of the boundaries.

The deformations estimated with the GT are also non-invertible (e.g. crossing lines in Fig. 1 over the North part of the cold front). To ensure its invertibility, a similar smoothing step has been applied. The same filter length $L = 12$ grid points is used for all four variables and vertical levels. This is made to keep the same setup for ST and GT.

As mentioned in the introduction, this study is not dealing with the balance relationships. The ARPEGE data assimilation scheme uses the formulation of Derber and Bouttier (1999). All variables are approximately uncorrelated by using the balance operator. This debalancing

procedure removes the balanced part of the temperature and of the divergence. Vorticity and specific humidity are left unaltered.

4. Results

From the 90 member ensemble, the anisotropy index O and the total length scales L_t are computed for every level and every (unbalanced) variable, both in physical (deformed) and computational (inverse) spaces.

4.1. Anisotropy

Background error correlations are more isotropic when the anisotropy index O is lower. In order to present a synthetic view of the results, the anisotropy index has been averaged over the horizontal. The vertical profiles are presented in Fig. 5. On average, the error correlations show pronounced anisotropy in the lower stratosphere (between 100 and 200 hPa). They are more isotropic at the top of the

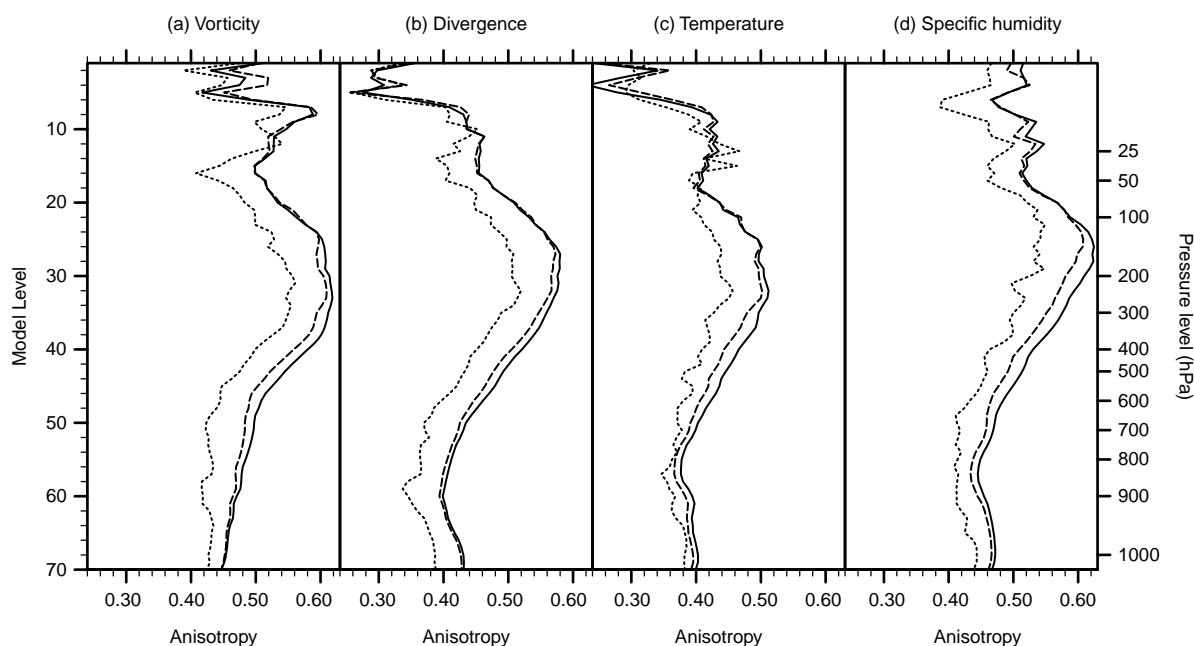


Fig. 5. Vertical profiles of the anisotropy O , for (a) vorticity, (b) divergence, (c) temperature, and (d) specific humidity, for the ensemble in physical space (solid lines) and in computational space after applying the inverse of the deformation estimated with the ST (dotted lines) and with the GT (dashed lines). O is averaged over the whole domain (except near boundary points) on the 7th of November 2011 at 12UTC.

boundary layer (between 850 and 900 hPa) than at the surface for all variables but vorticity. The surface and in particular the orography is indeed likely to have an influence on the anisotropy of the background errors.

Figure 5 shows that on average both the GT and the ST have a positive impact at least in the troposphere (up to 150 hPa), in the sense that the mean anisotropy is lower in inverse space than in physical (deformed) space. The strongest improvement for both algorithms is located in the higher troposphere (around level 35, approximately 300 hPa). According to the anisotropy vertical profile of the raw ensemble, this layer is also the place where the anisotropy itself is largest. The decrease of the mean anisotropy is systematically higher for the ST than for the GT (apart from model level 15 for temperature). For temperature, using the ST, the anisotropy O at level 35 is decreasing from 0.5 to 0.42 (equivalent to a 14% reduction of length of the major axis of an ellipse), meanwhile with the GT O is only decreasing to 0.46 (equivalent to a 7% reduction of the length of the major axis of an ellipse).

The efficiency of the GT in diminishing the mean anisotropy is dependent on the vertical level. Close to the surface and in the stratosphere above level 25 (= 100 hPa) the GT has a neutral impact. In the troposphere between level 55 (≈ 800 hPa) and level 30 (≈ 200 hPa), it clearly has a greater efficiency. Maybe it has a connection with the fact that the QG assumptions ($R_O \ll 1$, with R_O the Rossby

number) may be more valid at those levels than at the surface or the tropopause. In agreement with Semple (2001), this study highlights that the use of the GT in modelling background error correlations may be physically appealing in frontal areas of extratropical cyclones. As measured by our metric, it is however less efficient than the ST.

According to the Fig. 5, the ST and GT (in smaller manner) are on average decreasing the anisotropy over the domain. But this result does not give information on the local behaviour. In Fig. 6, anisotropy maps of the raw ensemble are presented and the undeformed ensemble in transformed space after the ST and the GT. The average decrease with GT and ST is confirmed especially around the frontal area. While the cold front structure is still visible after the GT, for the ST, it shows a clear decrease of O on the bottom left part of the domain with an almost complete isotropisation of the south part of the cold front. Elsewhere in the domain, the anisotropy is also decreasing but in a smaller way.

4.2. Heterogeneity

If the correlations were homogeneous over the domain, the standard deviations of the oblateness index and of the total length scale would be reduced to zero. Figure 7 presents the standard deviation of the oblateness index O over the domain. For levels 35–65 (from ≈ 250 hPa to the boundary

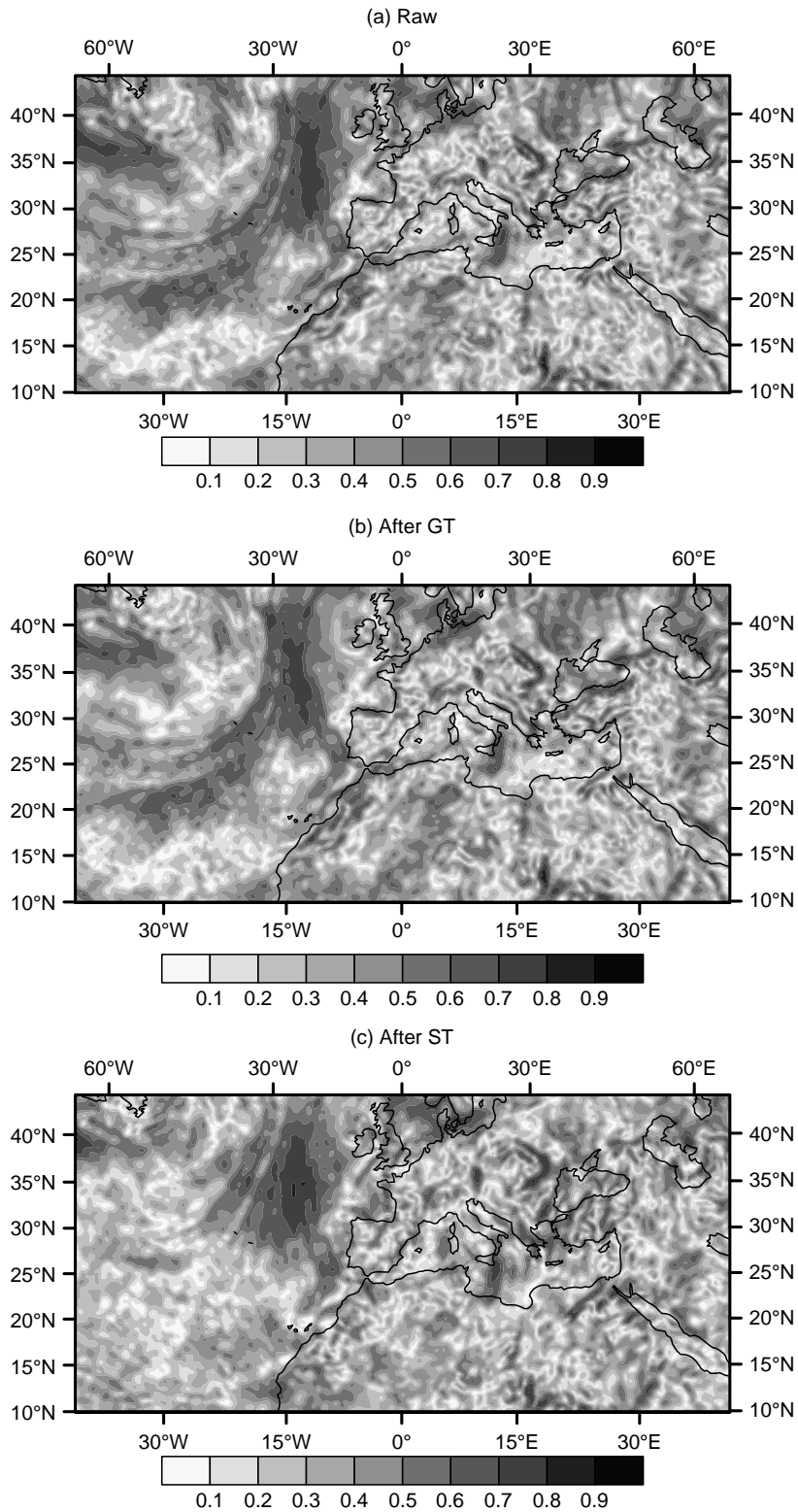


Fig. 6. Maps of anisotropy for (a) the raw ensemble, (b) the ensemble undeformed with the GT, and (c) the ensemble undeformed with the ST. The variable is temperature on the 7th of November 2011 at 12UTC at level 60 (≈ 900 hPa). Near boundaries points are also avoided. One notices that for the raw ensemble the background coast lines are exact. But for the two other panels, since the ensemble has been undeformed, they are not corresponding to the foreground and are just drawn to simplify the recognition of the features with the top panel.

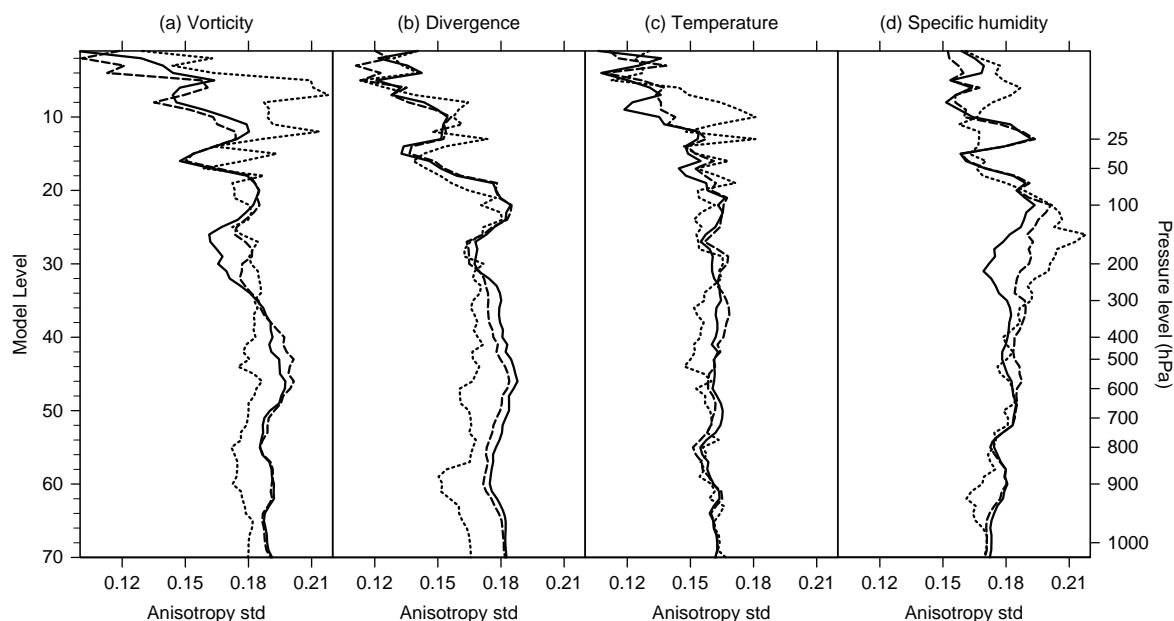


Fig. 7. Vertical profiles of the standard deviation of anisotropy O , for (a) vorticity, (b) divergence, (c) temperature, and (d) specific humidity, for the ensemble in physical space (solid lines) and in transformed space after applying the inverse of the deformation estimated with the ST (dotted lines), or after the GT (dashed lines). Results are averaged over the whole domain (except near boundary points) on the 7th of November 2011 at 12UTC for ARPEGE model.

layer), the standard deviation is globally diminished after the ST and also after the GT, yet in a smaller manner.

Figure 8 represents the standard deviation of the total length scale for each vertical level and variable. As regards the ST, the undeformed ensemble has lower total length scale standard deviations as well as the anisotropy standard deviations (Fig. 7) from the surface to the mid-troposphere. In this case this diminution is the highest for temperature at the surface, with a decrease of the standard deviation from 72 to 54 km. This is illustrated in Fig. 9 with the simplified representation with ellipses of the correlations in physical space and in transformed space after the ST. Large total length scales of both the South-East and the West of the domain are clearly diminished by the ST, giving overall less variation of the length scales over the domain. Despite an increase of the ellipses size along the frontal area, the ellipses size are more homogeneous over the domain for the ensemble in computational (inverse) space.

At higher levels (above 600 hPa), the standard deviation of the total length scale is increased by the ST for every variable (Fig. 8); therefore, the improvement brought by the ST on homogeneity is limited to lower tropospheric levels (under 700 hPa). Elsewhere, the decrease of anisotropy and its variations are accomplished at the detriment of the heterogeneity of the total length scale.

The GT does not really affect the standard deviation of the anisotropy (Fig. 7). However, the standard deviation of the total length scale increases at all vertical levels for all

variables but humidity (Fig. 8). This shows that the GT does introduce some spurious inhomogeneities in the modelled correlations in a systematic way (e.g. for all variables and vertical levels).

4.3. Diagnostics in the stratosphere

In stratospheric levels (levels ≤ 25), the decrease of standard deviation of the anisotropy index is no longer systematic with the ST, with a large increase, for example at level 25 (≈ 150 hPa) for the specific humidity variable. Moreover, the total length scale standard deviation is increasing for most levels and variables (not shown). A different way of looking at these results is proposed in Fig. 10. The mean anisotropy is clearly decreasing, but the standard deviation does not. Yet, the distribution of anisotropies over the domain appears better in computational space than in physical space, with less values of strong anisotropies.

It is clear that applying the inverse deformation is not removing all the anisotropy or the heterogeneity. Perhaps using the deformation approach together with another heterogeneous modelling approach could be useful [for instance the wavelet approach from Deckmyn and Berre (2005) or the inhomogeneous recursive filters from Purser et al. (2003b)]. Progress could be expected also by going to a fully three-dimensional computation of the statistical deformation.

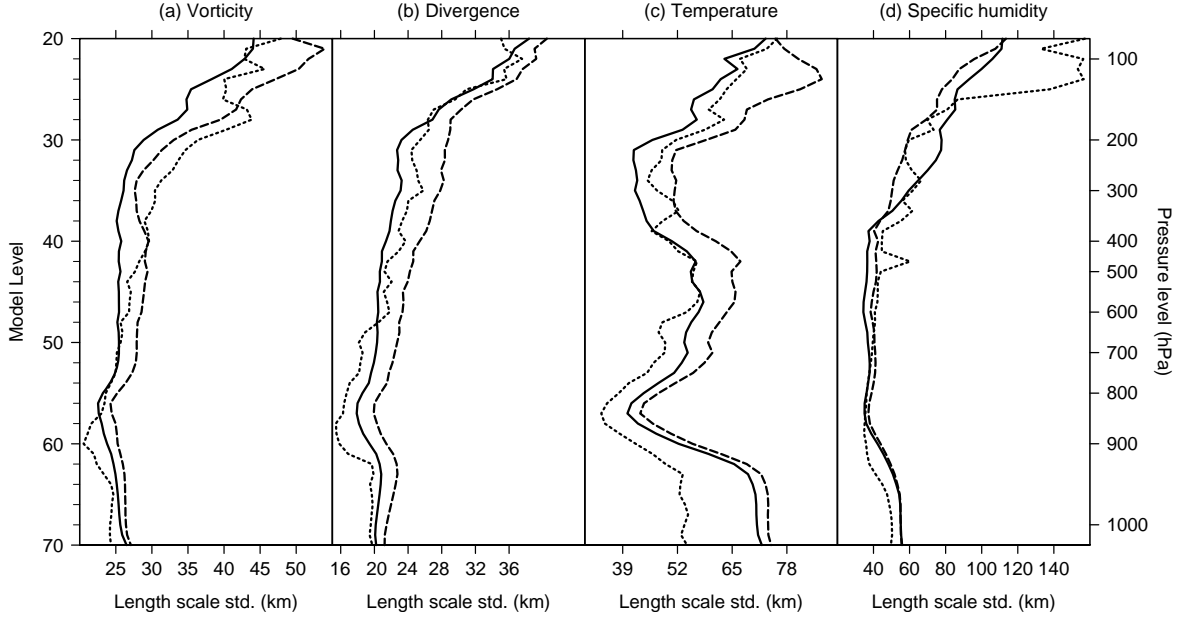


Fig. 8. Vertical profile of standard deviation of total length scale L_t , for (a) vorticity, (b) divergence, (c) temperature, and (d) specific humidity, for the ensemble in physical space (solid line) and in transformed space after applying the inverse of the deformation estimated with the ST (dotted line), or after the GT (dashed line). Near boundary points were also neglected. Stratospheric model levels from 1 to 19 are not represented.

As regards the GT, for every variable and every level, the standard deviation of the total length scale is increasing except for a few levels for the specific humidity. This makes this transform probably less useful than the ST for the modelling of background error correlations.

4.4. Comparing the deformations between variables

The deformation acts on unbalanced variables to represent the anisotropies and heterogeneities of their horizontal cor-

relations. Another possibility is to use the same deformation for all variables. This may allow commuting the spatial deformation with the balance transform. We have found close similarities between the statistical deformations estimated for the different variables. In Fig. 11, the Jacobian of the deformation for all four variables studied is represented. The features modelled by the deformation are located at the same places and almost have the same shape. Only the amplitude is variable. Despite using unbalanced (i.e. mostly uncorrelated) variables, the algorithm is detecting the same

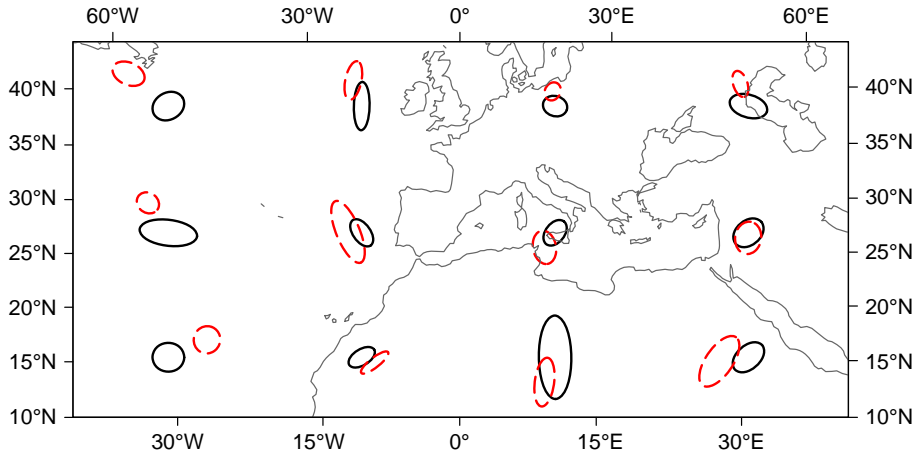


Fig. 9. Superposition of simplified elliptic visualisation of the correlation of the raw ensemble (solid lines) and of the ensemble in transformed space after ST (dashed lines). Correlation tensors are computed at level 60 (≈ 900 hPa) for temperature. For each position (4×5), ellipses for the raw ensemble and for the ensemble in transformed space are not superposed since error fields have been displaced by \mathbf{d}^{-1} .

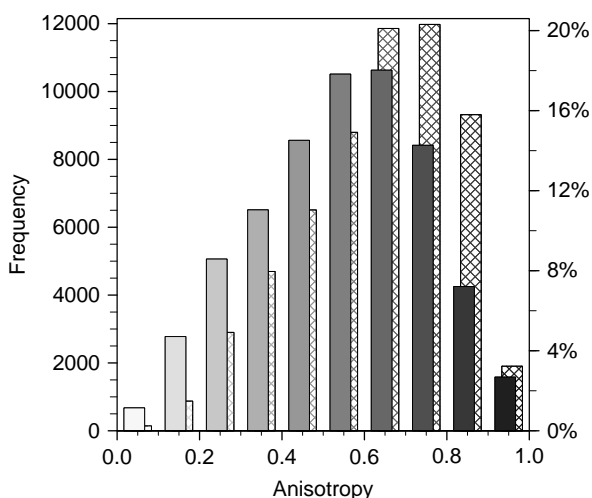


Fig. 10. Distribution histogram of anisotropy O over the domain for the raw ensemble (hatched bars, background) and the ensemble undeformed with the ST (filled bars, foreground), for specific humidity, on the 7th of November 2011 at 12UTC for ARPEGE model at the stratospheric level 25 (≈ 150 hPa).

convergent and divergent areas. This result would support the fact that the anisotropy of structure functions are similar enough such that the same deformation could be used for all variables. This conclusion has however to be mitigated for high stratospheric levels and surface level

where the deformation varies more between each unbalanced variable (not shown).

5. Conclusions

This study compares two deformations (coordinate changes) proposed for the modelling of the background error correlations in meteorological data assimilation. This comparison uses a large ensemble of data assimilations to make this comparison on a case study. The accuracy of the deformation and so the quality of the correlation function modelling is quantified by calculating the decrease of the anisotropy and heterogeneity of correlations between the raw ensemble of forecasts with its undeformed representation in transformed space.

The ST is shown to systematically decrease the averaged anisotropy for every level and every control variable. The biggest improvements of the algorithm are located in the higher troposphere. In contrast, the improvements brought by the GT are lower and limited to the frontal area of an extra-tropical cyclone. Overall, both algorithms are diminishing the heterogeneity of the anisotropy. The ST is decreasing the heterogeneity of the total length scale in the lower half of the troposphere, but increasing it at higher levels. It may be that there is a ‘competition’ into getting more isotropic correlation functions versus more

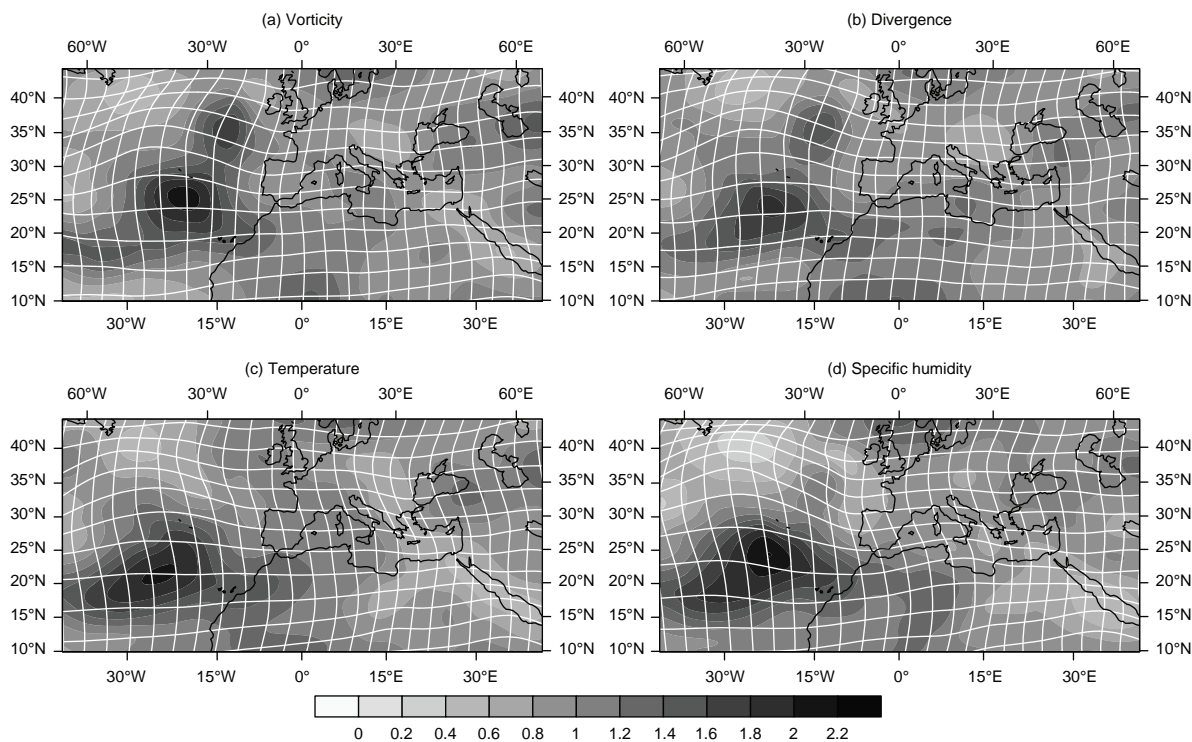


Fig. 11. Jacobian determinant of the deformation (grey shades) and the regular grid deformed by the deformation \mathbf{d} (white lines, as Fig. 4) estimated by ST for (a) vorticity, (b) divergence, (c) temperature, and (d) specific humidity, at level 60 (≈ 900 hPa).

homogeneous ones. The GT is increasing the heterogeneity for every level and variable.

It is possible that the deformed, flow-dependent horizontal correlation may interact with the multivariate aspects of the parameter transform. In spite of the fact that the variables were previously unbalanced by the parameter transform, the statistical deformations estimated for all variables are very similar. Only the amplitude of the deformation varies between variables. This would support the idea of estimating a single deformation for all variables, and possibly to try to commute the application of the deformation operator with the multivariate transform. This would mean that the total and balanced increment would then be stretched by the grid change. This might lower this interaction effect but clearly this has to be studied in much more detail.

In this study, the statistical deformation is designed to model the anisotropy and heterogeneity of correlations over the horizontal domain. The extension to three dimensions is possible when original correlations are also stationary over the vertical or at least separable. This is not the case of the background error model used here, which makes the problem more complex. However, we expect that the method could be applied to localisation functions rather than correlations. Most localisation schemes are indeed using separable formulations. The spatial deformations estimated from the ST could therefore be used in ensemble assimilation approaches for building heterogeneous and anisotropic localisations, as illustrated in Michel (2013a). This will be the topic of future studies.

6. Acknowledgements

This study benefited from the support of the MISTRAL-HYMEX research program and from the RTRA STAE foundation within the framework of the FILAOS project. The authors thank Benjamin Ménétrier for his scientific advice. The careful readings of Thibaut Montmerle, Tom Auligné and Philippe Arbogast also proved very useful to improve the manuscript.

References

- Bannister, R. N. 2008. A review of forecast error covariance statistics in atmospheric variational data assimilation. II: modelling the forecast error covariance statistics. *Q. J. Roy. Meteorol. Soc.* **134**(637), 1971–1996.
- Benjamin, S. G. 1989. An isentropic meso α -scale analysis system and its sensitivity to aircraft and surface observations. *Mon. Weather Rev.* **117**(7), 1586–1603.
- Berre, L. 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Mon. Weather Rev.* **128**, 644–667.
- Berre, L. and Desroziers, G. 2010. Filtering of background error variances and correlations by local spatial averaging: a review. *Mon. Weather Rev.* **138**(10), 3693–3720.
- Bouttier, F. 1993. The dynamics of error covariances in a barotropic model. *Tellus A.* **45**(5), 408–423.
- Bouttier, F. 1994. A dynamical estimation of forecast error covariances in an assimilation system. *Mon. Weather Rev.* **122**, 2376–2390.
- Chaboureaud, J.-P., Nuissier, O. and Claud, C. 2012. Verification of ensemble forecasts of Mediterranean high-impact weather events against satellite observations. *Nat. Hazards Earth Syst. Sci.* **12**(8), 2449–2462.
- Clerc, M. and Mallat, S. 2002. The texture gradient equation for recovering shape from texture. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(4), 536–549.
- Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M. and co-authors. 1998. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation. *Q. J. Roy. Meteorol. Soc.* **124**(550), 1783–1807.
- Deckmyn, A. and Berre, L. 2005. A wavelet approach to representing background error covariances in a limited-area model. *Mon. Weather Rev.* **133**, 1279–1294.
- Derber, J. and Bouttier, F. 1999. A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A.* **51**(2), 195–221.
- Desroziers, G. 1997. A coordinate change for data assimilation in spherical geometry of frontal structures. *Mon. Weather Rev.* **125**(11), 3030–3038.
- Ducrocq, V., Belamari, S., Boudevillain, B., Bousquet, O., Cocquerez, P. and co-authors. 2013. HyMeX, les campagnes de mesures: focus sur les événements extrêmes en Méditerranée. *La Météorologie.* **80**, 37–47.
- Fisher, M. 2004. Generalized frames on the sphere, with application to the background error covariance modelling. In: *Proceedings of ECMWF Seminar on Developments in Numerical Methods for Atmospheric and Ocean Modelling*, Reading, UK, pp. 87–101.
- Hamill, T. M., Snyder, C. and Morss, R. E. 2002. Analysis-error statistics of a quasi-geostrophic model using three-dimensional variational assimilation. *Mon. Weather Rev.* **130**(11), 2777–2790.
- Hoskins, B. J. 1975. The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.* **32**(2), 233–242.
- Hoskins, B. J. and Bretherton, F. P. 1972. Atmospheric frontogenesis models: mathematical formulation and solution. *J. Atmos. Sci.* **29**(1), 11–37.
- Mallat, S. 2008. *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, Elsevier, Burlington, MA.
- Ménétrier, B., Montmerle, T., Berre, L. and Michel, Y. 2014. Estimation and diagnosis of heterogeneous flow-dependent background error covariances at convective scale using either large or small ensembles. *Q. J. Roy. Meteorol. Soc.* **140**, 2050–2061.
- Michel, Y. 2013a. Estimating deformations of random processes for correlation modelling in a limited area model. *Q. J. Roy. Meteorol. Soc.* **139**, 534–547.

- Michel, Y. 2013b. Estimating deformations of random processes for correlation modelling: methodology and the one-dimensional case. *Q. J. Roy. Meteorol. Soc.* **139**, 771–783.
- Nordeng, T. E. 1998. *A Simple and Efficient Method to Obtain Flow Dependent Structure Functions for Objective Analysis of Weather Elements*. HIRLAM4, Technical Report, no. 36, August 1998, c/o Met Éireann, Glasnevin Hill, Dublin 9, Ireland.
- Pailleux, J., Geleyn, J.-F. and Legrand, E. 2000. La prévision numérique du temps avec les modèles ARPÈGE et ALADIN-Bilan et perspectives. *La Météorologie*. **30**, 32–60.
- Pereira, M. B. and Berre, L. 2006. The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Weather Rev.* **134**(9), 2466–2489.
- Perrin, O. and Senoussi, R. 1999. Reducing non-stationary stochastic processes to stationarity by a time deformation. *Stat. Probab. Lett.* **43**, 393–397.
- Purser, R. J., Wu, W.-S., Parrish, D. F. and Roberts, N. M. 2003a. Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: spatially homogeneous and isotropic Gaussian covariances. *Mon. Weather Rev.* **131**(8), 1524–1535.
- Purser, R. J., Wu, W.-S., Parrish, D. F. and Roberts, N. M. 2003b. Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: spatially inhomogeneous and anisotropic general covariances. *Mon. Weather Rev.* **131**(8), 1536–1548.
- Raynaud, L., Berre, L. and Desroziers, G. 2011. An extended specification of flow-dependent background error variances in the Météo-France global 4D-Var system. *Q. J. Roy. Meteorol. Soc.* **137**(656), 607–619.
- Raynaud, L., Berre, L. and Desroziers, G. 2012. Accounting for model error in the Météo-France ensemble data assimilation system. *Q. J. Roy. Meteorol. Soc.* **138**(662), 249–262.
- Semple, A. 2001. *A Meteorological Assessment of the Geostrophic Co-ordinate Transform and Error Breeding System When used in 3D Variational Data Assimilation*. NWP Tech Rep, UK, 357.
- Snyder, C., Hamill, T. M. and Trier, S. B. 2003. Linear evolution of error covariances in a quasigeostrophic model. *Mon. Weather Rev.* **131**(1), 189–205.
- Varella, H., Berre, L. and Desroziers, G. 2011. Diagnostic and impact studies of a wavelet formulation of background-error correlations in a global model. *Q. J. Roy. Meteorol. Soc.* **137**(658), 1369–1379.

3 | Diagnostics de non-Gaussianité des erreurs de prévisions et d'analyses aux échelles convectives

*Les développements et les résultats présentés dans ce chapitre ont été soumis au journal *Nonlinear Processes in Geophysics* dans un article intitulé « *Diagnosing non-Gaussianity of forecast and analysis errors in a convective scale model* » (Legrand et al., 2015).*

Sommaire

3.1	Contexte de l'étude et résumé de l'article	61
3.1.1	Contexte	61
3.1.2	Méthodologie	61
3.1.3	Résultats	62
3.1.4	Perspectives	62
3.2	Version originale de l'article publié	63

3.1. Contexte de l'étude et résumé de l'article

3.1.1. Contexte

Le but de cette deuxième partie du travail de thèse, est de proposer une documentation de la non-Gaussianité (NG) des distributions d'erreurs d'ébauche présentes dans le système d'assimilation d'AROME. Ce dernier étant basé sur un système variationnel reposant sur un formalisme où les erreurs sont supposées additives, non biaisées, et Gaussiennes (cf section 1.2), une telle étude permet de diagnostiquer les variables ou les situations météorologiques pour lesquelles ces hypothèses sont mises à mal.

Cette étude se limite à la description des résultats de diagnostics de NG et à leur sensibilité à la présence de nuages. Certains travaux tels que Auligné et al. (2011) et Michel et al. (2011) ont en effet montré que les processus non linéaires associés aux phénomènes diabatiques étaient une source importante de NG.

3.1.2. Méthodologie

Les distributions d'erreurs d'ébauche sont échantillonnées pour chaque variable du modèle, à partir d'un grand ensemble d'assimilations (cf section 1.4.4). La NG des distributions est alors indexée par la valeur du test univarié de D'Agostino (D'agostino et al., 1990). Une caractéristique de cet indice est que sa valeur ne dépend que de deux paramètres de forme de la distribution testée : son asymétrie et son aplatissement. Pour les systèmes de grande taille telle que la PNT, seule une caractérisation incomplète des PDFs d'erreurs est accessible avec des tests univariés (Bocquet et al., 2010), c'est à dire qu'aucun lien entre les points de grille alentours, entre plusieurs échéances, ou entre différentes variables n'est testé. Comme le remarque Tarantola (2005), l'échantillonnage plus complet d'une PDF de grande dimension nécessite des techniques d'explorations spécifiques afin de couvrir de façon assez dense les zones où les

probabilités ne sont pas négligeables.

L'échantillonnage est effectué à partir d'un ensemble d'assimilations de 90 membres du modèle AROME-France. Les caractéristiques de cet ensemble ont été décrites au chapitre 1. La situation météorologique étudiée est celle du 4 novembre 2011. Elle est caractérisée par la présence d'un épisode Cévenol touchant le sud-est de la France et la vallée du Rhône, ainsi que d'un front froid abordant les côtes bretonnes.

3.1.3. Résultats

Cette étude retrouve qu'en moyenne, pour la plupart des niveaux verticaux, c'est l'humidité spécifique (q) qui s'éloigne le plus du modèle Gaussien, de manière similaire à ce qui a déjà été noté à l'échelle globale (Holm et al., 2002; Dee & Da Silva, 2003). Au dessus de la couche limite, pour la température (T) et le vent (zonal, U , et méridien, V), les distributions d'erreurs sont proches de la Gaussianité. Dessous, par contre, la NG de ces dernières variables augmente en se rapprochant de la surface. Selon nos diagnostics, c'est T qui est la variable modèle la moins Gaussienne pour les premiers niveaux de surface. De plus, à partir de l'étude de carte de NG, de grandes similarités sont observées entre les structures de fortes NG et certaines structures météorologiques actives. C'est notamment le cas pour l'épisode Cévenol et pour le front froid qui sont associés à des processus nuageux et précipitant.

Durant l'étape de propagation de l'état de l'atmosphère par le modèle AROME, la NG augmente. L'augmentation est très rapide dans les six premières heures, puis se ralentit. Grâce à l'utilisation de diagnostics hétérogènes, un comportement différent est observé pour la NG des zones nuageuses par rapport à celle des zones de ciel clair. Les couches nuageuses sont associées à une NG élevée pour q et T . Ceci est particulièrement vrai à la base et au sommet du nuage, où les erreurs de déplacement vertical et les processus diabatiques sont présents. Pour les zones de ciel clair, la NG de T augmente très vite dans les basses couches. Ceci peut être lié aux processus turbulents et radiatifs très présents dans la couche limite.

Durant le cycle d'assimilation, après une augmentation rapide durant l'étape de propagation, la NG diminue largement durant l'étape d'assimilation. Cette diminution est d'autant plus grande que la zone est fortement contrainte par les observations. Contrairement aux variables modèles U et V , la vorticité ζ et la partie non-balancée de la divergence η_y sont les variables de contrôle les plus éloignées du modèle Gaussien. Les parties non balancées de T et q sont, elles, plus proches de la Gaussianité que T et q totales.

3.1.4. Perspectives

Cette étude met en exergue deux pistes pour améliorer la prise en compte de la non-Gaussianité. La première piste est de travailler au niveau des variables dynamiques de contrôle de l'assimilation. Vu leur forte NG, il peut être intéressant d'utiliser des techniques de « Gaussianisation » telle que cela est fait pour l'humidité spécifique (Holm et al., 2002; Dee & Da Silva, 2003), ou de rediscuter le choix des variables de contrôle dynamiques.

Dans un cadre idéalisé, une étude faite par Lawson & Hansen (2005) montre que les erreurs de déplacement engendrent de la NG. En se basant sur ce constat, la deuxième piste de travail consiste justement à corriger ces erreurs de déplacement pour tenter de prendre en compte au moins une partie de la NG. L'étude des erreurs de déplacement fait l'objet des chapitres 4.

3.2. Version originale de l'article publié



Diagnosing non-Gaussianity of forecast and analysis errors in a convective-scale model

R. Legrand, Y. Michel, and T. Montmerle

Centre National de Recherches Météorologiques (CNRM), Toulouse, France

Correspondence to: R. Legrand (raphael.legrand@meteo.fr)

Received: 2 June 2015 – Published in Nonlin. Processes Geophys. Discuss.: 18 July 2015

Revised: 27 November 2015 – Accepted: 4 December 2015 – Published: 26 January 2016

Abstract. In numerical weather prediction, the problem of estimating initial conditions with a variational approach is usually based on a Bayesian framework associated with a Gaussianity assumption of the probability density functions of both observations and background errors. In practice, Gaussianity of errors is tied to linearity, in the sense that a nonlinear model will yield non-Gaussian probability density functions. In this context, standard methods relying on Gaussian assumption may perform poorly.

This study aims to describe some aspects of non-Gaussianity of forecast and analysis errors in a convective-scale model using a Monte Carlo approach based on an ensemble of data assimilations. For this purpose, an ensemble of 90 members of cycled perturbed assimilations has been run over a highly precipitating case of interest. Non-Gaussianity is measured using the K^2 statistics from the D’Agostino test, which is related to the sum of the squares of univariate skewness and kurtosis.

Results confirm that specific humidity is the least Gaussian variable according to that measure and also that non-Gaussianity is generally more pronounced in the boundary layer and in cloudy areas. The dynamical control variables used in our data assimilation, namely vorticity and divergence, also show distinct non-Gaussian behaviour. It is shown that while non-Gaussianity increases with forecast lead time, it is efficiently reduced by the data assimilation step especially in areas well covered by observations. Our findings may have implication for the choice of the control variables.

1 Introduction

In data assimilation, the analysis step may be seen as finding a maximum likelihood of the probability density functions (PDFs) of the state x given the available observations y and a background state (usually a short range forecast). The usual Bayesian formulation yields (Kalnay, 2003)

$$P_a(x|y) \propto P_o(y|x)P_b(x), \quad (1)$$

where P_a , P_b , and P_o respectively are the PDFs of analysis, background errors (*a priori* PDF), and observation errors. For high-dimensional systems, to specify those PDFs as multivariate Gaussian is a natural choice for variables that may approximately verify the central limit theorem (Bocquet et al., 2010). Thus, up to now most operational Numerical Weather Prediction (NWP) centres have relied on variational assimilation schemes that are Gaussian or corrections to a Gaussian analysis-based strategy.

The time integration of the model nonlinear dynamics leads inevitably to non-Gaussian forecast errors (Bocquet et al., 2010). For instance, the highly nonlinear processes involved in clouds and precipitation are known to give non-Gaussian background errors (Auligné et al., 2011). Some authors have reported on displacement errors of meteorological features that turn into non-Gaussian background errors (Lawson and Hansen, 2005). Keeping the Gaussian formalism in this case may yield unrealistic analyses that are distorted (Ravela et al., 2007).

In NWP, the analysis of humidity may be the most problematic with respect to non-Gaussianity (NG). This is due to the condensation effects near saturation and the intrinsic positivity of humidity. The choice of the control variable for humidity is a long-standing debate (Dee and da Silva, 2003). Specific humidity exhibits NG but is rather weakly

correlated (in average) to other variables. Relative humidity has been found to be more Gaussian but has stronger cross-covariances with temperature that are state-dependent and difficult to model. It still has skewed distribution near condensation or in dry conditions. The solution adopted in several operational centres is to use a normalized relative humidity variable. The normalization factor is the standard deviation of the relative humidity error, stratified according to the analysed relative humidity itself. The asymmetries in PDFs are also accounted for through a nonlinear transformation. This scheme has been implemented through several variants both in global (Holm et al., 2002; Ingleby et al., 2013) and in limited area models (Gustafsson et al., 2011).

The 4D-Var (4-dimensional variational) algorithm commonly used in NWP (e.g. Rabier et al., 2000) has some ability to handle nonlinearities. It solves for what would be the most probable state in Eq. (1) in the Gaussian case, by minimizing a nonquadratic cost function with nonlinearities in the model and in the observation operator mapping the model state to the observation space. The approach, known in the community as incremental 4D-Var (Courtier et al., 1994), is based on a form of truncated Gauss–Newton iterations. The problem is solved by minimizing a succession of inner-loop quadratic optimization problems with increasing horizontal resolutions, in which the model is simplified and linearized around the state adjusted by the previous outer-loop iteration (Laroche and Gauthier, 1998).

The PDF of observation errors is also non-Gaussian in general. In NWP, quality controls are performed to exclude observations that are outliers compared to the model and using statistical knowledge (Lorenc, 1986). Unfortunately, this can be erroneous and a more flexible framework has been introduced, for instance, by Anderson and Järvinen (1999). It explicitly computes the probability of gross error for each observation, given the preliminary analysis from the outer loops. The weight of each observation is smoothly decreased with increased likelihood for gross error. More recently, this scheme has been replaced by the use of a Huber norm (Tavolato and Isaksen, 2014). The NG of observation errors is out of the scope of this paper.

The main goal of this paper is to rely on a Monte Carlo approach to document the spatial variations of non-Gaussianities of background and of analysis errors for a particular meteorological case, in the context of convective-scale NWP. For this purpose, a large ensemble of perturbed cycled assimilations has been set up with the AROME-France¹ model. The perturbations simulate the evolution of the true background and analysis errors (Houtekamer et al., 1996; Fisher, 2003; Berre et al., 2006). Local and spatially averaged diagnostics of NG may help to find out for which variables and/or in which areas efforts could be made to improve Gaussian assumptions in the assimilation algorithm

¹Application de la Recherche à l’Opérationnel à Mésos-Echelle (Seity et al., 2011).

or, for instance, to help designing advanced data assimilation schemes taking into account displacement errors (Ravela et al., 2007).

The paper is organized as follows: Sect. 2 presents the univariate D’Agostino test for NG (D’Agostino et al., 1990) and evaluates its efficiency on some specified PDFs. Section 3 describes the ensemble from which the NG is diagnosed. This ensemble is composed of assimilations and forecasts performed by the AROME-France model for a highly precipitating event over the Mediterranean sea, of interest for the HyMeX (HYdrological cycle Mediterranean EXperiment) campaign (Ducrocq et al., 2013). Results of the NG diagnostics are then documented. After an overview for model prognostic variables, time evolution of NG is discussed. The dependence of NG to physical nonlinear processes is then described by making use of geographical masks based on cloud contents. In Sect. 4, the impact of the data assimilation process on NG is studied by comparing diagnostics performed on both background and analysis errors and by computing diagnostics in the control space of the minimization. Conclusions are given in Sect. 5.

2 An index of non-Gaussianity

In NWP, dimensions of the state and observation vectors, including satellite and radar, are huge (respectively around 10^8 and 10^5 in AROME-France). As mentioned in Bocquet et al. (2010) only the simpler statistical tests of Gaussianity are tractable for such high dimensional problems. Therefore, while it is the Gaussianity of the global joint PDF that matters in Eq. (1), only univariate marginal PDF testing for NG is diagnosed in this paper. Spatial variations and the average of local values may however give an insight of non-Gaussian behaviours for the meteorological case treated here.

2.1 D’Agostino test

The D’Agostino test (hereafter K^2 test; D’Agostino et al., 1990) is a univariate statistical test where the deviation from Gaussianity is detected from the PDF’s skewness and kurtosis. The skewness is a measure of the asymmetry of the PDF about its mean. Positive (negative) values are associated with a median of the PDF smaller (larger) than its mean and with a large right (left) tail. For instance, a negative skewness for specific humidity at some point indicates that more than the half of the ensemble is more humid than the mean value of the ensemble. The kurtosis measures the peakedness of the distribution (Thode, 2002). A PDF with larger tails and a narrow modal peak has a large kurtosis.

The theoretical skewness and kurtosis are respectively estimated over an ensemble by the sample third (G_3) and fourth (G_4) standardized moments. They are defined given a sample

$x_{i=1..N_s}$ of size N_s and its sample mean \bar{x} as

$$G_3 = \frac{m_3}{m_2^{\frac{3}{2}}} = \frac{\frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})^3}{\left[\frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})^2 \right]^{\frac{3}{2}}}, \quad (2)$$

$$G_4 = \frac{m_4}{m_2^2} = \frac{\frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})^4}{\left[\frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - \bar{x})^2 \right]^2}, \quad (3)$$

with m_2 , m_3 and m_4 the sample second-, third-, and fourth-order moments. These quantities estimate the theoretical skewness and kurtosis of the distribution. For a Gaussian PDF, skewness is zero and kurtosis equals 3. Thus, the sample skewness and kurtosis defined above could be used to detect deviation from Gaussianity, yet their convergence to normality with ensemble size is slow. As reported in Tables 3.1 and 3.2 of Thode (2002), the normality is reached with sufficient accuracy typically for ensemble sizes of the order of ~ 5000 . For smaller ensemble sizes (more suitable to NWP), it has been suggested to transform these quantities into $f_3(G_3)$ and $f_4(G_4)$ respectively, in order to remedy this situation (D'Agostino, 1970; Anscombe and Glynn, 1983). f_3 is defined as

$$\begin{aligned} A &= G_3 \times \sqrt{\frac{(N_s + 1)(N_s + 3)}{6(N_s - 2)}}, \\ B &= 3 \frac{(N_s^2 + 27N_s - 70)(N_s + 1)(N_s + 3)}{(N_s - 2)(N_s + 5)(N_s + 7)(N_s + 9)}, \\ C &= \sqrt{2(B - 1)} - 1, \\ D &= \sqrt{C}, \\ E &= \frac{1}{\sqrt{\ln(D)}}, \\ F &= \frac{A}{\sqrt{\frac{2}{C-1}}}, \end{aligned}$$

$$f_3(G_3) = E \times \ln(F + \sqrt{F^2 + 1}),$$

and f_4 is defined as

$$\begin{aligned} O &= G_4 \times \frac{N_s(N_s + 1)}{(N_s - 1)(N_s - 2)(N_s - 3)} - 3 \frac{(N_s - 1)}{(N_s + 1)}, \\ P &= \frac{24N_s(N_s - 2)(N_s - 3)}{(N_s + 1)^2(N_s + 3)(N_s + 5)}, \\ Q &= \frac{(N_s - 2)(N_s - 3)}{(N_s + 1)(N_s - 1)\sqrt{P}} \times O, \\ R &= \frac{6(N_s^2 - 5N_s + 2)}{(N_s + 7)(N_s + 9)} \sqrt{\frac{6(N_s + 3)(N_s + 5)}{N_s(N_s - 2)(N_s - 3)}}, \\ S &= 6 + \frac{8}{R} \left[\frac{2}{R} + \sqrt{1 + \frac{4}{R^2}} \right], \\ T &= \frac{1 - \frac{2}{S}}{1 + Q\sqrt{\frac{2}{S-4}}}, \end{aligned}$$

$$f_4(G_4) = \frac{1 - \frac{2}{9S} - T^{\frac{1}{3}}}{\sqrt{\frac{2}{9S}}}.$$

While positive (negative) values of $f_3(G_3)$ point out distributions with a median smaller (higher) than the mean and with a longer right (left) tail, positive (negative) values of $f_4(G_4)$ mean that distribution tails are heavier (lighter) than Gaussian distributions, with also a larger (smaller) modal peak.

$f_3(G_3)$ and $f_4(G_4)$ statistics are then combined to produce an omnibus test K^2 , able to detect deviations from normality due to either skewness or kurtosis:

$$K^2 = f_3^2(G_3) + f_4^2(G_4). \quad (4)$$

When testing a Gaussian distribution, asymptotic values for the three criteria ($f_3(G_3)$, $f_4(G_4)$, and K^2) are respectively $f_3(G_3) = 0$, $f_4(G_4) = 0$, and $K^2 = 2$. Using finite sampling with ensemble size $N_s > 20$ (Thode, 2002), $f_3(G_3)$ and $f_4(G_4)$ could be both assumed to follow a Gaussian law with a zero mean and a unity variance. In this case, K^2 follows approximately a χ^2 distribution with 2° of freedom. Confidence intervals at 95 % are then given by $f_3(G_3) \in [-1.96; 1.96]$, $f_4(G_4) \in [-1.96; 1.96]$, and $K^2 \in [0; 5.991]$. Because G_3 and G_4 are uncorrelated but not independent, K^2 does not follow an exact χ^2 distribution, and the confidence interval is slightly different. Using a right-tailed unilateral testing at 95 % for $N_s = 100$, the critical value of K^2 is 6.271 instead of 5.991.

2.2 Evaluation

The efficiency of the K^2 test can be evaluated by measuring its probability of detection (POD) for the Gaussian hypothesis H_0 . For a sample known to be from a non-Gaussian PDF, the POD gives the probability that the test accurately rejects H_0 . The best result is $\text{POD} = 1$.

The POD of the K^2 test is estimated from N_{xp} independent experiments. For each experiment, K^2 is computed from N_s elements sampled from a known distribution. Depending on the K^2 value, H_0 is accepted or rejected. When the known distribution is non-Gaussian, POD is given by the frequency of H_0 rejections over the N_{xp} experiments.

The POD is estimated for three non-Gaussian distributions: uniform, log-normal, and a Gaussian mixture. The Gaussian mixture is defined through its PDF as $P(x) = w_1 P_1(x) + w_2 P_2(x) + w_3 P_3(x)$ with P_1 , P_2 , and P_3 , three Gaussian distributions with zero mean and respectively 0.1, 0.05, and 0.02 as chosen standard deviations. The chosen weights are given by $(w_1, w_2, w_3) = (0.2, 0.5, 0.3)$. The representation of the shapes of these three distributions is given in Fig. 1a, alongside the Gaussian distribution.

PODs are estimated over $N_{xp} = 10^5$ experiments. For both tests, different ensemble sizes N_s are tested ($N_s = 20, 30, 40, 50, 60, 70, 80, 90, 100, 200$). Results of this ideal case are shown in Fig. 1b. The log-normal distribution is the easiest one to discriminate from the Gaussian distribution, yielding

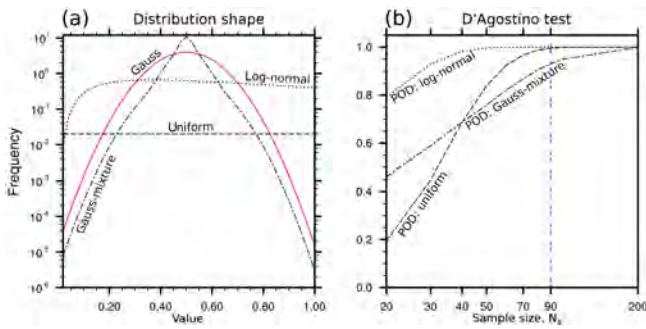


Figure 1. (a) Three non-Gaussian distributions on which PODs have been estimated: uniform distribution, log-normal distribution, and Gaussian mixture (see text for description). (b) POD for K^2 test. PODs are computed over $N_{xp} = 10^5$ one-dimensional experiments for different sample sizes N_s .

the highest POD that reaches almost 1 as soon as the ensemble size is above 40. For the two others, the non-Gaussian distributions (uniform and Gaussian-mixture) K^2 test is only correctly discriminating from Gaussianity (with $\text{POD} > 0.8$) when $N_s > 70$. For $N_s = 90$, which corresponds to the ensemble size for the real data set composed of AROME-France forecasts (see Sect. 3), POD values are over 0.9 for all three non-Gaussian distributions. In conclusion, the K^2 test is able to correctly discriminate NG for the ensemble size considered in this paper.

A review of other well-established tests for Gaussianity are presented in Bocquet et al. (2010), such as the measure of entropy (kullback (1959), used in geophysics by Pires et al. (2010)), or the univariate Anderson–Darling goodness-of-fit test (Anderson and Darling, 1954). The latter has been also tested in the same framework and the performances proved to be very similar to the ones of the K^2 test. When comparing the results, obtained over the ensemble (Sect. 3), these two tests also give very similar results; e.g. they indicate the same areas of NG over $\approx 90\%$ of the domain. However, measuring skewness and kurtosis may be more informative and may be of interest for some assimilation schemes that account for skewness (Hodyss, 2012). Also, describing the values of K^2 has the advantage of preventing the results from depending on the chosen confidence level.

3 Diagnosis of the non-Gaussianity of AROME forecast errors

3.1 An AROME-France ensemble for a high-precipitating case

AROME-France is an operational nonhydrostatic model covering France with a 2.5 km horizontal resolution at the time of the experiments. Its lateral boundary conditions are given

by the global model ARPEGE². Assimilation steps are done every 3 h with a 3D-Var scheme and make use of a comprehensive set of observations such as conventional, satellite or Doppler radar data (see Seity et al. (2011) for more details).

The simulation of background and analysis errors is achieved by using a Monte Carlo sampling, called an ensemble data assimilation (EDA) in the context of NWP. A 90-member EDA is first run for the global model (AEARP, Berre and Desroziers, 2010). Each EDA member is based on a 4D-Var cycled assimilation which uses perturbed observations and a perturbed background, in order to simulate the error evolution (Berre et al., 2006). Observation perturbations are constructed as random draws of the specified observation error covariance matrix, and background perturbations result from the forecast evolution of previous analysis perturbations and from their inflation at the end of each forecast (Raynaud et al., 2012). This global ensemble provides perturbed boundary conditions to an ensemble of perturbed 3D-Vars for AROME-France, as described in Ménétrier et al. (2014). True background errors are then approximated by the deviations of the perturbed backgrounds from the ensemble mean. A few cycles (typically four) are necessary to reach a regime where the spread of the ensemble is representative of the true error spread; these cycles are discarded from the diagnostics presented below.

The case of interest is 4 November 2011 between 00:00 and 06:00 UTC (universal time coordinated). A strong southerly convergent flow occurs at low levels over southern France (Fig. 2). Warm and moist air from the Mediterranean sea is advected over land, which triggers deep convection. Those high intensity events, named cévenol events, are studied by the HyMeX research program (Ducrocq et al., 2014). Associated precipitations are visible all along the Rhone Valley, with local maxima exceeding 25 mm h^{-1} . Also, associated with a low pressure area over the north-east Atlantic (not shown), a cold active front extending from the bay of Biscay to the eastern British coast is sweeping north-west of France with locally strong precipitations.

3.2 Vertical profiles of NG

The vertical profiles of quantities related to NG are shown in Fig. 3 for 3 h forecasts of different variables, namely zonal (U) and meridian (V) winds, temperature (T) and specific humidity (q). On average, except near the surface, q is the variable that shows the largest deviation from Gaussianity, confirming results obtained at the global scale (Holm et al., 2002). From 850 to 350 hPa, q is indeed characterized with an increase of the deviation from Gaussianity. As shown in Fig. 3b, this NG is partly explained by negative values of the skewness, highlighting a left-tailed PDF of the background

²Action de Recherche Petite Échelle Grande Échelle (Pailleux et al., 2000).

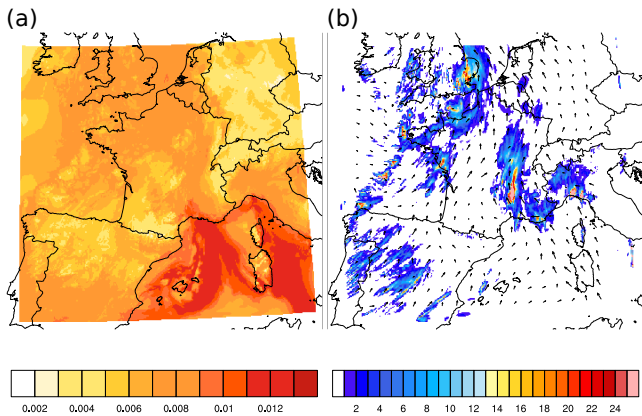


Figure 2. (a) Specific humidity (q , kg kg^{-1}) and (b) surface cumulative precipitation (mm h^{-1}) overlaid with winds vector at model level 52 (≈ 920 hPa). Maps are given for one member of the AROME-France 3 h forecasts ensemble, valid at 03:00 UTC on 4 November 2011.

errors, meaning that many values are more humid than the ensemble mean.

In the troposphere, K^2 is increasing while the q mean content, displayed in Fig. 3d, is largely decreasing. Values at higher levels, where q is almost nonexistent, may however be taken with caution. Below 850 hPa, K^2 is peaking around 960 hPa. Above 850 hPa, the wind components and T remain close to Gaussianity. Below, however, all variables have significant deviation from Gaussianity, especially T for which high values of K^2 are found at ground level, making of it the less Gaussian variable in the boundary layer.

3.3 Horizontal structures of NG

The range, defined as the difference between the 95th and the 5th percentiles, could be used to describe roughly the horizontal spatial variability for each vertical level. Vertical profiles of ranges of K^2 , $f_3(G_3)$, and $f_4(G_4)$ (not shown) have large similarities between each other and with the shapes of K^2 profiles displayed in Fig. 3a. They include in particular two maxima in the boundary layer and in the high troposphere for q and larger values towards the surface for T . Ranges are much larger for the four variables (approximately four times as large) than the respective mean values given in Fig. 3, implying a large spatial variability for the three NG diagnostics. An example of the horizontal structures of NG is given for q in the boundary layer by Fig. 4. They have large similarities with the meteorological coherent structures, as the southerly convergent flow over the south of France and the active cold front aloft the north-west of France are associated with high values of K^2 .

Supporting the conclusion drawn from Fig. 3, transformed skewness $f_3(G_3)$ is mainly negative (corresponding to left-tailed distributions) over the domain and has a larger contribution than transformed kurtosis $f_4(G_4)$ in large K^2 values.

Over the Mediterranean Sea, the skewness represents on average 70% of K^2 .

It may be interesting to compare NG with the variance of the ensemble, as K^2 is defined from standard third and fourth standardized moments avoiding any scale effect. As displayed in Fig. 5, the variance does not coincide with overall NG, even if it happens that Gaussian areas may coincide with regions of low variance.

NG of the surface pressure is not shown in this study since, according to our diagnostics, it is a mainly Gaussian variable (averaged K^2 around 2.7). High values of K^2 appear around the cold front and the convergence area but they are very localized and of smaller amplitude compared to the other model variables.

3.4 Time evolution of non-Gaussianity

For each member of the ensemble, 18 h forecasts have been run from the analyses performed at 00:00 UTC on 4 November 2011. This allows diagnosing NG every 6 h during the first 18 h of integration. The corresponding vertical profiles are shown in Fig. 6 for the two most non-Gaussian variables according to Sect. 3.2: q and T .

In order to get insights into the processes that may be involved in NG development, the diagnostics have been separately computed for cloudy and for clear sky areas, following a similar approach to that of Montmerle and Berre (2010) and Michel et al. (2011), in which precipitating masks have been used. Grid points over the domain are separated in two bins: “cloudy” or “clear-sky” points. The cloudy bin defines grid points whose vertically integrated simulated cloud water exceeds 0.1 g kg^{-1} for a majority of ensemble members (i.e. more than 45 members for the 90-member ensemble). The other points are classified as clear sky. The percentage of clear-sky points being 3–5 times larger (not shown) than the detected cloudy points, similarities between clear-sky profiles, and profiles averaged over the whole domain (as plotted in Fig. 3) are apparent.

During the first 6 h of forecasts, NG quickly increases. For q , all tropospheric model levels are affected. For T , starting from a fairly Gaussian profile, increase of NG is mainly affecting the boundary layer and higher levels remain close to Gaussianity. During the following 12 h (from 6 to 18 h forecast), changes of NG are smaller for both variables. Those results support that NG in the background may rather come from nonlinear processes acting on nearly Gaussian PDFs instead of linear processes acting non-Gaussian PDFs.

It is interesting to note that different behaviours can be found for diagnostics computed over cloudy and clear-sky areas. For q , NG is mainly found in cloudy areas, where K^2 quickly reaches values above 8, with two peaks around 900 and 700 hPa. The altitude of the lower peak rises with forecast ranges, while the amplitude of the higher one increases. According to Fig. 6c, which displays the time average of the mean cloud contents, this evolution of NG in cloudy areas

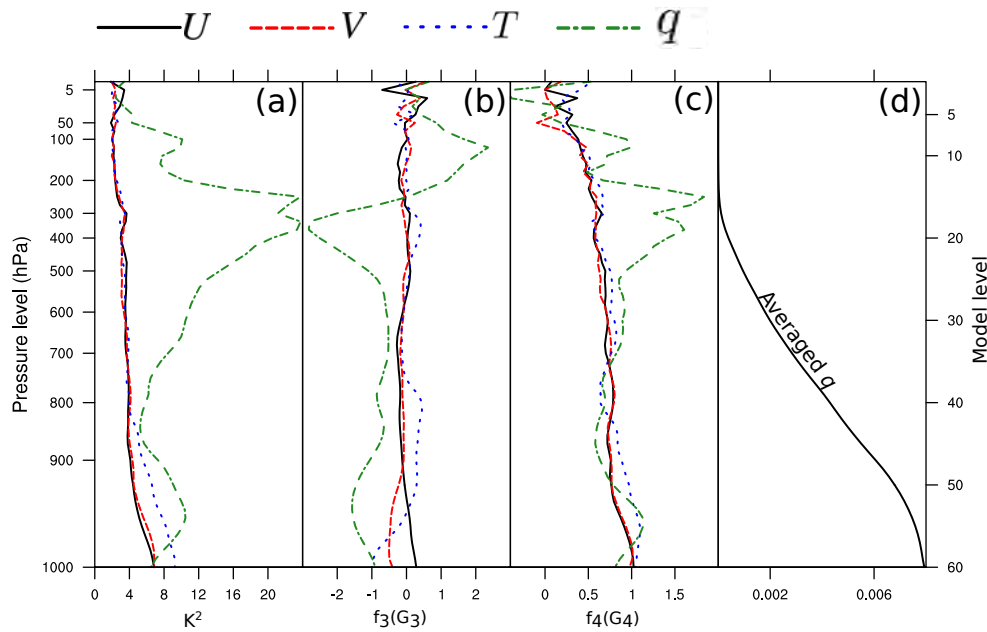


Figure 3. Vertical profiles of (a) K^2 , (b) transformed skewness $f_3(G_3)$, (c) transformed kurtosis $f_4(G_4)$, and (d) q (kg kg^{-1}) for one member of the ensemble. For each level, values are averaged over the horizontal domain. Profiles are computed from the 90-member ensemble of AROME-France 3 h forecasts valid at 03:00 UTC on 4 November 2011. Profiles in (a), (b), and (c) are given for four model variables: U , V , T , and q .

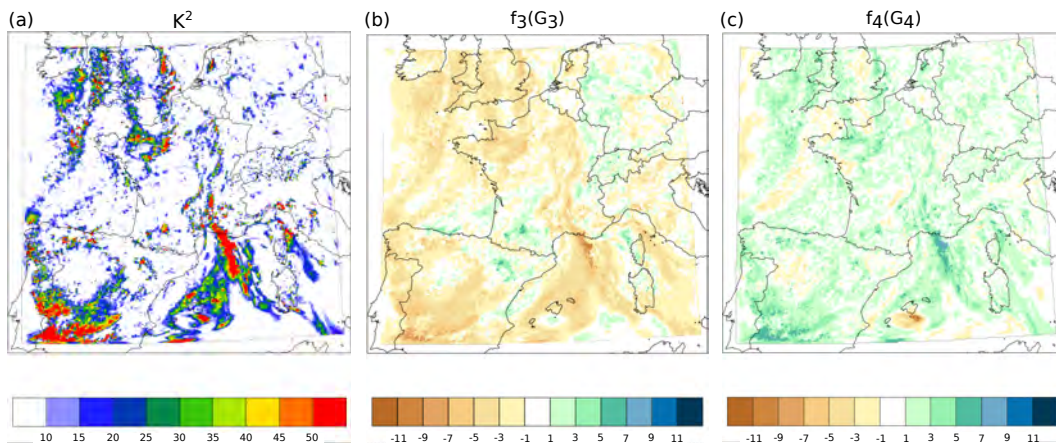


Figure 4. (a) K^2 , (b) transformed skewness $f_3(G_3)$, and (c) transformed kurtosis $f_4(G_4)$ for q at model level 52 (≈ 920 hPa), computed from the 90-member ensemble of AROME-France 3 h-forecast valid at 03:00 UTC on 4 November 2011.

is likely due to nonlinear processes such as the vertical displacement error of cloud base and top within the ensemble and possibly the diabatic processes. In surface layers, K^2 for T quickly increases especially for clear air areas where turbulent and radiative processes occur. After 12 h, NG is more spread vertically within clouds, probably because of diabatic processes. For T and q , diabatic processes are good candidates to produce NG because of intrinsic thresholds in cloud physics (e.g. moisture saturation) and nonlinear processes like turbulence on cloud-top.

For the wind components, behaviours close to T have been found but with smaller amplitude (not shown): NG increases mainly in the boundary layer in clear-sky areas and may be due to nonlinear turbulent processes.

4 Non-Gaussianity in the data assimilation process

Based on comparisons of NG diagnostics between successive background and analysis errors, this section focuses on the evolution of NG through cycled 3D-Var assimilations. Analysis errors will be treated for both model and control

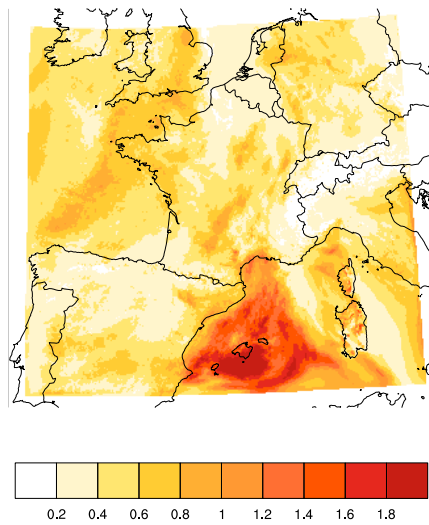


Figure 5. Background-error standard deviations of q (g kg^{-1}) for the model AROME-France, at model level 52 (≈ 920 hPa). Standard deviations are estimated from the 90-member ensemble of 3 h forecasts, valid at 03:00 UTC on 4 November 2011.

variables. The link between assimilated observations and NG reduction will be shown.

4.1 Overview

An overview of the NG evolution during the analysis process is given in Fig. 7 that shows averaged K^2 profiles for the analysis and the background errors computed for two consecutive assimilation/3 h forecast steps. Comparable results are found for the two cycles, confirming the increase of NG during the model integration and highlighting the substantial reduction of NG during the assimilation process, especially for levels where NG grows quickly. Values of K^2 are indeed brought back to much more Gaussian values, even in the lower levels for both q and T and in the higher troposphere for q .

Geographical variations of NG are illustrated in Fig. 8. As in Fig. 7, the NG of the background and of the following 3 h forecast is similar. The largest decreases of NG between background and analysis error match areas with a large analysis increment, in particular where radar data are assimilated (Fig. 8d). The analysis increment being a linear function of the innovation vector in model space (observation minus background), its Gaussianity is insured by a rough selection applied beforehand to the observations, allowing us to remove outliers (e.g. for radar data, Caumont et al., 2010; Wattrelot et al., 2014). Some NG areas remain though, especially in areas where the background is less constrained by observations (e.g. above Spain and above the sea). However, most areas where NG has been reduced thanks to the data assimilation process recover their NG nature after 3 h of model integration.

4.2 Non-Gaussianity in control space

Previous results are documenting the NG of four model prognostic variables: U , V , T , and q . As it is detailed in Brousseau et al. (2011), the assimilation scheme in AROME-France is based on a 3D-Var whose control variables are the vorticity ζ , the unbalanced divergence η_u , the unbalanced temperature and surface pressure $(T, P_s)_u$, and the unbalanced specific humidity q_u . These control variables are linked to the model variables following the multivariate formalism of Berre (2000), which is based on the decomposition of the background error covariance matrix in spatial operators and balance transforms. Since the minimization is performed in the control space, NG diagnostics have also been computed for these control variables.

4.2.1 Overview

Vertical profiles of NG for control variables are presented in Fig. 9. Unlike the zonal and meridian winds, ζ and η_u are strongly non-Gaussian over the whole troposphere, whereas T_u and q_u display much more Gaussian profiles.

Negative values of $f_3(G_3)$ below 800 hPa for η_u (Fig. 9b) denote a larger spread of the distribution below the mean, probably due to the occurrence of low level convergence. At mid-troposphere, error distributions of all four variables are near symmetric. Except for q_u , distributions in tropospheric levels remain symmetric and the K^2 index is mainly explained by the kurtosis (Fig. 9c).

Those results agree with one of the conclusions of Ménétrier et al. (2015). These authors describe an algorithm to find the optimal truncation dedicated to sample covariances filtering. This algorithm has two variants. The first one assumes a Gaussian PDF for the background perturbations while the second one does not. Their study indicates that, at convective scale, the Gaussian variant is accurate for T_u and q_u , but the more general non-Gaussian variant has to be used for ζ and η_u , which are significantly non-Gaussian variables in agreement with our study. To go further on this topic, NG diagnostics have been computed for the spatial first-order derivative of T . While T is in average a locally nearly Gaussian variable (see Fig. 3a), its spatial differentiation largely increases the NG (not shown), up to the order of magnitude found in Figs. 9a and 10 for ζ and η . This supports the attribution to differentiation for at least a part of the NG displayed for the dynamical control variables.

While very similar, the horizontal structures of K^2 for ζ and η_u are noisier compared to the other variables, with very small scale and intense signals (not shown). Maps mostly follow the land–sea mask, with high values of K^2 over sea and low values over land. Aloft, NG follows meteorological active structures (cold front and cévenol event).

As for Fig. 7, diagnostics of NG for vorticity ζ and total divergence η have been computed before and after the assimilation step (not shown). While NG of levels higher than

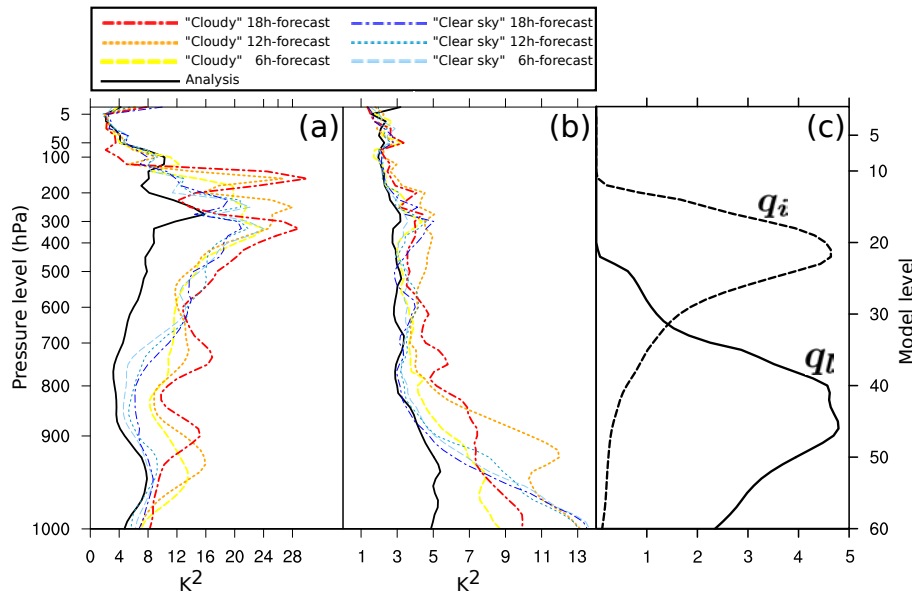


Figure 6. Time evolution of the vertical profiles of K^2 for (a) q and (b) T computed for (thick and hot colours) cloudy and (thin and cold colours) clear-sky points (see text). (c) Vertical profiles of liquid cloud q_l (solid line) and ice cloud q_i (dashed line) contents (g kg^{-1}). The cloud contents are averaged over the domain and over times from 06:00 to 18:00 UTC, every 6 h. Initial cloud water profile is null because the hydrometeors are not cycled. Consequently the initial time profiles of K^2 are common for the two bins. Profiles have been computed from 00:00 to 18:00 UTC, every 6 h, using forecasts initialized with analysis states valid on 4 November 2011, 00:00 UTC.

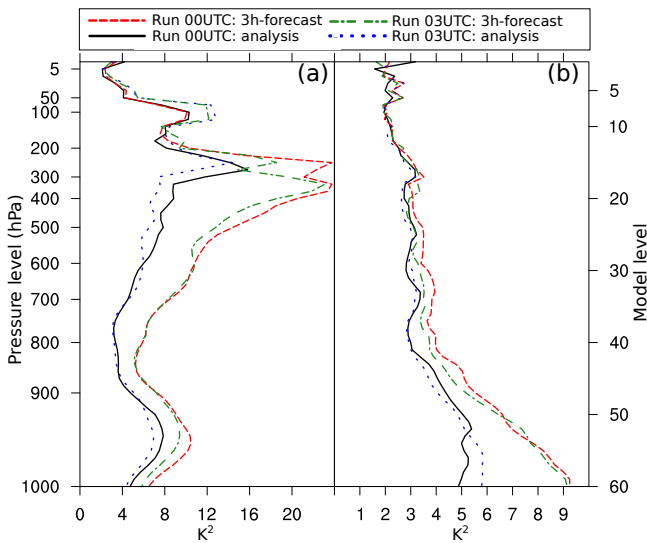


Figure 7. Vertical profiles of K^2 on background and analysis errors for (a) q and (b) T , for two successive cycled assimilation/3 h forecast steps starting at 00:00 UTC on 4 November 2011.

900 hPa are almost unchanged, the averaged K^2 of ζ and η is systematically lower for the analysis than the background state in the boundary layer. However, the order of magnitude of the decrease is much smaller than for T and q , and the dynamical variables ζ and η remain by far much more non-Gaussian.

4.2.2 Non-Gaussianity in the multivariate transform

To go further in the discussion on Gaussianity of the control variables, this section compares the K^2 values for total and unbalanced variables.

According to Fig. 10, the debalancing process is not really affecting the NG for the divergence, except at lower levels where K^2 is slightly decreasing while keeping large values. K^2 values remain 2–3 times larger for the divergence (total or unbalanced) than for T and q from the surface to the mid-troposphere. On the contrary, NG decreases significantly for T and q during the debalancing process. Changes mainly appear in the boundary layer for T . For q , changes appear for every model level, especially in the boundary layer and below the tropopause. From the surface to 750 hPa, the NG of q_u is equal to or smaller than the NG of T_u .

5 Conclusions

It is suggested to use the K^2 value from the D'Agostino test for diagnosing the local NG of a NWP system at convective scale. This diagnostic is computed from the univariate sample skewness and kurtosis from a 90-member ensemble. Even if checking local NG is not sufficient to describe deviations from the global Gaussian hypothesis for AROME-France background and analysis errors, spatial variations and averages of local diagnostics give an insight of non-Gaussian

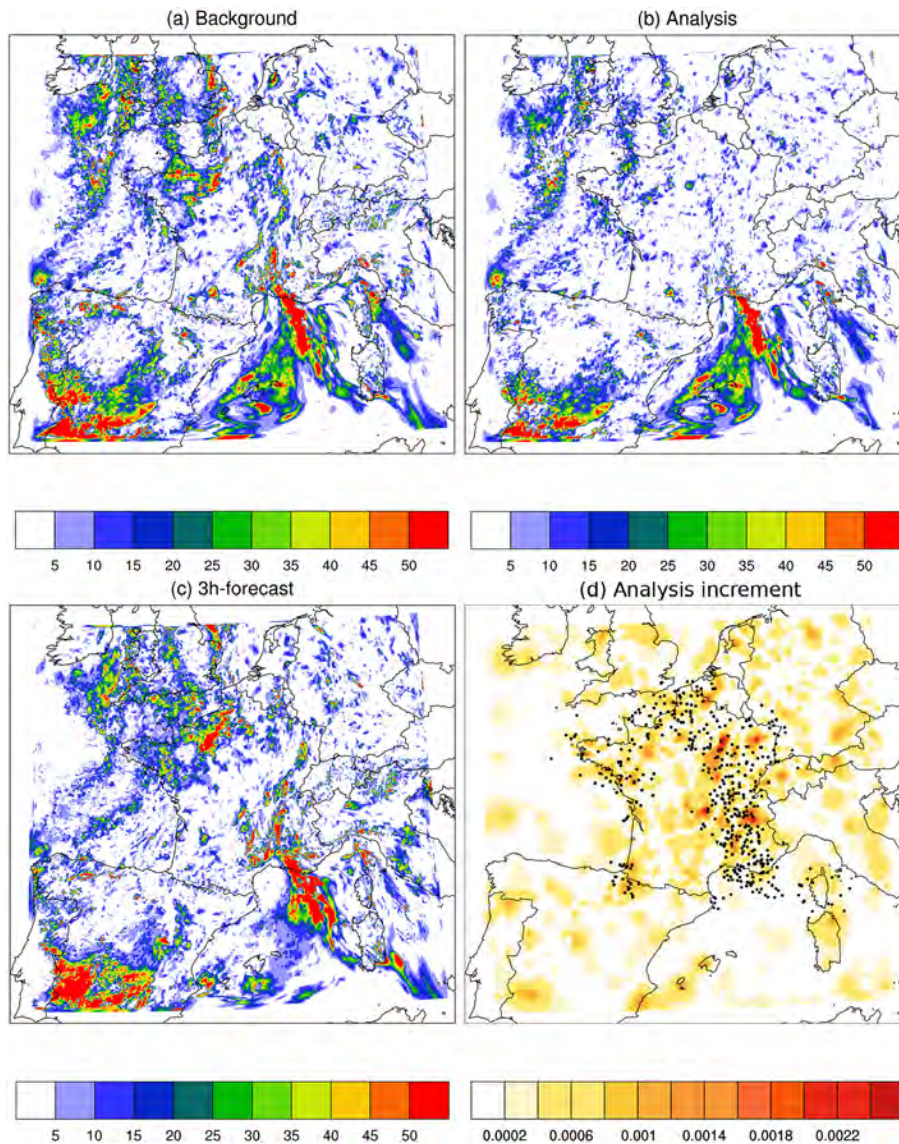


Figure 8. K^2 for q at level 52 (≈ 920 hPa) for (a) the background, (b) the analysis, and (c) the following 3 h forecast starting at 03:00 UTC on 4 November 2011. (d) Corresponding analysis increment (kg kg^{-1}) with positions of radar precipitation observations assimilated.

behaviours for a case study characterized by a cévenol event and an active cold front.

According to our diagnostic, among model variables, q has the largest deviation from Gaussianity, with a maximum of amplitude near the tropopause and in the boundary layer. Deviation from Gaussianity for U , V , and T only appears in the boundary layer. With an heterogeneous diagnostic, NG has been separately diagnosed for cloudy points and clear-sky points. For q , cloud covering leads to higher NG, especially at the bottom and at the top of the cloud layer. In clear-sky situations, surface processes are expected to enlarge K^2 for T , in a larger manner than for cloudy points. Studying time evolution through forecast ranges, NG is mainly increasing during the first 6 h. The 3D-Var assimilation appears to effi-

ciently reduce the growing NG of the forecast, especially in well-observed areas. Finally, among control variables of the assimilation, ζ and η_u deviate from Gaussianity in a larger manner than T_u and q_u , which are much more Gaussian than their balanced counterparts.

Despite the fact that this work is attributing non-Gaussian behaviours to well-known nonlinear processes, such as the microphysical or boundary layer processes, it is not precisely addressing the cause of NG. However, two important questions on variational data assimilation are highlighted. First, regarding control variables of the assimilation, according to our diagnostic, the most non-Gaussian variables are the vorticity and the divergence. Yet, the main efforts have been put on “Gaussianization” of specific humidity (e.g. Holm et al.,

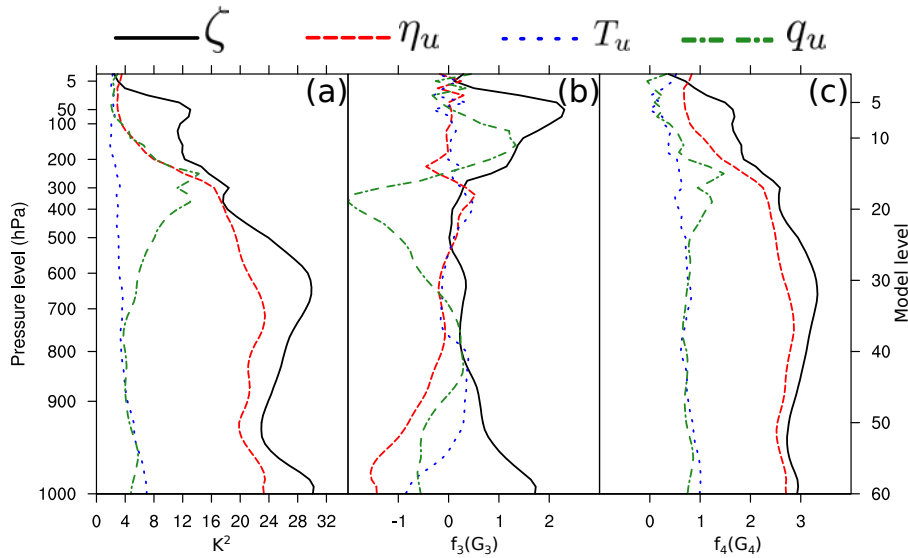


Figure 9. Vertical profiles of (a) K^2 , (b) transformed skewness $f_3(G_3)$, and (c) transformed kurtosis $f_4(G_4)$. For each level, values are averaged over the horizontal domain. Profiles are computed from the 90-member ensemble of AROME-France 3 h forecasts valid at 03:00 UTC on 4 November 2011. Profiles in (a), (b), and (c) are given for four control variables: ζ , η_u , T_u , and q_u .

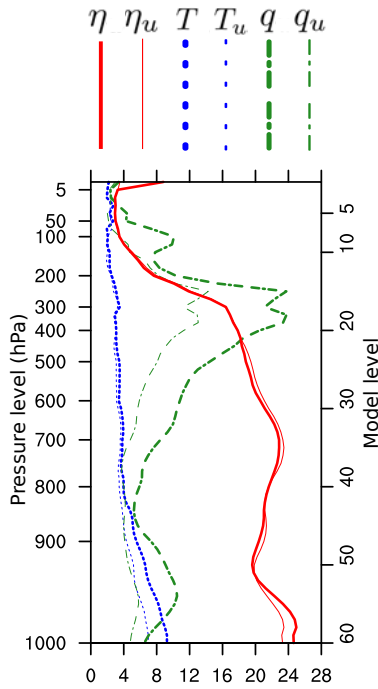


Figure 10. Comparison of K^2 vertical profiles of model variables (thick lines) and control variables (thin lines). Profiles are computed from 3 h forecasts valid at 03:00 UTC on 4 November 2011.

2002), but the discussion may also be focused on vorticity and divergence, either with a Gaussianization of those variables or with a discussion on the possibility to use other dynamical variables. Second, with the cloud mask approach, cloud layers have been associated with high values of NG.

This study uses an ensemble at convective scale that does not include model error either in the analysis or in the forecast steps. It is possible that conclusions would be different if stochastic noise drawn explicitly from a Gaussian distribution is added to the model states during the forecasts, as stated by Lawson and Hansen (2004). Also, this study is actually a part of work focused on the correction of displacement errors. Since displacement errors are identified to cause NG (Lawson and Hansen, 2005), diagnostics of NG may be used to evaluate improvements in the current amplitude error correction step (3D-Var) brought by a displacement error correction (Ravela et al., 2007). This will be examined in future work.

Acknowledgements. This work has been supported by the French Agence Nationale de la Recherche (ANR) via the IODA-MED Grant ANR-11-BS56-0005 and by the MISTRALS/HyMeX program. The authors thank Benjamin Ménétrier, Gérald Desroziers, and Loïk Berre for their scientific advice and their careful readings that proved very useful to improve the manuscript.

Edited by: O. Talagrand

Reviewed by: C. Pires and two anonymous referees

References

Anderson, E. and Järvinen, H.: Variational quality control, *Q. J. Roy. Meteorol. Soc.*, 125, 697–722, doi:10.1002/qj.49712555416, 1999.
 Anderson, T. W. and Darling, D. A.: A test of goodness of fit, *J. Am. Stat. Assoc.*, 49, 765–769, 1954.

- Anscombe, Francis J. and Glynn, William J.: Distribution of the kurtosis statistic b_2 for normal samples, *Biometrika*, 70, 227–234, 1983.
- Auligné, T., Lorenc, A., Michel, Y., Montmerle, T., Jones, A., Hu, M., and Dudhia, J.: Toward a new cloud analysis and prediction system, *B. Am. Meteorol. Soc.*, 92, 207–210, 2011.
- Berre, L.: Estimation of synoptic and mesoscale forecast error covariances in a limited-area model, *Mon. Weather Rev.*, 128, 644–667, 2000.
- Berre, L. and Desroziers, G.: Filtering of background error variances and correlations by local spatial averaging: a review, *Mon. Weather Rev.*, 138, 3693–3720, 2010.
- Berre, L., Ecaterina Ștefănescu, S., and Belo Pereira, M.: The representation of the analysis effect in three error simulation techniques, *Tellus A*, 58, 196–209, 2006.
- Bocquet, M., Pires, C. A., and Wu, L.: Beyond Gaussian statistical modeling in geophysical data assimilation, *Mon. Weather Rev.*, 138, 2997–3023, 2010.
- Brousseau, P., Berre, L., Bouttier, F., and Desroziers, G.: Background-error covariances for a convective-scale data-assimilation system: AROME–France 3D-Var, *Q. J. Roy. Meteorol. Soc.*, 137, 409–422, 2011.
- Caumont O., Ducrocq V., Wattrelot E., Jaubert G., and PRADIER-VABRE S.: 1D+ 3DVar assimilation of radar reflectivity data: A proof of concept, *Tellus A*, 62, 173–187, 2010.
- Courtier, P., Thépaut, J.-N., and Hollingsworth, A.: A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. Roy. Meteorol. Soc.*, 120, 1367–1387, 1994.
- D’Agostino, R. B.: Transformation to normality of the null distribution of G_1 , *Biometrika*, 57, 679–681, 1970.
- D’Agostino, R. B., Belanger, A., and D’Agostino Jr, R. B.: A suggestion for using powerful and informative tests of normality, *Am. Stat.*, 44, 316–321, 1990.
- Dee, D. P. and da Silva, A. M.: The choice of variable for atmospheric moisture analysis, *Mon. Weather Rev.*, 131, 155–171, 2003.
- Ducrocq, V., Belamari, S., Boudevillain, B., Bousquet, O., Cocquerez, P., Doerenbecher, A., Drobinski, P., Flamant, C., Labatut, L., Lambert, D., Nuret, M., Richard, E., Roussot, O., Testor, P., Arbogast, P., Ayrat, P.-A., Van Baelen, J., Basdevant, C., Boichard, J.-L., Bourras, D., Bouvier, C., Bouin, M.-N., Bock, O., Braud, I., Champollion, C., Coppola, L., Coquillat, S., Defer, E., Delanoë, J., Delrieu, G., Didon-Lescot, J.-F., Durand, P., Estournel, C., Fourrié, N., Garrouste, O., Giordani, H., Le Coz, J., Michel, Y., Nuissier, O., Roberts, G., Saïd, F., Schwarzenboeck, A., Sellegri, K., Taupier-Letage, I., and Vandervaere J.-P.: HyMeX, les campagnes de mesures: focus sur les événements extrêmes en Méditerranée, *Société météorologique de France, Paris, France, La Météorologie*, 80, 37–47, 2013.
- Ducrocq, V., Braud, I., Davolio, S., Ferretti, R., Flamant, C., Jansá, A., Kalthoff, N., Richard, E., Taupier-Letage, I., Ayrat, P.-A., Belamari, S., Berne, A., Borga, M., Boudevillain, B., Bock, O., Boichard, J.-L., Bouin, M.-N., Bousquet, O., Bouvier, C., Chiggiato, J., Cimini, D., Corsmeier, U., Coppola, L., Cocquerez, P., Defer, E., Delanoë, J., Di Girolamo, P., Doerenbecher, A., Drobinski, P., Dufournet, Y., Fourrié, N., Gourley, J. J., Labatut, L., Lambert, D., Le Coz, J., Marzano, F. S., Molinié, G., Montani, A., Nord, G., Nuret, M., Ramage, K., Rison, B., Roussot, O., Saïd, F., Schwarzenboeck, A., Testor, P., Van-Baelen, J., Vincen-
- don, B., Aran, M., and Tamayo, J.: HyMeX-SOP1, the field campaign dedicated to heavy precipitation and flash flooding in the northwestern Mediterranean, *B. Am. Meteorol. Soc.*, 95, 1083–1100, doi:10.1175/BAMS-D-12-00244.1, 2014.
- Fisher, M.: Background error covariance modelling, in: *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean*, ECMWF, 45–63, 2003.
- Gustafsson, N., Thorsteinsson, S., Stengel, M., and Holm, E.: Use of a nonlinear pseudo-relative humidity variable in a multivariate formulation of moisture analysis, *Q. J. Roy. Meteorol. Soc.*, 137, 1004–1018, doi:10.1002/qj.813, 2011.
- Hodyss, D.: Accounting for skewness in ensemble data assimilation, *Mon. Weather Rev.*, 140, 2346–2358, 2012.
- Holm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A., and Thépaut, J.-N.: Assimilation and modelling of the hydrological cycle: ECMWF’s status and plans, *European Centre for Medium-Range Weather Forecasts (ECMWF), Technical Memorandum*, 2002.
- Houtekamer, P., Lefaiivre, L., Derome, J., Ritchie, H., and Mitchell, H. L.: A system simulation approach to ensemble prediction, *Mon. Weather Rev.*, 124, 1225–1242, 1996.
- Ingleby, N., Lorenc, A., Ngan, K., Rawlins, F., and Jackson, D.: Improved variational analyses using a nonlinear humidity control variable, *Q. J. Roy. Meteorol. Soc.*, 139, 1875–1887, 2013.
- Kalnay, E.: *Atmospheric modeling, data assimilation, and predictability*, Cambridge University Press, p. 341, 2003.
- Kullback, S.: *Information theory and statistics*, New York, Wiley, 395 pp., 1959.
- Laroche S. and Gauthier, P.: A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow, *Tellus A*, 50, 557–572, 1998.
- Lawson, W. G. and Hansen, J. A.: Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth, *Mon. Weather Rev.*, 132, 1966–1981, 2004.
- Lawson, W. G. and Hansen, J. A.: Alignment error models and ensemble-based data assimilation, *Mon. Weather Rev.*, 133, 1687–1709, 2005.
- Lorenc, A. C.: Analysis methods for numerical weather prediction, *Q. J. Roy. Meteorol. Soc.*, 112, 1177–1194, 1986.
- Ménétrier, B., Montmerle, T., Berre, L., and Michel, Y.: Estimation and diagnosis of heterogeneous flow-dependent background-error covariances at the convective scale using either large or small ensembles, *Q. J. Roy. Meteorol. Soc.*, 140, 2050–2061, 2014.
- Ménétrier, B., Montmerle, T., Michel, Y., and Berre, L.: Linear filtering of sample covariances for ensemble-based data assimilation, Part I: optimality criteria and application to variance filtering and covariance localization, *Mon. Weather Rev.*, 143, 1622–1643, 2015.
- Michel, Y., Auligné, T., and Montmerle, T.: Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables, *Mon. Weather Rev.*, 139, 2994–3015, 2011.
- Montmerle, T. and Berre, L.: Diagnosis and formulation of heterogeneous background-error covariances at the mesoscale, *Q. J. Roy. Meteorol. Soc.*, 136, 1408–1420, 2010.

- Pailleux, J., Geleyn, J.-F., and Legrand, E.: La prévision numérique du temps avec les modèles ARPÈGE et ALADIN-Bilan et perspectives, La Météorologie, 2000.
- Pires, C. A., Talagrand, O. and Bocquet, M.: Diagnosis and impacts of non-Gaussianity of innovations in data assimilation, *Phys. D*, 239, 1701–1717, 2010.
- Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A.: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics, *Q. J. Roy. Meteorol. Soc.*, 126, 1143–1170, 2000.
- Ravela, S., Emanuel, K., and McLaughlin, D.: Data assimilation by field alignment, *Phys. D*, 230, 127–145, 2007.
- Raynaud, L., Berre, L., and Desroziers, G.: Accounting for model error in the Météo-France ensemble data assimilation system, *Q. J. Roy. Meteorol. Soc.*, 138, 249–262, 2012.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective-scale operational model, *Mon. Weather Rev.*, 139, 976–991, 2011.
- Tavolato, C. and Isaksen, L.: On the use of a Huber norm for observation quality control in the ECMWF 4D-Var, *Q. J. Roy. Meteorol. Soc.*, doi:10.1002/qj.2440, 2014.
- Thode, H. C.: Testing for Normality, in: Vol. 164 of *Statistics: Textbooks and Monographs*, Marcel Dekker, New York, p. 368, 2002.
- Wattrelot, E. and Caumont, O. and Mahfouf, J.-F.: Operational implementation of the 1D+ 3D-Var assimilation method of radar reflectivity data in the AROME model, *Mon. Weather Rev.*, 142, 1852–1873, 2014.

4 | Correction des erreurs de déplacement en PNT

L'objet de ce chapitre est de fournir un cadre variationnel pour la correction des erreurs de déplacement. Une revue des différentes méthodes utilisées en PNT est exposée à cette occasion. Le parti-pris est ici de pousser le plus loin possible l'analogie avec les méthodes variationnelles pour la correction d'erreurs d'amplitude exposées au chapitre 1, pour proposer une méthode d'estimation quasi-linéaire quasi-gaussienne des erreurs de déplacement.

Sommaire

4.1	Introduction	77
4.2	Méthodes d'estimation existantes en PNT	80
4.3	Approche Bayésienne	83
4.3.1	Cadre Bayésien total	83
4.3.2	Algorithme à deux étapes (Ravela et al., 2007)	83
4.3.3	Hypothèse Gaussienne	84
4.4	Estimation quasi-linéaire quasi-Gaussienne	85
4.4.1	Algorithme de Gauss-Newton	85
4.4.2	Rôle du gradient dans \mathcal{J}	86
4.4.3	Préconditionnement	87
4.5	Description technique de la méthode implémentée	87
4.5.1	Généralités	87
4.5.2	Déroulement et contrôle de la minimisation	91
4.6	Test sur des cas idéalisés	95
4.6.1	Translation d'un champ idéalisé	95
4.6.2	Translation d'un champ réel AROME-France	95
4.7	Application à un cas réel AROME-WMED	101
4.7.1	Situation météorologique étudiée : cas de Médicane	101
4.7.2	Estimation des erreurs de déplacement	101
4.8	Conclusions	109

4.1. Introduction

Le problème d'estimation des erreurs de déplacement est connu en traitement d'image sous le terme d'estimation du flot optique. Ce dernier est plus généralement défini comme la comparaison de deux représentations d'un même phénomène, en déterminant un champ de déplacement permettant de passer d'une représentation à l'autre. L'estimation du champ de déplacement est uniquement basée sur les variations d'amplitude entre les points colocalisés composant les représentations (e.g. luminance ou intensité des pixels pour une photo, valeurs aux points de grille pour un champ modèle). Toutes les méthodes d'estimation du flot optique se basent sur la conservation d'un invariant géométrique ou photométrique (Mémin, 2003). Seules les méthodes basées sur un invariant photométrique sont présentées dans ce chapitre. En plus de permettre une estimation dense du champ de déplacement, elles ont l'avantage d'éviter l'étape de détection des "cibles" géométriques à suivre d'une représentation à l'autre. Par exemple, la conservation d'un invariant photométrique F entre deux représentations successives aux temps

t et $t + \Delta t$, notées \mathbf{I}_t et $\mathbf{I}_{t+\Delta t}$, s'exprime telle que :

$$F(\mathbf{I}_{t,[s+\mathbf{d}]}) = F(\mathbf{I}_{t+\Delta t}) \quad (4.1)$$

avec $\mathbf{I}_{t,[s+\mathbf{d}]}$, le déformation du champ d'intensité lumineuse \mathbf{I}_t alignée à $\mathbf{I}_{t+\Delta t}$ grâce au champ de déplacement $\mathbf{d} = (\mathbf{d}^x, \mathbf{d}^y)^T$ (déplacement bidimensionnelle dans les deux directions de l'espace image, x et y). Horn & Schunck (1981) proposent l'utilisation d'une formulation différentielle de l'Eq.4.1 définie par la conservation de la dérivée Lagrangienne de \mathbf{I} . Cette dernière relation peut être formulée comme l'équation d'advection de $F(\mathbf{I})$ par un déplacement \mathbf{d} telle que :

$$\frac{\partial F(\mathbf{I})}{\partial t} + \nabla F(\mathbf{I})\mathbf{d} = 0$$

Sous forme matricielle $\nabla F(\mathbf{I})$ est de taille $N \times 2N$ (avec N la dimension de $F(\mathbf{I})$). Il est composé de la concaténation de deux matrices diagonales de taille $N \times N$, dont chacune des diagonales est définie comme les gradients de $F(\mathbf{I})$ dans les deux directions de l'espace : $\nabla_x F(\mathbf{I})$ et $\nabla_y F(\mathbf{I})$. La formulation différentielle est linéaire par rapport à \mathbf{d} mais a le désavantage de n'être valide que pour des faibles déplacements (Mémin, 2003).

En PNT, un premier cas d'application de l'estimation du flot optique est d'utiliser une succession d'images satellites afin de suivre les mouvements des structures nuageuses et en déduire une mesure assimilable du vent (AMV, e.g Leese et al., 1971; Nieman et al., 1997; Holmlund, 1998; Bormann et al., 2012). Les AMVs sont déduites du mouvement de cibles atmosphériques telles que les nuages ou les structures cohérentes de vapeur d'eau, avec une méthode utilisant les corrélations croisées entre plusieurs images satellites successives. L'avantage majeur des AMVs est de permettre l'observation dense d'une variable pronostique, le vent, dans des zones dénuées d'observations conventionnelles, telles qu'au dessus des océans. Pour être assimilable par le modèle, la position verticale des vents estimés doit être fixée. La méthode la plus simple consiste à comparer la température de brillance observées et les températures prévues par le modèle pour en déduire l'altitude des nuages opaques observées (méthode IRW, Nieman et al., 1993). La détermination de l'altitude est une source importante d'erreurs pour ces observations, avec des incertitudes généralement comprises entre 70 hPa et 110 hPa (pour une autre méthode dite "d'interception de l'eau", Lee & Salonen, 2015).

La correction des erreurs de déplacement est un deuxième cas d'application où cette fois les mouvements sont estimés dans l'espace du modèle, puisque le but est d'aligner l'état de l'atmosphère prévu \mathbf{x}_b avec l'état de l'atmosphère observé \mathbf{y}_o . Contrairement au cas précédent, les représentations à comparer (\mathbf{y}_o et \mathbf{x}_b) ne sont pas de même nature (ils peuvent par exemple être des représentations valides au même instant ou décalées dans le temps). Mais comme cela est introduit dans le chapitre 1, la comparaison est rendue possible grâce à l'opérateur d'observation \mathcal{H} . La conservation d'un invariant photométrique F entre les observations \mathbf{y}_o et l'ébauche déplacée par \mathbf{d} , $\mathbf{x}_{b,[s+\mathbf{d}]}$, s'exprime telle que

$$F(\mathbf{y}_o) = F\left(\mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]})\right) \quad (4.2)$$

Cette relation est rendue non linéaire par rapport à \mathbf{d} , par la non linéarité de \mathcal{H} et celle de la fonction $\mathbf{d} \mapsto \mathbf{x}_{b,[s+\mathbf{d}]}$. L'invariant photométrique classiquement utilisé en PNT est l'amplitude en chaque point d'observation (pour une photo en noir et blanc, ce serait le niveau de gris de

chaque pixel). L'Eq.4.2 devient :

$$\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b, [\mathbf{s}+\mathbf{d}]}) = 0 \quad (4.3)$$

Cette relation est appelée par la suite conservation de la luminance (CL) en référence aux méthodes d'estimation du flot optique utilisées en traitements d'images.

La résolution du système donné par la CL ne suffit pas pour une estimation dense des erreurs de déplacement. En effet, le problème est mal posé dans le sens où il ne vérifie pas les trois conditions données par Hadamard (1902) (Souopgui et al., 2012) : la solution existe, qu'elle soit unique et qu'elle soit stable (de petites variations dans les conditions initiales ou dans les termes de sources n'entraînent que de petites variations de la solution, Maz'ya & Shaposhnikova, 2012). Notamment l'hypothèse d'unicité de la solution est transgressée par un nombre d'observations bien inférieur au nombre de degrés de liberté du problème (taille de la variable de contrôle \mathbf{d}). En plus du fait que seulement certaines variables et certains points de grille (pdg) sont observés, cette différence est d'autant plus marquée pour les erreurs de déplacement. En effet pour chaque variable et pour chaque point (i, j, k) d'une grille tridimensionnelle, positionné par $\mathbf{s}(i, j, k)$, il est envisageable d'associer un déplacement dans les trois dimensions de l'espace, $\mathbf{d}(i, j, k) = (\mathbf{d}(i, j, k)^x, \mathbf{d}(i, j, k)^y, \mathbf{d}(i, j, k)^z)$. La taille de la variable de contrôle \mathbf{d} est alors trois fois plus grande que celle de l'ébauche. Pour palier à ce problème, une contrainte de régularisation de la forme du déplacement est ajoutée. On pourra citer quatre exemples classiques de contrainte, où l'on cherche à minimiser un ou plusieurs des quatre paramètres suivant :

- la norme de \mathbf{d}
- la norme du gradient de \mathbf{d}
- la norme de la divergence et/ou du rotationnel de \mathbf{d}
- la norme du gradient de la divergence et/ou du rotationnel de \mathbf{d} .

L'utilisation de contraintes de régularité a également l'avantage de rendre la solution plus robuste, puisqu'elle doit permettre de prendre en compte les erreurs d'observation, et de relaxer l'hypothèse de CL qui n'est pas vérifiée dans toutes les situations (Mémin, 2003). Ce dernier point apparaît par exemple dans les cas de "non-prévision" où la structure observée n'est pas présente dans l'ébauche ou qui a une forme trop différente.

La section suivante présente une revue des différentes méthodes pour estimer les erreurs de déplacement en PNT. Au moins trois critères permettent de les différencier entre elles :

1. le type de contrainte appliquée au déplacement. Notamment si cette contrainte est forte ou faible (contrainte par pénalisation), et si elle affecte le déplacement lui-même ou ses dérivées successives. De plus, le choix de la contrainte ayant une influence sur les propriétés du déplacement estimé, l'objectivité de la contrainte est aussi à prendre en compte. Moins la contrainte sera caractérisée par des choix *a priori*, plus l'algorithme sera objectif et sera capable d'estimer une large variété d'erreurs de déplacement.
2. l'utilisation de l'équation de conservation 4.3, ou d'une version linéarisée. Cette distinction a une importance car dans le deuxième cas, la contrainte du mouvement apparent ne reste valable que pour de faibles déplacements autour du point de linéarisation.
3. la méthode numérique utilisée pour inverser le système exprimant la CL et la contrainte du déplacement. L'utilisation de techniques multi-échelles est notamment à prendre en compte.

Les critères de différenciation sur la nature des observations utilisées ainsi que leurs répartitions spatiales ne sont pas explicités. La raison est que, pour cette thèse, aucune observation réelle n'a été utilisée pour aligner l'ébauche. Tous les résultats sont montrés pour le cadre idéal où les observations sont en fait un vecteur d'état du modèle (i.e où \mathcal{H} est l'identité).

4.2. Méthodes d'estimation existantes en PNT

Cette section expose les principaux algorithmes de corrections d'erreurs de déplacement en PNT. Ce sont tous des algorithmes qui cherchent à vérifier l'Eq.4.3 de la CL, en minimisant une fonction coût \mathcal{J} . Dans cette section, seule l'estimation de \mathbf{d} est exposée. La résolution conjointe du problème de correction d'erreurs d'amplitude avec celui de correction des erreurs de déplacement est discutée dans la section 4.3.

Brewster (2003a) et Brewster (2003b)

Brewster (2003a) et Brewster (2003b) étudient l'impact de la correction des erreurs de déplacement (également appelées erreurs de phase ou erreurs position) sur un cas de convection profonde associée à la formation de plusieurs tornades. L'estimation de ces erreurs est faite grâce à l'assimilation de réflectivités (dBz) venant de radars météorologiques.

À notre connaissance, seule cette étude utilise une contrainte multiplicative, où \mathcal{J} est définie par

$$\mathcal{J} = \underbrace{\frac{\exp(|\mathbf{d}|l^{-1})}{(1 + |\mathbf{d}|l^{-1})}}_{\text{contrainte sur } \mathbf{d}} \underbrace{\|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]})\|_{\mathbf{R}}^2}_{\text{rappel aux observations}}$$

où \mathbf{R} est la matrice diagonale de covariance d'erreurs d'observation (qui ne contient donc que les variances d'erreurs d'observation), et l est une longueur caractéristique définie comme la demi-diagonale des sous-domaines définis plus bas. La contrainte multiplicative sur \mathbf{d} , qui utilise l'inverse de la fonction auto-régressive du second ordre (SOAR) de manière analogue à Thiebaut et al. (1990), permet d'éviter tout aliasing en rajoutant une contrainte faible sur la norme du déplacement. Cet algorithme utilise également une contrainte forte sur \mathbf{d} en imposant un déplacement constant pour chacun des pdg d'un même sous-domaine (translation locale). Le champ de déplacement total est obtenu par lissage du "patchwork" des déplacements restitués sur tous les sous-domaines. Ce type de contrainte permet de réduire considérablement le nombre de degré de liberté à estimer et d'utiliser des algorithmes de recherche tels que "la force brute", où, pour chaque sous domaine, la solution $\hat{\mathbf{d}}$ est obtenue en calculant \mathcal{J} pour toutes les valeurs de déplacement possibles (dans les limites du raisonnable). Pour éviter d'aboutir sur un minimum local de la fonction coût \mathcal{J} , une approche multi-échelles est développée, le même algorithme étant appliqué pour un découpage en sous-zones de plus en plus petites.

Les contraintes faibles additives sont toutefois plus couramment utilisées en PNT, comme cela est fait dans les méthodes exposées maintenant.

Beezley & Mandel (2008)

Dans cette étude estimant les erreurs de déplacement pour l'amélioration de la prévision des feux de forêts, la contrainte pénalise la norme du déplacement et de son gradient $\nabla \mathbf{d}$. \mathcal{J} est définie telle que :

$$\mathcal{J} = C_1 \|\mathbf{d}\| + C_2 \|\nabla \mathbf{d}\| + |\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]})|$$

où C_1 et C_2 sont deux paramètres à fixer *a priori* afin de donner plus ou moins d'importance à l'une ou l'autre contrainte. Avec $\mathbf{d}(i, j) = (\mathbf{d}(i, j)^x, \mathbf{d}(i, j)^y)$ le déplacement bidimensionnel,

en chaque point de grille (i, j) du domaine, $\|\mathbf{d}\|$ et $\|\nabla\mathbf{d}\|$ sont définis tels que :

$$\begin{aligned}\|\mathbf{d}\| &= \sum_{i,j} (|\mathbf{d}(i, j)^x| + |\mathbf{d}(i, j)^y|) \\ \|\nabla\mathbf{d}\| &= \sum_{i,j} \left(\left| \frac{\partial\mathbf{d}(i, j)^x}{\partial x} \right| + \left| \frac{\partial\mathbf{d}(i, j)^x}{\partial y} \right| + \left| \frac{\partial\mathbf{d}(i, j)^y}{\partial x} \right| + \left| \frac{\partial\mathbf{d}(i, j)^y}{\partial y} \right| \right).\end{aligned}$$

L'algorithme utilise une méthode itérative de descente d'échelle, avec à chaque itération une minimisation non-linéaire uni-dimensionnelle, dans chacune des directions horizontales. Un filtrage implicite des petites longueurs d'ondes du déplacement est effectué grâce au lissage préalable des champs à aligner. Une des idées importantes de cette étude est non seulement la Gaussianisation de l'ébauche suite au déplacement, mais également l'utilisation d'un ENKF pour analyser les erreurs de déplacement en parallèle des erreurs d'amplitude. Ceci permet de maintenir un ensemble d'assimilation perturbé en amplitude mais également en déplacement.

Pour cette méthode (ainsi que pour celle de Brewster, 2003a), les formulations de la contrainte de régularisation manquent d'objectivité puisque le fait de pénaliser la norme du déplacement ou son gradient est un choix que l'on fait *a priori* et dont on ne connaît pas l'efficacité. Il faut ajouter à ceci la paramétrisation empirique de C_1 et C_2 .

Ravela et al. (2007) et Ravela (2012)

D'autres choix de formulation *a priori* de la contrainte sont possibles. Ravela et al. (2007) proposent la minimisation d'une fonction coût \mathcal{J} utilisant une pénalisation du gradient et de la divergence du déplacement telle que

$$\mathcal{J} = \underbrace{\frac{w_1}{2}\|\nabla\mathbf{d}\|^2 + \frac{w_2}{2}\|\nabla \cdot \mathbf{d}\|^2}_{L(\mathbf{d})} + \|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b, [\mathbf{s}+\mathbf{d}]})\|_{\mathbf{R}}^2 \quad (4.4)$$

\mathbf{R} est la matrice de covariance d'erreurs d'observation et w_1 et w_2 sont deux paramètres à fixer *a priori*. De plus, avec $\mathbf{d}(i, j) = (\mathbf{d}(i, j)^x, \mathbf{d}(i, j)^y)$ le déplacement bidimensionnel en chaque point de grille (i, j) du domaine, $\|\nabla\mathbf{d}\|^2$ et $\|\nabla \cdot \mathbf{d}\|^2$ sont définis tels que

$$\|\nabla\mathbf{d}\|^2 = \sum_{i,j} \left(\left| \frac{\partial\mathbf{d}(i, j)^x}{\partial x} \right|^2 + \left| \frac{\partial\mathbf{d}(i, j)^y}{\partial y} \right|^2 \right) \text{ et } \|\nabla \cdot \mathbf{d}\|^2 = \sum_{i,j} \left(\left| \frac{\partial\mathbf{d}(i, j)^x}{\partial x} + \frac{\partial\mathbf{d}(i, j)^y}{\partial y} \right|^2 \right)$$

Il faut également citer les travaux de Hoffman et al. (1995), Hoffman & Grassotti (1996), Grassotti et al. (1999) (notamment repris par Nehrkorn et al., 2014) qui utilisent également des termes de contraintes *a priori*. Ces travaux ont en plus la particularité d'effectuer la minimisation dans l'espace de Fourier, afin de ne pas prendre en compte les petites échelles.

L'étude proposée par Ravela et al. (2007) est remarquable pour deux raisons. La première est l'utilisation d'un algorithme itératif pour résoudre le problème de minimisation de \mathcal{J} . En effet $\hat{\mathbf{d}}$ est supposé être le point fixe de la suite (\mathbf{d}_k) , définie par :

$$\frac{\partial L}{\partial \mathbf{d}}(\mathbf{d}_k) + \nabla \mathbf{x}_{b, [\mathbf{s}+\mathbf{d}_{k-1}]}^T \mathbf{H}^T \mathbf{R} (\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b, [\mathbf{s}+\mathbf{d}_{k-1}]})) = 0$$

avec L une fonction définie dans l'Eq.4.4. On rappelle que sous forme matricielle $\nabla \mathbf{x}_{b, [\mathbf{s}+\mathbf{d}_k]}$ est de

taille $N \times 2N$ (avec N la dimension de \mathbf{x}_b à déplacer). Il est composé de la concaténation de deux matrices diagonales de taille $N \times N$, dont chacune des diagonales est définie comme le gradient de l'ébauche déplacée dans une des deux directions de l'espace : $\nabla_x \mathbf{x}_{b,[s+\mathbf{d}_k]}$ et $\nabla_y \mathbf{x}_{b,[s+\mathbf{d}_k]}$. Le deuxième point important abordé par Ravela et al. (2007) est la dérivation Bayésienne complète du problème d'estimation des erreurs d'ébauche (amplitude et déplacement), avec notamment la proposition d'un algorithme à deux étapes pour la correction de $\hat{\mathbf{d}}$ au préalable de l'erreur d'amplitude $\hat{\epsilon}$. Ce point est repris dans la section 4.3.2.

Nehrkorn et al. (2003)

Le dernier type de contrainte présenté est celui qui se rapproche le plus du terme de rappel à l'ébauche J_b dans un 3D-Var par exemple. La fonction coût associée, \mathcal{J} , peut être définie telle que (formulation inspirée de Souopgui et al., 2012)

$$\mathcal{J} = \|\Phi(\mathbf{d})\|_{\mathbf{D}}^2 + \frac{1}{2} \|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]})\|_{\mathbf{R}}^2$$

avec Φ une fonction de \mathbf{d} ou de ses dérivées successives, et \mathbf{D} la matrice de covariance des erreurs de déplacement. Nehrkorn et al. (2003) utilisent une pénalisation du déplacement lui-même (Φ est l'identité). Ainsi \mathcal{J} est définie (sous sa forme développée) telle que

$$\mathcal{J} = \frac{1}{2} \mathbf{d}^T \mathbf{D}^{-1} \mathbf{d} + \frac{1}{2} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]}) \right)^T \mathbf{R}^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]}) \right) \quad (4.5)$$

Pour Nehrkorn et al. (2003), la pénalisation est appliquée aux coefficients de Fourier des champs de déplacement, et \mathbf{D} est alors modélisée diagonale dans l'espace spectral. L'algorithme de minimisation utilisé est documenté par Dennis et al. (1981). La pénalisation de type $\mathbf{d}^T \mathbf{D}^{-1} \mathbf{d}$ semble être une contrainte plus générale, permettant de remplacer certaines contraintes *a priori* qui pénalisent spécifiquement le gradient, la divergence, etc... En effet, dans au moins deux cas particuliers, il est possible d'associer $\mathbf{d}^T \mathbf{D}^{-1} \mathbf{d}$ à des contraintes *a priori* :

- si $\mathbf{D} = \mathbf{I}$, la pénalisation s'applique sur la norme euclidienne du déplacement
- si les fonctions de covariance utilisées pour définir \mathbf{D} sont de type exponentiel, c'est à dire si $\mathbf{D}(x, x') = \exp\left(-\frac{x-x'}{\sigma}\right)$ pour deux points x et x' d'un domaine unidimensionnel, alors $\mathbf{d}^T \mathbf{D}^{-1} \mathbf{d}$ est une combinaison linéaire de la norme euclidienne de \mathbf{d} et de ses dérivées spatiales du premier ordre (Béréziat & Herlin, 2008).

Un des points essentiels de Nehrkorn et al. (2003) est la volonté d'estimer objectivement les variances spectrales d'erreurs de déplacement par une méthode itérative.

Nehrkorn et al. (2015)

Dans cette étude, l'apport de la correction des erreurs de déplacement est appliquée au cas du cyclone Katrina du mois d'août 2005.

De la même manière que Nehrkorn et al. (2003), Nehrkorn et al. (2015) utilisent également une forme de contrainte basée sur une matrice de covariance \mathbf{D} . Le parti-pris a été de réutiliser le système WRFDA dédié à la correction d'erreurs d'amplitude, et d'utiliser pour les erreurs de déplacement, le même terme de contrainte que celui pour le vent. \mathbf{D} est alors fixée égale à la matrice \mathbf{B} pour le vent (aux unités près).

La méthode choisie durant cette thèse et définie dans la suite de ce chapitre, est également basée sur la formulation de \mathcal{J} donnée ici par l'Eq.4.5, avec un terme de contrainte caractérisé

par une matrice de covariance \mathbf{D} dont les fonctions de structure sont modélisées dans l'espace points de grille avec une forme quasi-Gaussienne et qui sont paramétrisées par des variances et des longueurs de corrélation d'erreurs de déplacement. Le mode de résolution est proche de celui de Nehrkorn et al. (2015), puisque les erreurs de déplacement seront estimés par une méthode variationnelle qui est quasi-linéaire et quasi-Gaussienne.

4.3. Approche Bayésienne

Cette section commence par dériver la formulation Bayésienne complète de l'étape d'assimilation, c'est à dire en considérant conjointement la correction des erreurs d'amplitude et de déplacement. Pour des raisons pratiques, cette formulation n'est cependant pas utilisée, au profit d'un algorithme à deux étapes où les erreurs de déplacement sont corrigées en amont des erreurs d'amplitude, toujours en utilisant une approche Bayésienne de type maximum de vraisemblance (Ravela et al., 2007). Cette dernière approche permet d'exprimer la Bayésianité de la formulation de l'Eq.4.5 lorsque l'hypothèse de Gaussianité des erreurs de déplacement est admise.

4.3.1. Cadre Bayésien total

On se place dans le cadre où le modèle d'erreur de l'ébauche est mixte. L'état d'analyse est estimé parmi les \mathbf{x} définis par :

$$\mathbf{x} = \mathbf{x}_{b,[s+d]} + \boldsymbol{\epsilon}$$

avec $\boldsymbol{\epsilon}$ les erreurs additives d'amplitude, \mathbf{d} les erreurs de déplacement, et $\mathbf{x}_{b,[s+d]}$ l'ébauche déplacée par \mathbf{d} . Avec une approche de type maximum de vraisemblance, le couple $(\hat{\boldsymbol{\epsilon}}, \hat{\mathbf{d}})$, solution du problème d'analyse, est défini comme le mode de $\mathcal{P}(\boldsymbol{\epsilon}, \mathbf{d} | \mathbf{y}_o)$. En suivant la règle de Bayes, cette dernière PDF peut être reformulée telle que :

$$\mathcal{P}(\boldsymbol{\epsilon}, \mathbf{d} | \mathbf{y}_o) \propto \mathcal{P}(\mathbf{y}_o | \boldsymbol{\epsilon}, \mathbf{d}) \times \mathcal{P}(\boldsymbol{\epsilon} | \mathbf{d}) \times \mathcal{P}(\mathbf{d})$$

où la forme de $\mathcal{P}(\boldsymbol{\epsilon} | \mathbf{d})$ dépend du déplacement \mathbf{d} (Ravela et al., 2007). Par exemple, si l'on fait l'hypothèse Gaussienne des erreurs additives d'amplitude, la distribution de $\boldsymbol{\epsilon}$ est donnée par :

$$\mathcal{P}(\boldsymbol{\epsilon} | \mathbf{d}) \propto \exp\left(-\boldsymbol{\epsilon}^T \mathbf{B}(\mathbf{d})^{-1} \boldsymbol{\epsilon}\right).$$

avec $\mathbf{B}(\mathbf{d})$ la matrice de covariance d'erreurs d'amplitude de l'ébauche qui dépend de \mathbf{d} . Ainsi l'estimation simultanée de \mathbf{d} et $\boldsymbol{\epsilon}$ est très coûteuse (e.g Aonashi & Eito, 2011), puisqu'elle nécessite de modifier $\mathbf{B}(\mathbf{d})$ pour chaque valeur de \mathbf{d} rencontrée. C'est pourquoi Ravela et al. (2007) proposent un algorithme "à deux étapes" qui estime \mathbf{d} et $\boldsymbol{\epsilon}$ en deux étapes distinctes et successives.

4.3.2. Algorithme à deux étapes (Ravela et al., 2007)

Actuellement, il y a un consensus entre toutes les études de PNT corrigeant les erreurs de déplacement, pour utiliser l'algorithme à deux étapes proposé par Ravela et al. (2007). Celui-ci est décomposé en une première phase de correction des erreurs de déplacement pour former une ébauche "alignée". Puis lors de la deuxième phase, c'est l'erreur d'amplitude de l'ébauche qui est corrigée à partir de l'ébauche alignée. En plus de permettre de ne spécifier qu'une seule fois la PDF d'erreurs d'amplitude d'ébauche $\mathcal{P}(\boldsymbol{\epsilon} | \mathbf{d})$ en fonction de $\hat{\mathbf{d}}$, l'intérêt de cet algorithme est de préparer l'étape de correction d'erreurs d'amplitude pour qu'elle soit de meilleure qualité. En effet, l'hypothèse Gaussienne habituellement utilisée est supposée avoir

une plus grande validité avec une ébauche alignée (Lawson & Hansen, 2005). De plus, on peut s'attendre à une plus grande proximité entre les observations et leurs équivalents simulés à partir d'ébauches alignées. Ceci implique également une plus grande Gaussianité des innovations (et donc d'un plus grand nombre d'observations passant le contrôle de qualité).

L'étape de correction des erreurs d'amplitude est présentée à la section 1.3, en utilisant l'ébauche alignée $\mathbf{x}_{b,[s+\hat{\mathbf{d}}]}$ au lieu de \mathbf{x}_b . $\hat{\mathbf{d}}$ est la solution de l'étape préliminaire d'estimation des erreurs de déplacement. Avec une approche de type maximum de vraisemblance, il est le mode de $\mathcal{P}(\mathbf{d}|\mathbf{y}_o)$. Avec les règles de Bayes, cette dernière PDF est décomposée telle que :

$$\mathcal{P}(\mathbf{d}|\mathbf{y}_o) \propto \mathcal{P}(\mathbf{y}_o|\mathbf{d}) \times \mathcal{P}(\mathbf{d})$$

L'estimation de $\hat{\mathbf{d}}$ revient alors à minimiser la fonction coût $\mathcal{J} = -\ln(\mathcal{P}(\mathbf{d}|\mathbf{y}_o))$.

4.3.3. Hypothèse Gaussienne

Terme de rappel aux observations

En respectant la CL, pour contraindre la norme de $\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]})$, la PDF d'erreurs d'observation sachant \mathbf{d} est modélisée par une distribution Gaussienne telle que :

$$\mathcal{P}(\mathbf{y}_o|\mathbf{d}) \propto \exp\left(-\frac{\|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]})\|_{\mathbf{R}}^2}{2}\right)$$

avec \mathbf{R} la matrice de covariance d'erreurs d'observation. Cette formulation implique qu'en plus d'être Gaussiennes, les erreurs d'observation sont supposées être non biaisées, et indépendantes des erreurs de déplacement \mathbf{d} .

Terme de contrainte du déplacement

De même il est possible de faire l'hypothèse Gaussienne pour les erreurs de déplacement (Nehrkorn et al., 2003, 2015). $\mathcal{P}(\mathbf{d})$ est alors définie selon une matrice de covariance d'erreurs de déplacement \mathbf{D} telle que

$$\mathcal{P}(\mathbf{d}) \propto \exp\left(-\frac{\|\mathbf{d}\|_{\mathbf{D}}^2}{2}\right)$$

Cette formulation indique également que les erreurs de déplacement estimées sont non biaisées. La validité de cette hypothèse reste à vérifier, néanmoins elle permet d'éviter l'utilisation d'une ébauche d'erreurs de déplacement qui est difficile à estimer dans la pratique.

Ainsi une approche de type maximum de vraisemblance associée à une hypothèse Gaussienne pour les erreurs d'observation et de déplacement, est équivalente à minimiser la fonction coût définie par $\mathcal{J} = \left(-\ln(\mathcal{P}(\mathbf{d}|\mathbf{y}_o))\right)$, donnée par l'Eq.4.5. De même que pour les méthodes variationnelles habituelles (e.g 3D-Var, 4D-Var) c'est un problème de type moindres carrés avec une fonction coût à minimiser qui est non-quadratique. Avec la même idée que Nehrkorn et al. (2015), l'analogie avec la correction des erreurs d'amplitude est poussée jusqu'à la résolution du problème d'optimisation. En effet, comme nous le verrons ci-dessous, une méthode de résolution quasi-linéaire de type Gauss-Newton est proposée pour estimer $\hat{\mathbf{d}}$.

4.4. Estimation quasi-linéaire quasi-Gaussienne

En plus de l'opérateur d'observation \mathcal{H} , la fonction coût \mathcal{J} donnée par l'Eq.4.5 a une deuxième source de non-linéarité : la fonction $\mathbf{d} \mapsto \mathbf{x}_{b,[s+\mathbf{d}]}$. Malgré cet éloignement supplémentaire à la quadraticité par rapport au problème d'estimation des erreurs d'amplitude, la même méthode quasi-linéaire de type Gauss-Newton est testée dans cette thèse pour l'estimation des erreurs de déplacement. Pour accroître la robustesse de la minimisation, un critère d'arrêt adaptatif testant la validité des modèles linéaires successifs a été implémenté.

4.4.1. Algorithme de Gauss-Newton

L'algorithme utilisé dans cette section est le même que celui de la section 1.3.3, mais appliqué aux déplacements.

Pour une estimation bidimensionnelle du déplacement (l'extension au tridimensionnel est aisée) les notations suivantes sont utilisées. Le déplacement \mathbf{d} est la concaténation du déplacement dans deux directions α, β , de l'espace pour tous les pdg, tel que $\mathbf{d} = (\mathbf{d}^\alpha, \mathbf{d}^\beta)$. La matrice de covariance d'erreurs de déplacement est définie en fonction du "vrai" déplacement \mathbf{d}_t telle que :

$$\mathbf{D} = E \left[(\mathbf{d} - \mathbf{d}_t) (\mathbf{d} - \mathbf{d}_t)^T \right] \quad (4.6)$$

$$= \begin{pmatrix} \mathbf{D}_{\alpha,\alpha} & \mathbf{D}_{\alpha,\beta} \\ \mathbf{D}_{\beta,\alpha} & \mathbf{D}_{\beta,\beta} \end{pmatrix} \quad (4.7)$$

avec $\mathbf{D}_{\alpha,\beta} = E \left[(\mathbf{d}^\alpha - \mathbf{d}_t^\alpha) (\mathbf{d}^\beta - \mathbf{d}_t^\beta)^T \right]$. L'algorithme de type Gauss-Newton est alors défini par

Algorithme de Gauss-Newton

1. $\mathbf{d}_0 = \mathbf{d}_{init}$ (initialisation)
 2. à faire jusqu'à la convergence (boucle externe sur k)
 - a) résoudre $\mathbf{F}(\mathbf{d}_k)^T \mathbf{F}(\mathbf{d}_k) \widehat{\delta \mathbf{d}}_k = -\mathbf{F}(\mathbf{d}_k)^T f(\mathbf{d}_k)$ (boucle interne)
 - b) $\mathbf{d}_{k+1} = \mathbf{d}_k + \widehat{\delta \mathbf{d}}_k$ (actualisation)
 3. $\widehat{\mathbf{d}} = \mathbf{d}_{k+1}$
-

avec

$$f(\mathbf{d}) = \begin{pmatrix} \mathbf{D}^{-\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \mathbf{d} \\ \mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}]}) \end{pmatrix}. \quad (4.8)$$

et \mathbf{F} , le jacobien de f (dérivation par rapport à \mathbf{d}), est alors défini par

$$\mathbf{F}(\mathbf{d}) = \begin{pmatrix} \mathbf{D}^{-\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \mathbf{I} \\ \mathbf{H} \nabla \mathbf{x}_{b,[s+\mathbf{d}]} \end{pmatrix} \quad (4.9)$$

\mathbf{I} est la matrice identité, et \mathbf{H} l'opérateur tangent linéaire de \mathcal{H} . On rappelle que sous forme matricielle $\nabla \mathbf{x}_{b,[s+\mathbf{d}]}$ est de taille $N \times 2N$ (avec N la dimension de \mathbf{x}). Il est composé de la

concaténation de deux matrices diagonales de taille $N \times N$ dont chacune des diagonales est définie comme le gradient de l'ébauche déplacée dans une des deux directions de l'espace : $\nabla_x \mathbf{x}_{b,[s+\mathbf{d}]}$ et $\nabla_y \mathbf{x}_{b,[s+\mathbf{d}]}$. Enfin il a été choisi d'initialiser l'algorithme avec un déplacement nul, $\mathbf{d}_{ini} = \mathbf{0}$.

Pour former \mathbf{F} à partir de f , \mathcal{H} est linéarisé selon :

$$\mathcal{H}(\mathbf{x} + \delta\mathbf{x}) \approx \mathcal{H}(\mathbf{x}) + \mathbf{H}\delta\mathbf{x}$$

et la fonction $\mathbf{d} \mapsto \mathbf{x}_{b,[s+\mathbf{d}]}$ est également linéarisée par un développement de Taylor à l'ordre 1 tel que :

$$\mathbf{x}_{[s+\delta\mathbf{d}]} = \mathbf{x} + \nabla\mathbf{x}\delta\mathbf{d} \quad (4.10)$$

en considérant $\delta\mathbf{d}$ comme une faible variation en déplacement. Avec ces deux relations de linéarisation, chaque boucle interne k de l'algorithme de Gauss-Newton est équivalente à la minimisation de la fonction coût quadratique J_k définie comme la linéarisation de \mathcal{J} autour de \mathbf{d}_k par :

$$J_k = \|\mathbf{d}_k + \delta\mathbf{d}_k\|_{\mathbf{D}}^2 + \|\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}_k]}) - \mathbf{H}\nabla\mathbf{x}_{b,[s+\mathbf{d}_k]}\delta\mathbf{d}_k\|_{\mathbf{R}}^2 \quad (4.11)$$

Par analogie avec les équations 1.13 et 1.14 définissant l'incrément en amplitude $\widehat{\delta\mathbf{x}}_k$, la forme analytique de son unique minimum $\widehat{\delta\mathbf{d}}_k$ est donnée par :

$$\widehat{\delta\mathbf{d}}_k = \left(\mathbf{D}^{-1} + \nabla_k^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \nabla_k\right)^{-1} \left[\nabla_k^T \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}_k]})\right) - \mathbf{D}^{-1} \mathbf{d}_k\right] \quad (4.12)$$

$$= \mathbf{D} \nabla_k^T \mathbf{H}^T \left(\mathbf{H} \nabla_k \mathbf{D} \nabla_k^T \mathbf{H}^T + \mathbf{R}\right)^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}_k]}) - \mathbf{R} \mathbf{H} \nabla_k \mathbf{D}^{-1} \mathbf{d}_k\right) \quad (4.13)$$

Deux formulations équivalentes sont proposées en utilisant Sherman & Morrison (1949) et Sherman & Morrison (1950). Une interprétation possible de l'opérateur gradient ∇_k (notation raccourcie de $\nabla\mathbf{x}_{b,[s+\mathbf{d}_k]}$) est d'être un opérateur de projection d'un vecteur de déplacement dans l'espace modèle. Il est possible de pousser l'analogie plus loin en assimilant $\mathbf{H}\nabla_k$ à un opérateur d'observation projetant la variable de contrôle \mathbf{d} dans l'espace des observations.

4.4.2. Rôle du gradient dans \mathcal{J}

Avec la même démarche que dans la section 1.4.1, le rôle des différents termes de \mathcal{J} peut être grossièrement décrit dans un cadre totalement linéaire unidimensionnel et univarié, par l'assimilation d'une unique observation positionnée exactement sur le j -ème point de grille. \mathcal{H} est alors exactement donné par la forme linéaire $\mathbf{x} \mapsto \mathbf{x}(j)$, avec \mathbf{x} un vecteur d'état, et $\mathbf{x}(j)$ son j -ème coefficient. De plus \mathbf{R} est ici un scalaire égal à σ_o^2 , la variance des erreurs d'observation. En reprenant l'Eq.4.13 avec $\mathbf{d}_k = \mathbf{0}$, la valeur de l'incrément en déplacement au m -ème point de grille, $\widehat{\delta\mathbf{d}}(m)$, est donnée par

$$\widehat{\delta\mathbf{d}}(m) = \widehat{\mathbf{d}}(m) - \mathbf{d}_{ini}(m) = \frac{\mathbf{D}(m,j)\mathbf{G}(j)}{\sigma_o^2 + \mathbf{D}(j,j)(\mathbf{G}(j))^2} \times \left[\mathbf{y}_o(j) - \mathbf{x}_b(j)\right] \quad (4.14)$$

La différence avec la section 1.4.1 est la présence de \mathbf{G} , le vecteur gradient de l'ébauche \mathbf{x}_b . Pour mieux comprendre son rôle, il peut être remarqué que la confiance que l'on a en

l'ébauche non déplacée est représentée (dans le cas linéaire) par $\frac{1}{\mathbf{D}(j,j)(\mathbf{G}(j))^2} ((\nabla^T \mathbf{D} \nabla)^{-1}$ sous forme matricielle). Plus le gradient est faible, plus cette confiance est forte. Dans ce cas, le déplacement est proche de \mathbf{d}_{ini} ($= \mathbf{0}$ ici). Dans le cas extrême où $\mathbf{G}(j) = 0$, on a $\widehat{\delta \mathbf{d}} = \mathbf{0}$ i.e les observations ne sont pas prises en compte car toute la confiance est mise sur l'ébauche qui reste inchangée. On peut également remarquer que plus un champ est lisse (gradient faible), moins les erreurs de déplacement entraineront des différences importantes en amplitude avec les observations, et plus l'estimation de $\widehat{\mathbf{d}}$ sera délicate. Cette problématique liée à la présence de faibles gradients s'apparente au problème de champs homogènes pour lesquels l'information donnée par deux images ayant un faible contraste est inexploitable. Les autres termes s'interprètent de la même manière que dans la section 1.4.1.

4.4.3. Préconditionnement

Pour la minimisation de J_k définie par l'Eq.4.11, il a été choisi d'utiliser un preconditionnement par \mathbf{D} tel que :

$$\boldsymbol{\lambda} = \mathbf{D}^{-1} \delta \mathbf{d}_k. \quad (4.15)$$

Pour une méthode de résolution itérative de type Gauss-Newton, un tel changement de variable permet d'accélérer la convergence vers la solution du problème. Le système preconditionné est alors donné par

$$\left(\mathbf{I} + \nabla_k^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \nabla_k \mathbf{D} \right) \boldsymbol{\lambda} = \nabla_k \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+\mathbf{d}_k]}) \right) - \mathbf{D}^{-1} \mathbf{d}_k \quad (4.16)$$

Son inversion par une méthode BCG (Gürol et al., 2013) est abordée à la section 4.5.2.

4.5. Description technique de la méthode implémentée

Après avoir formalisé le problème variationnel d'estimation des erreurs de déplacement et dérivé une méthode quasi-linéaire, quasi-Gaussienne pour le résoudre, cette section rassemble tous les choix techniques et simplifications supplémentaires qui ont été faits durant cette thèse.

4.5.1. Généralités

Actuellement, l'algorithme quasi-linéaire quasi-Gaussien d'estimation des erreurs de déplacement est implémenté sous une première version simplifiée. Tout d'abord, on rappelle qu'il utilise en première approche une formulation où l'opérateur d'observation est l'identité : aucune observation réelle n'est utilisée puisque $\widehat{\mathbf{d}}$ est estimé en alignant deux vecteurs d'état dans l'espace modèle. Dans la pratique, avec un EDA, un des membres de l'ensemble est pris comme référence (notée \mathbf{x}_r). Le but est alors d'aligner un autre membre n du même ensemble, $\mathbf{x}_{b,n}$ (noté simplement \mathbf{x}_b par la suite, sauf quand le numéro du membre est utile), avec \mathbf{x}_r . Une limitation de cette simplification est que les erreurs dues à la nature et la localisation des observations, ainsi qu'aux différences de densité du réseau d'observation ne sont pas prises en compte. Ce point devra être abordé dans une étude future.

De plus, en suivant la formulation proposée jusqu'à maintenant dans ce chapitre, les déplacements sont bidimensionnels et sont estimés sur un niveau modèle donné. Mais l'estimation de déplacements verticaux serait une évolution intéressante, notamment pour la prévision des couches nuageuses dont l'altitude peut être entachée d'erreurs (Martinet et al., 2013).

Déplacement de l'ébauche

De manière similaire à Desroziers (1997), $\mathbf{x}_{b,[s+d]}$ est calculé à partir de x_b , sur une grille fixe et régulière, avec un algorithme itératif défini tel que

Algorithme du point fixe

1. $\mathbf{d}_{inv,0} = \mathbf{d}$
 2. à faire jusqu'à la convergence (boucle sur k)
 - a) déformation de la grille régulière telle que la position des pdg est déplacée par $-\mathbf{d}_{inv,k}$
 - b) $\mathbf{d}_{inv,k+1}$ est l'interpolation de \mathbf{d} sur la grille déformée
 3. interpolation de \mathbf{x}_b sur la grille régulière qui aura été déformée par $-\mathbf{d}_{inv,k+1}$
-

Toutes les interpolations utilisées sont linéaires. Ne permettant pas d'avoir une précision meilleure que la résolution de la grille modèle, le critère d'arrêt de la boucle 2 de l'algorithme de point fixe est alors défini par :

$$\max |\mathbf{d}_{inv,k+1} - \mathbf{d}_{inv,k}| < 1 \text{ point de grille}$$

Dans la pratique assez peu d'itérations sont nécessaires pour arriver à la convergence. Dans le cas idéal d'une translation, seule l'initialisation est nécessaire. Sinon dans la plupart des cas, une dizaine d'itérations suffisent à atteindre la convergence.

De la même manière, le critère d'arrêt de la boucle externe de l'algorithme de type GN, décrit à la section 4.4.1, est fixé selon la valeur de l'incrément en déplacement tel que

$$\max |\mathbf{d}_{k+1} - \mathbf{d}_k| < 5 \text{ point de grille} \quad (4.17)$$

Au lieu de 1 point de grille, la valeur limite de 5 pdg a été choisie pour réduire les coûts de calcul. Malgré cet allègement, ce choix est difficilement implémentable en opérationnel puisqu'il ne fixe pas à l'avance le nombre d'itérations dans la boucle externe. Néanmoins il permet de ne pas arrêter l'algorithme avant la convergence et de pouvoir ainsi documenter l'ordre de grandeur du nombre d'itérations nécessaire.

Conditions aux limites

Les conditions aux bords du domaine sont choisies pour respecter deux caractéristiques :

- la bipériodicité de tous les champs, pour permettre l'utilisation de filtres spectraux utilisés à la section 4.5.2.
- tous les déplacements doivent respecter le principe du "rien ne sort, rien ne rentre".

Pour le premier point, étant donné qu'AROME est un modèle spectral utilisant déjà des champs bipériodisés, aucune manipulation supplémentaire n'est faite. On rappelle que la périodisation des champs est rendue possible par l'agrandissement du domaine au Nord et à l'Est avec une zone d'extension (zone E) non physique de quelques pdg. Les bords Est et Ouest ainsi que Nord et Sud sont raccordés par une extrapolation de type spline cubique (Haugen & Machenhauer, 1993; Radnóti, 1995). Bien que le domaine soit bipériodisé, les dérivations spatiales ne

sont pas faites de façon exacte dans l'espace spectral, et sont calculées par différences finies avec un schéma décentré d'ordre 1 (dit "upwind").

Afin de respecter le deuxième point qui est une contrainte forte sur le déplacement, il a été choisi de délimiter un domaine "intérieur" centré sur le domaine total et d'estimer les erreurs de déplacement uniquement sur ce domaine intérieur. Les déplacements sont ensuite linéairement extrapolés dans la zone complémentaire dite "zone tampon" de telle sorte qu'aucun déplacement n'induit une déformation de l'ébauche hors du domaine total. Notamment, le déplacement doit s'annuler sur les pdg les plus excentrés. La Fig.4.1 présente un exemple d'extension dans la zone tampon d'une translation zonale de 10 pdg (dans le domaine intérieur). Analytiquement, le déplacement dans la zone tampon de largeur L_t est exprimé par une extrapolation linéaire entre le déplacement sur le dernier point de grille du domaine intérieur et la valeur nulle sur le dernier point de grille du domaine total. Par exemple dans le cas 1D, pour la zone tampon délimitée par $j = 1..L_t$, $\mathbf{d}(j)$ le déplacement en chaque point j est donné par

$$\mathbf{d}(j) = \frac{\mathbf{d}(L_t + 1)}{L_t} (j - 1)$$

Pour assurer la propriété "rien ne rentre rien ne sort", toujours en utilisant une extrapolation linéaire, la valeur du déplacement maximum autorisé sur les bords du domaine intérieur est égale à la largeur de la zone tampon. L_t est fixée de façon empirique à 50 pdg.

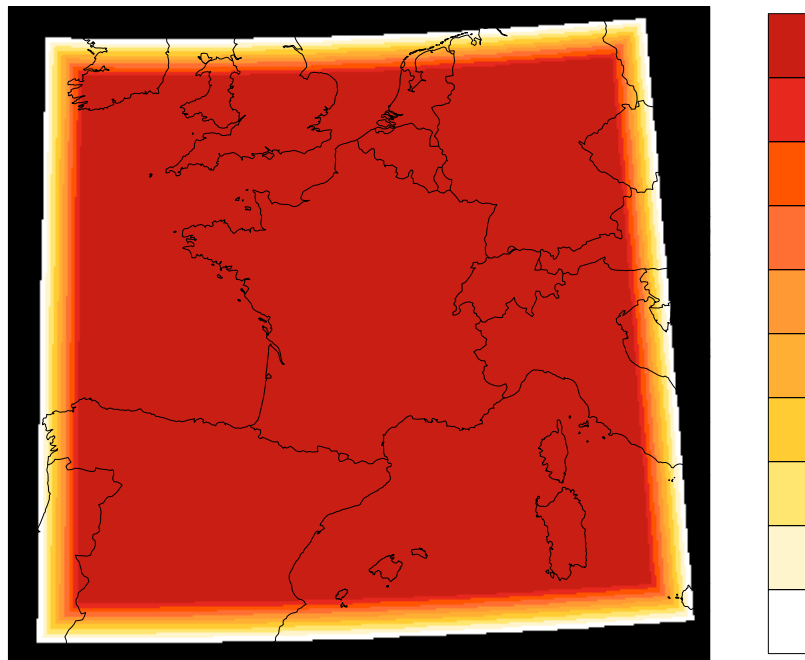


Figure 4.1. – Champ de déplacement zonal, défini comme constant et égal à 10 pdg sur tout le domaine intérieur, et extrapolé dans la zone tampon (la zone noire n'appartient pas au domaine de calcul). Cet exemple est donné pour le domaine AROME-France.

Modélisation des opérateurs de covariance **D** et **R**

Comme cela a déjà été expliqué à la section 1.4 pour **B**, les matrices **D** et **R** sont utilisées sous forme d'opérateurs simplifiés. Pour l'estimation des erreurs de déplacement, les choix suivants sont faits. Tout d'abord les corrélations d'erreurs d'observation (termes hors-diagonaux

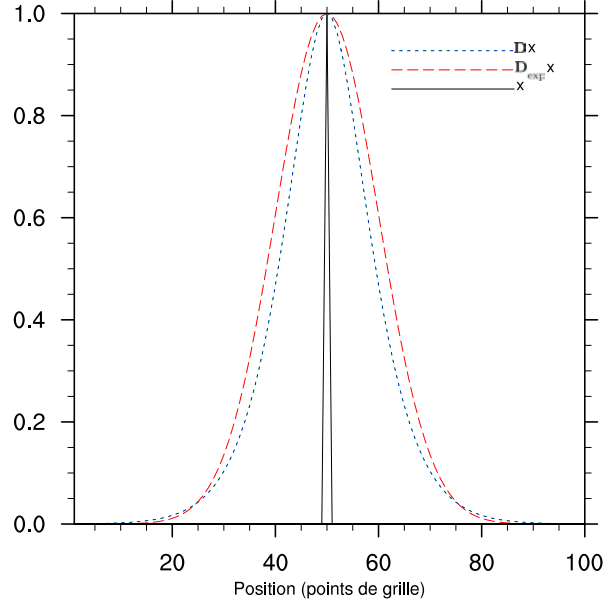


Figure 4.2. – Résultats du produit matrice vecteur entre un vecteur d'état (\mathbf{x} , impulsion au point de grille 50, ligne noire) et une matrice de covariance modélisée par un filtre récursif d'ordre 1 avec 3 passes, et une longueur de filtre $L_d = 10$ pdg (tirets bleus), ou par une matrice construite explicitement par des fonctions de covariances de forme Gaussienne caractérisées également avec des longueurs de corrélation de 10 pdg. Les deux modélisations ont ici des variances toutes deux égales à un point de grille.

de \mathbf{R}) sont négligées, comme cela est fait en opérationnel pour ARPEGE et AROME. De plus, dans notre cas où les observations sont en fait un état de référence du modèle, les statistiques d'erreurs associées sont assimilées à celles des erreurs d'ébauche. De même que pour la matrice \mathbf{B} opérationnelle d'AROME, les variances sont choisies homogènes sur le domaine (tous les coefficients diagonaux de \mathbf{R} sont constants et égaux à σ_o^2 , $\mathbf{R} = \sigma_o^2 \mathbf{I}$). σ_o^2 est estimée à partir des AEAROs (cf section 1.4.4) de 90 membres, comme la moyenne spatiale pour chaque niveau modèle (symbolisée par " $\langle \cdot \rangle$ ") des variances d'erreurs d'amplitude échantillonnées sur l'ensemble :

$$\sigma_o^2 = \left\langle \frac{1}{N_s - 1} \sum_{n=1}^{N_s} (\mathbf{x}_{b,n} - \bar{\mathbf{x}}_b)^2 \right\rangle \quad (4.18)$$

avec $\mathbf{x}_{b,n=1..N_s}$ les différents membres de l'ensemble, et $\bar{\mathbf{x}}_b$ la moyenne sur l'ensemble définie telle que

$$\bar{\mathbf{x}}_b = \frac{1}{N_s} \sum_{n=1}^{N_s} \mathbf{x}_{b,n}$$

De plus, \mathbf{D} est modélisée avec un filtre récursif du premier ordre, composé de trois passes (Purser et al., 2003). L'opérateur est totalement caractérisé par une variance σ_d^2 , et une longueur de corrélation L_d , toutes deux communes aux deux dimensions de l'espace et à tous les pdg. Un exemple unidimensionnel de la fonction de structure associée est donné par la Fig.4.2. Toujours pour simplifier le problème, il a été choisi de négliger les corrélations entre les déplacements dans les deux directions de l'espace. Dans la représentation matricielle de l'Eq.4.6, les deux blocs non-diagonaux sont nuls et les deux blocs diagonaux sont égaux.

4.5.2. Déroulement et contrôle de la minimisation

Pour minimiser la fonction coût non quadratique \mathcal{J} donnée par l'Eq.4.5, l'algorithme itératif de type Gauss-Newton, présenté à la section 4.4.1, est utilisé. Pour augmenter la robustesse de l'algorithme, une approche multi-échelles est implémentée. À chaque résolution de travail, des critères d'arrêt sont spécifiés dans les boucles internes afin de s'assurer de rester dans le domaine de validité des modèles linéaires successifs de \mathcal{J} .

Algorithme multi-échelles

L'algorithme multi-échelles est défini tel que :

Algorithme de Gauss-Newton multi-échelles

1. $\hat{\mathbf{d}}_0 = \mathbf{d}_{init}$ (initialisation)
 2. à faire pour toutes échelles $T_{l=1..N_T}$ (boucle sur les échelles l)
 - $\mathbf{d}_0 = \hat{\mathbf{d}}_{l-1}$ (initialisation par le résultat associé à T_{l-1})
 - à faire jusqu'à la convergence (boucle externe sur k)
 - a) résoudre $\tilde{\mathbf{F}}_l(\mathbf{d}_k)^T \tilde{\mathbf{F}}_l(\mathbf{d}_k) \widehat{\delta \mathbf{d}}_k = -\tilde{\mathbf{F}}_l(\mathbf{d}_k)^T \tilde{f}_l(\mathbf{d}_k)$ (boucle interne)
 - b) $\mathbf{d}_{k+1} = \mathbf{d}_k + \widehat{\delta \mathbf{d}}_k$ (actualisation)
 - $\hat{\mathbf{d}}_l = \mathbf{d}_{k+1}$ (solution associée à T_l)
 3. $\hat{\mathbf{d}} = \hat{\mathbf{d}}_{N_T}$
-

Les opérateurs \tilde{f}_l et $\tilde{\mathbf{F}}_l$ sont les équivalents aux opérateurs f et \mathbf{F} des équations 4.8 et 4.9, pour chacune des résolutions de travail $T_{l=1..N_T}$. En reprenant tous les choix spécifiques à cette thèse, \tilde{f}_l est défini tel que :

$$\tilde{f}_l(\mathbf{d}) = \begin{pmatrix} \mathbf{D}^{-\frac{1}{2}} & 0 \\ 0 & \sigma_{o,l} \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{d} \\ \tilde{\mathbf{x}}_{r,l} - \tilde{\mathbf{x}}_{b,l[s+\mathbf{d}]} \end{pmatrix}.$$

et $\tilde{\mathbf{F}}_l$ tel que :

$$\tilde{\mathbf{F}}_l(\mathbf{d}) = \begin{pmatrix} \mathbf{D}^{-\frac{1}{2}} & 0 \\ 0 & \sigma_{o,l} \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{I} \\ \nabla \tilde{\mathbf{x}}_{b,l[s+\mathbf{d}]} \end{pmatrix}$$

Aucune contrainte forte n'est imposée pour lisser les erreurs de déplacement. Seuls les champs à aligner \mathbf{x}_r et \mathbf{x}_b sont lissés avec un filtre spectral passe-bas spécifiant les résolutions de travail $T_{l=1..N_T}$ pour donner $\tilde{\mathbf{x}}_{r,l}$ et $\tilde{\mathbf{x}}_{b,l}$. Si l'on revient au cas réel où \mathcal{H} n'est pas l'identité, ce type d'approche ne pourra pas être adopté car cela nécessiterait le filtrage spatial des observations, ce qui est difficilement justifiable. Il faudra suivre par exemple Nehrkorn et al. (2014), où la contrainte forte sur la résolution de travail est appliquée directement sur l'incrément $\widehat{\delta \mathbf{d}}_k$. L'approche qui a été choisie pour cette thèse, même si elle ne correspond pas à des objectifs opérationnels, peut notamment permettre de documenter les erreurs de déplacement et leurs statistiques d'erreurs en fonction de la troncature spectrale des champs à aligner.

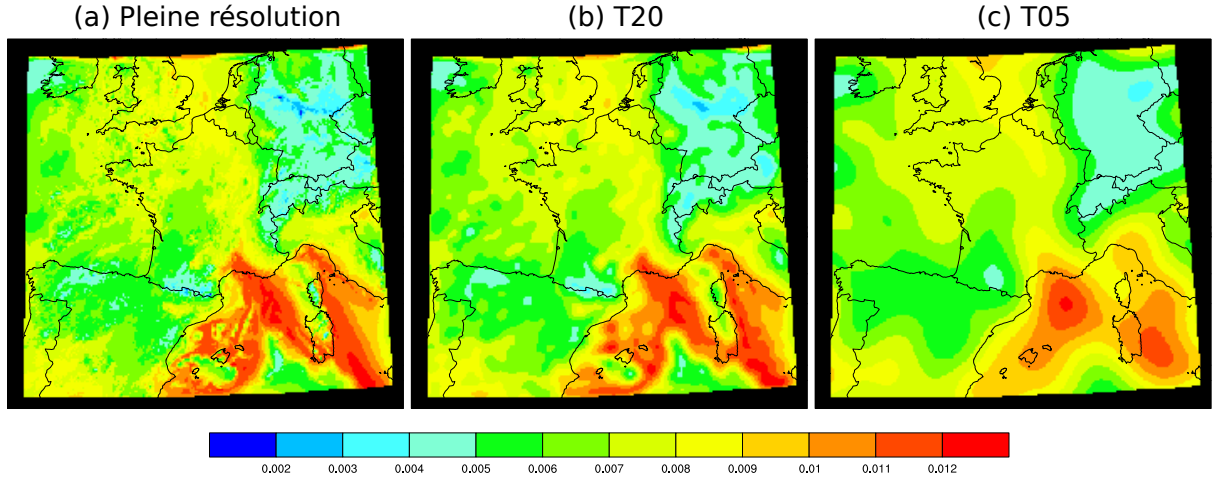


Figure 4.3. – Prévision AROME-France à 3h d’échéance d’humidité spécifique (g.kg^{-1}) au niveau modèle 52 (≈ 950 hPa), filtrée selon différentes troncatures : (a) pleine résolution, (b) T20, et (c) T05. La prévision est valide le 4 Novembre 2011, à 03 UTC.

Pour chacune des échelles l , la fonction coût non linéaire à minimiser lors de l’étape 2 de l’algorithme de Gauss-Newton multi-échelles, est notée \mathcal{J}_l et est définie par

$$\mathcal{J}_l = \frac{1}{2} \mathbf{d}^T \mathbf{D}^{-1} \mathbf{d} + \frac{1}{2\sigma_{o,l}^2} \left| \tilde{\mathbf{x}}_{r,l} - \tilde{\mathbf{x}}_{b,l,[s+\mathbf{d}]} \right|^2 \quad (4.19)$$

Pour chaque boucle interne k , \mathcal{J}_l est modélisée par linéarisation autour de \mathbf{d}_k par $J_{l,k}$ définie par :

$$J_{l,k} = \frac{1}{2} \left(\mathbf{d}_k + \delta \mathbf{d}_k \right)^T \mathbf{D}^{-1} \left(\mathbf{d}_k + \delta \mathbf{d}_k \right) + \frac{1}{2\sigma_{o,l}^2} \left| \tilde{\mathbf{x}}_{r,l} - \tilde{\mathbf{x}}_{b,l,[s+\mathbf{d}_k]} - \nabla \tilde{\mathbf{x}}_{b,l,[s+\mathbf{d}_k]} \delta \mathbf{d}_k \right|^2 \quad (4.20)$$

Pour cette thèse, il a été choisi de travailler avec $N_T = 2$, où les troncatures sont fixées à $T_1 = 5$ (notée T05 par la suite) et $T_2 = 20$ (notée T20 par la suite). Avec une résolution horizontale de 2.5 km et un domaine de taille 720×750 pdg, T05 correspond à des tailles caractéristiques autour de 360 km et T20 à des tailles caractéristiques autour de 90 km. La Fig.4.3 représente un exemple de carte AROME-France d’humidité spécifique brute ou lissée selon T05 et T20. Ce réglage est en accord avec le reste de la littérature puisque toutes les études portées sur l’estimation des erreurs de déplacement en PNT se restreignent à une estimation de $\hat{\mathbf{d}}$ pour les plus grandes échelles. Par exemple Aonashi & Eito (2011) se limitent à des tailles caractéristiques autour de 70 km, et Nehr Korn et al. (2014) autour de 105 km.

Par choix, les opérateurs \mathbf{D} et \mathbf{R} sont communs à toutes les résolutions de travail $T_{l=1..N_T}$ avec notamment σ_o^2 qui est définie par l’Eq.4.18 à partir de l’ensemble brut à pleine résolution. À chaque résolution de travail T_l , pour que la dispersion de l’ensemble filtré soit cohérente avec la valeur de σ_o^2 , chaque membre est inflaté par le coefficient θ_l défini tel que :

$$\theta_l = \frac{\sigma_o}{\sqrt{\left\langle \frac{1}{N_s-1} \sum_{n=1}^{N_s} (\tilde{\mathbf{x}}_{b,l,n} - \bar{\tilde{\mathbf{x}}}_{b,l})^2 \right\rangle}}$$

avec $\bar{\tilde{\mathbf{x}}}_{b,l}$ la moyenne sur l'EDA composé des membres filtrés $\tilde{\mathbf{x}}_{b,l,n=1..N_s}$. Ce choix de ne garder qu'une seule valeur de variance d'erreurs d'ébauche quelque soit la résolution de travail, simplifie l'implémentation mais surtout permet de garder une certaine variabilité spatiale malgré le filtrage spectral avec $\theta_{l=1..N_T} > 1$. Ceci ajoute une certaine robustesse à l'algorithme d'estimation de $\hat{\mathbf{d}}$ (cf section 4.4.2). Les profils verticaux de σ_o utilisées pour les quatre variables modèles (vent zonal U , vent méridien V , température T , et humidité spécifique q) sont représentés par la Fig.4.4 (ces résultats sont cohérents avec la figure 2.12 de Ménétrier, 2014).

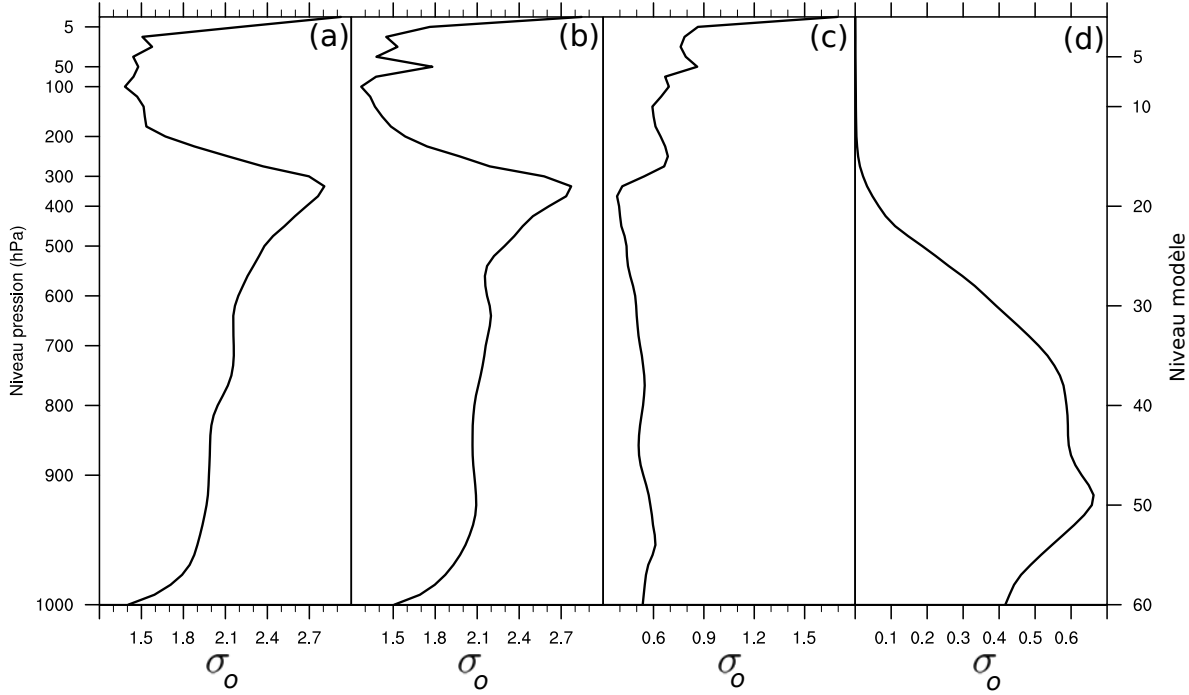


Figure 4.4. – Profils verticaux des σ_o AROME-France définies par l'Eq.4.18 pour quatre variables : (a) U (m.s^{-1}), (b) V (m.s^{-1}), (c) T (K), et (d) q (g.kg^{-1}). Ces écarts types sont estimés à partir de l'AEARO-France composée de 90 prévisions à 3h d'échéance, valides le 4 Novembre 2011 à 03 UTC.

Boucle interne

Avec un préconditionnement en \mathbf{D} (cf section 4.4.3), il a été choisi d'utiliser une dérivation de l'algorithme du gradient conjugué de type BCG (Gürol et al., 2013). En suivant les étapes formulées par cette dernière référence et par le chapitre 5 de Nocedal & Wright (2006), le système composé des équations 4.15 et 4.16 est reformulé tel que :

$$\left\{ \begin{array}{l} \mathbf{A}\boldsymbol{\lambda} = \mathbf{b} \\ \boldsymbol{\lambda} = \mathbf{D}^{-1}\boldsymbol{\delta d} \\ \mathbf{A} = \left(\mathbf{I} + \nabla^T \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \nabla \mathbf{D} \right) \\ \mathbf{b} = \nabla^T \mathbf{H}^T \mathbf{R}^{-1} \left(\mathbf{y}_o - \mathcal{H}(\mathbf{x}_{b,[s+d]}) \right) - \mathbf{D}^{-1} \mathbf{d} \end{array} \right.$$

(dans la suite de cette section pour ne pas surcharger les notations, tous les indices numérotent les itérations d'une boucle interne, et n'ont aucun lien avec la numérotation des itérations des boucles externes k , ou avec la numérotation des résolutions de travail l). Il est inversé de façon

itérative tel que :

Algorithme de BCG (largement inspiré de Gürol et al., 2013)

1. initialisation :
 - $\delta \mathbf{d}_0 = 0$ (incrément initial, espace modèle)
 - $\lambda_0 = 0$ (incrément initial, espace de contrôle)
 - $\mathbf{r}_0 = \mathbf{b}$ (résidu initial, espace de contrôle)
 - $\mathbf{z}_0 = \mathbf{D}\mathbf{r}_0$ (résidu initial, espace modèle)
 - $\mathbf{p}_0 = \mathbf{z}_0$ (direction de descente initiale, espace modèle)
 - $\mathbf{h}_0 = \mathbf{r}_0$ (direction de descente initiale, espace de contrôle)
 2. à faire jusqu'à la convergence (boucle interne sur i)
 - a) $\alpha_i = \frac{\mathbf{r}_i^T \mathbf{D}\mathbf{r}_i}{\mathbf{r}_i^T \mathbf{A}^T \mathbf{D}\mathbf{r}_i}$ (minimisation selon la direction de descente)
 - b) $\delta \mathbf{d}_{i+1} = \delta \mathbf{d}_i + \alpha_i \mathbf{p}_i$ (incrément actualisé, espace modèle)
 - c) $\lambda_{i+1} = \lambda_i + \alpha_i \mathbf{h}_i$ (incrément actualisé, espace de contrôle)
 - d) $\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i \mathbf{A}\mathbf{r}_i$ (résidu actualisé, espace de contrôle)
 - e) $\mathbf{z}_{i+1} = \mathbf{D}\mathbf{r}_{i+1}$ (résidu actualisé, espace modèle)
 - f) $\beta_i = \frac{\mathbf{r}_{i+1}^T \mathbf{D}\mathbf{r}_{i+1}}{\mathbf{r}_i^T \mathbf{r}_i}$ (recherche d'une direction de descente conjuguée)
 - g) $\mathbf{p}_{i+1} = \mathbf{z}_{i+1} + \beta_i \mathbf{p}_i$ (direction de descente actualisée, espace modèle)
 - h) $\mathbf{h}_{i+1} = \mathbf{r}_{i+1} + \beta_i \mathbf{h}_i$ (direction de descente actualisée, espace de contrôle)
 3. $\widehat{\delta \mathbf{d}} = \delta \mathbf{d}_{i+1}$
-

En pratique, la librairie M1CG1 développée par l'INRIA et documentée par Gilbert (2011), modifiée pour prendre en compte l'implémentation du produit entre l'opérateur \mathbf{D} et un vecteur, et l'ajout de critères d'arrêt adaptatifs présentés maintenant.

Pour chaque boucle externe, le critère d'arrêt défini par l'Eq.4.17 est basé sur une limitation induite par l'utilisation d'interpolations linéaires, mais ne fixe pas *a priori* le nombre d'itérations que l'on s'autorise. De la même manière, étant donné que l'on n'a pas d'idée *a priori* du nombre d'itérations nécessaires à la convergence de l'algorithme BCG, trois critères d'arrêt, tous adaptatifs, sont proposés :

- un premier critère d'arrêt, basé sur la limitation induite par les interpolations linéaires, est formulé tel que

$$\max(|\delta \mathbf{d}_{i+1} - \delta \mathbf{d}_i|) < 1 \text{ point de grille}$$

- un deuxième critère d'arrêt teste la validité de la linéarisation effectuée pour chaque boucle interne. Il compare les variations de la fonction coût totale avec celles de sa version linéarisée. Le critère est défini selon un seuil $\beta \geq 0$ tel que

$$\frac{\mathcal{J}_l(\mathbf{d}_k + \delta \mathbf{d}_{i+1}) - \mathcal{J}_l(\mathbf{d}_k + \delta \mathbf{d}_i)}{J_{l,k}(\mathbf{d}_k + \delta \mathbf{d}_{i+1}) - J_{l,k}(\mathbf{d}_k + \delta \mathbf{d}_i)} < \beta \quad (4.21)$$

Pour cette équation, les indices associés à la résolution de travail et à la numérotation des itérations des boucles externes sont spécifiés, car il est important de noter que ce sont les versions filtrées du problème total qui sont testées (\mathcal{J}_l définie par l'Eq.4.19 et

$J_{l,k}$ définie par l'Eq.4.20). Pour cette étude, le choix le plus simple $\beta = 0$ est fait i.e \mathcal{J}_l croît alors que $J_{l,k}$ décroît (par construction). D'autres valeurs positives pourraient aussi être efficaces mais n'ont pas été testées. Une interprétation de l'Eq.4.21 est de dire que \mathcal{J}_l a été modélisée par $J_{l,k}$ et que le critère d'arrêt est défini à l'itération i de la boucle interne comme le rapport entre la décroissance observée et la décroissance prévue de \mathcal{J}_l . Contrairement au premier critère d'arrêt, celui-ci n'est pas un critère de convergence, mais indique lors de son déclenchement que la modélisation linéaire actuelle du problème n'est plus valide.

4.6. Test sur des cas idéalisés

Des résultats de correction des erreurs de déplacement sont maintenant présentés pour deux cas idéalisés de translation, afin de vérifier le bon comportement de la méthode présentée précédemment. Les résultats exposés ici n'ont donc pas valeur de documentation des erreurs de déplacement. Les valeurs de σ_d et L_d utilisées pour caractériser \mathbf{D} ont été fixées arbitrairement : $\sigma_d = 10$ pdg (25km), et $L_d = 120$ pdg (300km).

4.6.1. Translation d'un champ idéalisé

Pour ce premier cas test, la référence \mathbf{x}_r n'est pas un champ modèle mais une quasi-Gaussienne bidimensionnelle construite avec le même opérateur que \mathbf{D} défini avec $\sigma_d = 1$ et $L_d = 120$ pdg. Par construction, \mathbf{x}_b est une translation de 40 pdg vers le nord-est de \mathbf{x}_r . Pour tous les pdg (i,j) du domaine intérieur, le déplacement à estimer est donc de -40 pdg pour les deux directions de l'espace. Les étapes de ce premier cas test sont représentées par la Fig.4.5. Le résultat est satisfaisant avec un bon alignement entre \mathbf{x}_r et $\mathbf{x}_{b,[s+\hat{\mathbf{d}}]}$ où $\hat{\mathbf{d}}$ est proche de la translation de 40 pdg vers le sud-ouest. Avec le même constat que dans la section 4.4.2, c'est dans les zones où le gradient de \mathbf{x}_b est faible (au centre de la Gaussienne, et loin des valeurs maximum) que le déplacement est sous estimé.

La Fig.4.6 rassemble les valeurs de la fonction coût non-quadratique $\mathcal{J}_{l=1,2}$, définie par l'Eq.4.19, tout au long de la minimisation. Avec ce premier cas test, la boucle externe est composée de 5 boucles internes, toutes à la première résolution de travail T05. Le nombre cumulé d'itérations dans les boucles internes successives est de 6. En plus de noter que la fonction coût non-linéaire décroît bien à toutes les itérations, on peut d'abord remarquer que dans chaque boucle interne, le nombre d'itérations de BCG est très faible (une ou deux itérations). L'étude des critères d'arrêts indique que toutes les boucles internes ont été arrêtées par le déclenchement du troisième critère d'arrêt défini par l'Eq.4.21, indiquant l'obsolescence du modèle linéaire $J_{k,l}$ de \mathcal{J} lorsque l'incrément en déplacement devient trop grand. Le deuxième point important est qu'aucune itération à la résolution de travail T20 n'a été nécessaire pour raffiner $\hat{\mathbf{d}}$. Ceci peut s'expliquer par des champs à aligner et des déplacements à estimer qui sont tous de grandes échelles.

4.6.2. Translation d'un champ réel AROME-France

Ce deuxième cas test, représenté par la Fig.4.7, est identique au premier sauf que \mathbf{x}_r est maintenant un champ réel du modèle AROME-France (prévision à 3 heures de q dans la couche limite, d'un des membres de l'AEARO mentionnée à la section 1.4.4). De même que pour le premier cas test, les résultats obtenus sont encourageants avec $\hat{\mathbf{d}}$ proche de la translation de 40 pdg vers le sud-ouest. Une première chose à remarquer est la bonne décroissance du champs de

déplacement sur les bords du domaine (dans la zone tampon), permettant d'éviter la sortie du domaine de certains pdg déplacés. De la même manière que pour le cas précédent, un moins bon alignement est visible pour les zones de plus faible gradient (par exemple au Nord-Est du domaine). À l'inverse, dans les zones caractérisées par de forts gradients comme sur la Méditerranée entre la Corse et les Baléares, la méthode paraît beaucoup plus efficace.

Selon la Fig.4.8, la minimisation a été plus coûteuse avec 18 boucles internes totalisant 45 itérations cumulées. Contrairement au premier cas test, la moitié des itérations ont été faites à T20 et ces itérations ont permis d'améliorer la qualité de $\widehat{\mathbf{d}}$. En effet, d'après la Fig.4.9 beaucoup plus de points de grille vérifient $\{\widehat{\mathbf{d}}^x = -40 \pm 10 \text{ pdg}, \widehat{\mathbf{d}}^y = -40 \pm 10 \text{ pdg}\}$ lorsque les deux résolutions T05 et T20 sont utilisées (66% des pdg contre 49% lorsque seule T05 est utilisée). On remarque de plus que cette amélioration est valable pour les deux directions de l'espace x et y. L'apport de T20 par rapport à T05 est attribué aux petites échelles contenues dans les champs de q d'AROME-France.

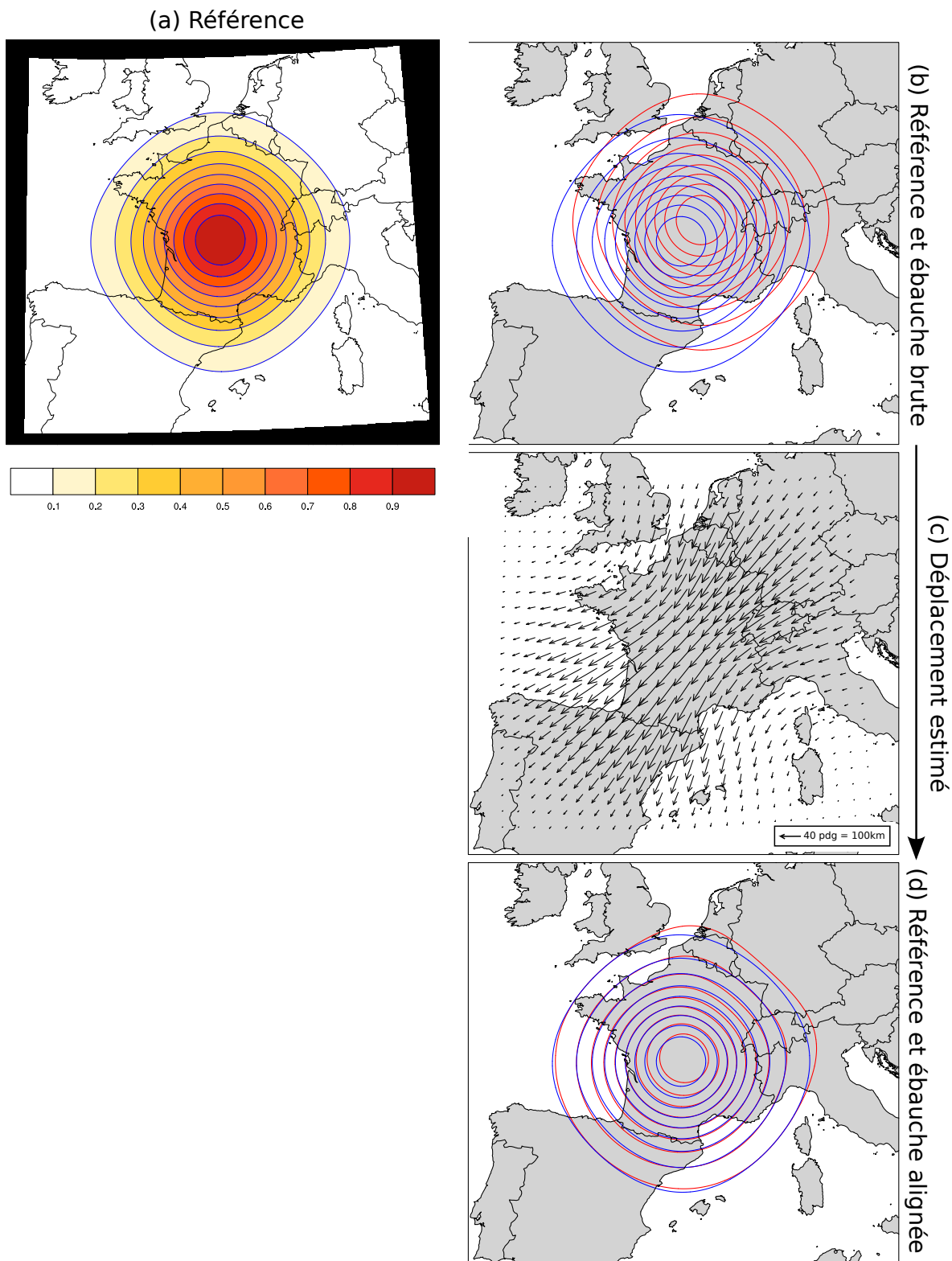


Figure 4.5. – Schéma de l'étape de d'alignement du premier cas test. (a) la référence \mathbf{x}_r est représentée par des isolignes bleues et avec des niveaux de couleurs. Sur la colonne de droite : (b) représente l'ébauche \mathbf{x}_b à aligner (isolignes rouges) et la référence \mathbf{x}_r (isolignes bleues), (d) l'ébauche alignée (isolignes rouges) et \mathbf{x}_r (isolignes bleues) et (c) le champ d'erreurs de déplacement estimé pour aligner \mathbf{x}_b , représenté par des vecteurs marquant l'emplacement de certains pdg avant et après l'alignement. La référence, l'ébauche, et l'ébauche alignée sont représentées à la résolution de travail T20.

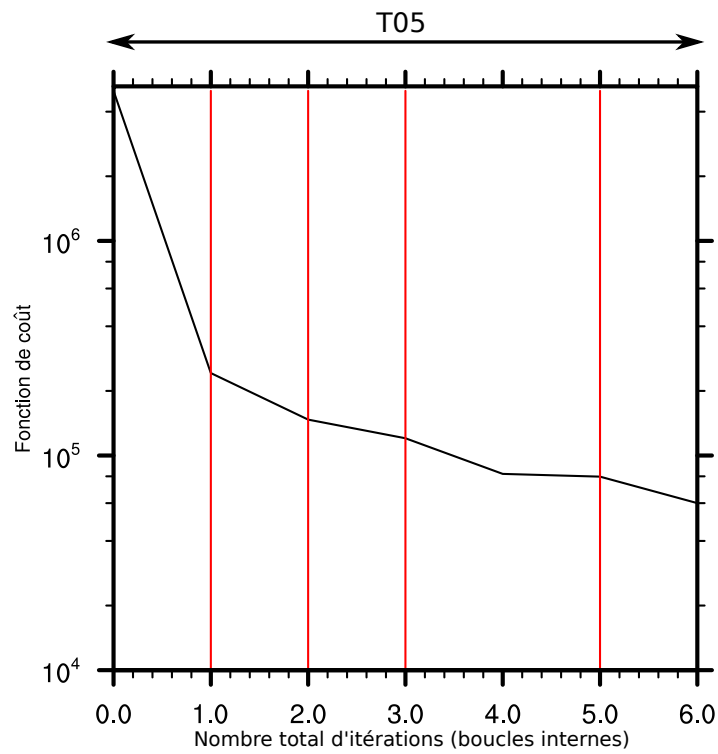


Figure 4.6. – Valeurs de la fonction $\mathcal{J}_{l=1,2}$, définie par l'Eq.4.19, tout au long de la minimisation. Ici, seule \mathcal{J}_1 est représentée (associée à T05) puisqu'aucune itération n'a été nécessaire à T20 (voir texte). Les barres verticales rouges délimitent les boucles internes entre elles.

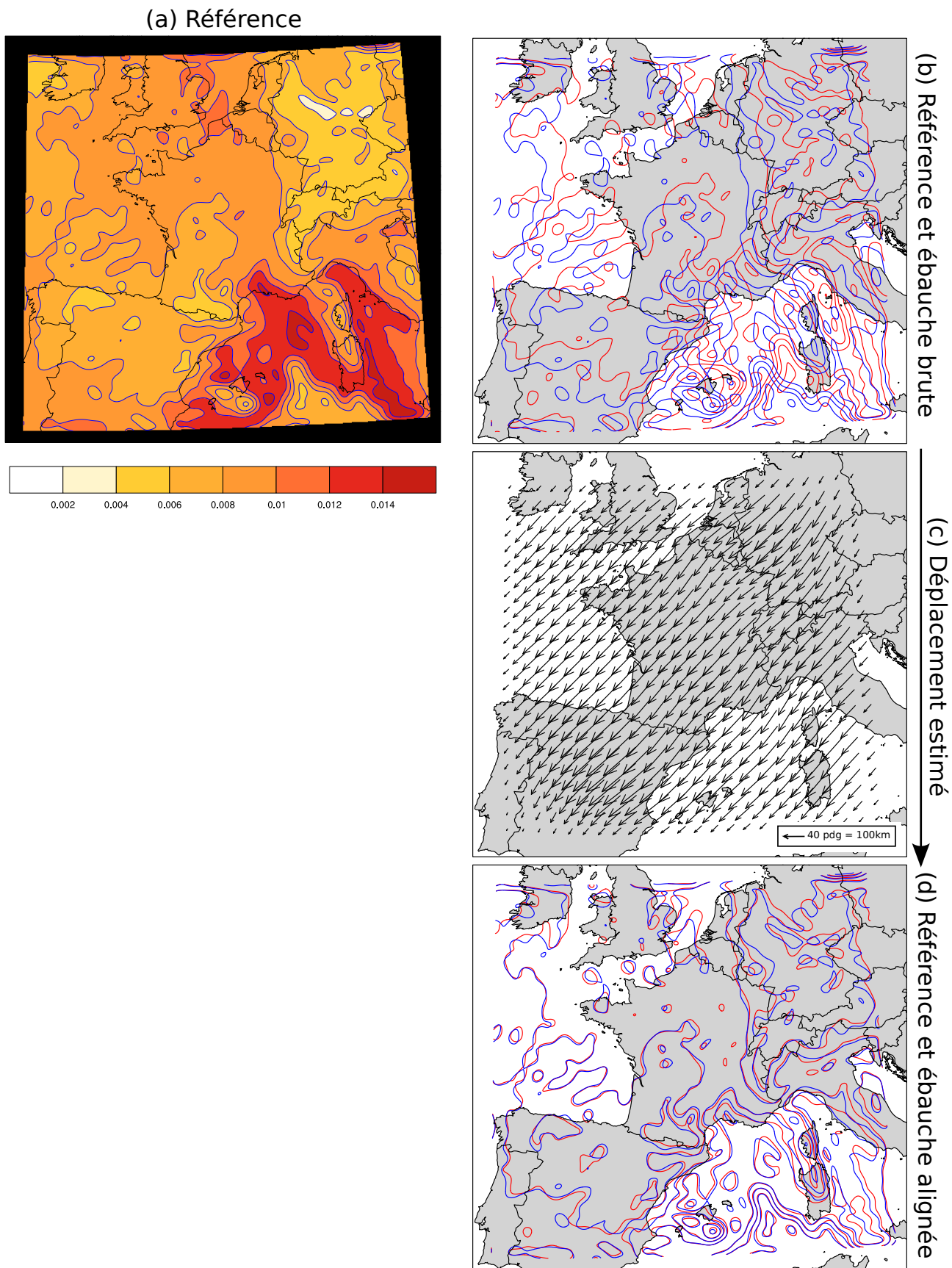


Figure 4.7. – Schéma d’alignement du deuxième cas test : idem que la Fig.4.5) mais en prenant comme référence x_r , une prévision AROME-France à 3h de q au niveau 52 (≈ 950 hPa), valide le 04/11/11 à 03 UTC.

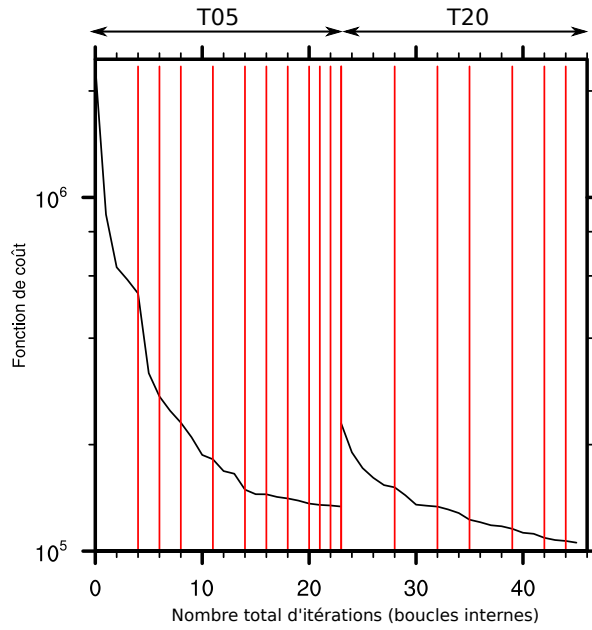


Figure 4.8. – Même représentation que la Fig.4.6 mais pour le deuxième cas test. Cette fois les deux résolutions de travail, T05 et T20, ont été utilisées.

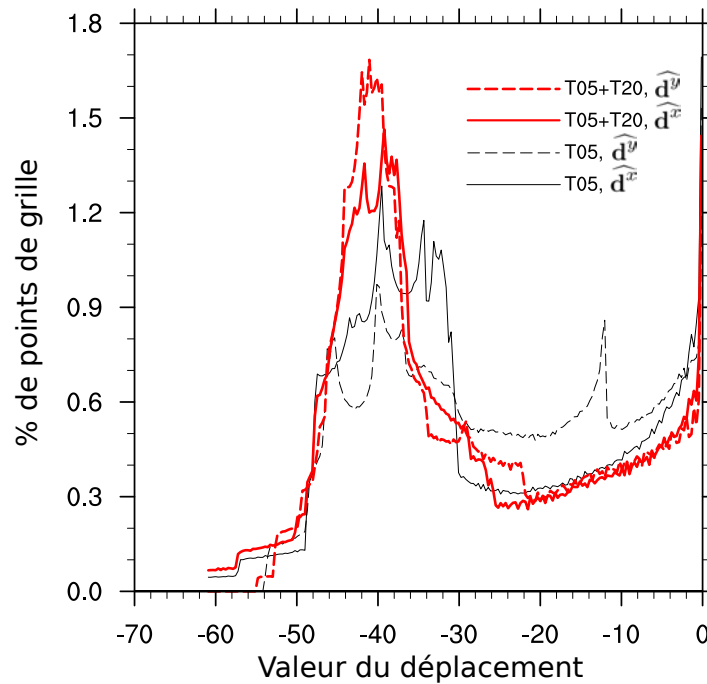


Figure 4.9. – Fonction de répartition des pdg du domaine (en %) selon leurs valeurs de $\widehat{\mathbf{d}}^x$ (lignes continues) et $\widehat{\mathbf{d}}^y$ (lignes pointillées). Ces fonctions sont calculées pour deux configurations. (noire) "T05" : l'algorithme d'estimation est arrêté à la fin de la minimisation à T05. (rouge) "T05+T20" : l'algorithme utilise également la deuxième résolution de travail T20.

4.7. Application à un cas réel AROME-WMED

La méthode est maintenant testée sur un cas réel, où la référence et l'ébauche sont deux membres de l'AEARO-WMED (cf section 1.4.4).

4.7.1. Situation météorologique étudiée : cas de Médicane

La simulation choisie pour tester la méthode d'estimation des erreurs de déplacement est une simulation AROME-WMED pour la situation du 8 novembre 2011 à 03 UTC. Elle a été choisie pour son cas de Médicane (centre dépressionnaire à cœur chaud, apparaissant en méditerranée, qui possède une dynamique qui est proche de celle des cyclones tropicaux, Chaboureau et al., 2012). Des cumuls importants de précipitations sont associés à ce système dépressionnaire (Fig.4.10), avec notamment sur les côtes italiennes de la Ligurie, des intensités allant jusqu'à 60 mm.h^{-1} . L'étude de ces épisodes fortement précipitants fait partie d'un des objectifs de la campagne d'étude HYMEX (cf section 1.1.2, Drobinski et al., 2014). Une deuxième zone active est présente au Nord-Ouest du domaine, associée à un minimum local de pression.

4.7.2. Estimation des erreurs de déplacement

Pour l'alignement des champs réels, l'écart type d'erreurs de déplacement σ_d est fixé à 1 pdg soit 2.5 km. Faute de temps, une étude de sensibilité à la valeur de ce paramètre n'a pas pu être entreprise. Le poids relatif du σ_d par rapport au σ_o paraît notamment être un point important à investiguer. On notera toutefois que Nehr Korn et al. (2015) montrent une faible dépendance

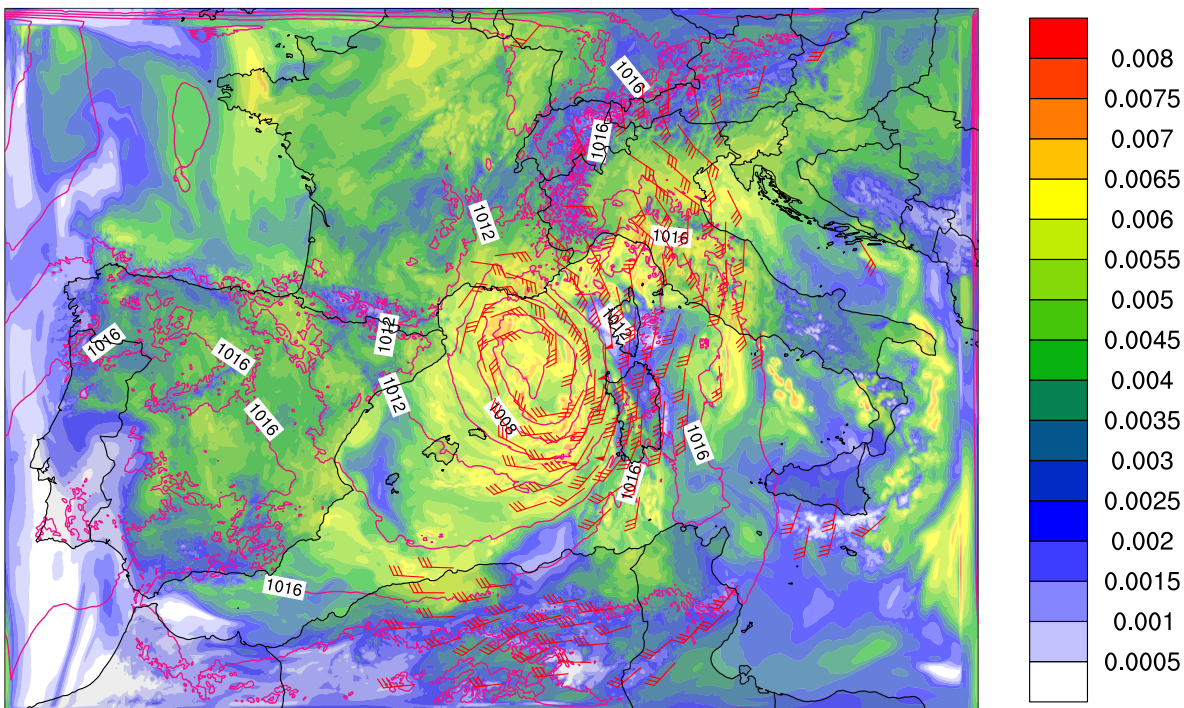


Figure 4.10. – Prévion AROME-WMED à 3h, valide le 8 novembre 2011 à 03 UTC : pression de surface au niveau de la mer (isolignes, hPa), vitesse du vent à 800 hPa (barbules, kt, seuillées à 30 kt), et q à 800 hPa (palette de couleur, kg.kg^{-1}).

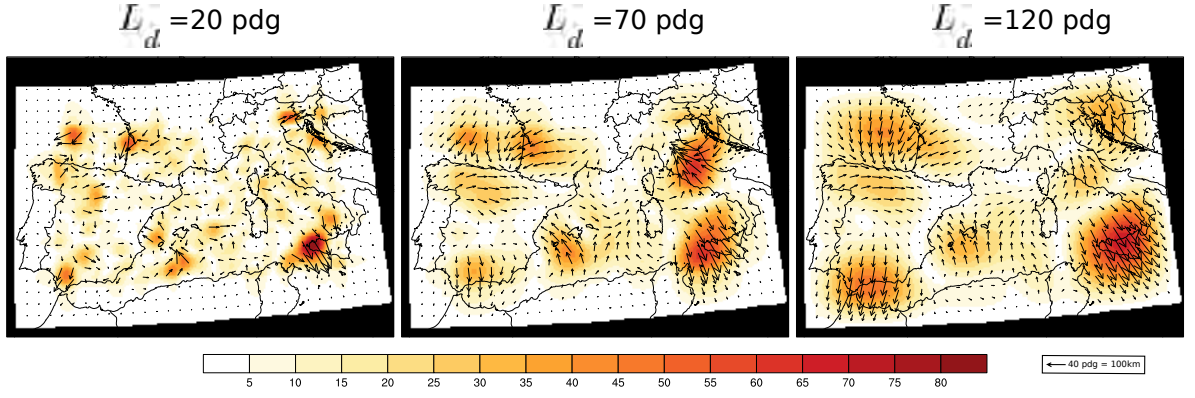


Figure 4.11. – Valeur du module et direction du déplacement pour différentes valeurs de L_d (pdg). Les résultats sont présentés pour q au niveau modèle 30 (≈ 600 hPa) pour un des membres de l’AEARO-WMED. Pour cette étude sur la sensibilité à L_d , la zone tampon a dû être agrandie à $L_t = 100$ pdg, à cause des forts déplacements en bord de domaine estimés pour $L_d = 20$ pdg.

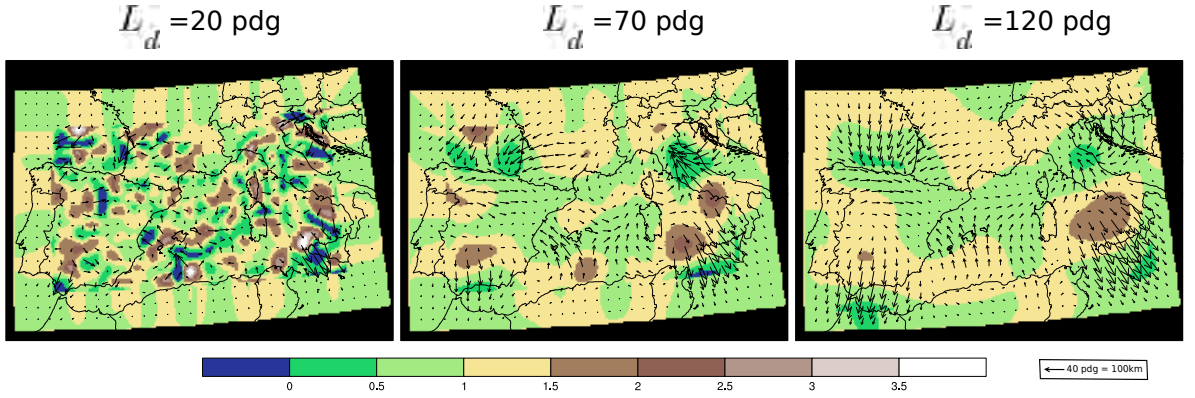


Figure 4.12. – Valeurs de \mathcal{I} (Eq.4.22) et déplacements estimés pour différentes valeurs de L_d (pdg) calculées pour les mêmes champs que ceux de la Fig.4.11.

de leurs résultats à la valeur de σ_d choisie.

L’utilisation de fortes valeurs de L_d augmente le pouvoir lissant de la contrainte de régularité sur $\hat{\mathbf{d}}$. Les incréments en déplacement sont alors de plus grandes échelles (Fig.4.11). Pour l’estimation de déplacement de plus petites échelles, une diminution de L_d pourrait être envisagée. Néanmoins, de faibles valeurs de L_d autorisent des structures de déplacement plus fines et plus intenses qui induisent des ébauches alignées qui peuvent être déformées de manière peu réaliste, avec notamment des points singuliers où le déplacement entraîne une déformation de l’ébauche qui n’est pas inversible. Afin de déterminer les valeurs de L_d qui conservent l’inversibilité des déplacements estimés, il est possible d’utiliser le diagnostic d’inversibilité \mathcal{I} (noté $\det(J_d)$ par Legrand & Michel, 2014), défini comme le déterminant du Jacobien du changement de coordonnées pour chaque point de grille (i, j) :

$$\mathcal{I}(i, j) = \frac{\partial \mathbf{d}_c^x(i, j)}{\partial x} \frac{\partial \mathbf{d}_c^y(i, j)}{\partial y} - \frac{\partial \mathbf{d}_c^x(i, j)}{\partial y} \frac{\partial \mathbf{d}_c^y(i, j)}{\partial x} \quad (4.22)$$

avec $(\mathbf{d}_c^x(i, j), \mathbf{d}_c^y(i, j))$ la nouvelle position de chaque point de grille (i, j) définie telle que :

$$\mathbf{d}_c^x(i, j) = i + \widehat{\mathbf{d}}^x(i, j) \text{ et } \mathbf{d}_c^y(i, j) = j + \widehat{\mathbf{d}}^y(i, j)$$

Une valeur négative de \mathcal{I} indique que le déplacement est non inversible. La Fig.4.12 représente des cartes de cet indice d'inversibilité pour trois valeurs de L_d . Alors que pour $L_d = 20$ pdg ou $L_d = 70$ pdg, des zones où \mathcal{I} est négatif apparaissent, l'allongement de la longueur de portée à $L_d = 120$ pdg permet à \mathcal{I} de rester positif sur le domaine. Ainsi, la longueur de corrélation L_d est fixée à $L_d = 120$ pdg soit 300 km, ce qui est en accord avec l'ordre de grandeur proposé par Nehr Korn et al. (2015).

Alignement de champs d'humidité spécifique

L'alignement de deux champs de q au niveau 30 (600 hPa) (cf Fig.4.13) montre tout d'abord que l'amplitude des déplacements estimés dépassent dans certaines zones les 68 pdg soit 170 km, ce qui largement supérieure à la valeur de $\sigma_d = 1$ pdg. Ceci semble indiquer que le poids relatif $\frac{1}{\sigma_d^2}$ est assez petit pour autoriser des déplacements d'assez grandes amplitudes pour permettre un meilleur alignement avec la référence. La qualité des déplacements estimés peut être quantifiée objectivement avec la diminution de l'écart quadratique entre la référence et l'ébauche lors de leur alignement (Souopgui et al., 2012). De manière équivalente, les variations de \mathcal{J}_o qui est le terme de rappel aux observations de l'Eq.4.5 (défini par $\mathcal{J}_o = \frac{1}{\sigma_o} |\mathbf{x}_{ref} - \mathbf{x}_{b,[s+d]}|^2$), peut donner la même information. D'après la Fig.4.14, l'étape d'alignement a permis de diminuer l'écart entre la référence et l'ébauche avec une diminution du \mathcal{J}_o à pleine résolution de l'ordre de 30%.

Néanmoins, lorsque l'on ne s'intéresse qu'à la minimisation à T05, on observe que la décroissance monotone du $\mathcal{J}_{o,1}$ à cette résolution n'empêche pas la croissance du \mathcal{J}_o en fin de minimisation. Ceci montre que la solution du problème à T05 n'est pas cohérente avec la solution obtenue à pleine résolution. En terme d'intensité du déplacement, les valeurs finales du terme de contrainte du déplacement, J_d , obtenues à T05 et T20 sont proches, malgré une diminution en début de la minimisation à T20. À ces résolutions, l'algorithme converge de manière satisfaisante vers la solution du problème, avec une décroissance monotone des fonctions coûts quadratiques linéaires et non-linéaires aux résolutions dégradées (cf Fig.4.14).

Alignement de champs de température

Comme on peut le voir sur la Fig.4.15.a, la température au niveau 30 (≈ 600 hPa) est un champ de plus grande échelle que l'humidité spécifique, mais sur lequel les effets des masques Terre-Mer et du relief sont marqués. Pour ce champ, dont les étapes d'alignement sont représentées par la Fig.4.15, la correction des erreurs de déplacement a pour effet d'augmenter la valeur du \mathcal{J}_o non-linéaire recalculés sur les champs déplacés à pleine résolution, de plus de 300% en fin de minimisation à T05 (Fig.4.16). La minimisation à T20 ne réussie pas à réduire complètement cet écart malgré une diminution de plus de 50% entre le \mathcal{J}_o initial et final. Dans le même temps, $\mathcal{J}_{o,l}$ diminue de façon monotone pour chaque phase de la minimisation. Comme pour le cas test précédent, ces incohérences montrent que la forme du déplacement estimé dépend fortement de l'échelle de résolution. L'utilisation d'une résolution trop dégradée entraîne même un échec de la méthode dans le sens où le \mathcal{J}_o diagnostiqué croît lors de la phase d'alignement.

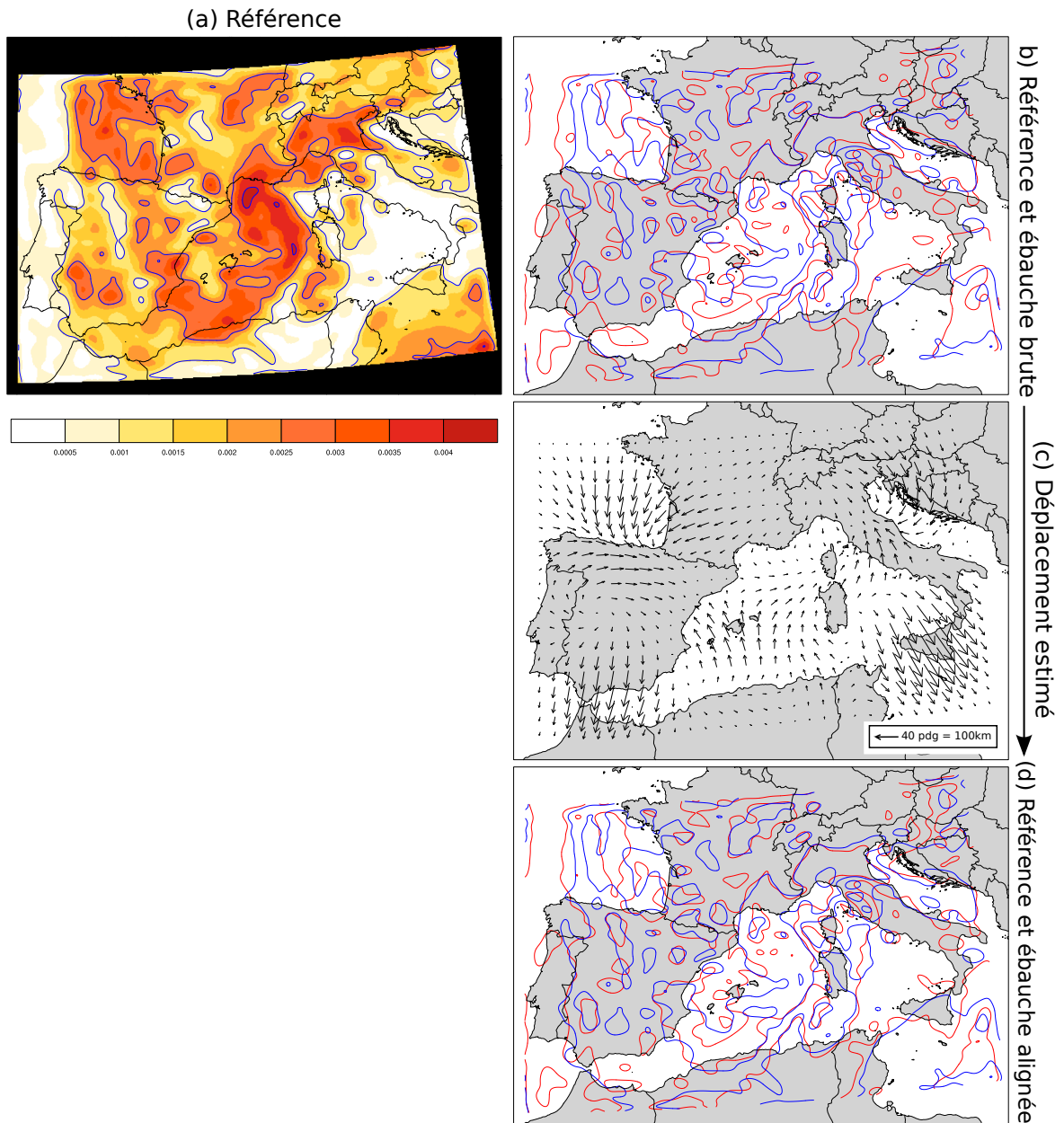


Figure 4.13. – Schéma d’alignement entre deux prévisions à 3h de q , au niveau 30 (≈ 600 hPa) qui sont des membres de l’AEARO-WMED. Les mêmes conventions que les figures 4.5 et 4.7) sont utilisées.

Afin de mieux comprendre d’où viennent ces incohérences, la Fig.4.17 représente les cartes de différences entre les écarts quadratiques entre l’ébauche alignée et la référence ($EQ(\hat{\mathbf{d}})$) et l’écart quadratique entre l’ébauche brute et la référence ($EQ(\mathbf{0})$). Si $EQ(\hat{\mathbf{d}}) - EQ(\mathbf{0})$ est négative cela implique que l’alignement a permis de rapprocher l’ébauche de la référence en diminuant les écarts d’amplitude. Ces cartes de $EQ(\hat{\mathbf{d}}) - EQ(\mathbf{0})$ sont montrées pour les champs à pleine résolution et les champs filtrés à la troncature de travail la plus élevée T20, correspondants ainsi respectivement aux fonctions coûts finales \mathcal{J}_o et $\mathcal{J}_{o,l}$ de la Fig.4.16. On observe qu’à la résolution dégradée T20, la différence $EQ(\hat{\mathbf{d}}) - EQ(\mathbf{0})$ est négative sur la majorité du domaine, alors qu’à pleine résolution, les zones côtières et surtout les zones de relief sont au contraire affectées par une forte augmentation des écarts quadratiques à cause de l’alignement.

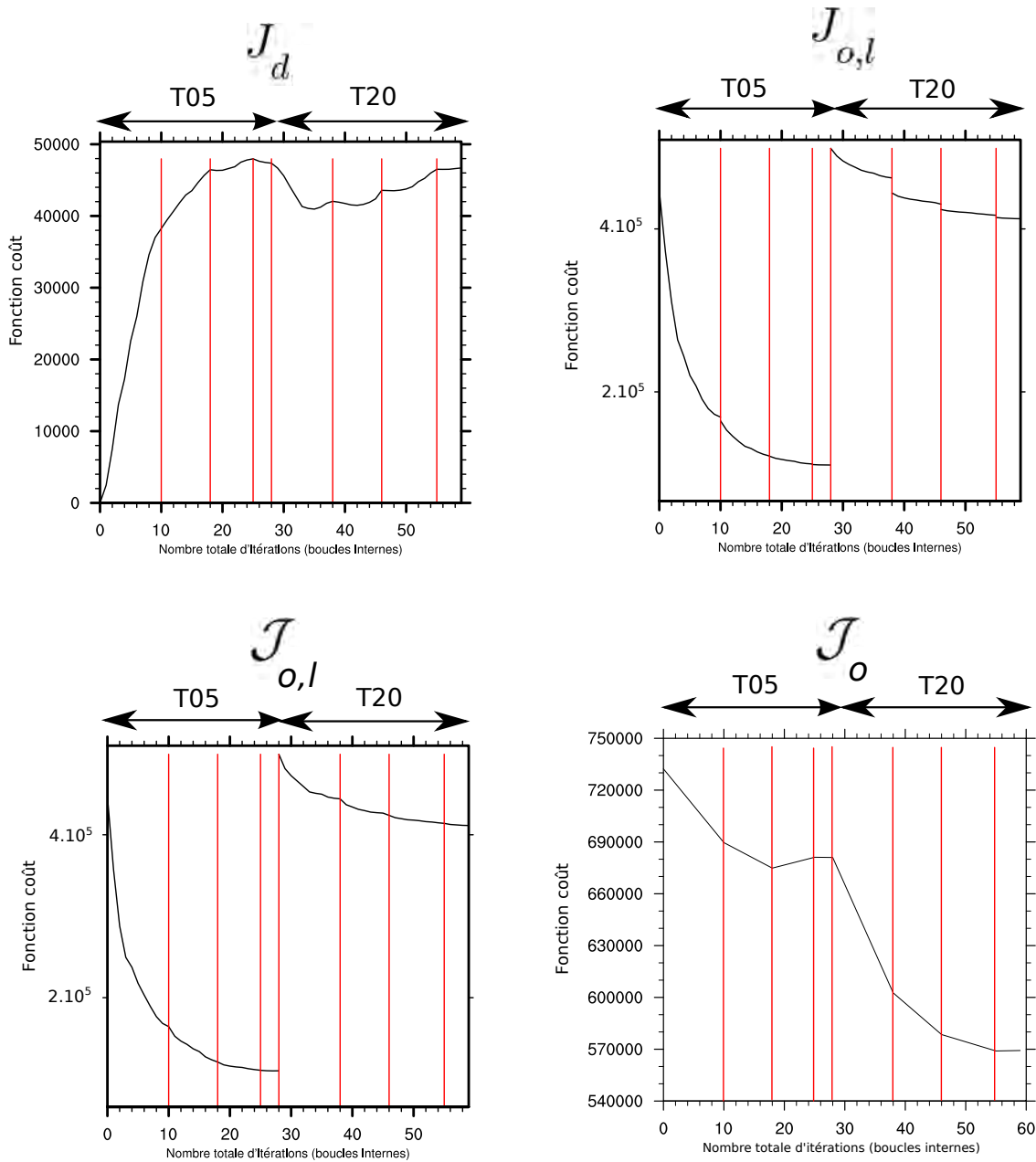


Figure 4.14. – Valeurs de J_d (en haut à gauche), de $J_{o,l}$ (en haut à droite), de $\mathcal{J}_{o,l}$ (en bas à gauche) et de \mathcal{J}_o (en bas à droite) durant la minimisation. Les résultats sont donnés pour q d'un des membres de l'ensemble AROME-WMED au niveau 30 (≈ 600 hPa). J_d est défini comme le terme de contrainte du déplacement (commun aux équations 4.5, 4.19 et 4.20), $J_{o,l}$ est défini comme le terme de droite de la fonction coût quadratique définie par l'Eq.4.20, $\mathcal{J}_{o,l}$ est défini par le terme de rappel aux observation de l'Eq.4.19, et \mathcal{J}_o est le terme de rappel aux observations de l'Eq.4.5 (c'est l'équivalent de $\mathcal{J}_{o,l}$ mais pour les champs à pleine résolution). Ces figures suivent les mêmes règles que les figures 4.6 et 4.8.

L'échec de la méthode semble dans ce cas provenir des effets de la surface sur la géométrie des coordonnées verticales hybrides qui dépendent de la pression de surface (cf section 1.1.1). La variable température est particulièrement touchée à cause de ses structures spatiales qui ont une grande similarité avec les aspérités du relief, même à des niveaux de moyenne et haute

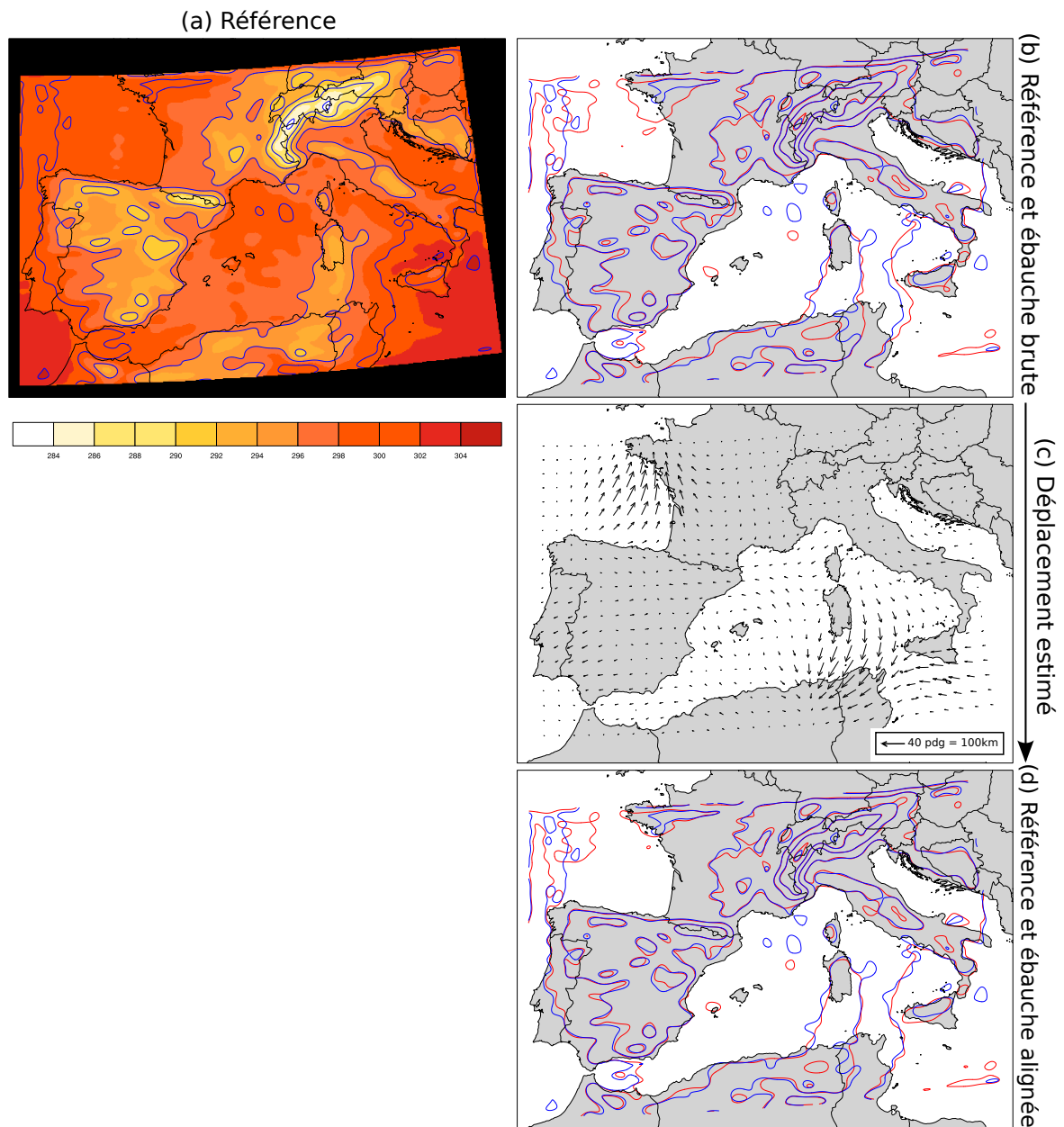


Figure 4.15. – Schéma d’alignement entre deux prévisions à 3h de T , au niveau 30 (≈ 600 hPa) qui sont des membres de l’AEARO-WMED. Les mêmes conventions que les figures 4.5 et 4.7) sont utilisées.

troposphère (cf Fig.4.15.a).

Dans la stratosphère, où l’influence de la surface n’est plus visible (les coordonnées hybrides suivent les isobares), ce comportement n’est plus observé (non montré). q est également touchée par cette incohérence mais semble l’être dans une moindre mesure (cf Fig.4.18), avec des zones où $EQ(\hat{\mathbf{d}}) - EQ(\mathbf{0})$ est positive qui ne coïncident pas avec les variations de relief ou le masque terre-mer.

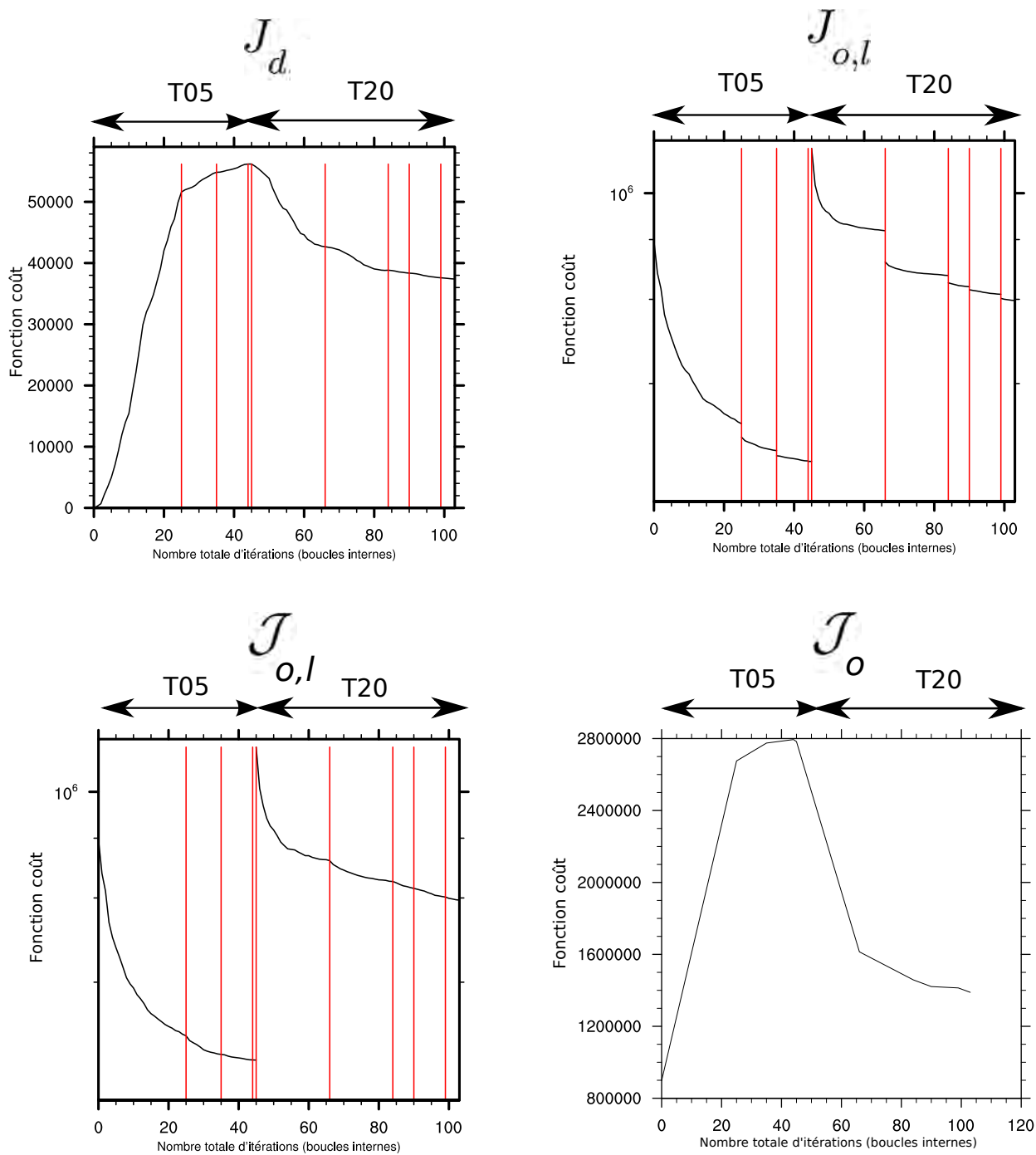


Figure 4.16. – Même représentation que la Fig.4.14 mais pour la variable T au niveau 30 (≈ 600 hPa).

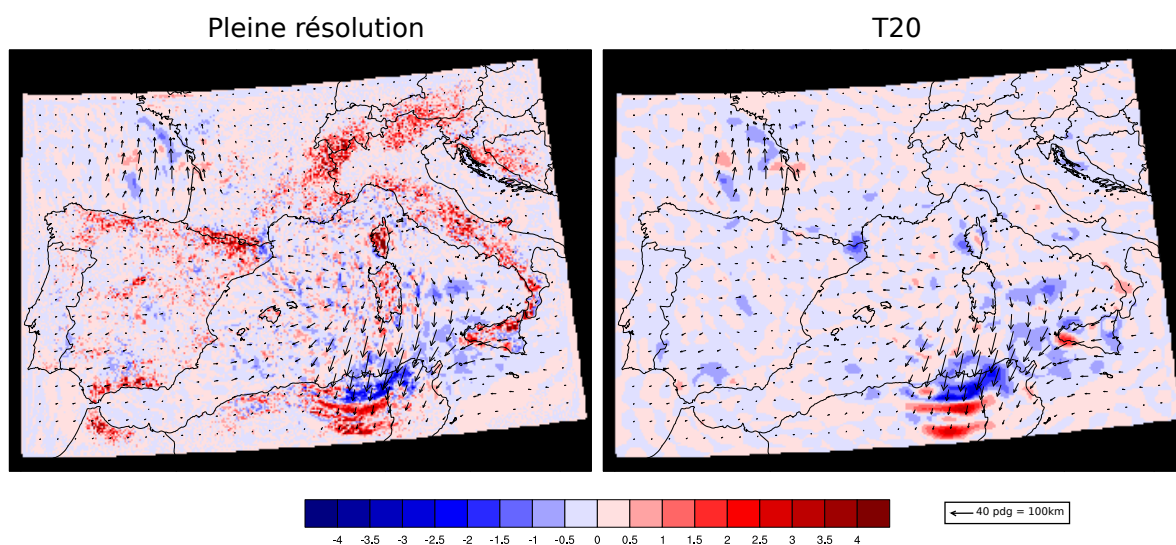


Figure 4.17. – Différences $EQ(\hat{\mathbf{d}}) - EQ(\mathbf{0})$ pour les champs bruts (gauche, "pleine résolution"), et les champs filtrés selon T20 (droite, "T20"). De même que pour les figures 4.5 et 4.7, une représentation vectorielle de $\hat{\mathbf{d}}$ est superposée. Les fortes valeurs positives sur le Nord-Est de l'Algérie sont dues au fait qu'elles appartiennent à la zone tampon qui n'est pas prise en compte lors de la minimisation. Résultats calculés pour la variable T , pour un des membres de l'AEARO-WMED au niveau modèle 30 (≈ 600 hPa).

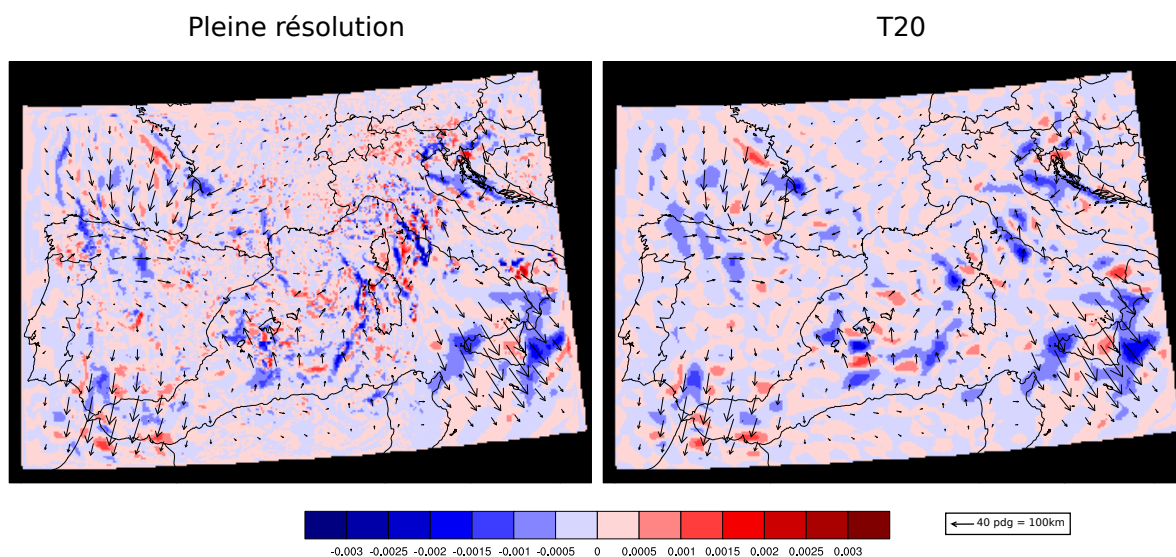


Figure 4.18. – Même représentation que la Fig.4.17, mais pour l'humidité spécifique q au niveau modèle 30 (≈ 600 hPa).

4.8. Conclusions

Dans ce chapitre, un système variationnel prenant en compte un modèle mixte d'erreurs d'ébauche (additives et de déplacement) est introduit. Grâce à une dérivation en deux étapes, les problèmes de correction d'erreurs d'amplitude (cf chapitre 1) et de déplacement peuvent être décomposés et résolus de manière séquentielle : les erreurs de déplacement sont corrigées au préalable des erreurs d'amplitude.

Parmi les études effectuées en PNT sur la correction des erreurs de déplacement, ce travail de thèse se rapproche des travaux de Nehrkorn et al. (2003) dans la formulation du problème variationnel avec un terme de contrainte formulé à partir d'une matrice de covariance d'erreurs de déplacement \mathbf{D} . Il se rapproche également de ceux de Nehrkorn et al. (2015) qui poussent plus loin l'analogie avec le problème de correction d'erreurs additives d'amplitude, afin de pouvoir utiliser des outils similaires lors de la minimisation. Il faut noter que, même pour des cas très simples tels que le premier cas test par exemple, les minimisations ont un nombre de boucles internes qui peut être important. La plus grande non-linéarité attendue de la fonction coût associée aux erreurs de déplacement \mathcal{J} (définie par l'Eq.4.5) par rapport à celle associée aux erreurs d'amplitude (définie par l'Eq.1.10) nécessite en effet des ré-actualisations plus fréquentes de ses modèles linéaires successifs.

La méthodologie présentée ici a tout d'abord été validée avec succès sur des cas idéalisés bidimensionnels. Pour des cas réels, la correction des erreurs de déplacement ne donne généralement pas lieu à des diminutions importantes des écarts entre la référence et l'ébauche. Plusieurs raisons peuvent être évoquées. La première est que les paramètres empiriques définissant la matrice \mathbf{D} sont probablement inadaptés aux champs à déplacer. Une approche statistique basée sur un ensemble de prévisions permettrait d'en avoir une estimation objective. Ensuite, cela est peut être tout simplement dû au fait que les erreurs de déplacement sont faibles entre les deux membres que l'on a aligné. Dans ce cas, l'extension de l'étude à d'autres situations où les erreurs de déplacement sont flagrantes devrait aboutir à de meilleurs résultats. Enfin les limitations dues aux effets du relief sont probablement trop importantes pour compenser les améliorations dans les autres parties du domaine. Pour cette deuxième hypothèse, l'application de la méthode dans une zone sans relief, ou sur des champs qui y sont moins sensibles, pourrait apporter un début de réponse.

Conclusions et perspectives

Cette thèse s'est attachée à utiliser des déformations spatiales dans un cadre ensembliste pour raffiner le modèle actuel d'erreur d'ébauche, soit en améliorant la représentation de la dépendance au flux des covariances d'erreurs d'amplitude, soit en corrigeant les erreurs de déplacement qui sont une des causes possibles de l'écartement à la Gaussianité des erreurs d'amplitude.

Le premier chapitre présente la dérivation habituelle des systèmes d'assimilation variationnelle où l'erreur d'ébauche est additive et suit une distribution Gaussienne (de moyenne nulle et avec une matrice de covariances \mathbf{B}). En suivant ce modèle additif Gaussien, le rôle de la matrice \mathbf{B} est essentiel puisqu'il est le seul opérateur à donner de l'information sur la confiance que l'on peut avoir vis-à-vis de l'ébauche. Ce chapitre met également en exergue que le modèle additif n'est pas restrictif en soi, mais qu'il n'est pas non plus le seul choix possible. Même si le modèle d'erreur utilisé n'est pas tenu de correspondre à la nature physique des sources d'erreurs (Lawson & Hansen, 2005), l'utilisation de modèles d'erreurs alternatifs peut rendre l'étape d'assimilation plus efficace et plus instructive sur les sources d'erreurs. Enfin, il est rappelé que la dérivation complète du problème d'assimilation avec une approche probabiliste dans un cadre Bayésien nécessite de faire des hypothèses sur la forme des PDFs d'erreurs pour aboutir au problème variationnel. L'hypothèse Gaussienne intervient donc très tôt, ce qui rend d'autant plus importante la discussion portant sur sa validité pour les systèmes de PNT.

La première partie du travail de thèse, qui a abouti à un article publié dans la revue *Tellus A*, a consisté à améliorer une modélisation de \mathbf{B} basée sur l'hypothèse diagonale spectrale, en prenant en compte objectivement l'anisotropie des corrélations d'erreurs d'ébauche à l'aide de déformations spatiales estimées statistiquement à partir d'un ensemble d'assimilation (Michel, 2013b,a). L'opérateur de déformation spatiale est d'abord estimé à partir d'un ensemble d'assimilation. Puis la déformation inverse est appliquée à un ensemble. L'efficacité de la déformation spatiale est quantifiée selon l'isotropisation et l'homogénéisation induites par la déformation inverse. Cette méthode statistique (ST) a été appliquée à une simulation réelle du modèle global de Météo-France ARPEGE, et ses résultats ont été comparés à ceux de la transformée géostrophique (GT Desroziers, 1997), qui vise à prendre en compte les effets de l'agéostrophisme. Les résultats montrent que la ST est plus efficace que la GT avec une prise en compte plus systématique de l'anisotropie des corrélations spatiales. On observe également que la ST, et dans une moindre mesure la GT, permettent également de diminuer l'hétérogénéité des anisotropies sur le domaine. Ensuite, il a été diagnostiqué que l'homogénéisation par la ST des longueurs de corrélation pour la partie basse de la troposphère n'est par contre pas vérifiée pour les niveaux supérieurs. Enfin, il a été observé que les structures spatiales des déformations estimées par la ST sont similaires pour les quatre variables de contrôle du modèle que sont la vortacité, la divergence, la température et l'humidité spécifique. Ceci pourrait justifier l'utilisation d'une unique déformation pour ces quatre variables.

La deuxième partie de ce travail de thèse, qui s'est également concrétisée par une publication dans le journal *Nonlinear Processes in Geophysics*, est une documentation de la non-Gaussianité (NG) diagnostiquée pour le modèle à échelle convective AROME-France. Les distributions d'erreurs additives de l'ébauche sont modélisées à l'aide d'un ensemble d'assimilation et leur NG est quantifiée selon K^2 , la valeur du test univarié de D'Agostino (D'agostino et al.,

1990). C'est un scalaire qui est défini pour chaque point de grille et qui est uniquement basé sur les paramètres de forme que sont l'asymétrie et l'aplatissement. Cette étude retrouve qu'en moyenne, pour la plupart des niveaux verticaux, c'est l'humidité spécifique qui s'éloigne le plus du modèle Gaussien, alors que la température, le vent zonal et méridien en sont proches au dessus de la couche limite. Dans cette dernière toutefois, la NG de ces variables augmente pour atteindre un maximum à la surface, notamment pour la température. Les structures de NG qui ont été observées ont de grandes similarités avec les structures météorologiques. De plus, il a été diagnostiqué que, durant un cycle d'assimilation, la NG augmente rapidement durant l'étape de propagation, notamment dans les zones nuageuses, et qu'elle diminue de manière significative lors de l'étape d'assimilation, notamment dans les zones fortement contraintes par les observations. Enfin, l'étude de la NG pour les variables de contrôle de l'assimilation a montré que les variables débalancées étaient plus Gaussiennes, et que ce sont les variables dynamiques, vorticité et divergence, qui sont les variables les moins Gaussiennes.

Dans ce contexte où les erreurs additives d'ébauche peuvent s'éloigner de l'hypothèse de Gaussianité, un modèle alternatif d'erreurs est investigué : les erreurs de déplacement. Plusieurs algorithmes de correction des erreurs de déplacement en PNT ont déjà été proposés dans la littérature. La méthode présentée dans le quatrième chapitre utilise une dérivation très proche de celle qui est employée pour la correction des erreurs d'amplitude. Avec un formalisme Bayésien, la méthode d'estimation et de correction des erreurs de déplacement est de type quasi-linéaire, et quasi-Gaussienne. Elle utilise également une formulation multi-échelles qui permet d'aligner les structures de grandes échelles avant de raffiner le champ de déplacement avec la prise en compte d'échelles plus fines. La méthode est utilisée dans le cadre idéalisé où l'on cherche à aligner deux champs modèles entre eux. Pour des cas simples de translation, la méthode réussit à estimer correctement les déplacements. Pour les champs réels AROME, les résultats sont plus mitigés avec une diminution de l'écart quadratique entre la référence et l'ébauche qui n'est pas systématique. Il a été montré que, pour certains champs comme la température, les effets de relief dus à la coordonnée verticale expliquent très probablement en partie ces résultats. Dans ces zones particulières en effet, le déplacement final devrait être imposé nul, ce qui n'est pas le cas dans la méthode actuelle.

Perspectives

La modélisation de l'anisotropie et des inhomogénéités des corrélations d'erreurs par des méthodes de déformations spatiales a été utilisée dans le chapitre 2 en complément d'une modélisation diagonale spectrale. Une première extension pourrait être de quantifier son apport lorsqu'elle est utilisée en complément de la modélisation diagonale ondelette des corrélations actuelles d'ARPEGE (Fisher, 2003; Varella et al., 2011; Berre et al., 2015). Dans un cadre ensembliste, une autre utilisation serait la modélisation d'une dépendance au flux des fonctions de localisation utilisées pour filtrer les covariances d'erreurs ensemblistes. Enfin, un troisième axe de recherche possible sur ce thème est l'extension de la méthode à des déformations spatiales tridimensionnelles. Pour ces trois pistes de travail, l'estimation de la déformation pourra soit être faite par la ST, soit par la méthode d'estimation des déformations spatiales à partir des tenseurs de corrélations définie par Michel (2013b) et Michel (2013a).

L'étude proposée au chapitre 3 se concentre sur la NG des erreurs de prévisions et d'analyses selon K^2 . Cet indice diagnostique uniquement la déviation par rapport au modèle Gaussien sans fournir plus d'informations sur la forme réelle des distributions testées. Une piste de travail serait de reprendre les travaux de Kliewer et al. (2015) qui testent également la ressemblance avec le modèle de distribution de type log-normal. L'utilisation de ces diagnostics peut être

vu comme le précurseur d'études plus poussées qui soit s'intéressent à des systèmes d'assimilation ne faisant pas intervenir l'hypothèse Gaussienne (e.g. Fletcher, 2007), soit s'orientent vers l'utilisation de transformations qui augmenteraient le caractère Gaussien des erreurs (e.g. Holm et al., 2002).

Enfin et en ce qui concerne la correction des erreurs de déplacement, une des perspectives à court terme serait de documenter les statistiques de ce type d'erreurs obtenues à partir d'un grand ensemble de champs déplacés grâce à l'algorithme proposé. Le choix de la surface sur laquelle les champs seront déplacés doit cependant être étudié plus en détail, notamment au niveau des zones de relief qui, comme nous l'avons vu, peuvent engendrer des problèmes méthodologiques. Outre le fait que cette étude permettrait une documentation tridimensionnelle et multivariée des erreurs de déplacement, elle pourrait également servir de base à une modélisation plus objective et plus optimale de ses covariances qui composent la matrice \mathbf{D} . La NG de l'ensemble aligné pourrait également être diagnostiquée de manière analogue à ce qui a été présenté au chapitre 3, ce qui permettrait de déterminer la part de NG des erreurs d'amplitude pouvant être éventuellement prise en compte en amont par une correction des erreurs de déplacement. L'implémentation d'une telle approche séquentielle dans un des systèmes opérationnels de PNT à Météo-France (ARPEGE ou AROME), appliquée à des observations réelles, serait facilitée par le fait que ces deux types de corrections se basent sur des algorithmes similaires utilisant les mêmes outils numériques. L'exploitation d'une matrice \mathbf{D} tridimensionnelle permettrait dans ce cadre d'aller plus loin que les méthodes existantes qui appliquent un déplacement bidimensionnel commun à toutes les variables et à tous les niveaux.

A | Annexes

Sommaire

A.1 Évolutions des systèmes opérationnels de Météo-France en 2015	115
A.1.1 ARPEGE	115
A.1.2 AROME	115

A.1. Évolutions des systèmes opérationnels de Météo-France en 2015

Dans cette annexe, sont rassemblées les principales évolutions apportées aux systèmes de PNT ARPEGE et AROME utilisés en opérationnel. Ce complément n'est donné qu'à titre informatif, puisqu'aucun de ces changements n'a été pris en compte pour ces travaux de thèse.

A.1.1. ARPEGE

La résolution des prévisions proposées par ARPEGE a augmenté avec une résolution horizontale de 7km sur la métropole et 35km aux antipodes (équivalent à une résolution spectrale T1198). La répartition verticale de la grille est maintenant faite sur 105 niveaux, avec le niveau le plus proche de la surface positionné à 10 mètres. L'échéance maximum de prévision a également augmentée, en passant de 102h à 114h.

Lors de chaque analyse, la minimisation est toujours effectuée avec 2 itérations dans la boucle externe, mais avec chacune 40 itérations dans la boucle interne (au lieu de 25 et 30 précédemment). La paramétrisation de l'approche multi-incrementale a également changé avec des résolutions de travail plus élevées : T149 dans la première boucle interne, et T399 dans la seconde.

Enfin, la modélisation de **B** a également évolué. Les variances du jour sont maintenant estimées avec une AEARP composée de 25 membres au lieu de 6, sur une grille ayant une résolution horizontale de 40km au lieu de 50km. De plus une modélisation prenant en compte une certaine dépendance au flux météorologique des corrélations d'erreurs d'ébauche est implémentée avec une formulation diagonale dans une base d'ondelettes (Fisher, 2003; Varella et al., 2011; Berre et al., 2015).

A.1.2. AROME

La résolution des prévisions a fortement augmenté en passant de 2.5km à 1.3km sur l'horizontale, avec en plus un domaine plus grand qui a été étendu de 20% vers le nord. Sur la verticale, AROME compte désormais 90 niveaux au lieu de 60, avec un premier niveau à 5m de la surface. La répartition des niveaux verticaux favorise toujours une meilleure résolution dans les basses couches. Le plafond du modèle a, lui, été descendu à ≈ 10 hPa. L'échéance maximum de prévision a été rallongée, en passant de 30h à un peu moins de deux jours (42h).

Les cycles d'assimilation ont été raccourcis en passant de 3h à 1h. Chaque analyse (une toute les heures) assimile donc toutes les observations contenues dans une fenêtre d'assimilation de ± 45 min autour de l'heure du réseau.

Bibliographie

- Aonashi, K. & Eito, H. (2011). Displaced ensemble variational assimilation method to incorporate microwave imager brightness temperatures into a cloud-resolving model. *Journal of the meteorological Society of Japan*, 89, 175–194.
- Auligné, T., Lorenc, A., Michel, Y., Montmerle, T., Jones, A., Hu, M., & Dudhia, J. (2011). Toward a new cloud analysis and prediction system. *Bulletin of the American Meteorological Society*, 92(2), 207–210.
- Bannister, R. N. (2008a). A review of forecast error covariance statistics in atmospheric variational data assimilation. I : Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134(637), 1951–1970.
- Bannister, R. N. (2008b). A review of forecast error covariance statistics in atmospheric variational data assimilation. II : Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, 134(637), 1971–1996.
- Beezley, J. D. & Mandel, J. (2008). Morphing ensemble kalman filters. *Tellus A*, 60(1), 131–140.
- Bénard, P., Vivoda, J., Mašek, J., Smolíková, P., Yessad, K., Smith, C., Brožková, R., & Geleyn, J.-F. (2010). Dynamical kernel of the aladin–nh spectral limited-area model : Revised formulation and sensitivity experiments. *Quarterly Journal of the Royal Meteorological Society*, 136(646), 155–169.
- Benjamin, S. G. (1989). An isentropic meso- α -scale analysis system and its sensitivity to aircraft and surface observations. *Monthly weather review*, 117(7), 1586–1603.
- Berre, L. (2000). Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly weather review*, 128(3), 644–667.
- Berre, L. & Desroziers, G. (2010). Filtering of background error variances and correlations by local spatial averaging : A review. *Monthly Weather Review*, 138(10), 3693–3720.
- Berre, L., Ștefănescu, S. E., & Belo Pereira, M. (2006). The representation of the analysis effect in three error simulation techniques. *Tellus A*, 58(2), 196–209.
- Berre, L., Varella, H., & Desroziers, G. (2015). Modelling of flow-dependent ensemble-based background-error correlations using a wavelet formulation in 4D-Var at météo-france. *Quarterly Journal of the Royal Meteorological Society*.
- Bocquet, M., Pires, C. A., & Wu, L. (2010). Beyond gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Review*, 138(8), 2997–3023.
- Bormann, N., Salonen, K., Peubey, C., McNally, T., & Lupu, C. (2012). An overview of the status of the operational assimilation of AMVs at ECMWF. In *Proc Eleventh Int Winds Workshop* (pp. 20–24).
- Brewster, K. A. (2003a). Phase-correcting data assimilation and application to storm-scale numerical weather prediction. part i : Method description and simulation testing. *Monthly weather review*, 131(3), 480–492.
- Brewster, K. A. (2003b). Phase-correcting data assimilation and application to storm-scale numerical weather prediction. part ii : Application to a severe storm outbreak. *Monthly weather review*, 131(3), 493–507.
- Brousseau, P., Berre, L., Bouttier, F., & Desroziers, G. (2011). Background-error covariances for a convective-scale data-assimilation system : Arome–france 3D-Var. *Quarterly Journal of the Royal Meteorological Society*, 137(655), 409–422.

- Brousseau, P., Desroziers, G., Bouttier, F., & Chapnik, B. (2014). A posteriori diagnostics of the impact of observations on the arome-france convective-scale data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 140(680), 982–994.
- Bubnová, R., Hello, G., Bénard, P., & Geleyn, J.-F. (1995). Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the arpege/aladin nwp system. *Monthly Weather Review*, 123(2), 515–535.
- Béréziat, D. & Herlin, I. (2008). Solving ill-posed image processing problems using data assimilation. application to optical flow. *rapport de recherche INRIA n° 6477*.
- Chaboureaud, J.-P., Nuissier, O., & Claud, C. (2012). Verification of ensemble forecasts of mediterranean high-impact weather events against satellite observations. *Natural Hazards and Earth System Science*, 12(8), 2449–2462.
- Clerc, M. & Mallat, S. (2002). The texture gradient equation for recovering shape from texture. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4), 536–549.
- Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., & Pailleux, J. (1998). The ECMWF implementation of three-dimensional variational assimilation (3D-Var). i : Formulation. *Quarterly Journal of the Royal Meteorological Society*, 124(550), 1783–1807.
- Courtier, P., Thépaut, J.-N., & Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519), 1367–1387.
- D’agostino, R. B., Belanger, A., & D’Agostino Jr, R. B. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4), 316–321.
- Dee, D. P. & Da Silva, A. M. (2003). The choice of variable for atmospheric moisture analysis. *Monthly weather review*, 131(1), 155–171.
- Dennis, J. J. E., Gay, D. M., & Walsh, R. E. (1981). An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software (TOMS)*, 7(3), 348–368.
- Derber, J. & Bouttier, F. (1999). A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus A*, 51(2), 195–221.
- Desroziers, G. (1997). A coordinate change for data assimilation in spherical geometry of frontal structures. *Monthly weather review*, 125(11), 3030–3038.
- Drobinski, P., Ducrocq, V., Alpert, P., Anagnostou, E., Béranger, K., Borga, M., Braud, I., Chanzy, A., Davolio, S., Delrieu, G., et al. (2014). Hymex : a 10-year multidisciplinary program on the mediterranean water cycle. *Bulletin of the American Meteorological Society*, 95(7), 1063–1082.
- Ehrendorfer, M. (2008). *Review of issues concerning ensemble-based data assimilation techniques*.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5), 10,143–10,162.
- Evensen, G. (2003). The ensemble kalman filter : Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4), 343–367.
- Fisher, M. (2003). Background error covariance modelling. In *Seminar on Recent Development in Data Assimilation for Atmosphere and Ocean* (pp. 45–63).
- Fisher, M. & Anderson, E. (2001). Developments in 4d-var and kalman filtering. *Techn. Rep*, 347.
- Fletcher, S. J. (2007). Implications and impacts of transforming lognormal variables into normal variables in var. *Meteorologische Zeitschrift*, 16(6), 755–765.
- Fourrié, N., Bresson, E., Nuret, M., Jany, C., Brousseau, P., Doerenbecher, A., Kreitz, M.,

- Nuissier, O., Sevault, E., Bénichou, H., Amodei, M., & Pouponneau, F. (2015). Arome-wmed, a real-time mesoscale model designed for the hymex special observation periods. *Geoscientific Model Development*, 8(7), 1919–1941.
- Gelb, A. (1974). *Applied optimal estimation*. MIT press.
- Gilbert, J. C. (2011). M1cg1—a solver of symmetric linear systems by conjugate gradient iterations, using/building a bfgs/c-bfgs preconditioner.
- Grassotti, C., Iskenderian, H., & N., H. R. (1999). Fusion of surface radar and satellite rainfall data using feature calibration and alignment. *Journal of Applied Meteorology*, 38, 677–695.
- Gratton, S., Lawless, A. S., & Nichols, N. K. (2007). Approximate gauss-newton methods for nonlinear least squares problems. *SIAM Journal on Optimization*, 18(1), 106–132.
- Gürol, S., Weaver, A., Moore, A., Piacentini, A., Arango, H., & Gratton, S. (2013). B-preconditioned minimization algorithms for variational data assimilation with the dual formulation. *Quarterly Journal of the Royal Meteorological Society*.
- Haben, S., Lawless, A., & Nichols, N. (2011). Conditioning of incremental variational data assimilation, with application to the met office system. *Tellus A*, 63(4), 782–792.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 13(49-52), 28.
- Haugen, J. E. & MACHENHAUER, B. (1993). A spectral limited-area model formulation with time-dependent boundary conditions applied to the shallow-water equations. *Monthly Weather Review*, 121(9), 2618–2630.
- Hoffman, R. N. & Grassotti, C. (1996). A technique for assimilating ssm/i observations of marine atmospheric storms : Tests with ECMWF analyses. *Journal of Applied Meteorology*, 35, 1177–1188.
- Hoffman, R. N., Liu, Z., Louis, J.-F., & Grassoti, C. (1995). Distortion representation of forecast errors. *Monthly Weather Review*, 123(9), 2758–2770.
- Hollingsworth, A. & Lönnberg, P. (1986). The statistical structure of short-range forecast errors as determined from radiosonde data. part i : The wind field. *Tellus A*, 38(2), 111–136.
- Holm, E., Andersson, E., Beljaars, A., Lopez, P., Mahfouf, J.-F., Simmons, A., & Thépaut, J.-N. (2002). Assimilation and modelling of the hydrological cycle : ECMWF’s status and plans. *European Centre for Medium-range Weather Forecasts Technical Memorandum*, (383).
- Holmlund, K. (1998). The utilization of statistical properties of satellite-derived atmospheric motion vectors to derive quality indicators. *Weather and Forecasting*, 13(4), 1093–1104.
- Horn, B. K. & Schunck, B. G. (1981). Determining optical flow. *Artificial intelligence*, 17(1), 185–203.
- Hoskins, B. J. (1975). The geostrophic momentum approximation and the semi-geostrophic equations. *Journal of the Atmospheric Sciences*, 32(2), 233–242.
- Houtekamer, P. L. & Mitchell, H. L. (2001). A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Review*, 129(1), 123–137.
- Ingleby, N. B. (2001). The statistical structure of forecast errors and its representation in the met. office global 3-d variational data assimilation scheme. *Quarterly Journal of the Royal Meteorological Society*, 127(571), 209–231.
- Joly, A. & Descamps, L. (2009). La prévision et l’utilisation des incertitudes en météorologie. *La Houille Blanche*, (6), 42–51.
- Kalnay, E. (2003). *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press.
- Kliwer, A., Fletcher, S., Jones, A., & Forsythe, J. (2015). Identifying non-normal and lognormal characteristics of temperature, mixing ratio, surface pressure, and wind for data

- assimilation systems. *Nonlinear Processes in Geophysics Discussions*, 2, 1363–1405.
- Kucukkaraca, E. & Fisher, M. (2006). Use of analysis ensembles in estimating flow-dependent background-error variances. *Techn. Rep*, 492.
- Lafore, J. P., Stein, J., Asencio, N., Bougeault, P., Ducrocq, V., Duron, J., Fischer, C., Hérel, P., Mascart, P., Masson, V., et al. (1998). The meso-nh atmospheric simulation system. part i : adiabatic formulation and control simulations. In *Annales Geophysicae*, volume 16 (pp. 90–109).
- Laprise, R. (1992). The euler equations of motion with hydrostatic pressure as an independent variable. *Monthly weather review*, 120(1), 197–207.
- Lawless, A. (2010). A note on the analysis error associated with 3d-fgat. *Quarterly Journal of the Royal Meteorological Society*, 136(649), 1094–1098.
- Lawless, A., Gratton, S., & Nichols, N. (2005). An investigation of incremental 4d-var using non-tangent linear models. *Quarterly Journal of the Royal Meteorological Society*, 131(606), 459–476.
- Lawson, W. G. & Hansen, James, A. (2005). Alignment error models and ensemble-based data assimilation. *Monthly Weather Review*, 133(6), 1687–1709.
- Le Dimet, F.-X. & Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects. *Tellus A*, 38(2), 97–110.
- Le Moigne, P., Boone, A., Calvet, J., Decharme, B., Faroux, S., Gibelin, A., Lebeaupin, C., Mahfouf, J., Martin, E., Masson, V., et al. (2009). Surfex scientific documentation. note de centre du groupe de météorologie moyenne echelle, 87, météo-france, cnrm, toulouse, france, 211 pp.
- Lee, S. & Salonen, K. (2015). Assessment of AMVs from COMS in the ECMWF system. *Techn. Rep*, 757.
- Leese, J. A., Novak, C. S., & Clark, B. B. (1971). An automated technique for obtaining cloud motion from geosynchronous satellite data using cross correlation. *Journal of applied meteorology*, 10(1), 118–132.
- Legrand, R. & Michel, Y. (2014). Modelling background error correlations with spatial deformations : a case study. *Tellus A*, 66.
- Legrand, R., Michel, Y., & Montmerle, T. (2015). Diagnosing non-gaussianity of forecast and analysis errors in a convective scale model. *Nonlinear Processes in Geophysics Discussions*, 2, 1061–1090.
- Leith, C. (1974). Theoretical skill of monte carlo forecasts. *Monthly Weather Review*, 102(6), 409–418.
- Lorenc, A. C. & Rawlins, F. (2005). Why does 4d-var beat 3d-var? *Quarterly Journal of the Royal Meteorological Society*, 131(613), 3247–3257.
- Lorenc, Andrew, C. (1986). Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112(474), 1177–1194.
- Martinet, P. (2013). *Apport des observations IASI pour la description des variables nuageuses du modèle AROME dans le cadre de la campagne HyMeX*. PhD thesis, Toulouse, INPT.
- Martinet, P., Fourrié, N., Guidard, V., Rabier, F., Montmerle, T., & Brunel, P. (2013). Towards the use of microphysical variables for the assimilation of cloud-affected infrared radiances. *Quarterly Journal of the Royal Meteorological Society*, 139(674), 1402–1416.
- Maz’ya, V. & Shaposhnikova, T. (2012). *Jacques Hadamard, un mathématicien universel*. EDP sciences.
- Mémin, E. (2003). *Estimation du flot-optique : contributions et panorama de différentes approches*. PhD thesis, L’université de Rennes 1, Institut de Formation Supérieur en Informatique et en Communication.

- Ménétrier, B. & Montmerle, T. (2011). Heterogeneous background-error covariances for the analysis and forecast of fog events. *Quarterly Journal of the Royal Meteorological Society*, 137(661), 2004–2013.
- Ménétrier, B., Montmerle, T., Berre, L., & Michel, Y. (2014). Estimation and diagnosis of heterogeneous flow-dependent background error covariances at convective scale using either large or small ensembles. *Quarterly Journal of the Royal Meteorological Society*.
- Michel, Y. (2013a). Estimating deformations of random processes for correlation modelling in a limited area model. *Quarterly Journal of the Royal Meteorological Society*, 139(671), 534–547.
- Michel, Y. (2013b). Estimating deformations of random processes for correlation modelling : methodology and the one-dimensional case. *Quarterly Journal of the Royal Meteorological Society*, 139(672), 771–783.
- Michel, Y., Auligné, T., & Montmerle, T. (2011). Heterogeneous convective-scale background error covariances with the inclusion of hydrometeor variables. *Monthly Weather Review*, 139(9), 2994–3015.
- Montmerle, T. & Berre, L. (2010). Diagnosis and formulation of heterogeneous background-error covariances at the mesoscale. *Quarterly Journal of the Royal Meteorological Society*, 136(651), 1408–1420.
- Montmerle, T. & Faccani, C. (2009). Mesoscale assimilation of radial velocities from doppler radars in a preoperational framework. *Monthly Weather Review*, 137(6), 1939–1953.
- Ménétrier, B. (2014). *Utilisation d'une assimilation d'ensemble pour modéliser des covariances d'erreur d'ébauche dépendantes de la situation météorologique à échelle convective*. PhD thesis, Institut National Polytechnique de Toulouse.
- Nehrkorn, T., Hoffman, R. N., Grassotti, C., & Louis, J.-F. (2003). Feature calibration and alignment to represent model forecast errors : Empirical regularization. *Quarterly Journal of the Royal Meteorological Society*, 129(587), 195–218.
- Nehrkorn, T., Woods, B., Auligné, T., & Hoffman, R. N. (2014). Application of feature calibration and alignment to high-resolution analysis : Examples using observations sensitive to cloud and water vapor. *Monthly Weather Review*, 142(2), 686–702.
- Nehrkorn, T., Woods, B. K., Hoffman, R. N., & Auligné, T. (2015). Correcting for position errors in variational data assimilation. *Monthly Weather Review*, 143(4), 1368–1381.
- Nieman, S. J., Menzel, W. P., Hayden, C. M., Gray, D., Wanzong, S. T., Velden, C. S., & Daniels, J. (1997). Fully automated cloud-drift winds in nesdis operations. *Bulletin of the American Meteorological Society*, 78(6), 1121–1133.
- Nieman, S. J., Schmetz, J., & Menzel, W. P. (1993). A comparison of several techniques to assign heights to cloud tracers. *Journal of Applied Meteorology*, 32(9), 1559–1568.
- Nocedal, J. & Wright, S. (2006). Numerical optimization, series in operations research and financial engineering. *Springer, New York*.
- Noilhan, J. & Mahfouf, J.-F. (1996). The isba land surface parameterisation scheme. *Global and planetary Change*, 13(1), 145–159.
- Pailleux, J., Geleyn, J.-F., & Legrand, E. (2000). La prévision numérique du temps avec les modèles arpège et aladin-bilan et perspectives.
- Pannekoucke, O., Berre, L., & Desroziers, G. (2007). Filtering properties of wavelets for local background-error correlations. *Quarterly Journal of the Royal Meteorological Society*, 133(623), 363–379.
- Parrish, D. F. & Derber, J. C. (1992). The national meteorological center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, 120(8), 1747–1763.
- Pereira, M. B. & Berre, L. (2006). The use of an ensemble approach to study the background error covariances in a global nwp model. *Monthly weather review*, 134(9), 2466–2489.

- Purser, R. J., Wu, W.-S., Parrish, D. F., & Roberts, N. M. (2003). Numerical aspects of the application of recursive filters to variational statistical analysis. Part I : Spatially homogeneous and isotropic Gaussian covariances. *Monthly Weather Review*, 131(8), 1524–1535.
- Radnóti, G. (1995). Comments on “a spectral limited-area formulation with time-dependent boundary conditions applied to the shallow-water equations”. *Monthly weather review*, 123(10), 3122–3123.
- Rasmussen, C. E. & Williams, C. (2006). Gaussian processes for machine learning.
- Ravela, S. (2012). Quantifying uncertainty for coherent structures. *Procedia Computer Science*, 9, 1187–1196.
- Ravela, S., Emanuel, K., & McLaughlin, D. (2007). Data assimilation by field alignment. *Physica D : Nonlinear Phenomena*, 230(1), 127–145.
- Raynaud, L. (2010). *Application, validation et réglage d’une assimilation d’ensemble*. PhD thesis, Université Toulouse III - Paul Sabatier.
- Raynaud, L., Berre, L., & Desroziers, G. (2011). An extended specification of flow-dependent background error variances in the météo-france global 4D-Var system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 607–619.
- Raynaud, L., Berre, L., & Desroziers, G. (2012). Accounting for model error in the météo-france ensemble data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 138(662), 249–262.
- Rossby, C.-G. (1937). Isentropic analysis. *Bulletin of American Meteorological Society*, 18, 201–209.
- Sangster, W. E. (1960). A method of representing the horizontal pressure force without reduction of station pressures to sea level. *Journal of Meteorology*, 17(2), 166–176.
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., & Masson, V. (2011). The arome-france convective-scale operational model. *Monthly Weather Review*, 139(3), 976–991.
- Semple, A. (2001). A meteorological assessment of the geostrophic co-ordinate transform and error breeding system when used in 3d variational data assimilation. *NWP Tech Rep*, 357.
- Senkova, A., Rontu, L., & Savijärvi, H. (2007). Parametrization of orographic effects on surface radiation in hirlam. *Tellus A*, 59(3), 279–291.
- Sherman, J. & Morrison, W. J. (1949). Adjustment of an inverse matrix corresponding to changes in a given column or a given row of the original matrix. *The Annals of Mathematical Statistics*, 20(4), 620–624.
- Sherman, J. & Morrison, W. J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1), 124–127.
- Simmons, A. & Strüfing, R. (1981). *An energy and angular momentum conserving finite-difference scheme, hybrid coordinates and medium-range weather prediction*. European Centre for Medium Range Weather Forecasts.
- Snyder, C., Bengtsson, T., Bickel, P., & Anderson, J. (2008). Obstacles to high-dimensional particle filtering. *Monthly Weather Review*, 136(12), 4629–4640.
- Souopgui, I., Le Dimet, F.-X., & Vidard, A. (2012). From regularization to preconditioning, application to motion estimation.
- Tarantola, A. (2005). *Inverse problem theory and methods for model parameter estimation*.
- Thiebaut, H., Julian, P., & DiMego, G. (1990). Areal versus collocation data quality control. In *Intl. Symp. on Assimilation of Observations in Meteorology and Oceanography, Clermont-Ferrand, France, WMO* (pp. 255–260).
- Tikhonov, A. N. (1963). Regularization of incorrectly posed problems. In *Soviet Math. Dokl*,

- volume 4 (pp. 1624–1627).
- Trémolet, Y. (2006). Accounting for an imperfect model in 4d-var. *Quarterly Journal of the Royal Meteorological Society*, 132(621), 2483–2504.
- Trémolet, Y. (2007). Incremental 4D-Var convergence study. *Tellus A*, 59(5), 706–718.
- Van Leeuwen, P. J. (2009). Particle filtering in geophysical systems. *Monthly Weather Review*, 137(12), 4089–4114.
- Varella, H., Berre, L., & Desroziers, G. (2011). Diagnostic and impact studies of a wavelet formulation of background-error correlations in a global model. *Quarterly Journal of the Royal Meteorological Society*, 137(658), 1369–1379.
- Veersé, F. & Thépaut, J.-N. (1998). Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 124(550), 1889–1908.
- Wattrelot, E., Caumont, O., & Mahfouf, J.-F. (2014). Operational implementation of the 1d+3d-var assimilation method of radar reflectivity data in the arome model. *Monthly Weather Review*, 142(5), 1852–1873.
- Zupanski, D. (1997). A general weak constraint applicable to operational 4DVAR data assimilation systems. *Monthly Weather Review*, 125(9), 2274–2292.