HAL
open science

Mise en œuvre, diagnostic et optimisation des schémas d'assimilation de données

Gérald Desroziers

To cite this version:

Gérald Desroziers. Mise en œuvre, diagnostic et optimisation des schémas d'assimilation de données. Mathématiques [math]. Université Paul Sabatier - Toulouse III, 2007. tel-00525615

HAL Id: tel-00525615

https://theses.hal.science/tel-00525615

Submitted on 12 Oct 2010

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

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

Mémoire pour l'obtention du diplôme d'habilitation à diriger des recherches

présenté à

l'Université Paul Sabatier

par

Gérald Desroziers

MISE EN OEUVRE, DIAGNOSTIC ET OPTIMISATION DES SCHEMAS D'ASSIMILATION DE DONNEES

Habilitation soutenue le 11 Mai 2007

devant le jury composé de :

M. Mohamed Masmoudi	Président
M. Olivier Talagrand	Directeur de Recherche
M. Pierre Brasseur	Rapporteur
M. Jean-Noël Thépaut	Rapporteur
M. Robert Vautard	Rapporteur
M. Jean Pailleux	Examinateur

Remerciements

Je tiens à remercier tout particulièrement Olivier Talagrand, dont les recherches ont été une source d'inspiration majeure pour un certain nombre des travaux présentés dans ce mémoire, et aussi pour avoir accepté d'encadrer cette Habilitation à Diriger des Recherches.

Je remercie François Bouttier, Alain Joly et Pierre Bénard qui m'ont encouragé à soutenir cette habilitation.

Je suis également reconnaissant à Eric Brun et Joël Poitevin pour avoir soutenu et facilité mon habilitation au sein du CNRM.

Un grand merci à Jean Maziejewski qui, par sa disponibilité et son efficacité, a joué un rôle essentiel dans l'organisation de mon habilitation, en particulier pour préparer la soutenance.

Enfin que soient remerciés tous ceux qui de près ou de loin ont contribué à cette habilitation et notamment mes collègues du GMAP.

Table des matières

1 Curriculum Vitae **7**
1.1 Position actuelle . . . 7
1.2 Expérience professionnelle . . . 7

2 Liste des travaux et publications **9**
2.1 Articles publiés dans des revues à comité de lecture . . . 9
2.2 Contributions à des congrès ayant donné lieu à des actes . . . 10
2.3 Contributions à des congrès sans actes . . . 11
2.4 Rapports . . . 13
2.5 Activités d'encadrement . . . 14
2.5.1 Mémoires d'étudiants ayant permis de soutenir des diplômes . . . 14
2.5.2 Autres stages accompagnés d'un mémoire . . . 15
2.5.3 Thèses . . . 15
2.6 Activités d'enseignement . . . 15
2.7 Activités diverses . . . 16

3 Synthèse des travaux de recherche **17**
3.1 Introduction . . . 17
3.2 Formalisme de l'estimation linéaire statistique . . . 19
3.3 Changement de coordonnées géostrophiques . . . 21
3.4 Diagnostics d'optimalité du problème variationnel . . . 27
3.5 Ré-analyse 4D-Var de l'expérience FASTEX . . . 39
3.6 Mesure de l'impact des observations sur les analyses et les prévisions . . . 45
3.7 Diagnostics de cohérence dans l'espace des observations . . . 51
3.8 Utilisation des ondelettes pour les covariances d'erreur d'ébauche . . . 59
3.9 Perspectives . . . 63

4 Publications essentielles **73**

1 Curriculum Vitae

Gérald Desroziers
Né le 16 mai 1955 à Casablanca

Diplômes

1972 : Baccalauréat série C
1978 : Diplôme d'Ingénieur des Travaux de la Météorologie

1.1 Position actuelle

Ingénieur Divisionnaire des Travaux, dans l'équipe ALGO, dirigée par P. Bénard, du Groupe de Modélisation pour l'Assimilation et la Prévision (GMAP) du Centre National de Recherches Météorologiques (CNRM) de Météo-France.

1.2 Expérience professionnelle

De 1979 à 1982 :

Service "Etudes spéciales" (ES), à Paris, du Service Météorologique Métropolitain (SMM). Service dirigé par J.C. Tuffigo puis B. Strauss.

• Chargé d'études.

De 1983 à 1986 :

Equipe "Moyens Mobiles de Mesures Météorologiques" (4M), à Toulouse, du Centre National de Recherches Météorologiques (CNRM). Equipe dirigée par M. Payen.

• Participation et dépouillement de campagnes de mesures (FOS 83, FRONTS 84, HAPEX 86, FRONTS 87).

De 1987 à 1994 :

Equipe "Mésoéchelle Assimilation de Données" (MAD), à Toulouse, du CNRM. Equipe dirigée par M. Payen (jusqu'en 1988), puis par P. Bernardet.

• Mise en place de méthodes d'interpolation de données à l'aide de fonctions splines.

• Réglage des paramètres de ces interpolateurs splines par des méthodes statistiques (Validation Croisée).

• Développement du concept d'analyse dans l'espace géostrophique pour améliorer l'analyse des fronts.

• Ré-analyse de la campagne FRONTS 87 à l'aide du schéma par Interpolation Optimale CANARI associé au modèle global ARPEGE.

• Etude de la structure des erreurs de prévision du modèle ARPEGE en mode étiré.

• Intégration du changement de coordonnées géostrophiques dans ARPEGE.

De 1995 à 2001 :

Equipe "Recherche sur le Cyclogénèses et les Fronts" (RECYF), à Toulouse, du Groupe de Météorologie de Moyenne Echelle (GMME), du CNRM. Equipe dirigée par A. Joly.

• Préparation de l'expérience FASTEX.

• Application du 3D-Var ARPEGE à la ré-analyse du jeu de dropsondes de l'expérience FRONTS 92.

• Participation à la campagne FASTEX.

• Participation au développement du 4D-Var ARPEGE.

• Développement du concept d'erreur de représentativité due à la formulation incrémentale du 3D-Var ARPEGE.

• Concept d'optimalité de l'analyse et réglage des paramètres d'erreur.

• Ré-analyse 4D-Var de l'expérience FASTEX.

• Développement d'un opérateur d'observation "tourbillon potentiel".

De 2002 à 2006 :

Equipe ALGO, du Groupe de Modélisation pour l'Assimilation et la Prévision (GMAP), du Centre National de Recherches Météorologiques (CNRM), à Toulouse. Equipe dirigée par P. Bénard.

• Optimisation du schéma d'analyse 4D-Var.

• Réglage des statistiques d'erreur dans une assimilation variationnelle.

• Optimisation de l'étirement dans l'analyse ARPEGE.

• Modification de l'algorithme de minimisation.

• Mise en place d'une assimilation 3D-Var FGAT à "cut-off" très court.

• Travail sur la trajectoire de linéarisation du 4D-Var.

• Optimisation des Filtres digitaux dans le 4D-Var ARPEGE.

• Diagnostic de l'impact des observations sur les analyses et les prévisions.

• Utilisation des ondelettes pour les statistiques d'erreur de prévision.

• Diagnostic des statistiques d'erreur dand l'espace des observations.

2 Liste des travaux et publications

2.1 Articles publiés dans des revues à comité de lecture

Desroziers, G. and Lafore, J. P., 1993 : A coordinate transformation for objective frontal analysis, Mon. Wea. Rev., 121, 1531-1553.

Desroziers, G., 1997 : A coordinate change for data assimilation in spherical geometry of frontal structures, Mon. Wea. Rev., 125, 3030-3038.

Desroziers, G., Pouponneau, B., Thépaut, J.-N., Janiskova, M., and Veersé, F., 1999 : 4D-Var analyses of FASTEX situations using special observations, Q. J. R. Meteorol. Soc., 125, 3393-3414.

Cammas, J.-P., Pouponneau, B., **Desroziers**, G., and Santurette, P., 1999 : Triggering and first development phases of the FASTEX cyclone IOP 17 : an introductory survey from field data and operational analyses, Q. J. R. Meteorol. Soc., 125, 3393-3414.

Janiskova, M., Veersé, F., Thépaut, J.-N., **Desroziers**, G., and Pouponneau, B., 1999 : Impact of a simplified physical package in Four dimensional variational analyses of FASTEX situations, Q. J. R. Meteorol. Soc., 125, 2465-2485.

Desroziers, G. and Ivanov, S., 2001 : Diagnosis and adaptive tuning of information error parameters in a variational assimilation, Q. J. R. Meteorol. Soc., 127, 1433-1452.

Desroziers, G., Brachemi, O., and Hamadache, B., 2001 : Estimation of the representativeness error due to the incremental formulation of variational data assimilation, Q. J. R. Meteorol. Soc., 127, 1775-1794.

Desroziers, G., Hello, G., and Thépaut, J.-N., 2003 : A 4D-Var re-analysis of FASTEX, Q. J. R. Meteorol. Soc., 129, 1301-1315.

Chapnik, B., **Desroziers**, G., Rabier, F. and Talagrand, 0., 2004 : Properties and first application of an error statistics tuning method in a variational assimilation, Q. J. R. Meteorol. Soc., 130, 2253-2275.

Desroziers, G., Brousseau, P., and Chapnik, B., 2005 : Use of randomization to diagnose the impact of observations on analyses and forecasts, Q. J. R. Meteorol. Soc., 131, 2821-2837.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P., 2005 : Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. R. Meteorol. Soc., 131, 3385-3396.

Chapnik, B., **Desroziers**, G., Rabier, F., and Talagrand, O., 2006 : Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting, Q. J. R. Meteorol. Soc., 132, 543-565.

Fourrié, N., Marchal, D., Rabier, F., Chapnik, B., and **Desroziers**, G., 2006 : Impact study of the 2003 North Atlantic THORPEX Regional Campaign, Q. J. R. Meteorol. Soc., 132, 275-295.

Guerin, R., **Desroziers**, G., and Arbogast, P., 2006 : 4D-Var analysis of pseudo potential-vorticity observations, Q. J. R. Meteorol. Soc., 132, 1283-1298.

Pannekoucke, O., Berre, L. and **Desroziers**, G., 2006 : Wavelets expansion and its use as an adaptive filter, Q. J. R. Meteorol. Soc., Under revision.

Berre, L., Pannekoucke, O., **Desroziers**, G., and Martel, C., 2006 : A spectral diagnosis and filtering of sampling noise in ensemble background error standard deviations, Q. J. R. Meteorol. Soc., Under revision.

Poli, P., Moll, P., Rabier, F., **Desroziers**, G., Chapnik, B., Berre, L., Healy, S.B., Andersson, E., and El Guelai, F.-Z., 2006 : Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. Journal of Geophysical Research, Accepted for publication.

2.2 Contributions à des congrès ayant donné lieu à des actes

Desroziers, G., Bernardet, P., and Amodéi, L., 1990 : Analyse par fonctions splines des sondages verticaux de l'expérience Fronts 87. Atelier de Modélisation de l'Atmosphère, Toulouse.

Desroziers, G., Mathiot, V., and Orain, F., 1995 : A study of the ARPEGE forecast error covariances. Second WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, 13-17 March, Tokyo.

Desroziers, G., 1995 : Intégration dans ARPEGE d'une transformation dynamique des champs pour l'assimilation de données. Atelier de Modélisation de l'Atmosphère, 28-30 novembre, Toulouse.

Desroziers, G., Pouponneau, B., Janiskova, M., Thépaut, J.-N., Hello, G., and Wattrelot, E., 1999 : 4D-Var re-analysis of FASTEX data, Third WMO Symposium on Data Assimilation, 7-11 June, Québec.

Desroziers, G., Pouponneau, B., Janiskova, M., and Thépaut, J.-N., 1999 : 4D-Var re-analysis of FASTEX data, 8th Conference on Mesoscale Processes, 28 June-2 July, Boulder, published by the American Meteorological Society, 328-331.

Desroziers, G., Hello, G., Pouponneau, B., and Thépaut, J.-N., 1999 : Application du 4D-Var incrémental à la ré-analyse de l'expérience FASTEX. Atelier de Modélisation de l'Atmosphère, 30 novembre-2 décembre, Toulouse.

Thépaut, J.-N., Gauthier, P., Janiskova, M., Veersé, F., Desroziers, G., Moll, P., Bouyssel, F., Pouponneau, B., Hello, G., Torki, M., Peuch, A., and El Khatib, R., 1999 : Towards an operational implementation of a 4D-Var system at Météo-France. Proceedings of the 7th International ALADIN Workshop, 17-19 november, Ljubljana.

Desroziers, G., and Ivanov., S., 2000 : Réglage des paramètres d'erreur d'information dans une analyse variationnelle. Atelier de Modélisation de l'Atmosphère, 30 novembre-2 décembre, Toulouse.

Desroziers, G., Chapnik, B., Ivanov., S., and Rabier, F., 2001 : Adaptive tuning of observation error parameters in a variational data assimilation. 14th Conference on Numerical Weather Prediction, 30 July- 2 August, Fort Lauderdale, Florida.

Desroziers, G., Berre, L. Belo Pereira, M., 2002 : Etude de la variation géographique des corrélations d'erreur de prévision du modèle ARPEGE et de leur dépendance à la résolution locale du modèle. Atelier de Modélisation de l'Atmosphère, 17-19 décembre, Toulouse.

Belo Pereira, M., Berre, L., and Desroziers, 2002 : Estimation et étude des caovariances d'erreur de prévision d'ARPEGE/ALADIN. Atelier de Modélisation de l'Atmosphère, 17-19 décembre, Toulouse.

Berre, L., Boloni, G., Brozkova, R., Cassé, V., Desroziers, G., Fischer, C., Geleyn, J.-F., Horanyi, A., Raouindi, M., Sadiki, W., and Siroka, M., 2002 : Background error covariances for the 3D-Var of the ALADIN LAM. Hirlam Workshop on Variational Data Assimilation and Remote Sensing, 21-23 January, Helsinki.

Chapnik, B., Desroziers, G., Rabier, F., Talagrand, O., 2002 : Réglage de statistqiues d'erreurs dans une assimilation variationnelle. Atelier de Modélisation de l'Atmosphère, 17-19 décembre, Toulouse.

Desroziers, G., Berre, L., Chapnik, B. and Poli, P., 2005 : A simple method to diagnose and adapt observation and background errors. 4th WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, 18-22 April, Prague.

Desroziers, G., Brousseau, P., Chapnik, B., and Fourrié, N., 2005 : Use of randomization to diagnose the impact of observations on analyses and forecasts. 4th WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, 18-22 April, Prague.

Chapnik, B., Desroziers, G., Rabier, F. and Talagrand, O., 2005 : Evaluation of observations impact and tuning of their error statistics. 4th WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, 18-22 April, Prague.

Pannekoucke, O., Berre, L., and Desroziers, G., 2005 : Properties of wavelet expansions for the representation of heterogeneous background error covariances. 4th WMO International Symposium on Assimilation of Observations in Meteorology and Oceanography, 18-22 April, Prague.

2.3 Contributions à des congrès sans actes

Amodéi, L., Bernardet, P., Delage, C., and Desroziers, G., 1988 : Objective analysis of meteorological fields using spline functions : II. Linear constraints applied to horizontal wind determination. XIII EGS General Assembly, Bologna.

Amodéi, L., and Desroziers, G., 1989 : Smoothing spines applied to mesoscale analysis of cold fronts observations. XIV EGS General Assembly, Bologna.

Delage, C., and Desroziers, G., 1989 : Multidimensional spines applied to mesoscale analysis of cold fronts observations. V EGS IAMAP Assembly, Reading.

Desroziers, G., and Lafore, P., 1992 : A frontal analysis in geostrophic coordinates. XVII EGS General Assembly, Edinburgh.

Desroziers, G., 1994 : Variational frontal analysis including a coordinate transformation. 2 nd Workshop on Adjoint Applications in Dynamic Meteorology, Visegrad, Hungary.

Bernardet, P., Desroziers, G., 1994 : Frontal analysis and semi-geostrophic theory. Sixth Conference on Mesoscale Processes, 18-22 July, Portland, Oregon.

Desroziers, G., Nechad, B., Sadiki, W., and Thépaut, J.-N., 1996 : Variational Assimilation of the FRONTS 92 Dopsonde Dataset. 7th Conference on Mesoscale Processes, 9-13 September, Reading, England.

Desroziers, G., Nechad, B., Sadiki, W., and Thépaut, J.-N., 1997 : Variational Assimilation of the FRONTS 92 Dopsonde Dataset. 3rd Workshop on Adjoint Applications in Dynamic Meteorology, 16-20 June, Lennoxville, Canada.

Desroziers, G., Pouponneau, B., Thépaut, J.-N., Janiskova, M., and Veersé, F., 1998 : 4D-Var reanalysis of the FASTEX data. FASTEX Workshop, 27-30 April, Toulouse.

Bergot, T., Desroziers, G., Pouponneau, B., and Hello, G., 1998 : Adjoint-based targeted observations during FASTEX : impact and feasability studies. FASTEX Workshop, 27-30 April, Toulouse.

Bergot, T., Desroziers, G., Pouponneau, B., and Hello, G., 1998 :Adaptive observations during FASTEX : presentation and first results. Rossby-100 Symposium, 8-12 June, Stockholm.

Desroziers, G., and Pouponneau, B., 1998 : L'assimilation 4D-Var et FASTEX. Journée FASTEX, 20 novembre 1998, Trappes.

Desroziers, G., Pouponneau, B., Janiskova, M., Thépaut, J.-N., Hello, G., and Wattrelot, E., 1999 : FASTEX 4D-Var re-analysis. IUGG99, 19-23 July, Birmingham.

Desroziers, G., and Ivanov. S., 2000 : A posteriori diagnostics and adative tuning of error statistics. 4th Workshop on Adjoint Applications in Dynamic Meteorology, 26-30 June, Moliets.

Desroziers, G., and Arbogast, P., 2001 : Introduction of a potential vorticity operator in the French ARPEGE 4D-Var. 5th Workshop on Adjoint Applications in Dynamic Meteorology, 21-26 April, Tuscarora, Pennsylvania.

Desroziers, G., Chapnik, B., Rabier, F., and Talagrand, O., 2002 : Tuning of observation error parameters in a variational data assimilation. IMA Workshop on Data Assimilation in the Oceanic and Atmospheric Sciences. April 29-May 3, Minneapolis, Minnesota.

Desroziers, G., 2003 : Assimilation of meteorological observations at Météo-France. Applied Mathematics and Applications of Mathematics First Joint Conference, 10-13 February, Nice.

Desroziers, G., and Chapnik, B., 2005 : Use of randomization to diagnose the impact of observations on analyses and forecasts. 6th Workshop on Adjoint Applications in Dynamic Meteorology, 23-28 May, Acquafredda di Maratea.

Chapnik, B., Desroziers, G., Rabier, F. and Talagrand, O., 2004 : A new randomised computation method of the degree of freedom for signal : application to the tuning of error variances in data assimilation. 6th Workshop on Adjoint Applications in Dynamic Meteorology, 23 -28 May, Acquafredda di Maratea.

Desroziers, G., 2005 : ARPEGE 4D-VAR multi-incremental. ALADIN-HIRLAM workshop on code maintenance and data assimilation, Budapest, 14-18 November.

Desroziers, G., 2005 : Minimizers for variational data assimilation. ALADIN-HIRLAM workshop on code maintenance and data assimilation, Budapest, 14-18 November.

Desroziers, G., Berre, L., Chapnik, B., Loo, C. and Yessad, K., 2006 : Large least-squares problems in Meteorology. Sparse Days at CERFACS Meeting, June 15-16, Toulouse.

Desroziers, G., and Berre, L., 2006 : Verification of ensemble background error variance in observation space. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

Berre, L., Pannekoucke, O., Desroziers, G., Martel, C., and Stefanescu, S., 2006 : Diagnosis and filtering of sampling noise in ensemble background error standard deviations. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

Pannekoucke, O., Berre, L., and Desroziers, G., 2006 : Filtering properties of wavelets for the local background error correlations. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

2.4 Rapports

Desroziers, G., 1986 : Présentation d'un modèle d'analyse variationnelle des champs de vent à moyenne échelle. Application à l'expériance FOS 83. Note de travail de l'EERM No 169.

Bessemoulin, P., Desroziers, G., Payen, M., et Tarrieu, C., 1987 : Atlas des données Samer. Note de travail de l'EERM.

Desroziers, G., Bernardet, P., et Delage, C., 1990 : Utilisation des fonctions splines pour l'analyse de données. Note de travail du CNRM No 16.

Desroziers, G., 1993 : Rénalyse ARPEGE/CANARI de l'expérience FRONTS 87. Note de travail du GMME No 22.

Desroziers, G., Mathiot, V., Orain, F., and Bernardet, P., 1995 : Estimation locale des covariances des erreurs de prévision d'un modèle spectral sur la sphère : application au modèle à résolution variable ARPEGE. Note de travail du GMME No 34.

Desroziers, G., Nechad, B., Sadiki, W., and Thépaut, J.-N., 1997 : Analyse variationnelle du réseau de dropsondes de l'expéreince FRONTS 92 : application du 3D-Var ARPEGE et discussion de l'erreur due à la formulation incrémentale de l'analyse à partir d'une maquette 1D-Var sur le cercle. Note de travail du GMME, No 53.

2.5 Activités d'encadrement

2.5.1 Mémoires d'étudiants ayant permis de soutenir des diplômes

Ohl, P., et Testard, J.-M., 1986 : Méthodes d'interpolation d'un champ de vent à mésoéchelle. Etude d'un modèle d'analyse variationnelle. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, mai 1986.

Larichi, A., 1989 : Application du Krigeage à l'analyse d'un champ météorologique à mésoéchelle. Rapport de Stage du DESS Méthodes Informatiques et Modèles Mathématiques de l'Université Paul Sabatier, juin 1989.

Sombret, D., 1993 : Formulation et étude d'un changement de coordonnées pour l'analyse des discontinuités frontales. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, juin 1993.

Nechad, B., et Sadiki, W., 1996 : Analyse variationnelle à méso-échelle de l'expérience FRONTS 92. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, juin 1996 (co-encadrement).

Brachemi, O., Hamadache, B., 1997 : Estimation de l'erreur de représentativité due à la formulation incrémentale de l'analyse 3D-Var/ARPEGE. Rapport de stage du Mastère de l'Ecole Nationale de la Météorologie, décembre 1997.

Wattrelot, E., 1999 : Analyse 4D-Var de l'expérience FASTEX : étude de l'impact de l'allongement de la période d'assimilation. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, juin 1999 (co-encadrement).

Chapnik, B., 2001 : Optimisation des paramètres d'erreur d'une assimilation variationnelle dans le cadre de la description et de la prévision des cyclogénèses. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, juin 2001.

Guerin, R., 2004 : Assimilation 4D-Var du tourbillon potentiel : un outil pour la TSR. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie,juin 2004 (co-encadrement).

Pannekoucke, O., 2004 : Etude d'une formulation ondelette du terme de rappel à l'ébauche pour le 4D-Var ARPEGE. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, juin 2004 (co-encadrement).

Brousseau, P., 2005 : Evaluation de l'impact des observations dans un système d'assimilation-prévision. Rapport de Stage d'approfondissement des Elèves Ingénieurs de l'Ecole Nationale de la Météorologie, juin 2005 (co-encadrement).

Martel, C., 2005 : Une technique de filtrage des cartes d'écarts types des erreurs de prévision. Rapport de Stage Master M3I de l'Université Paul Sabatier, août 2005 (co-encadrement).

2.5.2 Autres stages accompagnés d'un mémoire

Bahja, A., Bouksim, H., Moujahid, M., et Ouzzine, O., 1992 : Comparaison de méthodes directes et itératives pour lé résolution des systèmes d'interpolation optimale ou de splines. Rapport de Projet de Modélisation de l'Ecole Nationale de la Météorologie, juillet 1992.

El Mahdaoui, F., Mathiot, V., Mestre, O., et Orain, F., 1993 : Détermination optimale de paramètres en analyse variationnelle. Rapport de Projet de Modélisation de l'Ecole Nationale de la Météorologie, juillet 1993.

Dubuisson, B., et Mondon, S., 1995 : Structure des erreurs de prévision de tourbillon potentiel dans ARPEGE, Rapport de Projet de Modélisation de l'Ecole Nationale de la Météorologie, novembre 1995 (co-encadrement).

Dahoui, M., El Ouazzany, F., et Eljohra, B., 1996 : Recherche d'équations de balance pour l'inversion du tourbillon potentiel. Rapport de Projet de Modélisation de l'Ecole Nationale de la Météorologie, juillet 1996 (co-encadrement).

Ivanov, S., 1999 : Utilisation de la méthode de validation croisée généralisée pour l'optimisation des paramètres d'une analyse variationnelle et l'évaluation de sa qualité intrinsèque. Rapport de stage d'un visiteur ukrainien, octobre 1998 à juillet 1999.

Dziedzic, A., 2004 : Optimisation des filtres digitaux dans le 4D-Var ARPEGE, Rapport de stage ALADIN, septembre 2004 (co-encadrement).

Delourme, B., 2005 : Evaluation des matrices de variances-covariances des erreurs d'observation, d'ébauche et d'analyse dans l'espace des observations. Rapport de Stage de 4ième année INSA, août 2005 (co-encadrement).

2.5.3 Thèses

Chapnik, B., Thèse débutée en août 2001, soutenue en avril 2005 et co-encadrée par O. Talagrand (Directeur de thèse) et F. Rabier. Titre de la thèse : Réglage des statistiques d'erreur en assimilation variationnelle. Mémoire de thèse de l'Université Paul Sabatier

Pannekoucke, O., Thèse débutée en août 2005 et co-encadrée par F. Rabier (Directrice de thèse) et L. Berre.

2.6 Activités d'enseignement

Cours d'Assimilation de Données aux Elèves Ingénieurs de l'Ecole Nationale de la Météorologie depuis 1994.

Cours d'Assimilation de Données aux étudiants du Master M3I de l'Université Paul Sabatier depuis 2004.

2.7 Activités diverses

Membre du Comité Scientifique de l'Action Assimilation du Programme LEFE de l'INSU.

Coordinateur de l'Algorithmique de l'Assimilation de Données au sein du GMAP au CNRM.

Co-organisation de l'Atelier de Modélisation de l'Atmosphère, Toulouse, 1992.

Co-organisation du Workshop on Adjoint Applications in Dynamic Meteorology, Maratéa, Italy, 2004.

Co-organisation du Colloque National sur l'Assimilation de Données, 9-10 mai, Toulouse, 2006.

Co-organisation du Workshop on Adjoint Applications in Dynamic Meteorology, Obergurgl, Austria, 2006.

3 Synthèse des travaux de recherche

3.1 Introduction

L'assimilation de données, telle qu'elle s'est développée en particulier en météorologie et en océanographie, désigne le processus par lequel on cherche à estimer de la manière la plus précise possible l'état atmosphérique ou océanique. Cette estimation est produite à partir d'observations effectuées plus ou moins régulièrement à la fois spatialement et temporellement, mais aussi en s'appuyant sur un modèle numérique capable de décrire de manière réaliste la physique du phénomène étudié.

En retour, les états estimés sont le plus souvent produits pour donner les conditions initiales d'un modèle de prévision numérique utilisé pour fournir des prévisions à plus ou moins longue échéance.

Ces états estimés peuvent être également utiles pour eux-mêmes, dans le but d'étudier *a posteriori* un phénomène particulier ou l'évolution durant une période plus ou moins longue de l'état atmosphérique ou océanique. Des programmes de ré-assimilation de données sur plusieurs dizaines d'années sont en particulier menés dans différents centres. Plus ponctuellement, des campagnes expérimentales peuvent également faire l'objet de ré-analyses. La ré-analyse de l'expérience FASTEX résumée dans ce rapport entre dans ce cadre.

La science de l'assimilation de données a largement évolué durant les dernières années et a fait l'objet d'une recherche intensive avec la mise en oeuvre d'algorithmes toujours plus sophistiqués mais néanmoins efficaces.

Les premiers efforts de conceptualisation du problème de l'assimilation ont en particulier été entrepris par Gandin (1963), qui a introduit les concepts statistiques sous-jacents à la majeure partie des algorithmes d'assimilation aujourd'hui mis en oeuvre. Une autre évolution importante a été donnée par l'introduction du formalisme variationnel (Lewis et Derber1985 ; Le Dimet et Talagrand 1986 ; Courtier et Talagrand 1987 ; Talagrand et Courtier 1987), qui est aujourd'hui mis en place opérationnellement dans la plupart des grands centres météorologiques.

Le formalisme de l'assimilation peut être appréhendé de multiples manières, conduisant à des algorithmes divers, correspondant à des approximations différentes mais présentant entre eux des liens aujourd'hui assez clairement établis.

La formulation la plus générale du problème de l'assimilation est vraisemblablement donnée par l'approche Bayesienne faisant intervenir la notion de probabilités conditionnelles (voir par exemple Lorenc 1986).

La majorité des algorithmes d'analyse, y compris les plus élaborés comme l'assimilation variationnelle quadri-dimensionnelle (4D-Var), reste encore intimement liés à la théorie de l'estimation, développée et appliquée dans d'autres domaines que la météorologie ou l'océanographie. Il est en particulier classique de faire remonter cette théorie à Gauss (1809) qui en avait dressé les premières lignes pour l'estimation de la position des planètes. Mais ce sujet n'est en fait devenu mature que dans les années 60 et 70 (voir en particulier Jazwinski 1970 ; Gelb 1974). D'autres travaux théoriques sur le sujet de la théorie de l'estimation ont vu plus récemment le jour, en particulier dans le domaine des sciences de la terre (Tarantola 1987 ; Daley 1991 ; Bennett 1992). Une introduction très complète à ce sujet peut être trouvée dans Cohn (1997). En fait, les méthodes couramment utilisées en assimilation de données s'apparentent généralement à la théorie de l'estimation linéaire. Une présentation très pédagogique de l'assimilation sous cet angle est en particulier donnée par Talagrand (1997).

La synthèse qui suit de mes travaux ou de ceux des étudiants, que j'ai (ou j'ai eu) le plaisir de co-encadrer, est organisée comme suit.

Dans la partie 3.2, le formalisme de l'assimilation de données sous l'angle de l'estimation linéaire statistique est rapidement rappelé.

Le travail sur la possibilité de parvenir à une meilleure représentation des covariances d'erreur ébauche par une transformation physique, empruntée à la théorie semi-géostrophique, est résumé dans la partie 3.3.

Le paragraphe 3.4 est consacré aux diagnostics d'optimalité des systèmes d'assimilation. Il comprend en particulier les résultats obtenus par B. Chapnik durant sa thèse.

La ré-analyse de l'expérience FASTEX en mode 4D-Var est décrite dans la partie 3.5.

Une approche pour mesurer l'impact des observations sur les analyses et les prévisions, sous l'angle de la réduction de variance d'erreur d'estimation, est résumée au paragraphe 3.6.

Une méthode récemment proposée pour mesurer la cohérence des systèmes d'analyse dans l'espace des observations est décrite dans la partie 3.7.

La possibilité d'utiliser les ondelettes pour parvenir à une meilleure description des covariances d'erreur d'ébauche est ensuite discutée (paragraphe 3.8). Cette étude correspond au travail de thèse d'O. Pannekoucke.

Quelques perspectives pour la poursuite des travaux de recherche en assimilation de données sont enfin données dans la dernière partie de cette synthèse (partie 3.9).

3.2 Formalisme de l'estimation linéaire statistique

Une manière assez peu classique mais immédiate d'obtenir l'analyse optimale au sens de la théorie de l'estimation linéaire statistique est de s'appuyer sur le formalisme de la régression linéaire multiple. Dans ce formalisme, on cherche à prévoir, au sens statistique du terme, la meilleure valeur de prédictands $\boldsymbol{x}$ à partir de prédicteurs $\boldsymbol{y}$. En s'appuyant sur la formule classique de la régression linéaire multiple, la meilleure prévision $\boldsymbol{x}^a$ est donnée par

$$\boldsymbol{x}^a - E(\boldsymbol{x}) = E(\boldsymbol{x}\boldsymbol{y}^T)E(\boldsymbol{y}\boldsymbol{y}^T)^{-1}(\boldsymbol{y} - E(\boldsymbol{y})). \tag{1}$$

Si l'on considère que l'analyse cherchée est obtenue par une correction $\boldsymbol{\delta}x$ à une ébauche, ou information *a priori*, donnée par une prévision $\boldsymbol{x}^b$ fournie par un modèle atmosphérique, l'analyse $\boldsymbol{x}^a$ s'écrit en fait

$$\boldsymbol{x}^a = \boldsymbol{x}^b + \boldsymbol{\delta}x.$$

En posant $\boldsymbol{x}^b = \boldsymbol{x}^t + \boldsymbol{\epsilon}^b$, où $\boldsymbol{x}^t$ est l'état vrai inconnu et $\boldsymbol{\epsilon}^b$ le vecteur des erreurs d'ébauche, une correction parfaite à l'ébauche serait $-\boldsymbol{\epsilon}^b$. La variable à estimer ou prédictand est donc ici $-\boldsymbol{\epsilon}^b$ et le vecteur des prédicteurs $\boldsymbol{d} = \boldsymbol{y}^o - H(\boldsymbol{x}^b)$, c'est-à-dire la différence entre les observations $\boldsymbol{y}^o$ et l'équivalent $H(\boldsymbol{x}^b)$ de ces observations pour l'ébauche ($\boldsymbol{d}$ est le vecteur des innovations dans le formalisme du Filtre de Kalman). A noter que c'est $\boldsymbol{d} = \boldsymbol{y}^o - H(\boldsymbol{x}^b)$ qui doit être choisi comme prédictand et non pas simplement $\boldsymbol{y}^o$ car il n'y a pas de lien entre les réalisations aléatoires des observations et celles des erreurs d'ébauche. En utilisant l'équation (1) de la régression linéaire multiple avec ce choix particulier de prédictands et prédicteurs, la forme optimale de la correction à apporter à l'ébauche est

$$\boldsymbol{\delta}x = (-\boldsymbol{\epsilon}^b)^a = E(-\boldsymbol{\epsilon}^b\boldsymbol{d}^T)E(\boldsymbol{d}\boldsymbol{d}^T)^{-1}\boldsymbol{d}, \tag{2}$$

en considérant que les erreurs d'ébauche et d'observation sont non biaisées et donc que $E(\boldsymbol{\epsilon}^b) = \boldsymbol{0}$ et $E(\boldsymbol{d}) = E(\boldsymbol{y}^o - H(\boldsymbol{x}^t) + H(\boldsymbol{x}^t) - H(\boldsymbol{x}^b)) = E(\boldsymbol{\epsilon}^o - \boldsymbol{H}\boldsymbol{\epsilon}^b) = \boldsymbol{0}$. D'autre part, la covariance des innovations s'écrit

$$E[\boldsymbol{d}\boldsymbol{d}^T] = E[\boldsymbol{\epsilon}^o(\boldsymbol{\epsilon}^o)^T] + \boldsymbol{H}E[\boldsymbol{\epsilon}^b(\boldsymbol{\epsilon}^b)^T]\boldsymbol{H}^T,$$

en utilisant la linéarité de l'opérateur d'espérance statistique E et le fait que les erreurs d'observation $\boldsymbol{\epsilon}^o$ et d'ébauche $\boldsymbol{\epsilon}^b$ sont (en général) dé-corrélées. En notant respectivement $\boldsymbol{R}$ et $\boldsymbol{B}$ les matrices de covariance des erreurs d'observation et d'ébauche, la covariance des innovations s'exprime donc simplement

$$E[\boldsymbol{d}\boldsymbol{d}^T] = \boldsymbol{R} + \boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T.$$

De même, la covariance croisée entre les erreurs de prévision $-\boldsymbol{\epsilon}^b$ et les innovations $\boldsymbol{d} = \boldsymbol{\epsilon}^o - \boldsymbol{H}\boldsymbol{\epsilon}^b$ s'écrit

$$E(-\boldsymbol{\epsilon}^b\boldsymbol{d}^T) = \boldsymbol{B}\boldsymbol{H}^T.$$

On retrouve ainsi l'expression classique de l'analyse

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

où $\boldsymbol{K}$ est la matrice de gain apparaissant dans le filtre de Kalman.

La manière la plus rapide d'obtenir l'expression (2) de la régression linéaire est d'appliquer un raisonnement heuristique exposé par exemple dans Der Mégréditchian (1993). En effet, on souhaite que le vecteur des prédicteurs $\boldsymbol{d}$ apporte le plus à la prévision de

$-\boldsymbol{\epsilon}^b$ et donc qu'il contribue le moins à l'erreur sur cette estimation. En d'autres termes, la matrice $\boldsymbol{K}$ de la régression $(-\boldsymbol{\epsilon}^b)^a = \boldsymbol{K}\boldsymbol{d}$ doit annuler la corrélation entre l'erreur $-\boldsymbol{\epsilon}^b - (-\boldsymbol{\epsilon}^b)^a = -\boldsymbol{\epsilon}^b - \boldsymbol{K}\boldsymbol{d}$ et $\boldsymbol{d}$, d'où

$$E((-\boldsymbol{\epsilon}^b - \boldsymbol{K}\boldsymbol{d})\boldsymbol{d}^T) = E(-\boldsymbol{\epsilon}^b\boldsymbol{d}^T) - \boldsymbol{K}E(\boldsymbol{d}\boldsymbol{d}^T) = \boldsymbol{0},$$

ce qui donne bien l'expression (2).

A noter que $\boldsymbol{\epsilon}^b + \boldsymbol{K}\boldsymbol{d}$ représente bien l'erreur d'analyse sur $\boldsymbol{x}$ puisque

$$\begin{aligned} \boldsymbol{x}^a &= \boldsymbol{x}^b + \boldsymbol{\delta}x \\ &= \boldsymbol{x}^t + \boldsymbol{\epsilon}^b + \boldsymbol{K}\boldsymbol{d}, \end{aligned}$$

et donc $\boldsymbol{\epsilon}^a = \boldsymbol{x}^a - \boldsymbol{x}^t = \boldsymbol{\epsilon}^b + \boldsymbol{K}\boldsymbol{d}$. La condition de décorrélation entre le prédictand $-(\boldsymbol{\epsilon}^b + \boldsymbol{K}\boldsymbol{d})$ et le prédicteur $\boldsymbol{d}$ signifie donc que, par construction, l'erreur d'analyse $\boldsymbol{\epsilon}^a$ doit être orthogonale au vecteur innovation $\boldsymbol{d}$.

En utilisant la formule de Sherman-Morrison-Woodbury, l'incrément d'analyse optimal s'écrit aussi sous la forme

$$\boldsymbol{\delta}x = (\boldsymbol{B}^{-1} + \boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^T\boldsymbol{R}^{-1}\boldsymbol{d}.$$

Il est facile de vérifier que cette expression de $\boldsymbol{\delta}x$ est aussi le minimum de la forme quadratique

$$J(\boldsymbol{\delta}x) = (\boldsymbol{\delta}x)^T\boldsymbol{B}^{-1}\boldsymbol{\delta}x + (\boldsymbol{d} - \boldsymbol{H}\boldsymbol{\delta}x)^T\boldsymbol{R}^{-1}(\boldsymbol{d} - \boldsymbol{H}\boldsymbol{\delta}x).$$

Cette expression n'est rien d'autre que la fonction coût de la formulation variationnelle de l'assimilation de données. On peut montrer que chercher le minimum de $J(\boldsymbol{\delta}x)$ revient à déterminer la correction $\boldsymbol{\delta}x$ à l'ébauche la plus probable, si à la fois les erreurs d'observations et d'ébauche suivent des distributions gaussiennes.

La fonction coût précédente correspond au problème 3D-Var de l'assimilation de données, mais le passage au 4D-Var s'obtient simplement en incluant dans l'opérateur d'observation le modèle de prévision lui-même.

La prise en compte des non-linéarités comprises dans l'opérateur d'observation H, incluant éventuellement le modèle, est obtenue en effectuant des remises à jour régulière de la trajectoire de linéarisation (Courtier *et al.*, 1994).

Dans le cas où l'opérateur d'observation est strictement linéaire, l'erreur d'estimation $\boldsymbol{\epsilon}^a = \boldsymbol{\epsilon}^b + \boldsymbol{K}\boldsymbol{d}$ s'écrit encore

$$\begin{aligned} \boldsymbol{\epsilon}^a &= \boldsymbol{\epsilon}^b + \boldsymbol{K}(\boldsymbol{\epsilon}^o - \boldsymbol{H}\boldsymbol{\epsilon}^b) \\ &= (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{\epsilon}^b + \boldsymbol{K}\boldsymbol{\epsilon}^o, \end{aligned}$$

et donc la covariance d'erreur d'estimation s'exprime

$$\begin{aligned} \boldsymbol{A} &= E(\boldsymbol{\epsilon}^a(\boldsymbol{\epsilon}^a)^T) \\ &= (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})E(\boldsymbol{\epsilon}^b(\boldsymbol{\epsilon}^b)^T)(\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})^T + \boldsymbol{K}E(\boldsymbol{\epsilon}^o(\boldsymbol{\epsilon}^o)^T)\boldsymbol{K}^T \\ &= (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\boldsymbol{B}(\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})^T + \boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^T \\ &= \boldsymbol{B} - \boldsymbol{K}\boldsymbol{H}\boldsymbol{B}, \end{aligned}$$

cette dernière formule indiquant comment l'incertitude sur l'état à estimer est réduite par le processus d'analyse.

On peut montrer que l'expression précédente de $\boldsymbol{x}^a$ est l'estimateur non biasé qui qui minimise la variance d'erreur. Cet estimateur correspond au Best Linear Unbiased Estimate (BLUE).

3.3 Changement de coordonnées géostrophiques

Introduction

Les méthodes d'assimilation utilisées en météorologie cherchent à combiner de manière optimale des observations et une ébauche. Elles reposent donc sur une description des statistiques d'erreur associées aux observations et au modèle assimilateur. Parce que ces statistiques ne sont pas parfaitement connues mais aussi pour rendre praticable la mise en place d'un schéma d'assimilation, des hypothèses simplificatrices sur la forme de erreurs de prévision sont généralement faites.

Le travail mené dans Desroziers (1997), correspondant au premier article joint à ce rapport, vise à obtenir une meilleure justification de ces hypothèses simplificatrices en formulant les covariances d'erreur d'ébauche dans un espace transformé.

Transformation géostrophique dans un schéma variationnel

La théorie semigéostrophique (Hoskins et Bretherton 1972) possède plusieurs propriétés remarquables. Une d'elles est qu'elle permet de s'affranchir des non-linéarités non-géostrophiques dans l'évolution de l'écoulement atmosphérique et d'éviter les chocs, comme les discontinuités frontales.

L'idée développée dans Desroziers (1997) est d'utiliser une coordonnée géostrophique pour l'analyse de données comme cela avait déjà été suggéré dans Desroziers et Lafore (1993) pour l'interpolation dans un plan vertical d'observations à haute densité à travers des fronts simulés et observés. Un des intérêts de ce changement de coordonnées est qu'il permet d'obtenir des corrélations des erreurs de prévision dépendantes de l'écoulement. Une autre manière d'obtenir cette dépendance à l'écoulement est d'utiliser une transformation isentrope (Benjamin *et al.* 1991), les deux approches pouvant être éventuellemnt combinées.

La formulation du changement de coordonnées est basée sur la transformation proposée par Hoskins (1975) suivant l'horizontale (la coordonnée verticale est inchangée dans la transformation) :

$$\begin{aligned} x_G &= x_R + \frac{1}{f} v_g \\ y_G &= x_R - \frac{1}{f} u_g, \end{aligned}$$

où f est le paramètre de Coriolis, u_g et v_g les composantes du vent géostrophique, et (x_R, y_R) et (x_G, y_G) les coordonnées respectivement physiques et géostrophiques d'un point. L'essence de la transformation imaginée par Hoskins et Bretherton (1972) est de se ramener dans un espace dans lequel la non-linéarité des termes d'advection des équations primitives disparaît. Cette non-linéarité qui permet la création des discontinuités frontales en un temps fini est reportée dans le changement de coordonnées inverse pour revenir dans l'espace physique.

Dans Desroziers (1997), on trouve une proposition d'extension de ce changement de coordonnées à la sphère. En effet, si $\boldsymbol{z}^r$ est la représentation dans l'espace physique, sa représentation $\boldsymbol{z}^g$ dans l'espace géostrophique est donnée par l'évaluation $z^g(G) = z^r(R)$ en tout point $G(\lambda_G, \theta_G)$ de la grille de discrétisation du nouvel espace, où $R(\lambda_R, \theta_R)$ est l'image de G dans l'espace réel (λ et θ désignent respectivement la longitude et la latitude d'un point sur la sphère). Une fois les coordonnées $R(\lambda_R, \theta_R)$ obtenues, z^g peut être obtenu par une simple interpolation $\boldsymbol{\mathcal{T}}$, dans l'espace physique, de $\boldsymbol{z}^r$: $\boldsymbol{z}^g = \boldsymbol{\mathcal{T}} \boldsymbol{z}^r$.

En considérant que cette transformation peut être ainsi construite, l'idée développée dans Desroziers (1997) est que les erreurs de prévision devraient être plus homogènes dans l'espace géostrophique. En effet, l'impact du changement de coordonnées géostrophiques est qu'il dilate les zones frontales et par conséquent les variations des champs dans ce nouvel espace deviennent plus régulières dans l'ensemble des directions. Ainsi, le terme de rappel à l'ébauche de la formulation variationnelle devrait plutôt être écrit sous la forme

$$J(\delta \boldsymbol{x}) = \frac{1}{2}[\boldsymbol{L}_{\mathcal{T}} \delta \boldsymbol{x}]^T \boldsymbol{B}^{-1} [\boldsymbol{L}_{\mathcal{T}} \delta \boldsymbol{x}],$$

où $\boldsymbol{L}_{\mathcal{T}}$ est la transformation d'un champ vers l'espace géostrophique incluant l'opérateur $\boldsymbol{\mathcal{T}}$ et éventuellement un retour dans l'espace physique si $\delta \boldsymbol{x}$ est spécifié dans l'espace spectral.

L'opérateur $\boldsymbol{L}_{\mathcal{T}}$ devrait dépendre de l'état analysé $\boldsymbol{x}$. Dans la formulation incrémentale proposée par Courtier *et al.* (1994), l'analyse est basée sur une paire de boucles imbriquées : à chaque itération de la boucle externe, la trajectoire du modèle est redéfinie avec le modèle complet, alors que les boucles internes utilisent un modèle simplifié pour l'analyse des incréments autour des trajectoires mises à jour. Ainsi, avec cette approche, il est possible d'introduire une redéfinition de l'opérateur $\boldsymbol{L}_{\mathcal{T}}$ à chaque itération externe.

La modification induite par l'opérateur $\boldsymbol{L}_{\mathcal{T}}$ peut être explicitée. En effet, la solution du problème d'estimation sans transformation peut être écrite

$$\delta \boldsymbol{x} = \boldsymbol{B}\boldsymbol{H}^T (\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1} \boldsymbol{d}.$$

Lorsque la transformation géostrophique est introduite, cette expression devient

$$\delta \boldsymbol{x} = \boldsymbol{B}_{\mathcal{T}} \boldsymbol{H}^T (\boldsymbol{H}\boldsymbol{B}_{\mathcal{T}}\boldsymbol{H}^T + \boldsymbol{R})^{-1} \boldsymbol{d}, \tag{3}$$

avec $\boldsymbol{B}_{\mathcal{T}} = \boldsymbol{L}_{\mathcal{T}} \boldsymbol{B} (\boldsymbol{L}_{\mathcal{T}})^T$.

Cette expression montre que l'analyse avec changement de coordonnées utilise implicitement des covariances d'erreurs de prévision dépendantes de l'écoulement et non isotropes, construites par une déformation par $\boldsymbol{L}_{\mathcal{T}}$ des covariances isotropes spécifiées dans $\boldsymbol{B}$.

Application à une situation météorologique réelle

Le changement de coordonnées geostrophiques a été introduit dans le modèle opérationnel global ARPEGE et appliqué à une situation de l'expérience européenne FRONTS 87 organisée pour étudier la dynamique de méso-échelle des fronts froids. Cette situation a été décrite en détail par Thorpe et Clough (1991) et était essentiellement caractérisée par un front bi-dimensionnel avec un fort gradient de température (Fig. 1). La situation se caractérise également par un fort jet d'altitude atteignant 50 m/s au-dessus de l'Irlande, avec un fort gradient cyclonique sur son bord ouest. Ce gradient cyclonique de vent et le gradient de température sont presque colocalisés. A l'opposé, la région d'air chaud est caractérisée par l'absence de fortes variations de la température et des valeurs négatives du champ de tourbillon.

La transformation géostrophique est bâtie avec un champ de vent issu d'une prévision 6h correspondant à l'ébauche utilisée dans une analyse. Les composantes géostrophiques du vent sont approximées par les composantes de la partie rotationnelle du vent réel.

D'un autre côté, la transformation est progressivement relâchée vers l'équateur. La construction du changement de coordonnées sur chacune de ces surfaces nécessite la définition de l'ensemble complet des coordonnées (λ_R, μ_R) dans l'espace physique correspondant aux coordonnées (λ_G, μ_G) des points de grille de l'espace transformé fictif.

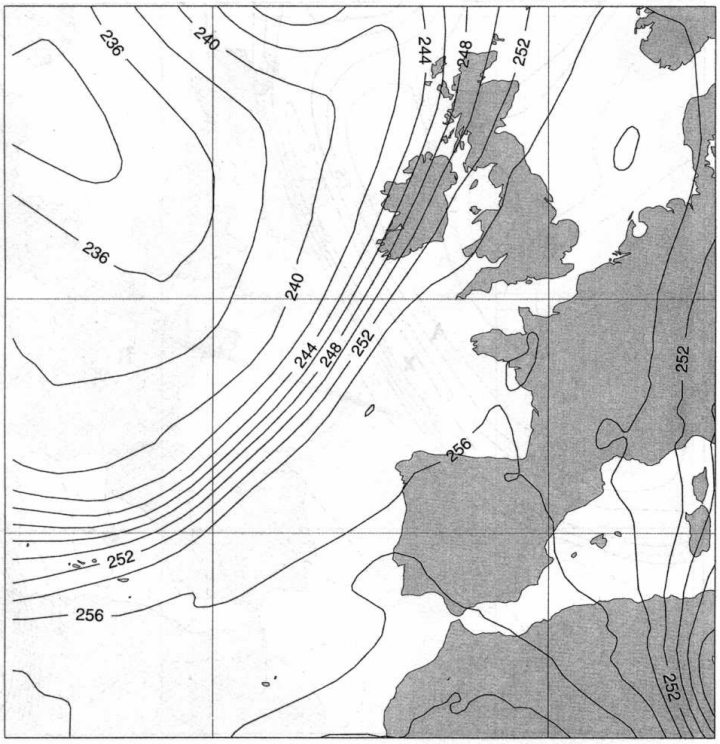

FIG. 1 – Prévision ARPEGE 6h du champ de température à 500 hPa pour le 9 janvier 1988 à 12H. L'écart entre les isolignes est de 2K.

L'intérêt de la transformation pour l'assimilation peut être étudié à partir de la déformation implicite que la transformation induit sur les fonctions de structure utilisées dans l'analyse.

Si l'on considère que la corrélation ρ^g des erreurs de prévision d'un champ est homogène dans l'espace transformé, alors en accord avec la relation (3), la forme de la fonction de corrélation correspondante ρ^r dans l'espace physique en un point donné $R_O(\lambda_{R_O}, \mu_{R_O})$ peut être simplé construite en écrivant pour tout point R

$$\rho^r[(\lambda_{R_O}, \mu_{R_O}), (\lambda_R, \mu_R)] = \rho^g[(\lambda_{G_O}, \mu_{G_O}), (\lambda_G, \mu_G)], \tag{4}$$

où $(\lambda_{G_O}, \mu_{G_O})$ et (λ_G, μ_G) sont les coordonnées transformées correspondant respectivement aux coordonnées dans l'espace physique $(\lambda_{R_O}, \mu_{R_O})$ et (λ_R, μ_R).

La déformation de la fonction de corrélation des erreurs de prévision de la température peut par exemple être étudiée. La fonction de corrélation est ici définie avec une portée horizontale de 300 km (Fig. 2).

Pour un point situé dans le front froid, la fonction de structure implicite induite par la transformation et définie par la relation (4) montre une structure anisotrope caractéristique avec une contraction à travers la discontinuité et un rapport d'anisotropie atteignant deux (Fig. 3). Pour le point B situé dans le secteur chaud de la perturbation, la transformation a, au contraire, un effet d'étirement avec un rapport d'anisotropie également proche de deux.

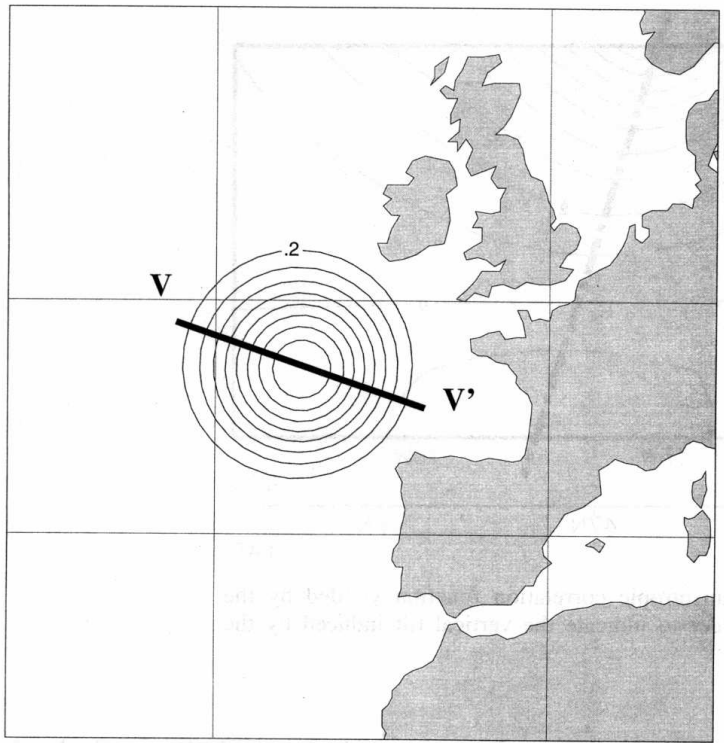

FIG. 2 – Représentation de la fonction de corrélation isotrope pour les erreurs de la température, basée en un point donné. Les contours sont tous les 0.1.

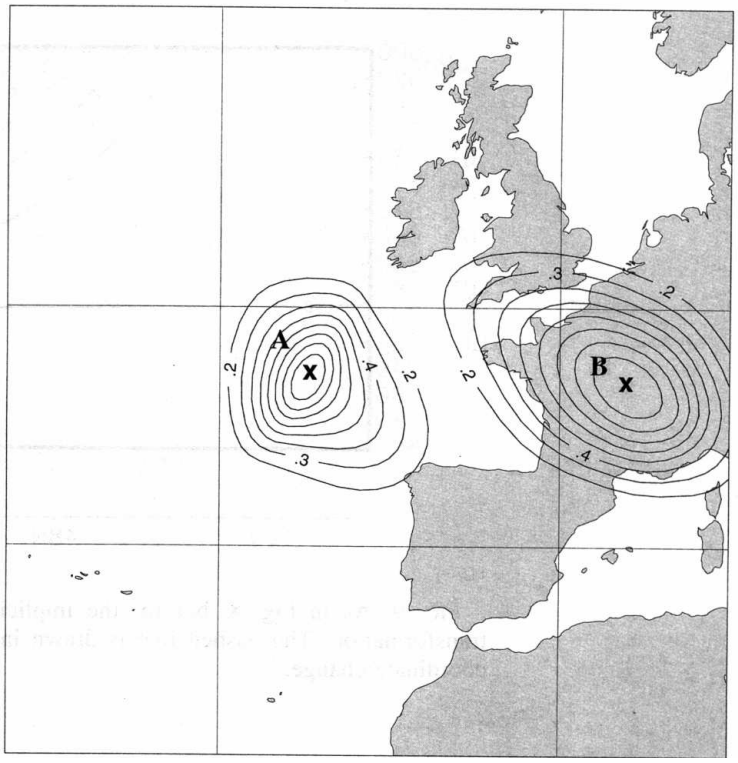

FIG. 3 – Comme dans la Fig. 2, mais pour la fonction de corrélation anisotrope implicite induite par la transformation aux points A et B.

Conclusion

Une transformation de type géostrophique a été proposée sur la sphère. Elle a été implémentée dans le cadre du système global ARPEGE. Elle permet d'obtenir des anisotropies réalistes tout en spécifiant des covariances isotropes dans l'espace géostrophique.

L'introduction d'une telle transformation dans l'analyse a été testée au UK Met Office et envisagée dans d'autres centres de prévision numérique.

3.4 Diagnostics d'optimalité du problème variationnel

Introduction

Le travail mené dans Desroziers et Ivanov (2001), correspondant au deuxième article joint à ce mémoire, est largement inspiré des idées développées dans Talagrand (1999) sur la valididation *a posteriori* des systèmes d'assimilation. Il a également été guidé par les travaux de Dee (1995) sur la recherche des paramètres d'une analyse par une approche par Maximum de Vraisemblance et les travaux de Wahba et Wendelberger (1980) et Wahba *et al.* (1995) sur la Validation Croisée Généralisée, traitant aussi de l'optimisation de paramètres dans un schéma d'analyse.

Espérance statistique de sous-parties de la fonction coût

En suivant Talagrand (1997), il est possible d'introduire un vecteur d'observations étendu $\boldsymbol{z}^o$, comprenant le vecteur des observations proprement dites $\boldsymbol{y}^o$, de dimension p, et le vecteur d'ébauche $\boldsymbol{x}^b$, de même dimension n que l'état vrai inconnu $\boldsymbol{x}^t$. On a donc $\boldsymbol{z}^o = \{(\boldsymbol{x}^b)^T(\boldsymbol{y}^o)^T\}^T$ et

$$\boldsymbol{x}^b = \boldsymbol{x}^t + \boldsymbol{\epsilon}^b,$$

où $\boldsymbol{\epsilon}^b$ est le vecteur des erreurs de prévision inconnues de covariance donnée par la matrice $\boldsymbol{B}$, et

$$\boldsymbol{y}^o = H(\boldsymbol{x}^t) + \boldsymbol{\epsilon}^o,$$

où $\boldsymbol{\epsilon}^o$ est le vecteur des erreurs d'observation également inconnues de covariance donnée par la matrice $\boldsymbol{R}$. Il est donc possible d'écrire $\boldsymbol{z}^o$ sous la forme

$$\boldsymbol{z}^o = \Gamma(\boldsymbol{x}^t) + \boldsymbol{\epsilon},$$

où Γ est un opérateur d'observation étendu et $\boldsymbol{\epsilon}$ le vecteur des erreurs d'ébauche et d'observation de dimension $n + p$.

Un résultat important souligné par Talagrand (1999) est que si J_i désigne un terme de la fonction coût J du problème d'estimation écrit sous forme variationnelle qui est la somme de p_i éléments, alors l'espérance statistique de J_i au minimum $\boldsymbol{x}^a$ de la fonction coût globale est donnée par

$$E\{J_i(\boldsymbol{x}^a)\} = 1/2\ \{p_i - Tr(\boldsymbol{S}_i{}^{-1/2}\boldsymbol{\Gamma}_i\boldsymbol{A}\boldsymbol{\Gamma}_i^T\boldsymbol{S}_i{}^{-1/2})\}, \tag{5}$$

où $\boldsymbol{\Gamma}_j$ et $\boldsymbol{S}_i$ désignent respectivement l'opérateur d'observation et la matrice de covariance associés à ces p_i éléments et $\boldsymbol{A}$ la matrice de covariance d'erreur d'estimation.

En particulier, pour $\boldsymbol{\Gamma}_i = \boldsymbol{I}_n$ et $\boldsymbol{S}_i = \boldsymbol{B}$, on obtient

$$E\{J^b(\boldsymbol{x}^a)\} = 1/2\ Tr(\boldsymbol{HK}),$$

et pour $\boldsymbol{\Gamma}_i = \boldsymbol{H}$ et $\boldsymbol{S}_i = \boldsymbol{R}$,

$$E\{J^o(\boldsymbol{x}^a)\} = 1/2\ \{p - Tr(\boldsymbol{HK})\},$$

et donc

$$E\{J(\boldsymbol{x}^a)\} = E\{J^b(\boldsymbol{x}^a)\} + E\{J^o(\boldsymbol{x}^a)\} = p/2.$$

Ceci signifie, que si les statistiques des erreurs d'ébauche et d'observation sont bien spécifiées, alors la statistique des valeurs de la fonction coût globale à son minimimum devrait être égale à $p/2$. Ici p correspond aussi au nombre de degrés de liberté théoriques du

problème. Comme l'ont indiqué Bennett *et al.* (1993) et Talagrand (1999), ceci constitue un critère simple de validation *a posteriori* de l'analyse : un écart de l'espérance de J de la valeur $p/2$ indique une mauvaise spécification des statistiques de l'innovation.

Estimation de type Monte Carlo de $E\{J_i(\boldsymbol{x}^a)\} = 1/2\ \{p_i - Tr(\boldsymbol{S}_i{}^{-1/2}\boldsymbol{\Gamma}_i\boldsymbol{A}\boldsymbol{\Gamma}_i^T\boldsymbol{S}_i{}^{-1/2})\}$

Talagrand (1999) a indiqué que les expressions $Tr(\boldsymbol{S}_i{}^{-1/2}\boldsymbol{\Gamma}_i\boldsymbol{A}\boldsymbol{\Gamma}_i^T\boldsymbol{S}_i{}^{-1/2})$ peuvent être vues comme des mesures des contributions relatives des sous-ensembles z_i^o des observations généralisées (incluant l'ébauche) à la précision globale de l'analyse. Ces expressions sont difficiles à calculer explicitement puisque, en particulier, la matrice $\boldsymbol{A}$ n'est pas disponible dans une résolution variationnelle du problème de l'analyse. Il est cependant montré dans Desroziers et Ivanov (2001) qu'une estimation de ces expressions peut être obtenue par une méthode de type Monte Carlo. Cette méthode est inspirée de Girard (1987) et Wahba *et al.* (1995) qui ont proposé une méthode similaire pour le calcul du critère de Validation Croisée Généralisée.

En effet, on peut montrer que si les erreurs d'observation d'un sous-ensemble i d'observations sont décorrélées des erreurs du reste des observations, alors on a simplement

$$E\{J_i^o(\boldsymbol{x}^a)\} = 1/2\ \{p_i - Tr(\boldsymbol{K}\boldsymbol{\Pi}_i{}^T\boldsymbol{\Pi}_i\boldsymbol{H})\},$$

où $\boldsymbol{\Pi}_i$ est le projecteur qui permet de passer de l'ensemble des observations au sous-ensemble i d'observations et $\boldsymbol{K}$ la matrice de gain.

De même,

$$E\{J_i^b(\boldsymbol{x}^a)\} = 1/2\ Tr(\boldsymbol{\Pi}_i{}^T\boldsymbol{\Pi}_i\boldsymbol{K}\boldsymbol{H}),$$

où $\boldsymbol{\Pi}_i$ est le projecteur qui permet de passer de l'ensemble des éléments de $\boldsymbol{x}^b$ au sous-ensemble i d'éléments de $\boldsymbol{x}^b$.

On peut par ailleurs montrer qu'une estimation de $E\{J_i^o(\boldsymbol{x}^a)\}$ est donnée par

$$p_i - \delta\boldsymbol{y}_i^{oT}\boldsymbol{R}_i{}^{-1}\boldsymbol{\Pi}_i\boldsymbol{H}\boldsymbol{K}\delta\boldsymbol{y}^o,$$

où $\delta\boldsymbol{y}^o = \boldsymbol{R}^{1/2}\boldsymbol{\eta}^o$ est un vecteur de perturbations sur l'ensemble des observations, $\boldsymbol{R}_i$ est la matrice de covariance des erreurs d'observations du jeu i, et $\delta\boldsymbol{y}_i^o = \boldsymbol{\Pi}_i\delta\boldsymbol{y}^o$ est la projection de $\delta\boldsymbol{y}^o$ sur le sous-ensemble i d'observations.

Mais $\boldsymbol{K}\delta\boldsymbol{y}^o$ n'est rien d'autre que la perturbation $\delta\boldsymbol{x}^a(\delta\boldsymbol{y}^o)$ sur l'analyse produite par une perturbation sur les observations puisque $\delta\boldsymbol{x}^a(\delta\boldsymbol{y}^o) = \boldsymbol{K}\delta\boldsymbol{y}^o$ dans le cas linéaire. Ainsi une estimation par Monte Carlo de $E\{J_i^o\}$ est donnée par

$$E\{J_i^o\} \simeq p_i - \delta\boldsymbol{y}_i^{oT}\boldsymbol{R}_i{}^{-1}\boldsymbol{\Pi}_i\boldsymbol{H}\delta\boldsymbol{x}^a(\delta\boldsymbol{y}^o). \tag{6}$$

De même, une estimation par Monte Carlo d'une sous-partie de J^b est donnée par

$$E\{J_i^b\} \simeq \delta\boldsymbol{x}_i^{bT}\boldsymbol{B}_i{}^{-1}\boldsymbol{\Pi}_i\delta\boldsymbol{x}^a(\delta\boldsymbol{x}^b), \tag{7}$$

où $\delta\boldsymbol{x}^b = \boldsymbol{B}^{1/2}\boldsymbol{\eta}^b$ est un vecteur de perturbations sur l'ensemble du vecteur d'ébauche, $\boldsymbol{B}_i$ est la matrice de covariances des erreurs d'ébauche de la sous-partie i de $\boldsymbol{x}^b$, et $\delta\boldsymbol{x}_i^b = \boldsymbol{\Pi}_i\delta\boldsymbol{x}^b$ est la projection de $\delta\boldsymbol{x}^b$ sur le sous-ensemble i de $\boldsymbol{x}^b$.

Réglage des variances d'erreur d'observation et d'ébauche

La possibilité de calculer l'espérance statistique des sous-parties de la fonction coût nous a permis d'imaginer une procédure pour optimiser les statistiques des erreurs associées aux sous-parties du vecteur $\boldsymbol{z}^o$. En effet, la fonction coût initiale peut être ré-écrite sous la forme modifiée

$$J(\delta\boldsymbol{x}) = \sum_i \frac{1}{s_i^{b^2}} J_i^b(\delta\boldsymbol{x}) + \sum_i \frac{1}{s_i^{o^2}} J_i^o(\delta\boldsymbol{x}),$$

où $\delta\boldsymbol{x}$ est l'incrément d'analyse et les $s_i^{b^2}$ et $s_i^{o^2}$ sont respectivement des paramètres de pondération des erreurs d'ébauche et d'observation, supposées homogènes sur un sous-ensemble i. L'idée de la procédure de réglage des statistiques d'erreurs est alors de trouver les valeurs de ces pondérations telles que les valeurs de $\frac{1}{s_i^{b^2}} J_i^b(\delta\boldsymbol{x})$ et $\frac{1}{s_i^{o^2}} J_i^o(\delta\boldsymbol{x})$ soient proches de leurs espérances statistiques, c'est-à-dire de $1/2\, Tr({\boldsymbol{\Pi}_i}^T \boldsymbol{\Pi}_i \boldsymbol{K}\boldsymbol{H})$ et $1/2\, \{p_i - Tr(\boldsymbol{K}{\boldsymbol{\Pi}_i}^T \boldsymbol{\Pi}_i \boldsymbol{H})\}$ respectivement.

En définissant $S_i^b = 2J_i^b / Tr({\boldsymbol{\Pi}_i}^T \boldsymbol{\Pi}_i \boldsymbol{K}\boldsymbol{H})$ et $S_i^o = 2J_i^o / \{p_i - Tr(\boldsymbol{K}{\boldsymbol{\Pi}_i}^T \boldsymbol{\Pi}_i \boldsymbol{H})\}$, on est donc amené à chercher les valeurs de $s_i^{b^2}$ et $s_i^{o^2}$ telles que $s_i^{b^2} = S_i^b(\boldsymbol{s}^b, \boldsymbol{s}^o)$ et $s_i^{o^2} = S_i^o(\boldsymbol{s}^b, \boldsymbol{s}^o)$, où $\boldsymbol{s}^b$ et $\boldsymbol{s}^o$ sont les vecteurs contenant l'ensemble des pondérations recherchées.

Ce sont des équations non linéaires en $\boldsymbol{s}^b$ et $\boldsymbol{s}^o$, mais leur forme particulière suggère l'utilisation d'une méthode itérative de point fixe pour les résoudre, avec la séquence $(s_i^b)_{k+1} = (S_i^b(\boldsymbol{s}_k^b, \boldsymbol{s}_k^o))^{1/2}$ et $(s_i^o)_{k+1} = (S_i^o(\boldsymbol{s}_k^b, \boldsymbol{s}_k^o))^{1/2}$, où k est l'indice de l'itération de point fixe.

Application dans le 3D-Var ARPEGE

La méthode de réglage précédente a été appliquée avec succès pour le réglage des statistiques d'erreurs d'observation et d'ébauche dans le cadre d'une analyse simplifiée. Il apparaît en effet que la méthode de point fixe converge très rapidement et que le réglage obtenu après une seule itération, c'est-à-dire avec une seule analyse perturbée, soit déjà pertinent dans les cas favorables.

Un test de cette procédure a également été effectué dans le cadre fourni par le 3D-Var associé au modèle ARPEGE global, en simulant à la fois une ébauche et des observations correspondant à des covariances d'erreurs données, puis en cherchant à retrouver les variances réelles par la procédure précédente. On a considéré dans ce test que les valeurs exactes $\boldsymbol{x}^t$ des variables ARPEGE (température, tourbillon, divergence et humidité spécifique sur 31 niveaux et pression de surface) sont données par une prévision 6h particulière. Un vecteur d'erreur d'ébauche $\boldsymbol{\epsilon}^b$ est obtenu en appliquant la racine carrée de la matrice de covariance d'erreur d'ébauche $\boldsymbol{B}^{1/2}$ à un vecteur de nombres aléatoires (ceci peut être effectué en pratique en appliquant l'opérateur de changement de variables, $\boldsymbol{L} = \boldsymbol{B}^{1/2}$, introduit pour pré-conditionner le problème variationnel de l'analyse). Etant donné que la formulation du terme J^b est multivariée dans le 3D/4D-Var ARPEGE, en suivant en particulier l'approche développée par Bouttier *et al.* (1997), cette opération produit un vecteur $\boldsymbol{\epsilon}^b$ dans lequel les champs de température et de tourbillon sont en particulier corrélés. La Figure 4 montre le champ correspondant d'erreurs d'ébauche pour la température à 500hPa. Les observations simulées sont données par un ensemble de mesures de radiosondages avec des positions vraies (voir la Fig. 4) : les erreurs d'observation sont produites en utilisant les variances d'erreur opérationnelles pour le géopotentiel, la température, le vent et l'humidité spécifique. La variation de l'écart-type de ces paramètres est donnée par la courbe en trait plein dans la Fig. 5. Les observations fictives sont produites avec $\boldsymbol{y}^o = H(\boldsymbol{x}^t) + \boldsymbol{\epsilon}^o$, où $\boldsymbol{x}^t$ désigne la représentation fictive de l'état atmosphérique donnée

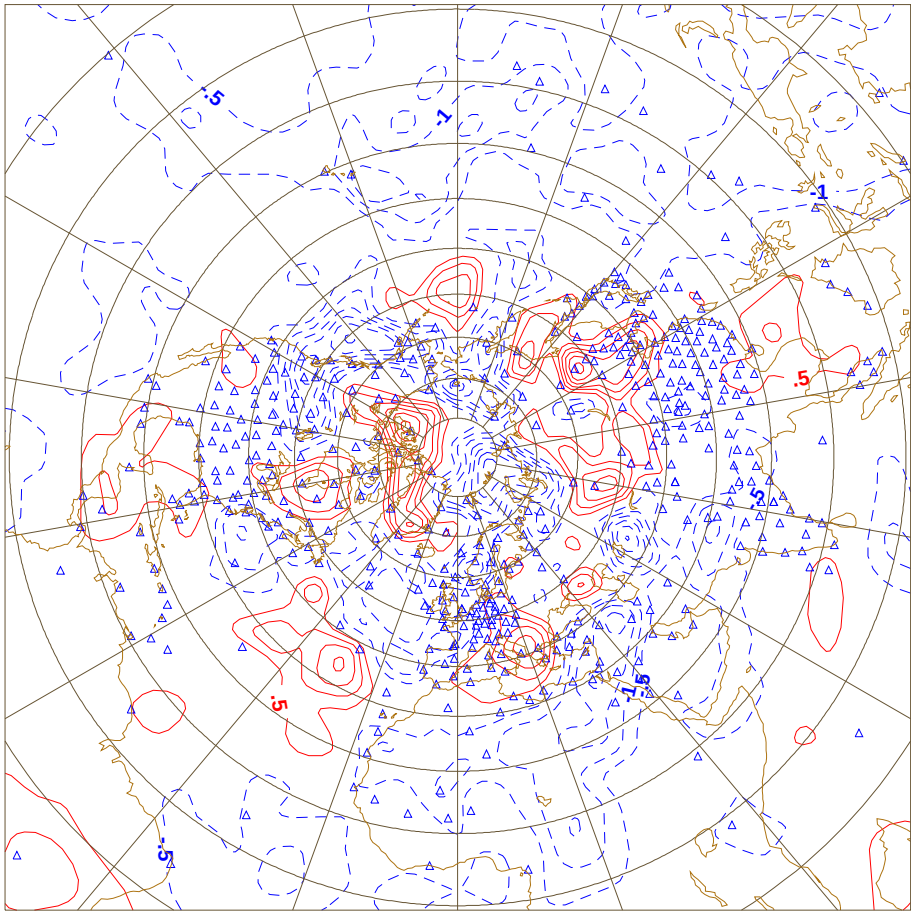

FIG. 4 – Exemple de réalisation d'un champ d'erreurs d'ébauche pour la température à 500 hPa ; les isothermes sont espacées de 0.5 K ; les triangles indiquent les positions des mesures de radiosondages.

par la prévision ARPEGE, H est l'opérateur associé aux observations de radiosondages et $\boldsymbol{\epsilon}^o$ un vecteur d'erreur d'observations.

La procédure de réglage décrite précédemment a été mise en place dans le 3D-Var ARPEGE pour optimiser les valeurs de variances d'erreurs associées aux radiosondes. Dans ce test, il n'y a pas de recherche simultanée des variances d'erreur d'ébauche : leurs valeurs sont gardées constantes au cours des itérations et égales à leurs valeurs optimales utilisées pour la simulation des erreurs d'ébauche. Ainsi, chaque itération de la procédure de réglage nécessite une paire d'analyses 3D-Var avec des observations non perturbées et perturbées. Les profils initiaux des paramètres d'erreur d'observation (courbe pointillée dans la Fig. 5) ont été choisis délibérément très différents des profils vrais : ils sont contants suivant la verticale pour le géopotenetiel, la température, le vent et l'humidité relative. La Fig. 5 montre que la convergence de l'algorithme itératif est extrêmement rapide : les profils obtenus après une seule itération (courbe en trait tireté-pointillé) sont très proches de ceux obtenus à l'itération 5 (courbe en trait tireté). De plus, la Fig. 5 montre également que les profils restitués sont très proches des profils opérationnels utilisés pour la simulation des observations (courbe en trait plein).

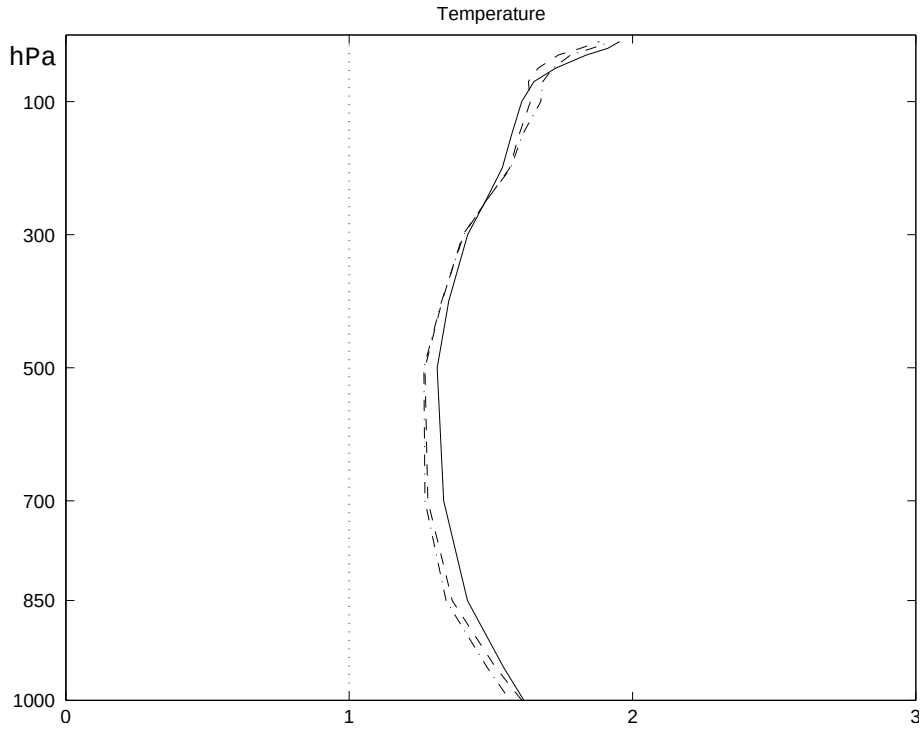

FIG. 5 – Profils d'erreurs de température utilisés pour la simulation des observations (trait plein), imposés au début de la procédure itérative (trait pointillé), après 1 itération (trait tireté-pointillé) et après 5 itérations (trait tireté).

Dans son travail de thèse, Chapnik (Chapnik 2005) a montré que les pondérations $s_i^{b^2}$ et $s_i^{o^2}$ respectant les critères définis précédemment sont aussi les solutions de la méthode du Maximum de Vraisemblance pour estimer ces paramètres (Dee et da Silva 1998). En effet, en supposant que les erreurs d'observations et d'ébauche sont gaussiennes, le vecteur d'innovation est également un vecteur aléatoire gaussien d'espérance nulle et de variance $\boldsymbol{D}$. Considérant un vecteur de paramètres $\boldsymbol{s}$ tel que $\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{s})$, la densité de probabilité conditionnelle d'une certaine réalisation $\boldsymbol{d}$, sachant que $\boldsymbol{s}$ est le vecteur de paramètres, est définie par la fonction de vraisemblance

$$f(\boldsymbol{d}, \boldsymbol{s}) = \frac{1}{\sqrt{(2\pi)^p det(\boldsymbol{D}(\boldsymbol{s}))}} \exp(-\frac{1}{2}\boldsymbol{d}^T \boldsymbol{D}(\boldsymbol{s})^{-1}\boldsymbol{d}).$$

L'estimateur au sens du Maximum de Vraisemblance du vecteur de paramètres $\boldsymbol{s}$ est défini comme le vecteur qui maximise $f(\boldsymbol{d}, \boldsymbol{s})$ à $\boldsymbol{d}$ donné. En pratique, il est plus aisé de considérer la fonction de Log-Vraisemblance

$$L(\boldsymbol{s}) = -\log(f(\boldsymbol{d}, \boldsymbol{s})) = \frac{p}{2}\log 2\pi + \frac{1}{2}\det(\boldsymbol{D}(\boldsymbol{s})) + \frac{1}{2}\boldsymbol{d}^T \boldsymbol{D}(\boldsymbol{s})^{-1}\boldsymbol{d}.$$

En dérivant cette fonction par rapport à $\boldsymbol{s}$, on montre que les valeurs de $\boldsymbol{s}$ qui minimisent L sont celles vérifiant les relations introduites plus haut (Desroziers et Ivanov 2001).

Dee et da Silva (1998) ont montré que la Hessienne de la fonction L peut être utilisée pour évaluer l'écart-type des paramètres et leur identifiabilité.

Dee et da Silva (1998) ont également montré que l'aptitude du Maximum de Vraisemblance à produire des estimations correctes des paramètres dépend de l'hypothèse qu'une partie du vecteur d'innovation est spatialement corrélée ($\boldsymbol{H}\boldsymbol{\epsilon}^b$) alors que l'autre ne l'est pas ($\boldsymbol{\epsilon}^o$). Cette hypothèse est également utilisée dans Hollingsworth et Lönnberg (1986) afin d'évaluer les variances d'erreurs d'ébauche et d'observations.

L'examen de la Hessienne de L montre aussi qu'il y a peu d'influence d'une partie non réglée des paramètres sur une autre partie : par exemple, il est possible de déterminer le coefficient de réglage relatif à un certain type d'observations sans régler simultanément la matrice $\boldsymbol{B}$ ou d'autres parties de la matrice $\boldsymbol{R}$.

Des difficultés peuvent cependant provenir de la robustesse des hypothèses faites sur la covariance des innovations $\boldsymbol{D}$: une mauvaise estimation *a priori* de la structure de corrélation contenue dans les matrices $\boldsymbol{B}$ ou $\boldsymbol{R}$ peut entrainer une mauvaise estimation des paramètres de réglage.

Des expériences ont été menées par Chapnik *et al.* (2004) dans un cadre simplifié afin de vérifier l'influence du nombre d'observations p sur les estimations et aussi d'étudier l'influence de la spécification d'une corrélation sous-optimale dans $\boldsymbol{B}$ ou $\boldsymbol{R}$.

Pour chacune de ces expériences, deux cents réglages différents ont été réalisés afin d'obtenir une statistique des coefficients de réglage. Pour chaque expérience, les écarts types d'erreur d'analyse ont été calculés afin d'obtenir l'impact du réglage sur la distance moyenne de l'analyse à la réalité.

Le cadre simplifié utilisé est celui déjà mis en place par Desroziers et Ivanov (2001), où des observations sont assimilées sur un domaine circulaire de périmètre 40 000 km (la taille du cercle équatorial). Les observations sont réparties régulièrement sur le domaine. L'ébauche est arbitrairement choisie égale à zéro ($\boldsymbol{x}^b = \boldsymbol{0}$). Le signal vrai est généré dans l'espace spectral avec une corrélation et une variance spécifiées dans la matrice $\boldsymbol{B}^t$:

$\boldsymbol{x}^t = -\boldsymbol{\epsilon}^b = -\boldsymbol{B}^{1/2}\boldsymbol{\eta}^b$, où $\boldsymbol{\eta}^b$ est un vecteur aléatoire gaussien et avec une corrélation gaussienne de longeur de corrélation 300 km.

En supposant une matrice de corrélation diagonale pour les erreurs d'observations $\boldsymbol{R}^t = \sigma^{o2}\boldsymbol{I}_p$ et un opérateur d'observation $\boldsymbol{H}$, le vecteur d'observations est construit avec $\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x}^t + \boldsymbol{R}^{t1/2}\boldsymbol{\eta}^o$, où $\boldsymbol{\eta}^o$ est un vecteur aléatoire gaussien. Il est alors possible de prescrire des matrices $\boldsymbol{B}$ et $\boldsymbol{R}$ mal spécifiées qui seront utilisées dans l'analyse. Dans les expériences présentées ici, les rapports entre les erreurs vraies et spécifiées sont $s_t^{o1/2} = 0.5$ pour l'erreur d'observation et $s_t^{b1/2} = 0.75$ pour l'erreur d'ébauche.

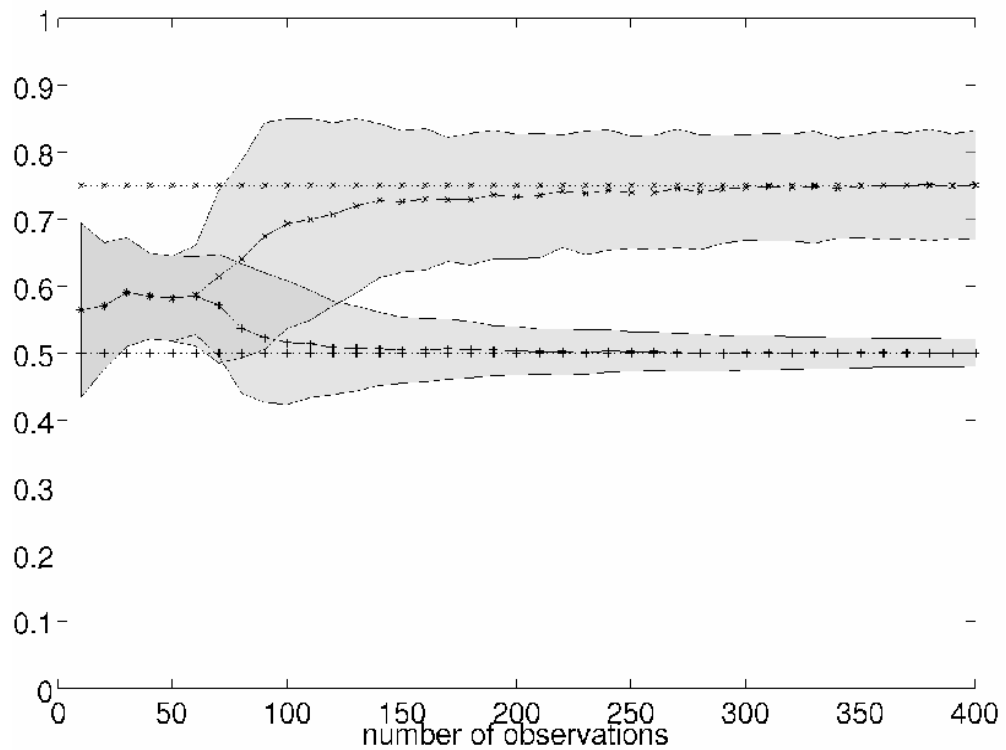

FIG. 6 – Influence du nombre d'observations sur les coefficients de réglage (coefficients portés en ordonnée ; les valeurs recherchées de $s^{o1/2}$ et $s^{b1/2}$ sont respectivement 0.5 et 0.75). L'abcisse est le nombre d'observations. Les courbes en trait plein avec les symboles "+" et "x" sont respectivement les moyennes des racines carrées de s^o et s^b calculés sur 200 expériences, alors que les courbes pointillées avec les mêmes symboles indiquent respectivement les valeurs des racines carrées de s^o et s^b. Les régions grisées se situent à moins d'un écart type (calculé sur les mêmes expériences) de la moyenne.

La Fig.6 montre le comportement de la méthode pour différents nombres d'observations. Pour chaque cas, on applique l'algorithme de point fixe sur deux cents réalisations aléatoires. On peut remarquer que pour les petits nombres d'observations, les comportements de s^o et s^b sont très semblables ($p \leq 50$). Dans ce cas, la corrélation d'erreur d'ébauche entre deux observations également réparties, et donc éloignées, est petite devant 1. On a alors $\boldsymbol{HBH}^T = \alpha\boldsymbol{R}$, où α est un coefficient de proportionalité. Dans ce cas, la méthode est incapable de faire la séparation entre le signal de l'erreur d'ébauche et celui des erreurs des observations. Pour chacune des deux cents réalisations les valeurs de s^o et s^b sont alors très proches.

Pour les valeurs intermédiaires de p ($50 \leq p \leq 150$) chaque coefficient se déplace progressivement vers sa valeur limite.

Pour les plus grandes valeurs de p l'estimation est près de la valeur vraie. Les écarts types des estimateurs décroissent avec le nombre d'observations, et on voit que l'estimation de s^o est plus précise que celle de s^b.

Ces résultats confirment que la densité spatiale des observations est un point clé de la qualité du réglage.

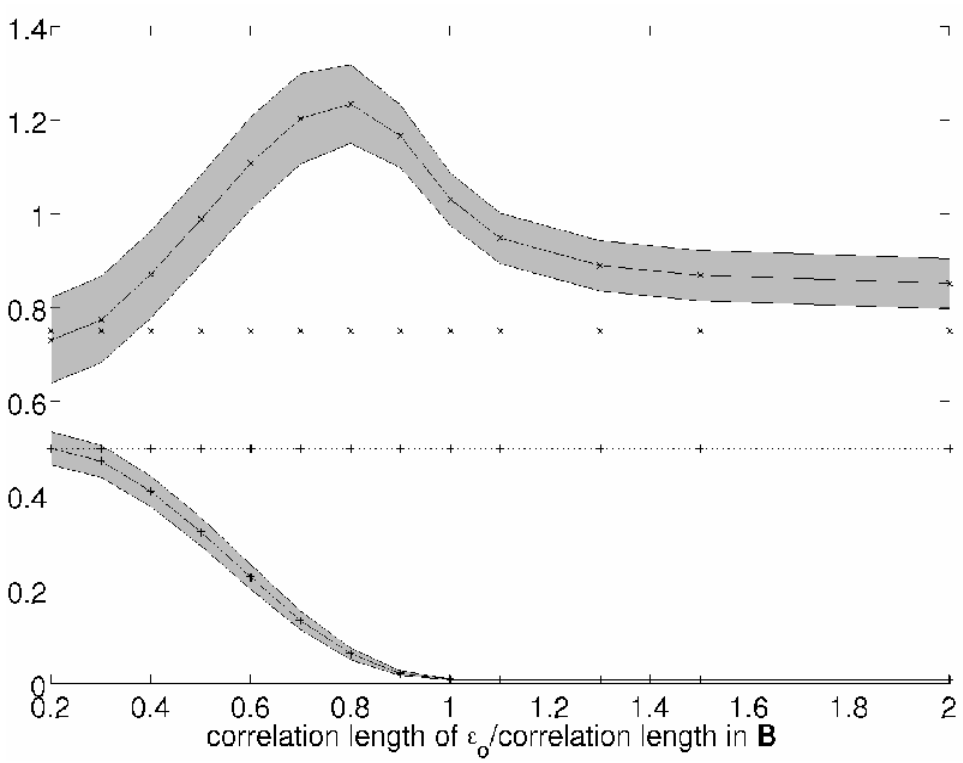

FIG. 7 – Influence d'une corrélation spatiale dans $\boldsymbol{\epsilon}^o$ sur les coefficients de réglage quand ils sont analysés avec une matrice $\boldsymbol{R}$ diagonale. Les coefficients de réglage sont portés en ordonnée (les valeurs recherchées de ${s^o}^{1/2}$ et ${s^b}^{1/2}$ sont respectivement 0.5 et 0.75). L'abscisse est le ratio " longueur de corrélation dans $\boldsymbol{\epsilon}^o$/ longueur de corrélation dans $\boldsymbol{\epsilon}^b$". Les conventions de tracé sont les mêmes que pour la Figure 6.

Les vraies observations peuvent avoir des erreurs spatialement corrélées, mais elles sont habituellement analysées avec un matrice $\boldsymbol{R}$ diagonale. Ce cas est simulé en générant 200 observations avec une erreur gaussienne spatialement corrélée alors qu'une matrice $\boldsymbol{R}$ diagonale est utilisée pour l'analyse. En réglant à la fois s^o et s^b, la Figure 7 montre une décroissance rapide de s^o avec la longueur de corrélation. Il est intéressant de noter qu'en faisant décroître σ^o l'algorithme agit de manière contraire à ce qui est en général fait pour prendre en compte des corrélations inconnues entre les erreurs d'observations (augmenter l'écart type, voir la discussion dans Liu, 2002). Les résultats ne changent pas en réglant s^o seul, que σ_b soit bien spécifié ou non (non montré).

Le réglage a un grand impact sur la variance de l'erreur d'analyse (Figure 7). La variance de l'analyse augmente avec la corrélation jusqu'à une valeur maximum, bien au-delà de la valeur pour l'analyse non-réglée. Dans cet exemple, l'analyse est vraiment dégradée par la procédure de réglage. Pour des observations faiblement ou non corrélées, la procédure améliore l'analyse.

Des expériences de réglage des écarts types d'erreur d'observation ont également été réalisées par Chapnik *et al.* (2006) dans le cadre du système de prévision numérique

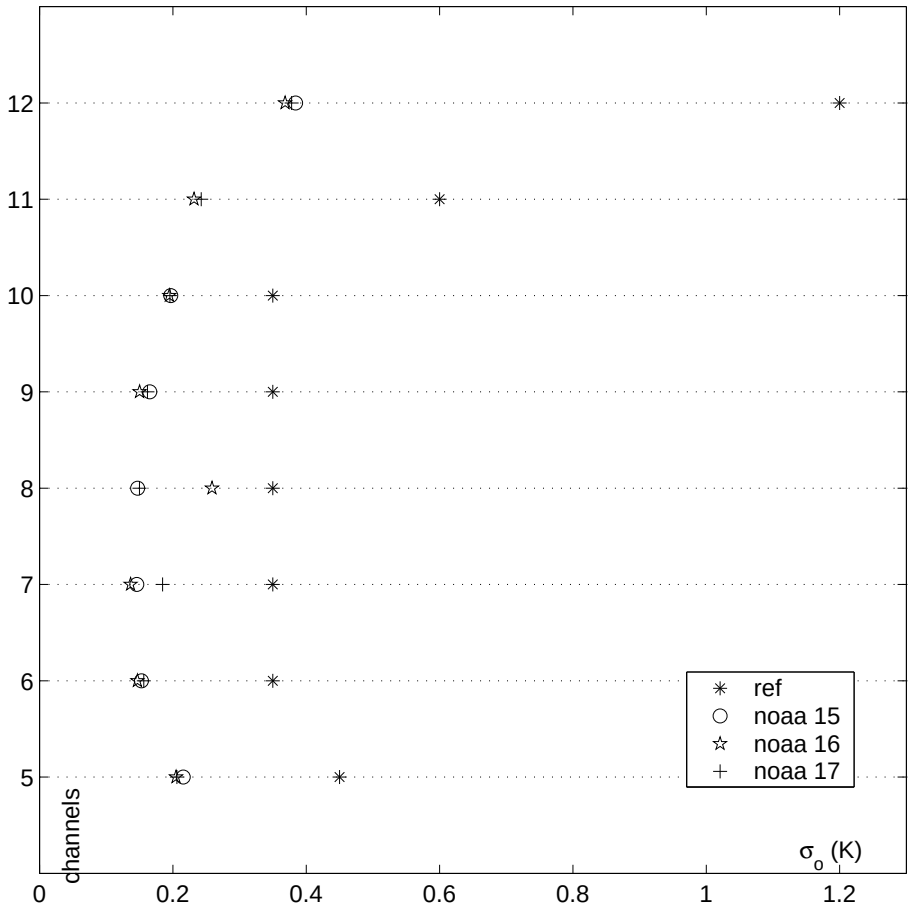

FIG. 8 – σ_o HIRS réglés pour NOAA 15 (symboles "o"), NOAA 16 (symboles "étoile") et NOAA 17 (symboles "+"). Le profil originel est le même pour les trois satellites (symboles "*").

français ARPEGE. ARPEGE est un modèle spectral global qui utilise un schéma d'assimilation variationnel 4D sauf pour les observations de surface, analysées avec un schéma d'interpolation optimale. Le choix a, en conséquence, été fait de n'inclure que les observations d'altitude et satellitaires. Le schéma d'analyse opérationnel est multi-incrémental, toutefois, afin de rester le plus près possible du cadre linéaire, on a omis la mise à jour de l'ébauche, ce qui revient à un simple 4D-Var incrémental. Le réglage a été réalisé sur des situations de cinq dates différentes : 04/04/2003, 09/04/2003, 14/04/2003, 19/04/2003 et 24/04/2003 à 00 UTC. Le choix d'utiliser des dates séparées par cinq jours est inspiré de Sadiki et Fischer (2004) qui ont mis en évidence la possibilité de corrélation temporelle si l'on emploie un écart plus court, ce qui invaliderait l'hypothèse ergodique implicitement faite ici.

Le réglage a été en particulier appliqué aux canaux NOAA AMSU-A. Sur la Figure 8 les valeurs réglées de σ_o pour les huit canaux assimilés sont montrées pour chacun des satellites NOAA 15, 16 et 17 (respectivement les marqueurs "o", "étoile" et "+") et comparées aux valeurs utilisées opérationnellement (identiques pour les trois satellites, symboles "*"). On peut remarquer les valeurs très similaires des écarts types pour les différents canaux, ce qui est un comportement attendu. Cette cohérence suggère que l'écart entre le canal 8 pour l'instrument de NOAA 16 et les deux autres ($\simeq$ 0.15 K pour NOAA 15 et 17 et 0.25 pour NOAA 16) reflète réellement une différence authentique que la méthode a été capable de détecter.

Application au calcul des DFS

Un critère intéressant pour l'évaluation d'un système d'assimilation de données est l'influence que le système donne aux observations. Cette influence peut être quantifiée comme la trace de la dérivée de l'analyse dans l'espace des observations par rapport aux vraies observations :

$$\begin{aligned} DFS &= Tr\ (\frac{\partial(\boldsymbol{H}\boldsymbol{x}^a)}{\partial \boldsymbol{y}}) \\ &= \sum_{\boldsymbol{y}_i \in\ observations} \frac{\partial(\boldsymbol{H}_i\boldsymbol{x}^a)}{\partial \boldsymbol{y}_i}. \end{aligned}$$

On pourra trouver plus d'informations sur la théorie et les propriétés du DFS en assimilation de données météorologique dans Rodgers (1996) ou Cardinali *et al.* (2004). Rabier *et al.* (2002) ont montré comment utiliser ce diagnostic pour la sélection des canaux IASI afin d'extraire l'information utile parmi l'immense quantité de données. Dans un cadre linéaire, le DFS est

$$DFS = Tr\ (\frac{\partial(\boldsymbol{H}\boldsymbol{x}^a)}{\partial \boldsymbol{y}}) = Tr\ (\boldsymbol{K}^T\boldsymbol{H}^T) = Tr\ (\boldsymbol{H}\boldsymbol{K}),$$

ce qui ne dépend que des éléments spécifiés du système d'assimilation. Ce diagnostic peut être particularisé pour un certain sous-ensemble d'observations ou même pour une seule observation tant que la matrice $\boldsymbol{R}$ demeure diagonale :

$$DFS_i = Tr\ (\boldsymbol{\Pi}_i\boldsymbol{H}\boldsymbol{K}\boldsymbol{\Pi}_i^T).$$

Talagrand (1999, 2004) a fourni un diagnostic objectif de la contribution de toute sous partie $\boldsymbol{z}_i$ du vecteur d'information en montrant que la quantité $Tr(\boldsymbol{S}_i^{-1/2}\boldsymbol{\Gamma}_i\boldsymbol{P}_a\boldsymbol{\Gamma}_i^T\boldsymbol{S}_i^{-1/2})$,

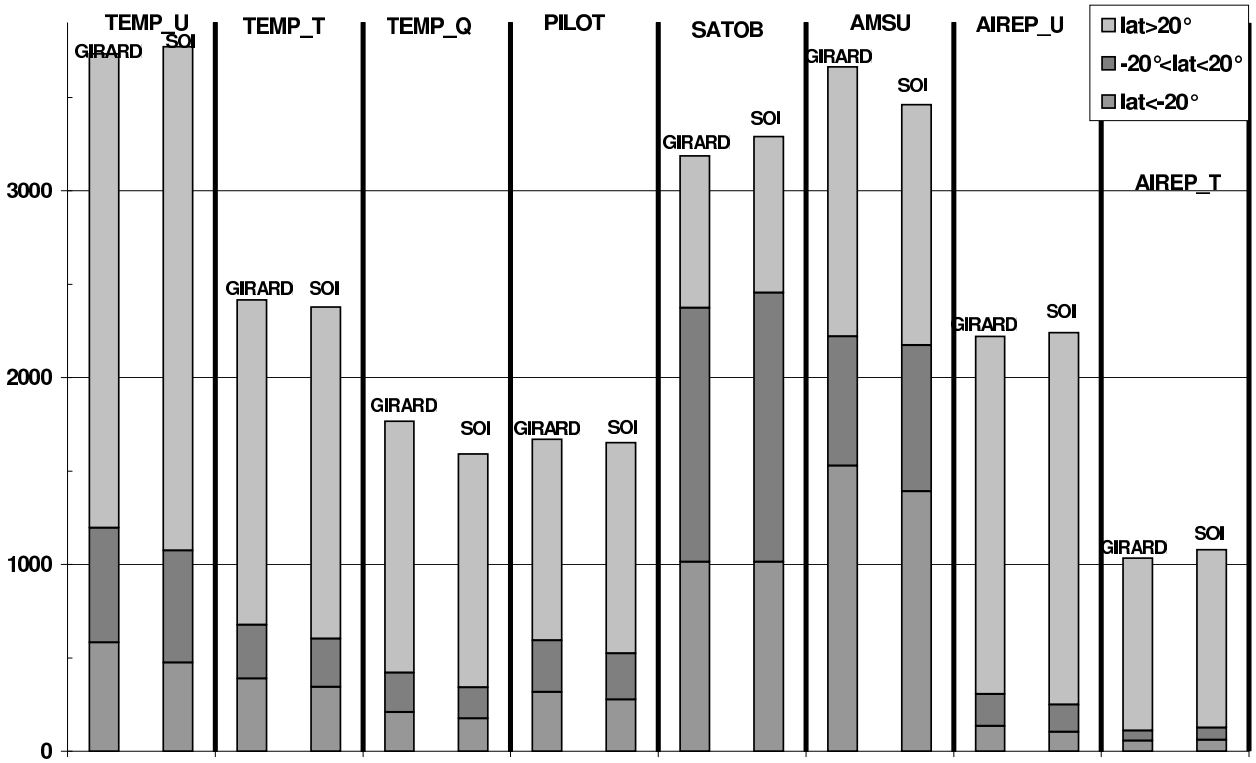

FIG. 9 – DFS pour différents jeux d'observations. Les jeux sont séparés par les traits noirs épais. Pour chaque jeu la barre de gauche a été calculée avec la méthode de Girard et la barre de droite avec la méthode IOS présentée dans Chapnik *et al.* (2006). Chaque barre est divisée en trois parties qui montrent les contributions des observations de l'hémisphère nord, des observations intertropicales et des observations de l'hémisphère sud. Les types TEMP U et AIREP U prennent en compte les composantes zonales et méridiennes du vent.

apparaissant dans l'expression (5) introduite plus haut, est la contribution du jeu d'information i à la précision globale, que les erreurs de ce jeu soient corrélées ou non aux autres erreurs. De plus, ce diagnostic est intrinsèque : il ne dépend pas du choix d'une base spécifique pour le vecteur de contrôle ou pour les observations. Dans le cas où le jeu d'information est un jeu i d'observations et sous la condition que les erreurs prescrites correspondantes ne montrent pas de corrélations avec le reste des erreurs du vecteur d'information (ce qui est le cas en pratique ici puisque la matrice prescrite $\boldsymbol{R}$ est diagonale), alors cette expression est égale au DFS : $Tr\ (\mathbf{\Pi}_i \boldsymbol{HK} \mathbf{\Pi}_i^T)$.

La Figure 9 montre une estimation du DFS pour huit types d'observations et paramètres. Les types d'observations utilisés ici sont ceux qui sont utilisés communément de manière opérationnelle. Il y a huit couples de barres, celles de gauche ont été calculées avec la méthode de Girard et celles de droite avec la méthode des Innovations Optimales Simulées (IOS) (Chapnik *et al.* 2006). Avec cette méthode, l'espérance des sous -parties de la fonction coût sont obtenues comme des statistiques de ces fonctions coût appliquées aux analyses obtenues avec des innovations simulées cohérentes avec les matrices $\boldsymbol{B}$ et $\boldsymbol{R}$ spécifiées dans le schéma d'analyse. L'examen des DFS montrent que les observations TEMP ont un grand impact, en particulier dans l'hémisphère nord, alors que pour AMSU-A, les contributions pour les deux hémisphères ont approximativement la même grandeur. Les SATOB contribuent le plus dans les régions intertropicales. Ces remarques sont cohérentes avec la répartition géographique des observations. On peut voir que les observations TEMP, prises dans leur ensemble, ont la plus large contribution au DFS. Cardinali *et al.* (2004) trouvent une plus grande contribution des données satellite, due au fait qu'une plus grande fenêtre d'assimilation a été utilisée pour le 4D-Var (12 heures au lieu de 6 heures dans la présente étude) et que l'écrémage géographique des observations est quatre fois moindre, ce qui augmente la contribution relative des données satellitaires alors que le même nombre d'observations TEMP est utilisé.

Conclusion

Nous avons montré que les critères de validation *a posteriori* proposés en particulier par Talagrand (1999) pouvaient être évalués numériquement par une méthode de Monte Carlo, basée sur une perturbation des observations. La possibilité d'évaluer ces critères permet d'autre part de régler certaines des statistiques des erreurs d'ébauche ou d'observations spécifiées dans le système d'assimilation. L'évaluation de ces critères permet également de mesurer l'impact respectif des observations dans l'analyse.

3.5 Ré-analyse 4D-Var de l'expérience FASTEX

Introduction

La mise en place opérationnelle du 4D-Var à Météo-France, à laquelle j'ai participé, s'est effectuée en juin 2000. Le développement de cet outil a été également l'occasion de produire une ré-analyse complète en mode 4D-Var de l'expérience FASTEX (Fronts and Atlantic Storm-Track EXperiment). Cette ré-analyse est décrite dans Desroziers *et al.* (2003), qui constitue le troisième article joint au manuscrit.

Motivation de la ré-analyse FASTEX en mode 4D-Var

Les objectifs principaux de FASTEX (Joly *et al.* 1997) étaient de vérifier ou de compléter les théories sur la cyclogénèse mais aussi de s'intéresser aux problèmes de prévisibilité des systèmes dépressionnaires et d'améliorer leur prévision opérationnelle qui est toujours une tâche difficile et importante. Le cas des tempêtes de Noël 1999 avec des impacts dévastateurs sur l'Europe de l'ouest et centrale sont les exemples les plus récents de tels phénomènes et de la difficulté de les prévoir à courte échéance (ces tempêtes ont en effet été prévues avec plus ou moins de succès par les différents centres de prévision numérique opérationnels).

FASTEX a été en particulier la première opportunité de tester les concepts d'observations adpatatives concentrées dans des régions où de petites erreurs d'analyse s'amplifient rapidement. Ces observations étaient fournies pendant FASTEX par des "dropsondes" larguées par des avions. D'autres mesures par dropsonde étaient également déployées dans les zones amont et aval de FASTEX de manière à observer les systèmes dépressionnaires depuis leur génèse jusqu'à leur stade mature. Un grand nombre de profils verticaux supplémentaires ont également été fournis par des sites terrestres ou des bateaux.

La phase terrain de FASTEX a pris place du 5 janvier au 27 février 1997 avec 19 Périodes Intensives d'Observation (POIs). Elle a produit un jeu unique d'observations des systèmes dépressionnaires des moyennes latitudes (Joly *et al.* 1999).

L'impact des données adaptatives ou "ciblées" de FASTEX a été étudié par Bergot (1999), Montani *et al.* (1999) et Szunyogh *et al.* (1999). Amstrup et Huang (1999) ont également examiné l'impact de toutes les données supplémentaires des radiosondes FASTEX.

Une étude pilote effectuée par Desroziers *et al.* (1999) sur une seule POI (POI 17) avait montré que la qualité des ré-analyses dépendait du schéma d'assimilation lui-même et qu'une assimilation 4D-Var améliorait de manière significative l'utilisation des données FASTEX caractérisées en particulier par leur répartition irrégulière dans le temps. Ces résultats ont conduit au projet d'une ré-analyse 4D-Var de l'expérience FASTEX. Cependant, la ré-analyse n'a été réalisée qu'après avoir collecté les données validées et corrigées des sondages FASTEX rendues disponibles dans la base de données FASTEX (Jaubert *et al* 1999).

Le 4D-Var est opérationnel à Météo-France depuis juin 2000 à la suite du CEPMMT qui avait implémenté un schéma similaire en novembre 1997. L'implémentation d'un tel schéma a été rendue possible par l'utilisation d'un formalisme multi-incrémental. La ré-analyse FASTEX a été la première ré-analyse d'une période aussi longue en mode 4D-Var.

L'utilisation du 4D-Var a rendu possible l'assimilation de toutes les données des sondages FASTEX alors que seulement une partie d'entre eux pouvaient être assimilée dans une ré-analyse parallèle 3D-Var également effectuée comme un test du gain apporté par le 4D-Var. La comparaison entre les scores des prévisions issues des ré-analyses FASTEX

4D-Var et 3D-Var ont été un argument supplémentaire pour rendre opérationnel le 4D-Var en juin 2000. Pour mesurer le poids des observations FASTEX dans les ré-analyses, une ré-analyse supplémentaire 4D-Var sans les observations FASTEX a également été effectuée.

Description de la ré-analyse FASTEX

Une description du 4D-Var développé à Météo-France et utilisé pour la ré-analyse FASTEX peut être trouvée dans Janisková *et al.* (1999) et Geleyn *et al.* (2001). La formulation 4D-Var développée à Météo-France inclut en particulier un terme supplémentaire de contrainte basé sur les filtres digitaux pour atténuer les oscillations parasites associées aux ondes de gravité (Gauthier and Thépaut 2001). Le modèle de prévision est le modèle opérationnel français ARPEGE (Courtier *et al.* 1991), qui est un modèle global spectral basé sur une géométrie étirée. La ré-analyse a été effectuée avec une résolution spectrale T199 sur la sphère étirée et un facteur d'étirement 3.5, correspondant approximativement à une résolution de 90 km près de Terre-Neuve, dans la zone amont FASTEX, et de 30km resolution près de l'Irlande dans zone aval FASTEX.

Pour réduire le coût du schéma 4D-Var, nécessitant environ 70 intégrations directes du modèle et de son adjoint, le formalisme appliqué pour la ré-analyse FASTEX est basé sur la formulation multi-incrémentale (Veersé and Thépaut 1998) où la résolution de l'incrément à l'ébauche est augmentée progressivement, avec des corrections finales effectuées à une résolution spectrale uniforme T95.

Comme mentionné précédemment, deux ré-analyses 4D-Var de l'ensemble de la période FASTEX ont en fait été réalisées : une analyse de référence avec seulement le jeu de données opérationnel et la ré-analyse proprement dite FASTEX avec l'ajout des observations FASTEX. Les données opérationnelles incluent toutes les données conventionnelles (SYNOP, AIREP, SATOB, TEMP, PILOT) mais aussi les données de radiances clarifiées TOVS dont l'utilisation directe est permise par le formalisme variationnel. Le jeu de données FASTEX correspond aux observations SYNOP-SHIP données par les bateaux FASTEX, aux données issues des bouées supplémentaires déployées durant l'expérience et aux profils TEMP fournis par les dropsondes larguées des avions FASTEX ou par les radiosondages additionnels effectués par des stations terrestres, par les quatre bateaux FASTEX ou par les bateaux ASAP. A noter que les bateaux FASTEX n'étaient pas stationnaires durant l'expérience mais se déplaçaient de manière à rester proches de la principale zone barocline. L'intérêt d'une telle procédure a été confirmée par une étude de simulation d'un système d'observation idéalisé (Fischer *et al.* 1998).

Tous les sondages FASTEX assimilés ont été extraits de la base de données FASTEX dans leur forme haute résolution et corrigée, puis mis au format conventionnel TEMP de manière à les rendre compatibles avec la résolution de l'analyse finale. Tous ces profils ont également été soigneusement vérifiés avant la ré-analyse (un contrôle de qualité supplémentaire intégré au schéma d'assimilation a également été effectué dans le cours de la ré-analyse).

Impact statistique des observations FASTEX

La résolution variable du modèle ARPEGE est relativement bien adaptée au design de l'expérience FASTEX : la résolution correspondante est suffisante ($\simeq$ 90 km) pour décrire les précurseurs de relativement grande échelle dans la zone amont près de Terre-Neuve et aussi suffisament fine pour décrire les structures de méso-échelle des systèmes matures dans la zone aval près de l'Irlande ($\simeq$ 30 km). Cependant, il faut noter que l'analyse

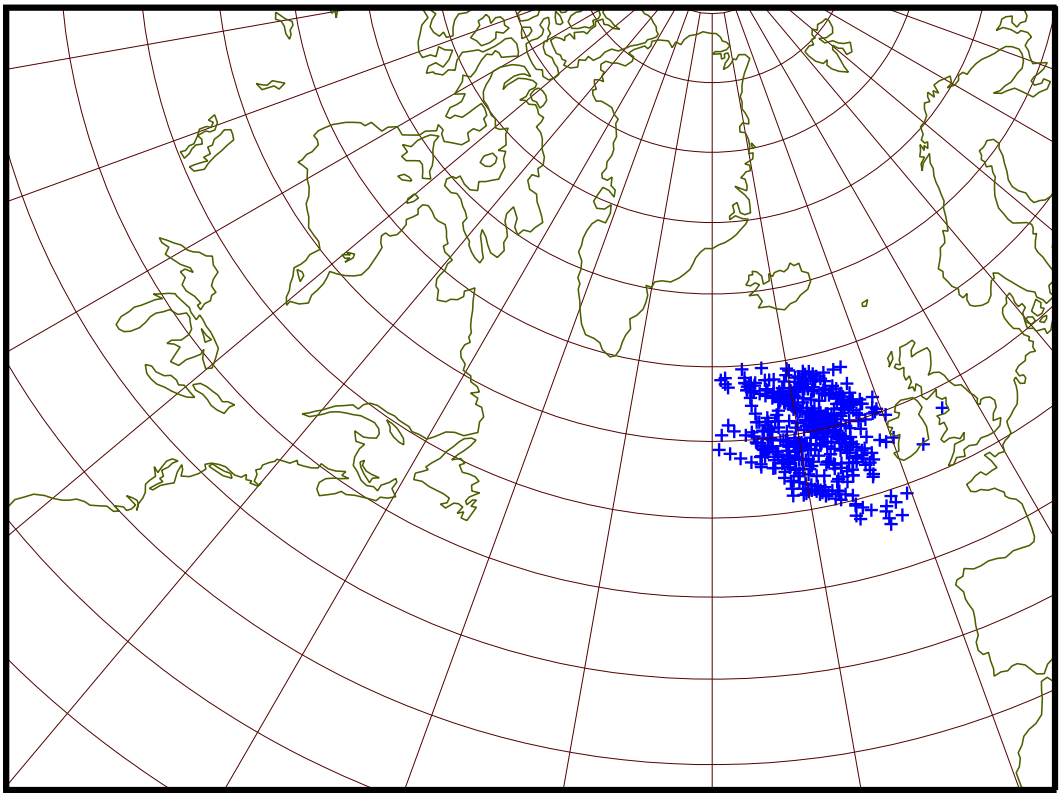

FIG. 10 – Positions des dropsondes larguées par l'avion britannique UK-C130.

n'est pas effectuée à pleine résolution. L'avion UK-C130 avait effectué un ensemble de vols fournissant un réseau de dropsondes avec une interdistance typique de l'ordre de la résolution d'ARPEGE dans cette zone (voir la Fig.10).

Les écarts quadratiques moyens entre les analyses et les observations dropsonde ont été calculés sur dix vols du UK-C130 effectués durant l'expérience et correspondant respectivement aux systèmes observés au cours des POIs 2, 7, 10, 11, 12, 15, 16, 17 and 19. Une description de ces vols peut être trouvée par exemple dans Joly *et al.* (1999). Il a été vérifié que les les ré-analyses FASTEX étaient plus proches des observations dropsonde UK-C130 que les ré-analyses opérationnelles. Ceci est attendu mais n'est pas une preuve suffisante de la qualité de ces analyses. La réduction de l'écart entre les analyses et les observations de température ou de vent est relativement homogène suivant la verticale. Son amplitude est d'environ 30% ce qui est tout à fait substantiel. Le meilleur accord des analyses FASTEX avec les observations est en fait déjà présent dans les champs d'ébauche correspondants, ce qui montre qu'il y a un impact cumulatif dans le temps des observations FASTEX dans le cycle d'assimilation et aussi que les analyses FASTEX sont de meilleure qualité que les analyses opérationnelles. Un critère synthétique pour évaluer l'impact des données FASTEX est donné par le calcul des écarts quadratiques moyens entre les deux analyses et les données UK-C130 de température, humidité relative et vitesse de vent, normalisés par leurs variances d'erreur. Ce critère correspond à l'évaluation de la fonction coût observation minimisée dans le schéma d'assimilation variationnel. La fonction coût observation est toujours plus faible pour l'analyse FASTEX que pour l'analyse opérationnelle et la réduction moyenne sur l'ensemble des vols est de l'ordre de 30% en accord avec la réduction observée sur les profils verticaux. La même fonction coût peut également être évaluée pour les ébauches correspondant respectivement aux analyses FASTEX et opérationnelles et montre que, dans la plupart des cas, les ébauches FASTEX sont plus

proches des observations que les ébauches opérationnelles avec une réduction de l'ordre de 10%.

La distance aux observations UK-C130 a également été calculée pour les prévisions 12 heures valides à l'instant de ces observations. Dans ce cas, le but est d'évaluer l'impact des observations FASTEX sur les prévisions dans la zone aval FASTEX, à l'ouest de l'Irlande, où les dropsondes UK-C130 étaient larguées. Les résultats (Fig. 11) montrent que la plupart des prévisions ont un meilleur score pour l'analyse FASTEX que pour l'analyse opérationnelle, avec quelques améliorations significatives (POI 19 en particulier). Ceci conduit à un impact moyen sensible comparable à celui obtenu pour les ébauches correspondant à une prévision 6 heures. Cependant l'impact moyen des données FASTEX sur la qualité des prévisions devient difficile à discerner à 24 heures d'échéance. A cette échéance, il y a environ autant de meilleures prévisions pour l'analyse FASTEX que pour l'analyse opérationnelle et les scores moyens deviennent semblables pour l'ensemble complet des vols. Cependant, il est probable que la zone de validation basée sur les dropsondes UK-C130 n'était pas bien adaptée pour évaluer les scores des prévisions à 24h. Alors que les vols UK-C130 étaient bien localisés pour étudier les dépressions à leur stade mature et pour vérifier l'impact des données FASTEX amont, il est probable qu'elles étaient placées trop à l'ouest pour identifier un développement en aval à cette échéance.

Une comparaison des prévisions aux observations situées dans un domaine plus aval a ainsi également été effectuée. Ce domaine (45N-60N,30W-10E) contient quelques uns des dropsondages UK-C130 précédents mais aussi les radiosondages effectués sur les Iles Britanniques et le nord de la France durant les POIs FASTEX (le nombre de ces sondages avaient été augmenté au cours de ces POIs). La comparaison entre les prévisions et ces observations a été effectuée durant les mêmes périodes que précédemment avec des vols UK-C130. Naturellement, le choix arbitraire de ce domaine est un compromis nécessaire pour capturer un signal dans la comparaison entre les prévisions FASTEX et opérationnelles (un domaine ou une période trop grand fait disparaître ce signal).

Le comportement des erreurs de prévision à 12 h obtenu avec ce nouveau critère est différent de celui obtenu avec la comparaison aux dropsondes UK-C130 : les erreurs de prévision sont plus faibles et apparaissent aussi plus régulières dans le temps mais aussi plus semblables entre les analyses FASTEX et opérationnelles. Ceci est dû à l'effet de moyennage d'un domaine de vérification plus grand. Mais le point le plus marquant est que l'impact moyen des observations FASTEX devient alors positif en comparant les prévisions 24 h aux radiosondes distribuées dans le domaine élargi. De plus, cet impact positif est encore présent pour les prévisions 36 h avec ce même domaine. La comparaison des prévisions pour chaque paramètre montre que l'amélioration moyenne des prévisions FASTEX peut être vue à la fois pour la température et le vent à 12h, mais est aussi discernable pour le vent à échéance 24 h and 36 h.

Conclusion

Une ré-analyse complète de la campagne FASTEX a été réalisée. Cette ré-analyse est disponible dans la base de données FASTEX et peut être utilisée en particulier pour l'étude des cyclogénèses dans l'Atlantique nord. Le schéma 4D-Var a permis de prendre en compte l'ensemble des données fortement asynchrones FASTEX, alors qu'un schéma de type 3D-Var ne peut prendre en compte qu'un sous-ensemble de ces données. L'impact positif des données FASTEX amont sur les prévisions à courte échéance a été mesuré. La comparaison des ré-analyses 4D-Var et 3D-Var a enfin été un argument supplémentaire pour la mise en place opérationnelle du 4D-Var à Météo-France en juin 2000.

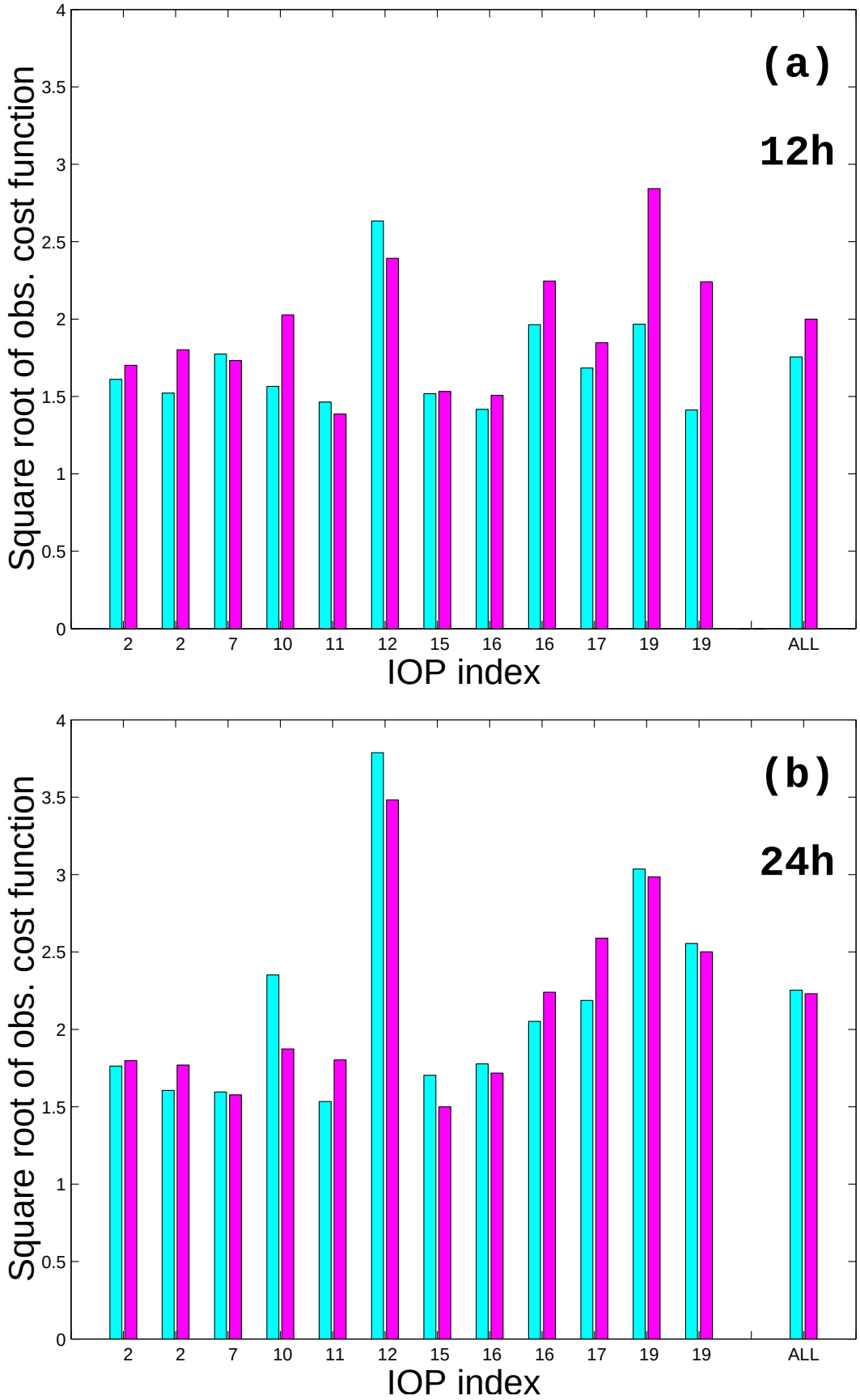

FIG. 11 – (a) racine carrée de la fonction coût observation mesurant la distance entre les prévisions 12h issues des analyses opérationnelles (barres foncées) et FASTEX (barres claires) et les observations UK-C130 de température, d'humidité relative et de vitesse du vent ; (b) comme (a) mais pour les prévisions 24 h.

3.6 Mesure de l'impact des observations sur les analyses et les prévisions

Introduction

Le travail décrit dans Desroziers *et al.* (2005), correspondant au quatrième article inclus dans ce mémoire, s'inscrit dans la suite des idées développées dans Desroziers et Ivanov (2001) et également du travail de thèse de B. Chapnik (2005). Dans sa thèse, B. Chapnik a en effet montré que les calculs par méthode de Monte Carlo de sous parties de la fonction coût du problème variationnel pouvaient être utilisés pour mesurer l'impact des observations sur les analyses. La même problématique est poursuivie dans Desroziers *et al.* (2005), mais avec la vision de la réduction de la variance d'erreur d'estimation.

Formalisme général de la réduction de variance d'erreur d'estimation

C'est un résultat classique en théorie de l'estimation linéaire (voir le paragraphe 3.2) que, si les matrices $\boldsymbol{B}$ et $\boldsymbol{R}$ sont bien spécifiées, alors la matrice de covariance d'erreur d'estimation s'écrit

$$\boldsymbol{A} = \boldsymbol{B} - \boldsymbol{KHB}.$$

Par conséquent, la réduction totale de variance d'erreur donnée par l'assimilation d'observations s'écrit

$$\boldsymbol{r} = Tr(\boldsymbol{B}) - Tr(\boldsymbol{A}) = Tr(\boldsymbol{KHB}), \tag{8}$$

où Tr désigne la trace d'une matrice. A noter que la reduction de variance d'erreur r est dépendante de la base dans laquelle on exprime ces erreurs ou en d'autres termes dépendante du choix d'une norme particulière.

La matrice $\boldsymbol{KHB}$ peut toujours être ré-écrite sous la forme

$$\boldsymbol{KHB} = \sum_i \boldsymbol{K}\boldsymbol{\Pi}_i{}^T\boldsymbol{\Pi}_i\boldsymbol{HB},$$

où $\boldsymbol{\Pi}_i$ est le projecteur qui permet de passer de l'ensemble des observations au sous-ensemble i d'observations.

La réduction associée au sous-ensemble i est donc

$$r_i = Tr(\boldsymbol{K}\boldsymbol{\Pi}_i{}^T\boldsymbol{\Pi}_i\boldsymbol{HB}),$$

avec

$$r = \sum_i r_i.$$

L'expression de r_i est similaire à l'expression (8), excepté que l'opérateur $\boldsymbol{\Pi}_i{}^T\boldsymbol{\Pi}_i$ est tout d'abord appliqué à $\boldsymbol{HB}$ de manière à sélectionner les observations associées au sous-ensemble i.

Transformation des matrices de covariance

Différents aspects de la réduction de variance d'erreur peuvent être examinés en utilisant une transformation $\boldsymbol{L}$ de l'erreur d'estimation $\boldsymbol{\epsilon}$ sur le vecteur d'état. Si $\boldsymbol{\epsilon}$ est transformé en $\boldsymbol{L\epsilon}$, alors la covariance du vecteur transformé devient $E(\boldsymbol{L\epsilon\epsilon}^T\boldsymbol{L}^T) = \boldsymbol{L}E(\boldsymbol{\epsilon\epsilon}^T)\boldsymbol{L}^T$.

En particulier, $\boldsymbol{B}$ and $\boldsymbol{A}$ deviennent respectivement $\boldsymbol{LBL}^T$ and $\boldsymbol{LAL}^T$.

De manière similaire, les réductions de variance pour l'ensemble complet et des sous-ensembles d'observations deviennent

$$r = Tr(\boldsymbol{LKHBL}^T)$$

et

$$r_i = Tr(\boldsymbol{LK}{\boldsymbol{\Pi}_i}^T\boldsymbol{\Pi}_i\boldsymbol{HBL}^T). \tag{9}$$

De nombreuses formes de la transformation $\boldsymbol{L}$ peuvent être imaginées. En particulier, $\boldsymbol{L}$ peut inclure une projection ou correspondre à une norme spécifique.

Un cas particulier est l'utilisation de la transformation $\boldsymbol{L} = \boldsymbol{B}^{-1/2}$, pour laquelle $\boldsymbol{B}$ est réduit à la matrice identité dans l'espace transformé, $\boldsymbol{A}$ à $\boldsymbol{B}^{-1/2}\boldsymbol{AB}^{-1/2}$ et la réduction de variance à $r = Tr(\boldsymbol{KH})$. Dans ce cas, on retrouve le fait déjà mentionné plus haut que

$$\begin{aligned} r_i &= Tr(\boldsymbol{LK}{\boldsymbol{\Pi}_i}^T\boldsymbol{\Pi}_i\boldsymbol{HBL}^T) \\ &= Tr(\boldsymbol{K}{\boldsymbol{\Pi}_i}^T\boldsymbol{\Pi}_i\boldsymbol{H}) \end{aligned}$$

est égal à l'espérance statistique de la sous-partie $J_i^o(\boldsymbol{x}^a)$ de la fonction coût associée au sous-ensemble d'observations i, si les matrices $\boldsymbol{B}$ et $\boldsymbol{R}$ sont correctement spécifiées (Talagrand, 1999).

La transformation $\boldsymbol{L}$ peut également inclure une prévision. Dans ce cas, les diagnostics sur l'analyse peuvent être étendus aux diagnostics sur les prévisions. Un opérateur de projection $\boldsymbol{P}$ sur une région de vérification peut aussi être introduit, ainsi qu'une métrique particulière $\boldsymbol{N}$ pour normaliser les erreurs de prévision dans cette zone. Dans ce cas, la variance de l'erreur de prévision devient

$$E(\| \boldsymbol{PM\epsilon} \|_N^2) = Tr(\boldsymbol{N}^{1/2}\boldsymbol{PM}E(\boldsymbol{\epsilon\epsilon}^T)\boldsymbol{M}^T\boldsymbol{P}^T\boldsymbol{N}^{1/2}),$$

où $\boldsymbol{\epsilon}$ correspond à un vecteur d'erreur sur les conditions initiales et $\boldsymbol{M}$ est le modèle linéaire-tangent associé au modèle de prévision.

Estimation par Monte Carlo des réductions d'erreur

On peut montrer que la réduction r_i associée à un jeu particulier d'observations peut être estimée par

$$r_i = \delta\boldsymbol{y}_i^{oT}{\boldsymbol{R}_i}^{-1}\boldsymbol{\Pi}_i\boldsymbol{HBL}^T\boldsymbol{LK}\delta\boldsymbol{y}^o,$$

où $\delta\boldsymbol{y}^o = \boldsymbol{R}^{1/2}\boldsymbol{\eta}^o$ est un vecteur de perturbations sur l'ensemble des observations et $\delta\boldsymbol{y}_i^o = \boldsymbol{\Pi}_i\delta\boldsymbol{y}^o$ est la projection de $\delta\boldsymbol{y}^o$ sur le sous-ensemble i d'observations.

Mais, comme pour le calcul des DFS dans la partie 3.4, $\boldsymbol{K}\delta\boldsymbol{y}^o$ n'est rien d'autre que la perturbation $\delta\boldsymbol{x}^a(\delta\boldsymbol{y}^o)$ sur l'analyse produite par une perturbation sur les observations puisque $\delta\boldsymbol{x}^a(\delta\boldsymbol{y}^o) = \boldsymbol{K}\delta\boldsymbol{y}^o$ dans le cas linéaire. Ainsi une estimation par Monte Carlo de r_i est donnée par

$$r_i \simeq \delta\boldsymbol{y}_i^{oT}{\boldsymbol{R}_i}^{-1}\boldsymbol{\Pi}_i\boldsymbol{HBL}^T\boldsymbol{L}\delta\boldsymbol{x}^a(\delta\boldsymbol{y}^o). \tag{10}$$

Il faut noter, que pour un opérateur donné $\boldsymbol{L}$, la chaîne d'opérations $\boldsymbol{HBL}^T\boldsymbol{L}\delta\boldsymbol{x}^a$, qui est la plus coûteuse, ne doit être effectuée qu'une seule fois. En particulier, la perturbation est effectuée sur l'ensemble des observations, mais l'évaluation de r_i pour tout sous-ensemble n'est effectuée que comme une dernière et peu coûteuse opération.

Si $\boldsymbol{L}$ contient l'intégration d'un modèle, alors $\boldsymbol{L}^T\boldsymbol{L}$ inclut l'intégration en avant du modèle linéaire-tangent $\boldsymbol{M}$ et une intégration en arrière du modèle adjoint correspondant $\boldsymbol{M}^T$.

Bien que l'expression précédente ait été dérivée dans le cas où l'incrément d'analyse est une combinaison linéaire des innovations, elle peut facilement être étendue à une analyse faiblement non-linéaire avec $\delta \boldsymbol{x}^a \simeq \boldsymbol{x}^a(\boldsymbol{y}^o + \delta \boldsymbol{y}^o) - \boldsymbol{x}^a(\boldsymbol{y}^o)$, où $\boldsymbol{x}^a(\boldsymbol{y}^o + \delta \boldsymbol{y}^o)$ et $\boldsymbol{x}^a(\boldsymbol{y}^o)$ sont respectivement les analyses avec des observations perturbées et non perturbées. Un exemple d'analyse faiblement non-linéaire est donné par le 4D-Var incremental où les non-linéarités apparaissent dans les mises à jour périodiques de la trajectoire avec le modèle complet (Courtier and Thépaut 1994, Rabier *et al.* 2000).

Application au schéma 4D-Var ARPEGE

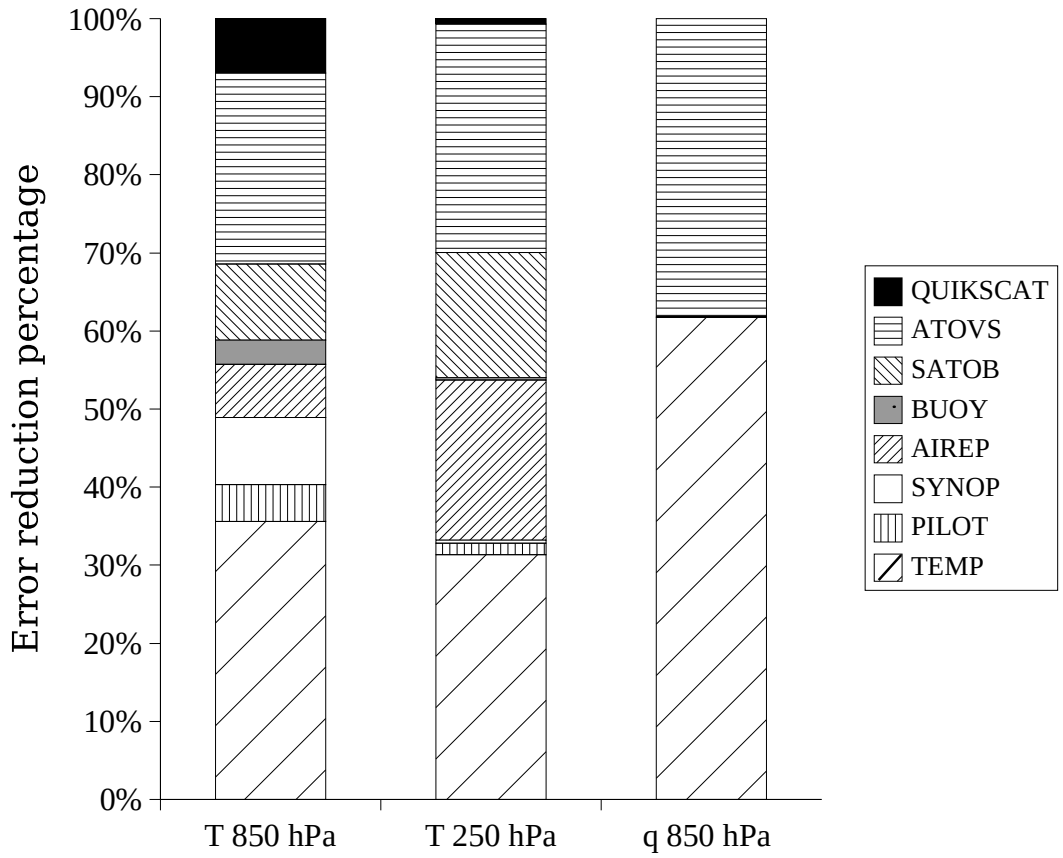

FIG. 12 – Pourcentage de réduction de variance d'erreur dû à différents jeux d'observations et pour trois champs différents (température à 850 HPa, température à 250 hPa, humidité specifique à 850 hPa).

Le calcul de la réduction d'erreur apportée par les observations a été implémenté dans le schéma français 4D-Var ARPEGE. Jusqu'à présent, seule la réduction d'erreur sur l'analyse a été étudiée. Le schéma 4D-Var ARPEGE utilise un grand nombre d'observations différentes (environ 300 000 observations sont utilisées dans chaque analyse).

La procédure de Monte Carlo décrite précédemment a été appliquée pour calculer les réductions totales ou partielles de variance d'erreur apportées par les observations. De manière à rendre les résultats plus faciles à interpréter, la possibilité d'effectuer une projection $\boldsymbol{P}$ sur un champ particulier a été mise en place. Ceci correspond au cas particulier où $\boldsymbol{L} = \boldsymbol{P}$ dans l'expression (10) de l'estimation par Monte Carlo de la réduction de variance d'erreur.

Dans une première étape, des tests ont été menés pour évaluer le nombre de perturbations nécessaire pour obtenir une estimation réaliste de la réduction de variance d'erreur

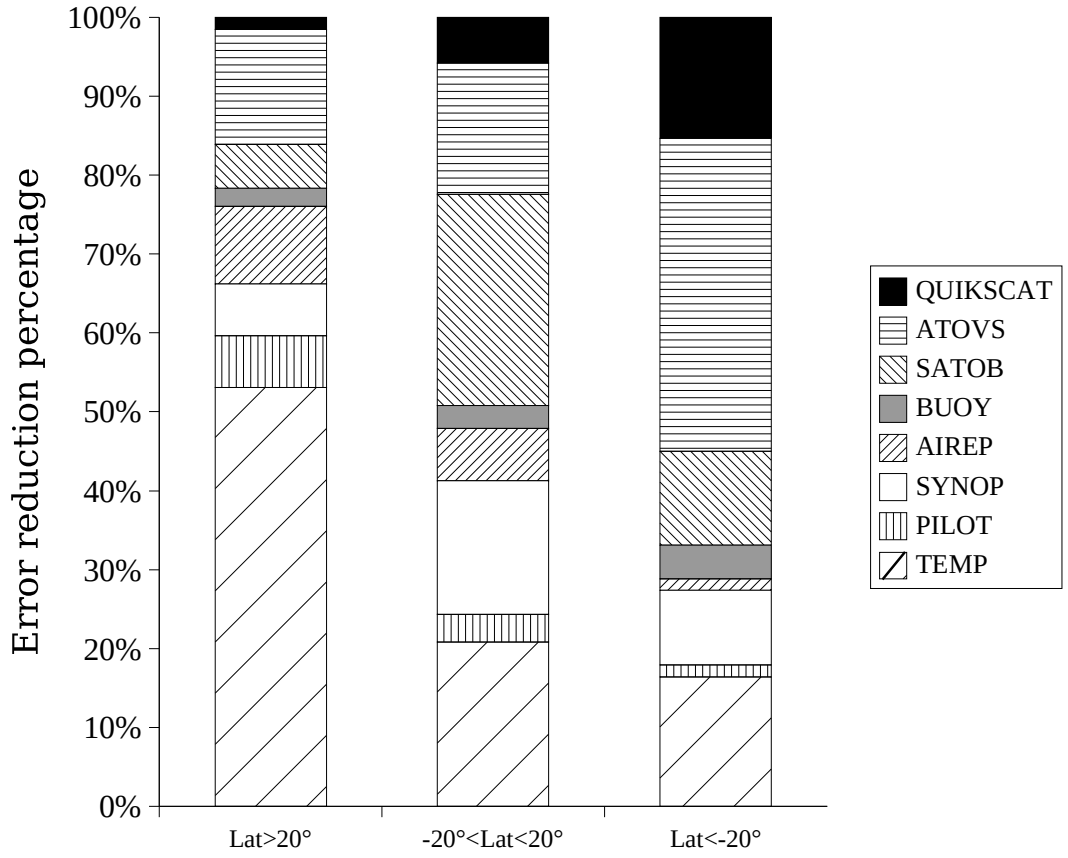

FIG. 13 – Pourcentage de réduction de variance d'erreur dû à différents jeux d'observations pour la température à 850 HPa et pour trois régions différentes.

apportée par les observations. Ils ont montré qu'environ cinq analyses perturbées sont nécessaires pour obtenir une estimation stabilisée de la réduction d'erreur. Ce nombre a également été estimé suffisant pour le calcul de la réduction d'erreur amenée par un sous-ensemble particulier d'observations.

Une autre validation des réductions de variance d'erreur obtenues a été effectuée en les comparant à la variance totale d'erreur de l'ébauche. Cette variance totale d'erreur de l'ébauche a été donnée par une autre procédure de Monte Carlo (Fisher and Courtier 1995, Andersson *et al.* 2000). Dans cette procédure, un ensemble de perturbations $\boldsymbol{\epsilon}^b = \boldsymbol{B}^{1/2}\boldsymbol{\eta}^b$, est construit, où $\boldsymbol{B}^{1/2}$ est l'opérateur correspondant à la racine carrée de la matrice de covariance d'erreur d'ébauche, et $\boldsymbol{\eta}^b$ est un vecteur de nombres aléatoires. Les variances d'erreur de l'ébauche sont déduites des statistiques du carré des perturbations d'ébauche $\boldsymbol{\epsilon}^b$. Le tableau 1 donne de telles variances totales d'erreur pour trois champs différents, et les réductions d'erreur correspondantes apportées par les observations. Dans la mesure où l'on ne s'intéresse à chaque fois qu'à un seul champ particulier, les variances sont calculées avec la norme euclidienne classique. Les pourcentages des réductions de variance d'erreur sont similaires pour la température à deux niveaux différents (850 hPa et 250 hPa) et paraissent raisonnables (de l'ordre de 30%). Le pourcentage plus faible de la réduction d'erreur pour l'humidité spécifique à 850 hPa est vraisemblablement lié au fait que l'analyse du champ d'humidité est en grande partie univariée et donc moins efficace en termes de réduction de l'erreur d'estimation.

Une fois la procédure de Monte Carlo vérifiée, il a été possible d'examiner les réductions d'erreur apportées par les différents jeux d'observations. Ceci a été fait pour trois champs

	Variance ébauche	Réduction de variance	Pourcentage de réduction
T à 850 hPa	20864 K^2	6558 K^2	31 %
T à 250 hPa	18373 K^2	6819 K^2	39 %
q à 850 hPa	0.0263 $(kg/kg)^2$	0.0038 $(kg/kg)^2$	14 %

TAB. 1 – Variance totale de l'ébauche et réduction totale de variance d'erreur pour la température à 850 hPa et 250 hPa et l'humidité spécifique à 850 hPa.

particuliers : la température à 850 hPa et 250 hPa et l'humidité specifique à 850 hPa. Les observations utilisées dans l'assimilation ARPEGE sont les suivantes :

- QUIKSCAT : vents de surface des scattéromètres (ou diffusiomètres),
- ATOVS : radiances des satellites en orbite polaire,
- SATOB : mesures de vitesse de vent déduites des déplacements de nuages donnés par l'imagerie des satellites géostationnaires,
- BUOY : observations à la surface des océans données par les bouées,
- AIREP : observations d'avions,
- SYNOP : observations de surface sur terre et sur mer,
- PILOT : observations de vent déduites de la poursuite manuelle des ballons pilot,
- TEMP : observations d'altitude de température, vitesse de vent et d'humidité des radiosondes.

La Fig. 12 montre la contribution de ces différents jeux d'observations à la réduction de l'erreur des champs sélectionnés. Pour le champ de température à 850 hPa, la contribution principale provient des TEMP et ATOVS. Ceci est logique puisque les TEMP fournissent des mesures de température et que les radiances sont supposées avoir un impact sur l'analyse de température. D'un autre côté, les SATOB, AIREP et PILOT apparaissent aussi apporter une contribution significative sur l'analyse de la température à 850 hPa, bien qu'ils ne correspondent qu'à des mesures de vitesse de vent. Ceci est une démonstration de l'impact de la formulation multivariée de l'analyse à travers la description des covariances entre les paramètres dans la matrice de covariance d'erreur d'ébauche $\boldsymbol{B}$ et aussi de l'impact de la dynamique dans le schéma 4D-Var.

La Fig. 12 montre également que les contributions des observations de surface BUOY et SYNOP s'annulent pour la température à 250 hPa, ce qui est de nouveau cohérent.

Finalement, il apparaît que l'analyse de l'humidité à 850 hPa est seulement influencée par les observations TEMP et ATOVS. Ceci est également logique puisqu'il n'y a pas de covariances entre l'humidité et les autres champs dans la matrice de covariance des erreurs d'ébauche dans ARPEGE (malgré le fait que l'on utilise ici un schéma 4D-Var) et puisque les TEMP et ATOVS sont la seule source d'information sur l'humidité dans l'ensemble des observations.

La Fig. 13 donne une autre vue de la contribution des sous-ensembles d'observations pour différentes régions. Les contributions apparaissent de nouveau cohérentes puisque, comme attendu, les données satellite ont proportionnellement plus d'impact dans l'hémisphère sud que dans l'hémisphère nord, ce qui est simplement lié au fait que les données in situ sont moins nombreuses dans l'hémisphère sud.

Conclusion

Une méthode de type Monte Carlo a été proposée pour estimer la réduction de variance d'erreur d'estimation due à la prise en compte des observations. Cette méthode de Monte Carlo est basée sur une perturbation des observations et est similaire à ce qui est

fait pour le calcul de l'espérance des sous parties de la fonction coût, qui donnent également une mesure de l'impact des observations dans l'analyse. Les deux critères peuvent donc être calculés par cette même procédure de Monte Carlo. L'intérêt du calcul de la réduction de la variance d'erreur d'estimation est qu'il peut être transposé au calcul de la réduction de l'erreur de prévision (et non pas seulement de l'erreur d'analyse) apportée par chaque type d'observations. Le calcul de la réduction de variance d'erreur d'estimation permet également d'obtenir la contribution d'un sous-ensemble des observations à un sous-ensemble de l'état analysé.

3.7 Diagnostics de cohérence dans l'espace des observations

Introduction

La possibilité de produire de nouveaux diagnostics d'optimalité d'une analyse a été développée dans Desroziers *et al.* (2005), qui correspond au cinquième article inclus dans ce mémoire. Ces diagnostics s'appuyent sur l'hypothèse que des schémas d'assimilation comme le 4D-Var incrémental restent proches de la théorie de l'estimation linéaire statistique. Ce travail s'inscrit encore une fois dans la ligne des idées traitées dans Desroziers et Ivanov (2001) et du thème de l'optimalité des systèmes d'assimilation mis en particulier en avant par Talagrand (1999).

Formalisme et interprétation géométrique

On a vu plus haut qu'un premier diagnostic classique de la cohérence d'une analyse est de comparer la covariance du vecteur d'innovation à la somme de la covariance des erreurs d'observation et des erreurs d'ébauche projetées dans l'espace des observations, ce qui s'écrit

$$E[\boldsymbol{d_b^o}(\boldsymbol{d_b^o})^T] = \boldsymbol{R} + \boldsymbol{HBH}^T, \tag{11}$$

(dans un souci d'homogénéisation des notations dans cette partie du rapport, on notera ici $\boldsymbol{d}_b^o$ le vecteur des différences *observations-moins-ébauche*). Cette relation fournit un test de cohérence global sur la spécification de ces covariances (Andersson 2003).

On a montré dans Desroziers *et al.* (2005) comment trois relations additionnelles peuvent être obtenues, fournissant des diagnostics séparés des statistiques d'erreur d'ébauche, d'observation et d'analyse.

La première relation est

$$E[\boldsymbol{d}_b^a(\boldsymbol{d}_b^o)^T] = \boldsymbol{HBH}^T, \tag{12}$$

où $\boldsymbol{d}_b^a$ est le vecteur des différences *analyse-moins-ébauche*.

Cette relation fournit un test de consistence séparé sur les covariances des erreurs d'ébauche dans l'espace des observations.

La deuxième relation est

$$E[\boldsymbol{d}_a^o(\boldsymbol{d}_b^o)^T] = \boldsymbol{R}, \tag{13}$$

où $\boldsymbol{d}_a^o$ est le vecteur des différences *observations-moins-analyse*.

Elle constitue un deuxième diagnostic additionnel qui fournit un test de cohérence séparé sur les covariances des erreurs d'observation.

La troisième relation est

$$E[\boldsymbol{d}_b^a(\boldsymbol{d}_a^o)^T] = \boldsymbol{HAH}^T, \tag{14}$$

où $\boldsymbol{A}$ est la matrice de covariance des erreurs d'analyse dans l'espace du modèle.

Elle constitue un troisième diagnostic additionnel, fournissant une information sur les covariances des erreurs d'analyse dans l'espace des observations.

Les relations (11) à (14) devraient donc être vérifiées dans une analyse cohérente (elles ne sont pas nécessaires à l'optimalité d'un schéma d'assimilation). Ces relations sont matricielles, ce qui signifie qu'elles devraient être également vraies pour les covariances croisées entre les différences associées à différentes observations.

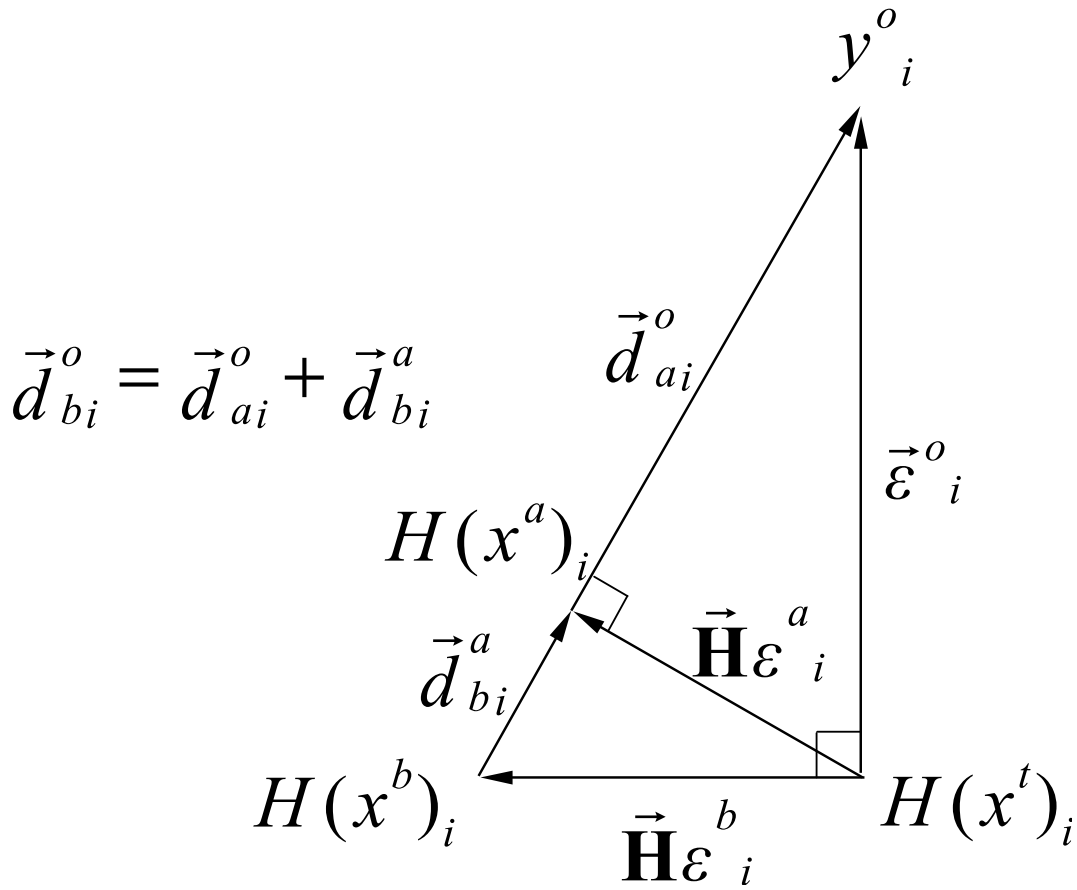

FIG. 14 – Représentation géométrique de l'analyse projetée sur un vecteur propre $\vec{\boldsymbol{v}}_i$ de la matrice $\boldsymbol{HK}$.

Une interprétation géométrique de ces relations peut être proposée dans l'espace des vecteurs propres $\boldsymbol{V}$ de la matrice $\boldsymbol{HK}$, tels que $\boldsymbol{HK} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T$, où $\boldsymbol{\Lambda}$ est la matrice diagonale des valeurs propres de $\boldsymbol{HK}$. Avec cette décomposition, le vecteur $\boldsymbol{d}^a_b$ des différences entre l'analyse et l'ébauche dans l'espace des observations peut être ré-écrit $\boldsymbol{d}^a_b = \boldsymbol{H}\delta\boldsymbol{x}^a = \boldsymbol{HK}\boldsymbol{d}^o_b = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T\boldsymbol{d}^o_b$. Par conséquent, la projection de $\boldsymbol{d}^a_b$ sur les vecteurs propres de $\boldsymbol{HK}$ est donnée par $\boldsymbol{V}^T\boldsymbol{d}^a_b = \boldsymbol{V}^T\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T\boldsymbol{d}^o_b = \boldsymbol{\Lambda}\boldsymbol{V}^T\boldsymbol{d}^o_b$. Si $\vec{\boldsymbol{d}}^a_{bi}$ et $\vec{\boldsymbol{d}}^o_{bi}$ désignent respectivement les projections de $\boldsymbol{d}^a_b$ et $\boldsymbol{d}^o_b$ sur un vecteur propre particulier $\vec{\boldsymbol{v}}_i$ de $\boldsymbol{HK}$, alors $\vec{\boldsymbol{d}}^a_{bi} = \lambda_i\vec{\boldsymbol{d}}^o_{bi}$, où λ_i est la valeur propre correspondante de $\boldsymbol{HK}$.

Dans la Figure 14, $H(\boldsymbol{x}^t)_i$, $\boldsymbol{y}^o{}_i$ et $H(\boldsymbol{x}^b)_i$ désignent respectivement les projections sur un vecteur propre particulier $\vec{\boldsymbol{v}}_i$ des équivalents vrais, observés et donnés par l'ébauche de $\boldsymbol{x}$ aux points d'observation. Le triangle $H(\boldsymbol{x}^b)_i$, $H(\boldsymbol{x}^t)_i$, $\boldsymbol{y}^o{}_i$ est rectangle en $H(\boldsymbol{x}^t)_i$ puisque les projections $\vec{\boldsymbol{\epsilon}^o}_i$ et $\vec{\boldsymbol{H}\boldsymbol{\epsilon}^b}_i$ des erreurs d'observation et d'ébauche sur le vecteur $\vec{\boldsymbol{v}}_i$ sont supposées décorrélées et donc orthogonales.

Cette orthogonalité est définie pour le produit scalaire particulier $< \vec{\epsilon}_1, \vec{\epsilon}_2 >= E(\epsilon_1\epsilon_2)$, où $\vec{\epsilon}_1$ and $\vec{\epsilon}_2$ sont deux vecteurs d'erreurs aléatoires, avec pour composantes ϵ_1 and ϵ_2 deux variables aléatoires. Ceci signifie que l'orthogonalité de $\vec{\boldsymbol{\epsilon}^o}_i$ et $\vec{\boldsymbol{H}\boldsymbol{\epsilon}^b}_i$ est définie d'un point de vue statistique. Avec cette définition du produit scalaire, l'angle entre deux vecteurs d'erreurs aléatoires est aussi directement lié à la corrélation des erreurs : un angle nul et un angle droit correspondent respectivement à une corrélation et dé-corrélation parfaites et un angle entre 0 et $\pi/2$ à une corrélation intermédiaire.

C'est un résultat classique (voir le paragraphe 3.2) que l'erreur d'analyse $\boldsymbol{H}\boldsymbol{\epsilon}^a$ est

également orthogonale au vecteur d'innovation $\boldsymbol{d}_b^o$, à nouveau d'un point de vue statistique. Puisque la Fig. 14 correspond à une projection sur un vecteur propre particulier $\vec{\boldsymbol{v}}_i$, il en résulte que $[H(\boldsymbol{x}^b)_i \vec{H}(\boldsymbol{x}^a)_i] = \vec{\boldsymbol{d}}_{b\,i}^a = \lambda_i \vec{\boldsymbol{d}}_{b\,i}^o$ et donc que $H(\boldsymbol{x}^a)_i$ est sur la ligne définie par le point $H(\boldsymbol{x}^b)_i$ et le vecteur $\vec{\boldsymbol{d}}_{b\,i}^o$. Le vecteur $\vec{\boldsymbol{H}\boldsymbol{\epsilon}}^a{}_i$ est donc perpendiculaire à $\vec{\boldsymbol{d}}_{b\,i}^o$ en $H(\boldsymbol{x}^a)_i$.

Par conséquent, l'application du théorème de Pythagore à ce triangle implique que $\|\vec{\boldsymbol{d}}_{b\,i}^o\|^2 = \|\vec{\boldsymbol{H}\boldsymbol{\epsilon}}^b{}_i\|^2 + \|\vec{\boldsymbol{\epsilon}}^o{}_i\|^2$, où la norme $\| \ \|$ est associée au précédent produit scalaire. Cette relation correspond au premier diagnostic classique (relation (11)) sur les statistiques du vecteur d'innovation.

L'application des théorèmes d'Euclide dans un triangle rectangle amène les trois relations supplémentaires suivantes : $< \vec{\boldsymbol{d}}_{b\,i}^a, \vec{\boldsymbol{d}}_{b\,i}^o >= \|\vec{\boldsymbol{H}\boldsymbol{\epsilon}}^b{}_i\|^2$, $< \vec{\boldsymbol{d}}_{a\,i}^o, \vec{\boldsymbol{d}}_{b\,i}^o >= \|\vec{\boldsymbol{\epsilon}}^o{}_i\|^2$, $< \vec{\boldsymbol{d}}_{b\,i}^a, \vec{\boldsymbol{d}}_{b\,i}^o >= \|\vec{\boldsymbol{H}\boldsymbol{\epsilon}}^a{}_i\|^2$, qui correspondent respectivement aux diagnostics des variances des erreurs d'ébauche (relation (12)), d'observation (relation (13)) et d'analyse (relation (14)).

Dans le cas ou les covariances des erreurs d'ébauche et d'observation sont homogènes, et la densité des observations uniforme, les vecteurs propres de la matrice $\boldsymbol{HK}$ sont les harmoniques spectrales (Fourier sur le plan, harmoniques sphériques sur la sphère). La Fig. 14 montre alors une projection sur une composante spectrale particulière. Ce cas est développé un peu plus loin.

Application aux analyses 4D-Var ARPEGE

Les diagnostics précédents fournissent potentiellement une information sur les covariances des erreurs d'observation, d'ébauche et d'analyse dans l'espace des observations. Une première application de ces diagnostics est de diagnostiquer les variances des erreurs d'observation et d'ébauche. En effet, pour tout sous-ensemble d'observations i avec p_i observations, il est possible de calculer simplement les quantités $(\widetilde{\sigma_i^b})^2 = (\boldsymbol{d}_b^a)_i^T (\boldsymbol{d}_b^o)_i / p_i = \sum_{j=1}^{p_i} (y_j^a - y_j^b)(y_j^o - y_j^b)/p_i$ et $(\widetilde{\sigma_i^o})^2 = (\boldsymbol{d}_a^o)_i^T (\boldsymbol{d}_b^o)_i / p_i = \sum_{j=1}^{p_i} (y_j^o - y_j^a)(y_j^o - y_j^b)/p_i$, où y_j^o est la valeur de l'observation j et y_j^b, y_j^a ses contre-parties respectives pour l'ébauche et l'analye. Les quantités $(\widetilde{\sigma_i^b})^2$ et $(\widetilde{\sigma_i^o})^2$ sont les valeurs diagnostiquées des erreurs d'ébauche et d'observation qui peuvent être différentes de leurs valeurs specifiées dans l'analyse. Elles correspondent à l'utilisation des relations (12) et (13) respectivement, mais seulement pour les éléments diagonaux de ces matrices. Ces calculs sont pratiquement sans coût et peuvent être effectuées, *a posteriori*, en utilisant une ou plusieurs analyses.

De tels calculs ont été effectués sur des analyses produites par le schéma opérationnel d'assimilation 4D-Var ARPEGE à Météo-France. Ce schéma d'assimilation 4D-Var est basé sur une formulation incrémentale (Courtier *et al.* 1994) et partage beaucoup d'aspects avec l'analyse 4D-Var du CEPMMT (Rabier *et al.* 2000). La Figure 15 montre un exemple des variances diagnostiquées des erreurs d'ébauche et d'observation pour des observations de vent données par les radiosondages dans l'hémisphère nord. Ces statistiques ont été calculées à partir de dix analyses 4D-Var. Elles montrent qu'à la fois les erreurs d'ébauche et d'observation semblent être sur-estimées dans l'analyse. Ce résultat est en accord pour les erreurs d'observations avec ce qui a été trouvé par Chapnik *et al.* (2005) en utilisant le diagnostic décrit au paragraphe 3.4. Les mêmes diagnostics ont été produits pour toutes les observations prises en compte dans le 4D-Var ARPEGE. L'ensemble de ces résultats confirme la sur-estimation des erreurs d'ébauche et d'observation pour la plupart des observations.

Réglage des variances d'erreur

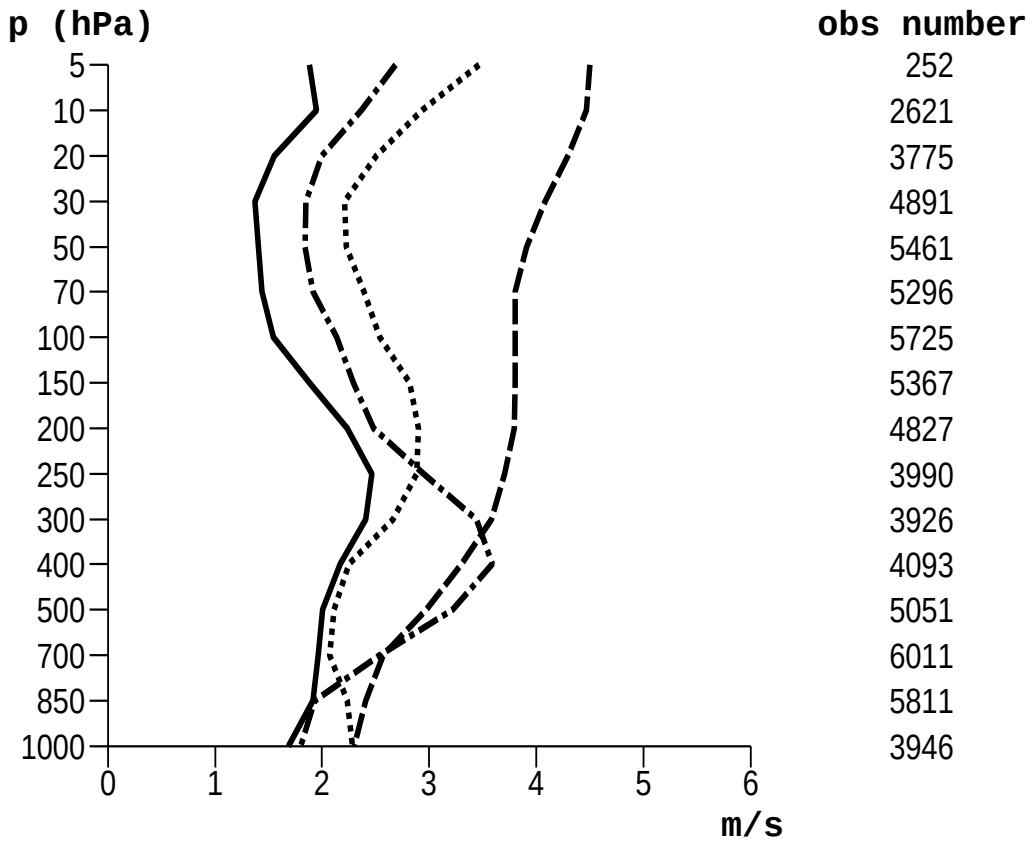

FIG. 15 – Profils verticaux des écarts-types diagnostiqués pour les erreurs d'ébauche (trait plein) et d'observation (trait pointillé) pour les observations des radiosondages dans l'hémisphère nord, comparés aux profils des écarts-types correspondants spécifiés pour les erreurs d'ébauche (trait tireté-pointillé) et d'observation (trait tireté). Toutes les valeurs sont en m/s. Les nombres d'observations utilisés pour calculer ces statistiques sont indiqués sur le côté droit de la Figure.

Puisque l'on sait que les erreurs d'ébauche et d'observation sont imparfaitement spécifiées dans une analyse opérationnelle, une méthode peut être envisagée pour les régler. Le principe d'une telle procédure de réglage est de trouver les valeurs de σ_i^b et σ_i^o, pour les différents sous-ensembles i d'observations telles que ces valeurs vérifient les relations $(\sigma_i^b)^2 = (\boldsymbol{d}_b^a)_i^T(\boldsymbol{d}_b^o)_i/p_i$ et $(\sigma_i^o)^2 = (\boldsymbol{d}_a^o)_i^T(\boldsymbol{d}_b^o)_i/p_i$. Ceci est un problème non-linéaire puisque les $(\boldsymbol{d}_b^a)_i$ et $(\boldsymbol{d}_a^o)_i$ dépendent eux-mêmes des valeurs de σ_i^b and σ_i^o. Cependant la forme de ces équations non-linéaires suggère l'utilisation d'une méthode itérative de point-fixe pour résoudre ce problème de réglage.

Cette procédure itérative est similaire à celle proposée par Desroziers et Ivanov (2001) pour résoudre le même type de problème mais avec un critère d'optimalité différent.

Un test préliminaire de l'algorithme précédent a été fait sur un problème simplifié donné par une analyse spectrale sur un domaine circulaire et aussi utilisé par Desroziers et Ivanov (2001). La longueur du domaine est de 40 000 km et la troncature égale à 200 correspondant à $n = 401$ coefficients spectraux. La matrice de covariance d'erreur d'ébauche $\boldsymbol{B}$ est construite dans l'espace spectral à partir d'une fonction de structure gaussienne dans l'espace physique de portée 300 km. Le problème est résolu avec $p = 401$ observations et une matrice diagonale de covariance d'erreur d'observation $\boldsymbol{R}$, en supposant que les erreurs d'observations sont décorrélées. A la fois les erreurs d'ébauche

et d'observation sont simulées en accord avec les matrices de covariance $\boldsymbol{B}$ and $\boldsymbol{R}$ et avec les écarts-types d'erreur homogènes $\sigma^b = 1$ et $\sigma^o = 2$.

	true	initial	iter. 1	iter. 2	iter. 3	iter. 4	iter. 5
σ_o	2	1	1.73	1.89	1.95	1.97	1.98
σ_b	1	2	1.41	1.19	1.10	1.07	1.03

TAB. 2 – Valeurs vraies et réglées de σ_o et σ_b au cours des itérations de point fixe.

Le tableau 2 montre que, partant des valeurs erronées $\sigma^b = 2$ et $\sigma^o = 1$, l'algorithme itératif de point fixe permet de retrouver une bonne approximation des vraies valeurs en seulement quelques itérations.

Interprétation spectrale des covariances diagnostiquées

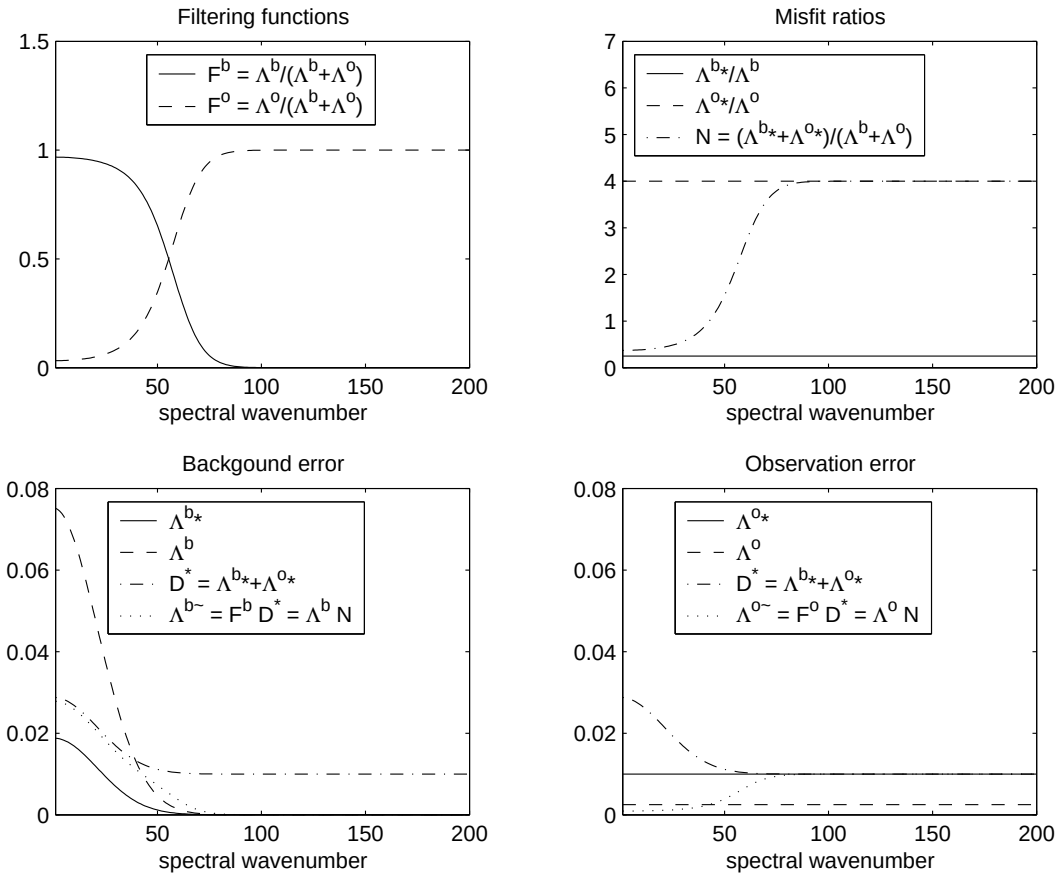

FIG. 16 – Interprétation spectrale des covariances diagnostiquées.

Les matrices covariance diagnostiquées dans les équations (12) et (13) peuvent être vues comme des covariances ajustées, notées respectivement $\boldsymbol{H}\tilde{\boldsymbol{B}}\boldsymbol{H}^T$ et $\tilde{\boldsymbol{R}}$. Elles s'expriment aussi

$$\begin{array}{rcccl} E[\boldsymbol{d}_b^a(\boldsymbol{d}_b^o)^T] & = & \boldsymbol{H}\tilde{\boldsymbol{B}}\boldsymbol{H}^T & = & \boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T \quad (\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1} \quad (\boldsymbol{H}\boldsymbol{B}^*\boldsymbol{H}^T + \boldsymbol{R}^*) \\ E[\boldsymbol{d}_a^o(\boldsymbol{d}_b^o)^T] & = & \tilde{\boldsymbol{R}} & = & \boldsymbol{R} \quad (\boldsymbol{H}\boldsymbol{B}\boldsymbol{H}^T + \boldsymbol{R})^{-1} \quad (\boldsymbol{H}\boldsymbol{B}^*\boldsymbol{H}^T + \boldsymbol{R}^*), \end{array}$$

où $\boldsymbol{B}^*$, $\boldsymbol{R}^*$ sont les vraies matrices de covariance.

En suivant Hollingsworth (1987) et Daley (1991, pages 125-128), on peut par exemple considérer les cas 1D ou 2D avec des covariances d'erreur d'ébauche et d'observation homogènes, et où la densité des observations est uniforme.

Dans ce cas, les différentes matrices de covariance ont les mêmes vecteurs propres, qui sont les modes spectraux : $\boldsymbol{HBH}^T = \boldsymbol{S}\boldsymbol{\Lambda}^{\boldsymbol{b}}\boldsymbol{S}^T, \boldsymbol{R} = \boldsymbol{S}\boldsymbol{\Lambda}^o\boldsymbol{S}^T$, où $\boldsymbol{S}$ est une matrice orthogonale dont les colonnes sont les vecteurs propres communs de $\boldsymbol{HBH}^T$ et $\boldsymbol{R}$, et $\boldsymbol{\Lambda}^b$, $\boldsymbol{\Lambda}^o$ sont les matrices diagonales qui contiennent les valeurs propres correspondantes. Les vecteurs propres sont les harmoniques spectrales (Fourier sur le plan, harmoniques sphériques sur la sphère), et $\boldsymbol{S}$ la transformation spectrale inverse. De plus, les valeurs propres correspondent aux variances des composantes spectrales

Des équations similaires s'écrivent pour les matrices de covariance diagnostiquées et exactes. On peut alors vérifier que

$$\boldsymbol{S}\tilde{\boldsymbol{\Lambda}}^b\boldsymbol{S}^T = \boldsymbol{S}\boldsymbol{\Lambda}^b(\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o)^{-1}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*})\boldsymbol{S}^T.$$

Le même type de relation peut être écrite pour la matrice diagnostiquée de covariance d'erreur d'observation. Ceci fournit les relations finales suivantes en termes de valeurs propres :

$$\begin{aligned}\tilde{\boldsymbol{\Lambda}}^b &= \boldsymbol{\Lambda}^b(\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o)^{-1}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}) \\ \tilde{\boldsymbol{\Lambda}}^o &= \boldsymbol{\Lambda}^o(\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o)^{-1}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}),\end{aligned}$$

où $\{\boldsymbol{\Lambda}^{b*}, \boldsymbol{\Lambda}^{o*}\}, \{\boldsymbol{\Lambda}^b, \boldsymbol{\Lambda}^o\}$ et $\{\tilde{\boldsymbol{\Lambda}}^b, \tilde{\boldsymbol{\Lambda}}^o\}$ désignent respectivement les valeurs propres des matrices exactes, spécifiées et diagnostiquées pour les erreurs d'ébauche et d'observation.

Un premier examen de ces relations montre que le spectre exact de variance de l'innovation $\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}$ est multiplié par deux rapports de filtrage, $\boldsymbol{F}^b = \frac{\boldsymbol{\Lambda}^b}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o}$ et $\boldsymbol{F}^o = \frac{\boldsymbol{\Lambda}^o}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o}$, pour donner les spectres de variances ajustées de $\boldsymbol{\epsilon}^b$ et $\boldsymbol{\epsilon}^o$:

$$\tilde{\boldsymbol{\Lambda}}^b = \frac{\boldsymbol{\Lambda}^b}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}) \text{ and } \tilde{\boldsymbol{\Lambda}}^o = \frac{\boldsymbol{\Lambda}^o}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}).$$

Les deux filtres sont représentés dans le panneau en haut à gauche de la Figure 16 pour l'exemple 1D discuté précédemment. En raison de la forme des deux types de corrélation d'erreur, $\boldsymbol{F}^b$ et $\boldsymbol{F}^o$ correspondent respectivement à une sorte de filtre passe-bas et passe-haut. Les panneaux du bas de la Fig. 16 illustrent le fait qu'appliquer $\boldsymbol{F}^b$, $\boldsymbol{F}^o$ à $\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}$ permet d'extraire respectivement ses composantes de grande et de petite échelle. Ils correspondent aux contributions respectives (estimées) des erreurs d'ébauche et d'observation. Ceci fournit des spectres ajustés de variances qui sont plus proches des spectres vrais que des spectres spécifiés. Ceci est également cohérent avec le fait que les équations indiquent que, si l'analyse est optimale, alors ces deux processus de filtrage doivent donner les spectres vrais de variance d'erreur.

Une autre vision complémentaire est de remarquer que les ajustements reviennent à multiplier chaque variance spectrale spécifiée par le rapport entre les variance exactes et spécifiées :

$$\tilde{\boldsymbol{\Lambda}}^b = \boldsymbol{\Lambda}^b\frac{\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o} \text{ et } \tilde{\boldsymbol{\Lambda}}^o = \boldsymbol{\Lambda}^o\frac{\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o}.$$

Les rapports d'ajustement parfaits devraient être plutôt respectivement $\frac{\boldsymbol{\Lambda}^{b*}}{\boldsymbol{\Lambda}^b}$ et $\frac{\boldsymbol{\Lambda}^{o*}}{\boldsymbol{\Lambda}^o}$. Mais, comme illustré par le panneau en haut à droite de la Fig. 16, le rapport de désaccord sur l'innovation $\boldsymbol{N} = \frac{\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*}}{\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o}$ est proche de $\frac{\boldsymbol{\Lambda}^{b*}}{\boldsymbol{\Lambda}^b}$ dans les grandes échelles (c'est-à-dire où

les erreurs d'ébauche ont leur plus grande amplitude), et est proche de $\frac{\mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{o}}$ dans une large bande de petites échelles (c'est-à-dire où les erreurs d'observation prédominent). En d'autres termes, le lien étroit entre $\boldsymbol{N}$ et $\frac{\mathbf{\Lambda}^{b*}}{\mathbf{\Lambda}^{b}}$ ou $\frac{\mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{o}}$, en fonction de l'échelle, indique que les ajustements sont faits de manière appropriée en fonction de l'échelle.

Les deux visions sont cohérentes, dans le sens qu'elles indiquent toutes les deux que les ajustements sont particulièrement pertinents lorsque les erreurs d'ébauche et d'observation tendent à prédominer pour des échelles différentes. Ceci est résumé par le fait que $\frac{\mathbf{\Lambda}^{b*}+\mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{b}+\mathbf{\Lambda}^{o}} \simeq \frac{\mathbf{\Lambda}^{b*}}{\mathbf{\Lambda}^{b}}$ dans les grandes échelles si à la fois $\mathbf{\Lambda}^{b*} \gg \mathbf{\Lambda}^{o*}$ et $\mathbf{\Lambda}^{b} \gg \mathbf{\Lambda}^{o}$ et $\frac{\mathbf{\Lambda}^{b*}+\mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{b}+\mathbf{\Lambda}^{o}} \simeq \frac{\mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{o}}$ dans les petites échelles si à la fois $\mathbf{\Lambda}^{b*} \ll \mathbf{\Lambda}^{o*}$ et $\mathbf{\Lambda}^{b} \ll \mathbf{\Lambda}^{o}$.

Les écarts-types d'erreur ajustés présentés précédemment pour le cas simplifié 1D sont simplement les racines carrées de la somme des différentes variances spectrales ajustées. La pertinence des ajustements de ces écarts-types est donc en accord avec les ajustements des variances spectrales.

De la discussion précédente, il apparaît en particulier que l'ajustement des variances des erreurs d'ébauche et d'observation est seulement pertinent si ces erreurs ont des structures différentes. En effet, l'application d'un tel ajustement ne fonctionnera pas dans le cas où les deux spectres (soit $\mathbf{\Lambda}^{b*}$ et $\mathbf{\Lambda}^{o*}$ ou $\mathbf{\Lambda}^{b}$ et $\mathbf{\Lambda}^{o}$) sont proportionnels. Ceci est également vrai pour les méthodes cherchant à diagnostiquer les covariances des erreurs de prévisions à partir des covariances du vecteur d'innovation (Hollingsworth and Lönnberg 1986).

Cependant, si les spectres des erreurs d'ébauche et d'observation sont suffisamment différents (à la fois pour les statistiques exactes et spécifiées), les ajustements seront capables de modifier les variances de manière correcte même si les corrélations ne sont pas parfaitement spécifiées.

Conclusion

Un nouveau jeu de diagnostics d'une analyse, basés sur les propriétés de l'estimation linéaire statistique, a été proposé. L'intérêt de ces diagnostics est qu'ils sont pratiquement gratuits et applicables à tout schéma d'analyse faiblement non-linéaire comme le 4D-Var incrémental. D'autre part, ces diagnostics permettent sous certaines conditions d'optimiser les statistiques des erreurs des différentes sources d'information. Les premières comparaisons des méthodes de réglage décrites dans Desroziers et Ivanov (2001) et Desroziers *et al.* (2005) montrent qu'elles produisent des résultats très similaires.

3.8 Utilisation des ondelettes pour les covariances d'erreur d'ébauche

Introduction

Le travail résumé ici correspond au travail de thèse d'O. Pannekoucke. Il s'apparente au travail réalisé dans Desroziers (1997) visant une meilleure spécification des covariances d'erreur d'ébauche mais avec un point de vue différent.

Le problème de la spécification des covariances d'erreur d'ébauche

L'estimation des covariances d'erreur d'ébauche est en effet toujours un problème particulièrement difficile puisque dans la pratique opérationnelle le vecteur d'ébauche est un vecteur de dimension $10^5 - 10^7$. Dans ce cas, il n'est pas seulement impraticable de manipuler une matrice de covariance de taille si importante, mais il est également impossible de la spécifier puisqu'il y a un manque d'information statistique (Dee 1995).

Pour surmonter cette difficulté, un modèle pour les statistiques d'erreur de prévision doit être défini. Un tel modèle repose souvent sur l'hypothèse que les corrélations sont homogènes et isotropes (Gaspari et Cohn 1999). Cette hypothèse est équivalente à considérer que la matrice de corrélation est diagonale dans l'espace spectral (Courtier *et al.* 1998), ce qui facilite la représentation des statistiques d'erreur.

Une approche pour déterminer la matrice de covariance d'erreur d'ébauche est d'utiliser un ensemble d'assimilations, obtenues par une perturbation des observations et éventuellement des paramétrisations du modèle de prévision (Houtekamer *et al.* 1996). Cette procédure a été récemment appliquée au CEPMMT et à Météo-France pour spécifier la composante stationnaire des covariances d'erreur (Fisher 2003 ; Belo Pereira et Berre 2005). Dans ce cas, les covariances sont calculées sur plusieurs semaines et l'hypothèse d'homogénéité est imposée.

Les hypothèses d'homogénéité et d'isotropie sont cependant une approximation grossière des erreurs réelles. Bouttier (1994) a en particulier montré que les échelles de corrélation dépendent de la situation météorologique et de la densité des observations.

En utilisant un ensemble d'assimilations, Belo Pereira et Berre (2005) ont montré de telles hétérogénéités et anisotropies dans les corrélations d'erreur d'ébauche en utilisant une nouvelle estimation des longueurs de corrélation.

Fisher (2003) a récemment introduit l'idée d'utiliser les ondelettes sur la sphère pour améliorer la représentation des corrélations d'erreur d'ébauche, et en particulier d'autoriser une part d'hétérogénéité dans la description de ces erreurs. Une telle formulation est maintenant utilisée opérationnellement au CEPMMT pour représenter des corrélations stationnaires mais inhomogènes, avec en particulier des corrélations plus longues dans les Tropiques. Une approche similaire a été considérée par Deckmyn et Berre (2005) pour un modèle en domaine limité.

La formulation ondelette

Les transformées ondelette directe et inverse d'un champ d'erreur d'ébauche $\boldsymbol{\epsilon}^b$ sont définies respectivement par (Fisher 2003, Pannekoucke *et al.* 2006) :

$$\hat{\boldsymbol{\epsilon}}_j^b = \boldsymbol{\epsilon}^b \otimes \boldsymbol{\psi}_j$$

et

$$\boldsymbol{\epsilon}^b = \sum_j \hat{\boldsymbol{\epsilon}}_j^b \otimes \boldsymbol{\psi}_j,$$

où $\otimes$ est le produit de convolution dans l'espace physique pour les deux expressions et $\boldsymbol{\psi}_j$ est une fonctions radiale associée à une échelle j. Chaque fonction $\boldsymbol{\psi}_j$ a une position et une échelle spécifique. Chaque coefficient $\hat{\boldsymbol{\epsilon}}^b_{j,x_j(i)}$ du champ transformé correspond à la valeur de l'erreur pour l'échelle j et la position $x_j(i)$ sur une grille dont la résolution dépend de j (la résolution est plus faible pour les grandes échelles).

Dans un cadre horizontal, l'approche diagonale ondelette pour $\boldsymbol{B}$ consiste à calculer les variances de ces coefficients $\hat{\boldsymbol{\epsilon}}^b_{j,x_j(i)}$. Comme chaque fonction ondelette contient une information sur à la fois la position et l'échelle, ces variances ondelette contiennent une information sur la forme locale des fonctions de covariance.

De plus, on peut montrer que cette approche diagonale ondelette revient à calculer un ensemble de moyennes locales et spatiales des fonctions de covariance. Elle apporte ainsi une manière de filtrer le bruit d'échantillonage présent dans les statistiques d'erreur d'ébauche.

Portées locales "climatologiques"

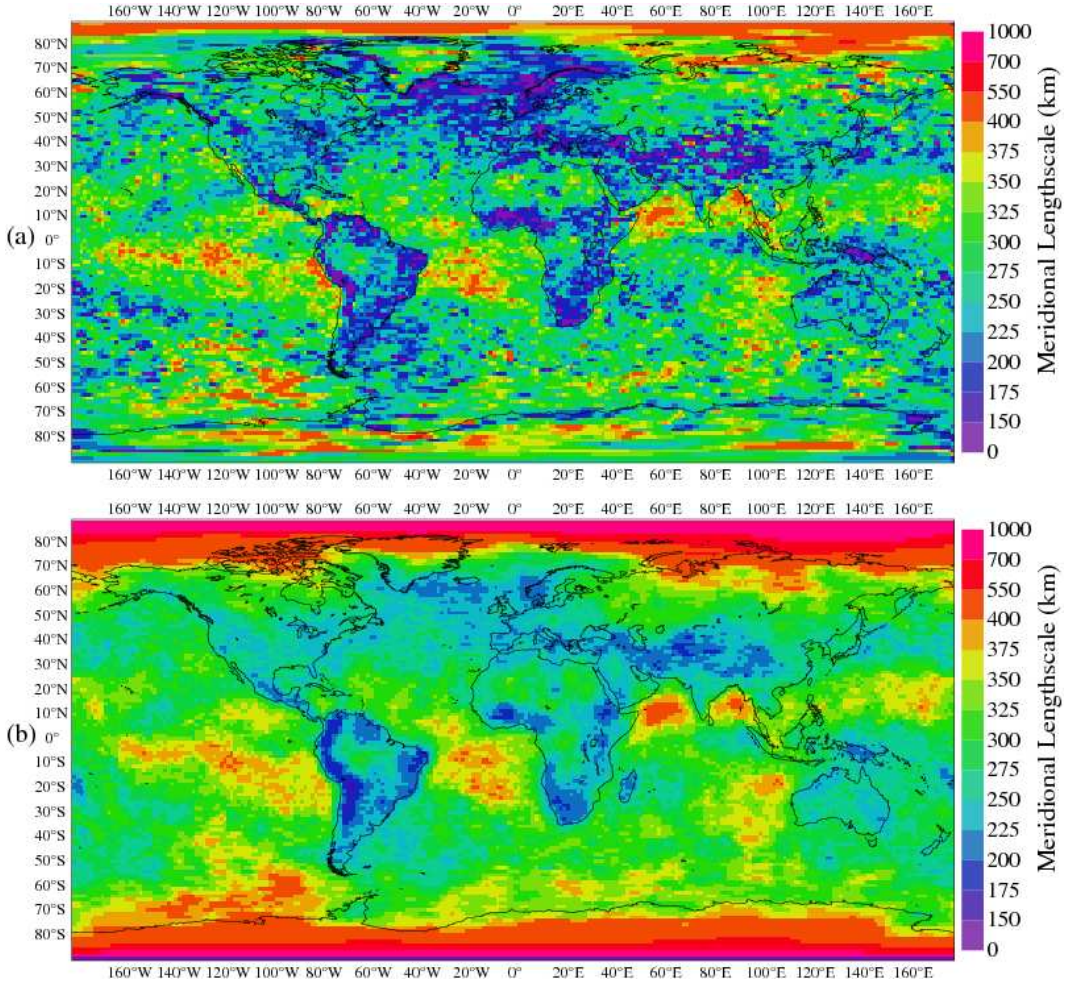

FIG. 17 – Portées méridiennes (en km) pour la pression de surface, moyennées sur une période de 46 jours et un ensemble de 6 membres. (a) : portées brutes. (b) : portées ondelettes.

Comme le montre la Figure 17, la méthode ensembliste fournit une information locale intéressante sur les corrélations "climatologiques", qui apparaissent être bien représentées par la formulation ondelettes. L'ensemble comporte ici 6 éléments et est obtenu par

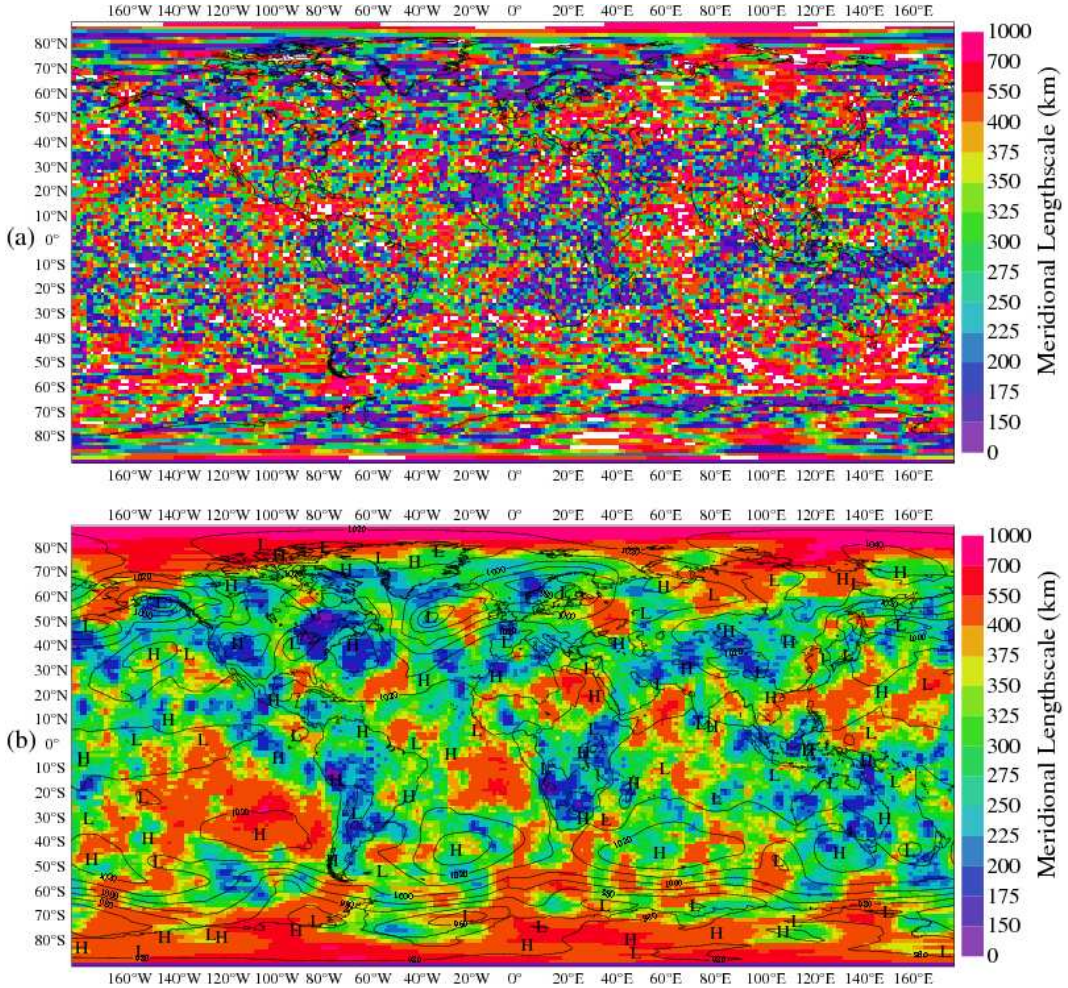

FIG. 18 – Portées méridiennes (en km) pour la pression de surface et pour le 10 février 2002 à 12H, calculées sur un ensemble de 6 membres. (a) : portées brutes. (b) : portées ondelettes superposées au champ d'ébauche de pression mer.

une perturbation des observations dans un cycle d'assimilations 4D-Var ARPEGE. Les longueurs de corrélation apparaissent être plus courtes dans l'hémisphère nord (riche en données) que dans l'hémisphère sud (pauvre en données). Des minima de longueur de corrélation peuvent être identifiés dans la région du rail des dépressions de l'Atlantique nord, et près de la zone de convergence inter-tropicale en Afrique de l'ouest. Les portées sont également plus courtes sur les régions montagneuses telles que l'Himalaya et les Andes, et dans la partie sud de l'Afrique. A l'opposé, quelques maxima de portées sont visibles sur les océans tropicaux.

Ces caractéristiques sur l'ensemble du globe confirment un des avantages majeurs de la formulation ondelettes (comparée par exemple avec une formulation spectrale) : elle autorise la représentation de variations dans les corrélations, telles que celles induites par les processus atmosphériques et par les contrastes en densité d'observations.

Il peut également être remarqué que les variations extrêmes tendent à être lissées dans la carte ondelette. Ceci correspond aux propriétés de filtrage de l'approche diagonale en ondelette. Ces propriétés de filtrage sont encore plus évidentes lorsque la taille de l'ensemble est réduite à $N_e = 6$ pour une date particulière.

Portées locales pour une date particulière

La Figure 18 montre les portées pour une date particulière, le 10 février 2002 à 12H. Les portée brutes (en haut) apparaissent très bruitées, en accord avec la très petite taille de l'ensemble (6 éléments). En contraste, la carte induite par les ondelettes (figure du bas, superposée avec le champ d'ébauche pour la pression mer) est relativement lisse et bien structurée. Les grandes valeurs apparaissent clairement sur les océans tropicaux, comme dans le cas climatologique mais d'une manière plus prononcée. Les portées courtes sont visibles sur terre. D'autres structures relatives à la situation météorologique locale peuvent également être identifiées, telles que les valeurs faibles au voisinage des dépressions des latitudes moyennes au-dessus des océans (voir par exemple les dépressions dans le Pacifique nord et près de la Scandinavie). Ceci est cohérent avec les résultats décrits par Thépaut *et al.* (1995) par exemple.

De telles différences entre les portées brutes et celles induites par les ondelettes, lorsque la taille d'ensemble est faible ($N_e = 6$) confirment l'idée que la formulation ondelette est capable de capturer et de représenter les variations de portée les plus significatives, à partir d'un petit ensemble de prévision, grace au moyennage spatial.

Conclusion

L'utilisation des ondelettes apparaît être une voie prometteuse pour la représentation des covariances des erreurs d'ébauche. Elles autorisent en effet la description de certaines variations caractéristiques des longueurs des corrélations, liées à la situation météorologique ou à la densité des observations. D'autre part, les ondelettes semblent fournir un outil adapté pour filtrer de manière adéquate le bruit d'échantillonnage lié à l'utilisation d'un ensemble de petite taille. Leur implémentation dans le schéma 4D-Var du CEPMMT a apporté une amélioration des scores de prévision, en particulier sur la zone inter-tropicale.

3.9 Perspectives

Bien que l'assimilation de données ait vu un développement très important durant les dernières années, en particulier dans le domaine de la métorologie et de l'océanographie, elle reste une discipline jeune avec beaucoup de questions toujours ouvertes.

Il est en particulier frappant de noter que la plupart des algorithmes ajourd'hui mis en place dans les grand centres météorologiques, comme le 3D-Var ou le 4D-Var, restent intimement liés à la théorie de l'estimation linéaire statistique. Cependant, malgré la simplicité apparente de cette approche, toutes ses propriétés n'ont pas encore été complètement explorées ou comprises. Comme le souligne Talagrand depuis un certain nombre d'années (voir par exemple Talagrand 1999), la théorie de l'estimation linéaire est basée sur un certain nombre de critères et par conséquent une manière de vérifier que les paramètres d'une analyse sont bien spécifiés est de vérifier *a posteriori* que les critères de l'estimation linéaire sont bien vérifiés. Une partie des travaux décrits dans la synthèse présentée ici a en fait été inspirée par cette idée, mais d'autres voies peuvent encore être explorées en suivant cet axe. Ainsi, un critère d'optimalité qui a commencé à être examiné dans le 4D-Var ARPEGE par Chapnik (2006) consiste à vérifier la décorrélation entre des incréments d'analyse successifs.

Une des difficultés de l'assimilation de données est de spécifier des statistiques correctes des erreurs d'observations et d'ébauche. On a montré que les critères d'optimalité liés à l'estimation linéaire pouvaient être utiles pour diagnostiquer et optimiser ces différentes statistiques et cette voie devra également être poursuivie. La méthode proposée dans Desroziers *et al.* (2005) fournit en particulier une approche pour estimer les corrélations éventuelles entre les erreurs d'observation. C'est une possibilité qu'il faudra considérer, dans la mesure où la non prise en compte de ces corrélations (en particulier pour les données satellite) est un point faible des systèmes d'analyse opérationnels comme le 4D-Var ARPEGE. La manière de représenter ces corrélations des erreurs d'observation dans un système de grande taille est par ailleurs examinée au CEPMMT (Fisher 2006).

Les dernières années ont vu en parallèle l'arrivée de méthodes ensemblistes dont un des avantages marquant est de pouvoir spécifier pour un coût raisonnable, c'est-à-dire avec un ensemble de taille limitée, une bonne approximation de la matrice de covariance des erreurs d'ébauche. Ces méthodes ensemblistes peuvent être couplées à un schéma d'assimilation variationnel déjà mis en place. C'est la voie qui est explorée actuellement à Météo-France, où l'on espère pouvoir déduire d'un petit ensemble d'analyses/prévisions les éléments les plus caractéristiques des statistiques des erreurs de prévision. Dans cette optique, on peut en particulier obtenir une description réaliste et dépendante de l'écoulement de la variance des erreurs de prévision. L'utilisation, dans un deuxième temps, des ondelettes pour la description de la matrice de covariance permettrait également d'obtenir des longueurs de corrélation variables et dépendantes de la situation. C'est l'objet du travail de thèse d'O. Pannekoucke mais qui devra être poursuivi pour obtenir cette spécification des covariances "du jour".

A noter également que les diagnostics issus des critères d'optimalité peuvent être utiles pour vérifier la pertinence de l'ensemble mis en place et en particulier que la variance de l'ensemble est correcte. Les diagnostics dans l'espace des observations proposés dans Desroziers *et al.* (2005) peuvent être utilisés dans ce but. Ces diagnostics sont en

effet essentiellement des diagnostics sur le vecteur innovation qui procure une source d'information essentielle sur l'erreur de prévision du modèle, comprenant la part venant de l'erreur d'analyse mais aussi la part liée à l'erreur introduite par le modèle.

Les travaux menés sur l'impact des observations sur les analyses et les prévisions ont montré que cet impact pouvait être obtenu en utilisant une méthode de type Monte Carlo dans laquelle les observations sont perturbées. C'est justement le principe des méthodes ensemblistes mises en place actuellement à Météo-France ou au CEPMMT. Un sous-produit assez immédiat de ce type d'ensemble est donc d'obtenir simplement l'impact des différents types d'observation dans le système d'assimilation. Ceci peut être important pour mesurer le poids respectif des observations conventionnelles par rapport aux observations satellite ou même de mesurer l'impact respectif des différents canaux de ces mesures satellitaires.

La mise en place de ces méthodes ensemblistes a fait apparaître de manière encore plus évidente le lien entre l'assimilation de données et la prévision d'ensemble utilisée pour documenter la prévisibilité des différentes situations météorologiques. Les ensembles d'analyses/prévisions mis en place dans le cadre de l'assimilation de données pour obtenir essentiellement une meilleure description des covariances d'erreur d'ébauche peuvent en effet être vus comme une méthode de prévision d'ensemble avec l'avantage que les perturbations sur l'analyse sont cohérentes avec l'incertitude sur l'erreur d'analyse. Il y a donc là une voie très prometteuse pour coupler les deux problématiques. L'application de ces méthodes ensemblistes à la prévision d'ensemble nécessite cependant de prendre en compte l'erreur de modèle. Dans ce but, une possibilité déjà explorée dans certains centres est de perturber la physique du modèle.

Per ailleurs, la plupart des grands centre météorologiques se sont lancés dans le développement de systèmes de prévision à échelle fine basés sur l'utilisation d'une physique élaborée, permettant en particulier de mieux décrire les phénomènes convectifs. Ces systèmes doivent naturellement comprendre un système d'assimilation propre. Le système à échelle fine AROME développé à Météo-France sera basé, du moins dans un premier temps, sur un schéma 3D-Var. Pour ces systèmes se pose en particulier la question d'une description adaptée des covariances d'erreur d'ébauche. La transposition des méthodes ensemblistes à ces systèmes peut être envisagée, mais la question de la définition de balances spécifiques des erreurs de prévision à cette échelle est également posée. L'application à petite échelle d'une méthode ensembliste basée sur une perturbation des observations pose cependant un problème dans la mesure où l'erreur modèle devient alors vraisemblablement importante dans la description fine des systèmes nuageux. Cette erreur de modèle devrait donc être également simulée dans l'ensemble.

De même, la possibilité et l'intérêt de mettre en place un schéma de type 4D-Var à échelle fine doivent être étudiés. En particulier, face aux fortes non-linéarités attendues à ces échelles il paraît pertinent d'examiner l'intérêt d'une approche de type quasi-continu, dans laquelle la durée de la période d'assimilation est progressivement allongée (Pires *et al.*, 1996). Le développement d'une physique simplifiée et régularisée utile à ces échelles constitue par aileurs un enjeu majeur. On peut également s'attendre à ce que l'hypothèse de modèle parfait, généralement faite dans les schémas 4D-Var à grande échelle, devienne plus criticable à petite échelle. La prise en compte de l'erreur modèle dans une formulation variationnelle devient alors nécessaire avec deux voies possibles : une avec contrôle de l'erreur de modèle dans l'espace du modèle (voir par exemple Trémolet 2005) et une autre

passant par l'utilisation d'une formulation duale du problème d'analyse, dans l'espace des observations (Amodéi 1995, Courtier 1997).

Références

[1] L. Amodéi. Solution approchée pour un problème d'assimilation de données météorologiques avec prise en compte de l'erreur de modèle. *Comptes Rendus de l'Académie des Sciences*, 321, série IIa :1087–1094, 1995.

[2] E. Andersson. Modelling of innovation statistics. pages 153–164, 2003. Proceedings of the ECMWF Workshop on Recent Developments in Data Assimilation for Atmosphere and Ocean, Reading, September 2003.

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

[4] S.G. Benjamin, K.A. Brewster, R. Brummer, B.F. Jewett, and T.W. Schlatter. An isentropic three-hourly data assimilation system using ACARS aircraft observations. *Mon. Wea. Rev.*, 119 :888–906, 1991.

[5] A.F. Bennett. *Inverse methods in physical oceanography.* Monographs on Mechanics and Applied Mathematics. Cambridge University Press, Cambridge, 1992. 346 pp.

[6] A.F. Bennett, L.M. Leslie, C.R. Hagelberg, and P.E. Powers. Tropical cyclone prediction using a barotropic model initialized by a generalized inverse method. *Mon. Wea. Rev.*, 121 :1714–1729, 1993.

[7] F. Bouttier. Sur la prévision de la qualité des prévisions météorologiques, 1994. Thèse de doctorat de l'Université Paul Sabatier à Toulouse , 240 pp.

[8] F. Bouttier, J. Derber, and M. Fisher. The 1997 revision of the J_b term in 3D/4D-Var. Technical Report 192, ECMWF Reading, 1997. 25 pp.

[9] C. Cardinali, S. Pezzuli, and E. Andersson. Influence matrix diagnostic of a data assimilation system. *Q. J. R. Meteor. Soc.*, 130 :2767–2786, 2004.

[10] B. Chapnik. Réglage des statistiques d'erreur en assimilation variationnelle, 2005. Thèse de doctorat de l'Université Toulouse III, 196 pp.

[11] B. Chapnik. Assessment of the optimality of an operational data assimilation system, 2006. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

[12] B. Chapnik, G. Desroziers, F. Rabier, and O. Talagrand. Properties and first application of an error statistics tuning method in variational assimilation. *Q. J. R. Meteor. Soc.*, 130 :2253–2275, 2004.

[13] B. Chapnik, G. Desroziers, F. Rabier, and O. Talagrand. Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. *Q. J. R. Meteor. Soc.*, 132 :543–565, 2006.

[14] S.E. Cohn. An introduction to estimation theory. *J. Met. Soc. Japan*, 75 :257–288, 1997.

[15] P. Courtier. Dual formulation of four-dimensional variational assimilation. *Q. J. R. Meteor. Soc.*, 123 :2449–2461, 1997.

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

[17] P. Courtier, C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas. The Arpege project at Météo-France. In *Proceedings of the ECMWF Workshop on Numerical methods in atmospheric models*, pages 193–231, Reading, 1991. 9-13 September.

[18] P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation. II : Numerical results. *Q. J. R. Meteor. Soc.*, 113 :1329–1347, 1987.

[19] P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteor. Soc.*, 120 :1367–1387, 1994.

[20] R. Daley. *Atmospheric data analysis.* Cambridge University Press, Cambridge, 1991. 460 pp.

[21] A. Deckmyn. A wavelet approach to representing background error covariances in a lam. *Mon. Wea. Rev.*, 133 :1279–1294, 2005.

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

[23] D. P. Dee and A. da Silva. Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I : Methodology. *Mon. Wea. Rev.*, 127 :1822–1834, 1999.

[24] G. DerMégréditchian. *Le traitement statistique des données multidimensionnelles. Tome II.* Ecole Nationale de la Météorologie, Toulouse, 1993. 260 pp.

[25] G. Desroziers. A coordinate change for data assimilation in spherical geometry of frontal structures. *Mon. Wea. Rev.*, 125 :3030–3038, 1997.

[26] G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Q. J. R. Meteor. Soc.*, 131 :3385–3396, 2005.

[27] G. Desroziers, P. Brousseau, and B. Chapnik. Use of randomization to diagnose the impact of observations on analyses and forecasts. *Q. J. R. Meteor. Soc.*, 131 :2821–2837, 2005.

[28] G. Desroziers, G. Hello, and J.-N. Thépaut. A 4D-Var reanalysis of FASTEX. *Q. J. R. Meteor. Soc.*, 129 :1301–1315, 2003.

[29] G. Desroziers and S. Ivanov. Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteor. Soc.*, 127 :1433–1452, 2001.

[30] G. Desroziers and J. P. Lafore. A coordinate transformation for objective frontal analysis. *Mon. Wea. Rev.*, 121 :1531–1553, 1993.

[31] G. Desroziers, B. Pouponneau, J.-N. Thépaut, M. Janisková, and F. Veersé. 4D-Var analyses of FASTEX situations using special observations. *Q. J. R. Meteor. Soc.*, 125 :3393–3414, 1999.

[32] C. Fischer, A. Joly, and F. Lalaurette. Error growth and kalman filtering within an idealized baroclinic flow. *Tellus*, 50A :596–615, 1998.

[33] M. Fisher. Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. Technical Report 397, ECMWF, 2003.

[34] M. Fisher. Accounting for correlated observation error in a variational analysis system, 2006. 7th Workshop on Adjoint Applications in Dynamic Meteorology, 8-13 October, Obergurgl, Austria.

[35] M. Fisher and P. Courtier. Estimating the covariance matrices of analysis and forecast error in variational data assimilation. Technical Report 220, ECMWF, 1995.

[36] L. S. Gandin. *Objective analysis of meteorological fields.* Gidrometeorologicheskoe Izdatel'stvo, Leningrad, 1963. Translated from Russian, Israel Program for Scientific Translations, Jerusalem, 1965, 242 pp.

[37] G. Gaspari and S.E. Cohn. Construction of correlation functions in two and three dimensions. *Mon. Wea. Rev.*, 125 :723–757, 1999.

[38] C.F. Gauss. Theoria motus corporum coelestium. 1809. Hamburg.

[39] P. Gauthier and J.-N. Thépaut. Impact of the digital filter as a weak constraint in the preoperational 4dvar assimilation system of météo-france. *Mon. Wea. Rev.*, 129(8) :2089–2102, 2001.

[40] A. Gelb. *Applied Optimal Estimation.* M.I.T Press, Cambridge, MA, 1974. 374 pp.

[41] J.-F. Geleyn, D. Banciu, M. Bellus, R. El Khatib, P. Moll, P. Saez, and J.-N. Thépaut. The operational 4D-Var data assimilation system of Météo-France : Characteristics and behaviour in the special case of the 99 xmas storms over france. pages 9–12, Fort Lauderdale, USA, 2001. Proceedings of 14th AMS conference on numerical weather prediction, 30 July-2 August.

[42] D. Girard. A fast Monte Carlo cross-validation procedure for large least squares problems with noisy data. Technical Report 687-M, IMAG, Grenoble, France, 1987.

[43] A. Hollingsworth. Objective analysis for numerical weather prediction. pages 11–59, 1987. In Short- and Medium-Range Numerical Weather Prediction, collected papers presented at WMO/IUGG NWP Symposium, Tokyo, 4-8 August, 1988 ed by T. Matsuno, Special Volume of the *J. Meteor. Soc. Japan.*

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

[45] B.J. Hoskins. The geostrophic momemtum approximation and the semi-geostrophic equations. *J. Atmos. Sci.*, 32 :233–242, 1975.

[46] B.J. Hoskins and F.P. Bretherton. Atmospheric frontogenesis models : mathematical formulation and solution. *J. Atmos. Sci.*, 29 :11–27, 1972.

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

[48] M. Janisková, J.-N. Thépaut, and J.-F. Geleyn. Simplified and regular physical parametrizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.*, 127(1) :26–45, 1999.

[49] G. Jaubert, C. Piriou, S.M. Loehrer, A. Petitpa, and J.A. Moore. Development and quality control of the FASTEX Data Archive. *Q. J. R. Meteor. Soc.*, 125 :3165–3188, 1999.

[50] A.H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, 1970. 376 pp.

[51] A. Joly, K.A. Browning, P. Bessemoulin, J.P. Cammas, M.A. Shapiro, G. Caniaux, J.P. Chalon S.A. Clough, R. Dirks, K.A. Emmanuel, L. Eymard, F. Lalaurette, R. Gall, T.D. Hewson, P.H. Hildebrand D. Jorgensen, R.H. Langlang, Y. Lemaitre, P. Mascart, J.A. Moore, P.O.G. Persson, F. Roux, M.A. Shapiro, C. Snyder, Z. Toth, and R.M. Wakimoto. Overview of the field phase of the Fronts and Atlantic Storm-Track Experiment (FASTEX) project. *Q. J. R. Meteor. Soc.*, 125 :3131–3163, 1999.

[52] F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations : Theoretical aspects. *Tellus*, 38A :97–110, 1986.

[53] J. M. Lewis and J. C. Derber. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, 37A :309–322, 1985.

[54] Z. Liu. Influence de la résolution d'observations sur l'assimilation de données, 2002. Thèse de doctorat de l'Université Paul Sabatier à Toulouse, 106 pp.

[55] A. C. Lorenc. Analysis methods for numerical weather prediction. *Q. J. R. Meteor. Soc.*, 112 :1177–1194, 1986.

[56] O. Pannekoucke, Loik Berre, and G. Desroziers. Wavelets expansion and its use as an adaptive filter. *Q. J. R. Meteor. Soc.*, 2006. Under revision.

[57] M. Belo Pereira and L. Berre. The use of an ensemble approach to study the background error covariances in a global NWP model. *Mon. Wea. Rev.*, 2005. To appear.

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

[59] F. Rabier, N. Fourrié, D. Chafaï, and P. Prunet. Channel selection method for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteor. Soc.*, 128 :1011–1032, 2002.

[60] F. Rabier, H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons. The ECMWF operational implementation of four-dimensional variational assimilation. I : Experimental results with simplified physics. *Q. J. R. Meteor. Soc.*, 126 :1143–1170, 2000.

[61] C.D. Rodgers. *Inverse Methods for Atmospheric Sounding Theory and Practice*. World Scientific Publishing, London, 2000. 256 pp.

[62] W. Sadiki and C. Fischer. A posteriori validation of a real data assimilation system. *Tellus*, 57A :21–34, 2005.

[63] O. Talagrand. Assimilation of observations, an introduction. *J. Met. Soc. Japan*, 75 :191–209, 1997.

[64] O. Talagrand. A posteriori verification of analysis and assimilation algorithms. In *Proceedings of the ECMWF Workshop on Diagnosis of Data Assimilation Systems*, pages 17–28, Reading, 1999. 2-4 November 1998.

[65] O. Talagrand. Objective validation and evaluation of data assimilation. In *Proceedings of the ECMWF Workshop on Diagnosis of Data Assimilation Systems*, pages 287–299, Reading, 2004. September 2004.

[66] O. Talagrand and P. Courtier. Variational assimilation of meteorological observations with the adjoint vorticity equation. I : Theory. *Q. J. R. Meteor. Soc.*, 113 :1311–1328, 1987.

[67] A. Tarantola. *Inverse Problem Theory : Methods for Data Filtering and Model Parameter Estimation*. Elvesier, 1987. 613 pp.

[68] J.-N. Thépaut, P. Courtier, G. Belaud, and G. Lemaître. Dynamical structure functions in 4D-Var : a case study. In *Proceedings of the Second International Symposium on Assimilation of Observations in Meteorology and Oceanography*, pages 129–134, Tokyo, 1995. 13-17 March.

[69] A.J. Thorpe and S.A. Clough. Mesoscale dynamics of cold fronts : structures described by dropsoundings in FRONTS 87. *Q. J. R. Meteor. Soc.*, 117 :903–941, 1991.

[70] Y. Trémolet. Accounting for an imperfect model in 4D-Var. Technical Report 477, ECMWF, Reading, 2005.

[71] F. Veersé and J.-N. Thépaut. Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q. J. R. Meteor. Soc.*, 124 :1889–1908, 1998.

[72] G. Wahba, D.R. Johnson, F. Gao, and J. Gong. Adaptative tuning of numerical weather prediction models : randomized GCV in three- and four-dimensional data assimilation. *Mon. Wea. Rev.*, 123 :3358–3369, 1995.

[73] G. Wahba and J. Wendelberger. Some new mathematical methods for variational objective analysis using splines and cross validation. *Mon. Wea. Rev.*, 108 :1122–1143, 1980.

4 Publications essentielles

Desroziers, G., 1997 : A coordinate change for data assimilation in spherical geometry of frontal structures, Mon. Wea. Rev., 125, 3030-3038.

Desroziers, G. and Ivanov, S., 2001 : Diagnosis and adaptive tuning of information error parameters in a variational assimilation, Q. J. R. Meteorol. Soc., 127, 1433-1452.

Desroziers, G., Hello, G., and Thépaut, J.-N., 2003 : A 4D-Var re-analysis of FASTEX, Q. J. R. Meteorol. Soc., 129, 1301-1315.

Desroziers, G., Brousseau, P., and Chapnik, B., 2005 : Use of randomization to diagnose the impact of observations on analyses and forecasts, Q. J. R. Meteorol. Soc., 131, 2821-2837.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P., 2005 : Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. R. Meteorol. Soc., 131, 3385-3396.

A Coordinate Change for Data Assimilation in Spherical Geometry of Frontal Structures

GÉRALD DESROZIERS

Météo-France, CNRM, Toulouse, France

1 October 1996 and 17 March 1997

ABSTRACT

In this note, it is discussed how a coordinate transformation based on semigeostrophic theory can be introduced in a global data assimilation scheme in order to improve the analysis of frontal structures. An expression of the coordinate change on the sphere is proposed and an iterative procedure is defined to perform this transformation. This algorithm also includes an iterative filtering of the wind field used for the coordinate change. This filtering is equivalent to a filtering of the wind in the transformed space that nearly keeps unchanged the corresponding high cyclonic vorticity values in physical space associated with cold fronts.

It is shown that the transformation can be built for a real meteorological case. The main impact of this transformation in the framework of data assimilation is that it enables the use of anisotropic forecast correlations that are flow dependent. As expected, the correlations in the vicinity of a front are found to be contracted in the horizontal but also vertically tilted.

1. Introduction

Assimilation techniques used in meteorology try to combine in an optimal way observations and forecasts. Both statistical (Gandin 1965) or variational (Lewis and Derber 1985; Le Dimet and Talagrand 1986) methods rely on a description of the statistics of forecast errors of the assimilating model. Because little is known about these statistics, but also in order to make the practical implementation of the analysis scheme feasible, some crude simplifications of the form of these forecast errors are usually done. In particular, the homogeneity and the isotropy of the forecast error correlation are most often assumed.

On the other hand, Thépaut et al. (1993) have shown that a 4D-Var algorithm implicitly uses flow-dependent structure functions although it does not make them explicitly evolve. On a real meteorological situation, they exhibited such inhomogeneous error covariances with a vertical tilt typical of a baroclinic structure. However, Thépaut et al. (1996) have also shown that this implicit deformation of the structure functions is dependent on the length of the assimilation period: this deformation is less important when the period is short. Another limitation is due to the fact that frontal zones are not precisely described with the resolution used by assimilating models.

Semigeostrophic theory (Hoskins and Bretherton 1972) has been found to possess many remarkable properties. One of them is that it allows one to get rid of parts of the nongeostrophic nonlinearities in the evolution of the atmospheric fluid and then to avoid shocks.

The idea to use the geostrophic coordinate for data analysis has been suggested and shown useful by Desroziers and Lafore (1993) for the 2D vertical interpolation of high density observations across simulated or observed fronts. Another way of achieving flow-dependent forecast error correlations is to use isentropic transformation (Benjamin et al. 1991).

Here, we show how a coordinate change inspired by the semigeostrophic theory could be implemented in a 3D assimilation scheme associated with an operational and global forecasting system. In particular, this implementation requires a proper expression of the coordinate change on the sphere. The aim of this note is to propose a formulation for this coordinate change and to describe its application to a real meteorological case.

In section 2, we describe the main ideas about the implementation of the coordinate transformation in a variational assimilation scheme. The proposed formulation of this coordinate change on the sphere is presented in section 3. The way it is iteratively built is discussed in section 4. Finally, an application of this transformation to a real case and its implied anisotropic structure functions are shown in section 5. Conclusions and perspectives are given in section 6.

Corresponding author address: Gérald Desroziers, Météo-France, CNRM/GMME, 42 Av. G. Coriolis, 31057 Toulouse Cedex, France.

© 1997 American Meteorological Society

2. Field transformation in a variational assimilation scheme

A 3D-Var assimilation scheme can be mathematically expressed as find the model description $\mathbf{c}$, which minimizes

$$J(\mathbf{c}) = J^b(\mathbf{c}) + J^o(\mathbf{c}) = \frac{1}{2}(\mathbf{c} - \mathbf{c}^b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{c} - \mathbf{c}^b) + \frac{1}{2}[\mathbf{H}(\mathbf{c}) - \mathbf{y}]^{\mathrm{T}} \times \mathbf{R}^{-1}[\mathbf{H}(\mathbf{c}) - \mathbf{y}]. \quad (1)$$

For a spectral model, $\mathbf{c}$ and $\mathbf{c}^b$ are the vectors of spectral coefficients of, respectively, the optimal model state and of the background information. Here, $\mathbf{y}$ stands for the vector of observations and $\mathbf{H}$ for the operator that provides the model forecast value, from the optimal state $\mathbf{c}$, at observation location. Here, $\mathbf{R}$ is the covariance matrix of observation errors and T denotes the transpose.

Here, $\mathbf{B}$ is the covariance matrix of background errors, which contains a huge amount of information. One simplification to this problem is given by the above-mentioned assumption of homogeneity of the horizontal correlations.

We propose to modify the background term J^b so that the simplification made on the form of the covariance matrix $\mathbf{B}$ could be better justified.

For the analysis of frontal structures, we suggest that it can be achieved by a field transformation based on the semigeostrophic theory. The three-dimensional formulation proposed by Hoskins (1975) is in the horizontal (the vertical coordinate is kept unchanged):

$$\left.\begin{aligned} x_G &= x_R + \frac{1}{f}v_g \\ y_G &= y_R - \frac{1}{f}u_g \end{aligned}\right\} \quad (2)$$

where f is the Coriolis parameter, u_g and v_g the components of the geostrophic wind, and (x_R, y_R) and (x_G, y_G) the respective real and geostrophic coordinates of a point. We will propose in the next section a corresponding transformation on the sphere. Then, for a description $\mathbf{z}^r$ of a field in real physical space, its representation $\mathbf{z}^g$ in geostrophic physical space is given by the evaluation $z^g(G) = z^r(R)$ at any point $G(\lambda_G, \theta_G)$ of the discretization grid in this new space, where $R(\lambda_R, \theta_R)$ is the image of G in real space (λ and θ stand, respectively, for the longitude and the latitude of a point on the sphere). Once the coordinates (λ_R, θ_R) are found, $\mathbf{z}^g$ can be computed by a simple interpolation $\mathcal{T}$, in physical space, of $\mathbf{z}^r$: $\mathbf{z}^g = \mathcal{T}\mathbf{z}^r$.

Assuming that this transformation can be built (this point is discussed in the following sections), we postulate that, since the forecast error correlations are likely to be more homogeneous in geostrophic space, then the background term, introduced in expression (1), should rather be written

$$J^b_{\mathcal{T}}(\mathbf{c}) = \frac{1}{2}[\mathcal{L}_{\mathcal{T}}(\mathbf{c} - \mathbf{c}^b)]^{\mathrm{T}}\mathbf{B}^{-1}[\mathcal{L}_{\mathcal{T}}(\mathbf{c} - \mathbf{c}^b)],$$

with $\mathcal{L}_{\mathcal{T}} = \mathcal{S}\mathcal{T}\mathcal{S}^{-1}$, and $\mathcal{S}$ and $\mathcal{S}^{-1}$ standing, respectively, for the direct and inverse spectral transform.

Note that the operator $\mathcal{T}$ should be dependent on the analyzed model state $\mathbf{c}$. However, some simplification to this nonlinear dependence can be found in the incremental approach proposed by Courtier et al. (1994). With this formulation, the assimilation procedure is based on a pair of nested loops: at each iteration of the outer loop, the model trajectory is redefined with the complete model, while the inner loop uses a simpler model for the analysis of the increments with respect to the redefined trajectory. Therefore, with this approach, it is possible to introduce a redefinition of the operator $\mathcal{T}$ at each iteration of the outer loop (such a procedure is still valid in 3D-Var with no model integration).

The modification induced by the operator $\mathcal{L}_{\mathcal{T}}$ on the form of the solution can be explicited.

We note $\mathbf{z} - \mathbf{z}^b = \mathcal{S}^{-1}(\mathbf{c} - \mathbf{c}^b)$ the vector of increments in gridpoint space. Then the solution of (1) can be written

$$\mathbf{z} - \mathbf{z}^b = \tilde{\mathbf{B}}\tilde{\mathbf{H}}^{\mathrm{T}}(\tilde{\mathbf{H}}\tilde{\mathbf{B}}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y} - \tilde{\mathbf{H}}\mathbf{z}^b), \quad (3)$$

where $\tilde{\mathbf{B}} = \mathcal{S}^{-1}\mathbf{B}\mathcal{S}$ is the homogeneous forecast error correlation matrix in gridpoint space and $\tilde{\mathbf{H}} = \mathbf{H}'\mathcal{S}$ with $\mathbf{H}'$ the linearization of $\mathbf{H}$ in the vicinity of the background $\mathbf{c}^b$.

If we include the transformation in the analysis scheme, and assuming here for the sake of simplicity that the transformation $\mathcal{T}$ is kept independent of c, then the expression (3) becomes alternatively

$$\mathbf{z} - \mathbf{z}^b = \tilde{\mathbf{B}}_{\mathcal{T}}\tilde{\mathbf{H}}^{\mathrm{T}}(\tilde{\mathbf{H}}\tilde{\mathbf{B}}_{\mathcal{T}}\tilde{\mathbf{H}}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y} - \tilde{\mathbf{H}}\mathbf{z}^b), \quad (4)$$

with $\tilde{\mathbf{B}}_{\mathcal{T}} = \mathcal{T}\tilde{\mathbf{B}}\mathcal{T}^{\mathrm{T}}$.

This expression shows that the analysis with the coordinate change implicitly makes use of flow-dependent anisotropic error correlations built from the deformation by $\mathcal{T}$, in physical space, of the isotropic error correlations specified in $\tilde{\mathbf{B}}$; this point will be documented, in a real meteorological case, in section 5.

3. A formulation of the transformation on the sphere

Semigeostrophic theory on the plane (Hoskins 1975) has been widely applied to the study of baroclinic waves and fronts. More recently, some authors have explored the possibility to extend this theory on a β-plane or on the sphere (Magnusdottir and Schubert 1991) in order to define a more general geostrophic balance. Here, we show how a coordinate transformation on the hemisphere, defined to be used for data assimilation prob-

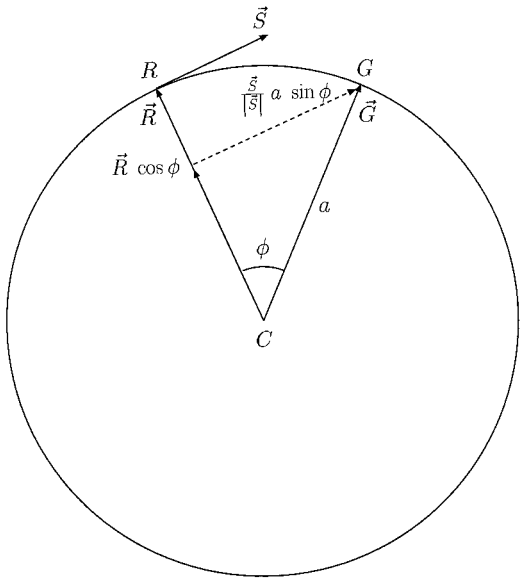

FIG. 1. Geometrical view of the $\mathbf{G} = \mathbf{R} + f^{-1}\mathbf{S}$ transformation on the sphere. Here, C and a are, respectively, the center and the radius of the earth. The positional shift, along the great circle passing by R and its image G, is equal to $a\phi$, with $\phi = f^{-1}|\mathbf{S}|$. From this figure, it is easy to check that the position of G is deduced from the position of R by the vectorial relation $\mathbf{G} = \mathbf{R} \cos\phi + (\mathbf{S}/|\mathbf{S}|)a \sin\phi$.

lems, can be retrieved by a simple extension of the relation (2) introduced in the previous section.

If we note $\mathbf{G}(\lambda_G, \theta_G)$ and $\mathbf{R}(\lambda_R, \theta_R)$ the vectors defining the respective geostrophic position G and real space position R of a same point, these relations can be written in a vector form $\mathbf{G} = \mathbf{R} + f^{-1}\mathbf{S}$, where $f^{-1}\mathbf{S}$ is the positional shift vector with components $f^{-1}(v_g, -u_g)$.

To transpose this relation to the sphere, we simply assume that the positional shift induced by this relation occurs along great circles defined by the vector $f^{-1}\mathbf{S}$.

With this assumption and if we now define the vectors $\mathbf{G}$ and $\mathbf{R}$ as the vectors passing by the center of the earth and the respective positions G and R of a same point in geostrophic and physical spaces, then it can be checked that the previous expression becomes on the sphere (see Fig. 1 for a geometrical interpretation)

$$\mathbf{G} = \mathbf{R} \cos\phi + \frac{\mathbf{S}}{|\mathbf{S}|} a \sin\phi, \tag{5}$$

where a is the earth radius and ϕ is the shift central angle defined by $\phi = (fa)^{-1}|\mathbf{S}|$.

Note that the problem of finding the image G in geostrophic space of R in real space is similar to the problem of the computation of the position of the origin point corresponding to a model grid point in a semi-Lagrangian scheme. Parts of the following developments are therefore inspired by the adaptation of such a scheme to the spherical geometry (Yessad 1995).

Conversely, given a point $G(\lambda_G, \theta_G)$ in geotrophic space, its image $R(\lambda_R, \theta_R)$ in real space can be determined, using the same shift vector $f^{-1}\mathbf{S}$ computed in R by the expression

$$\mathbf{R} = \mathbf{G} \cos\phi + \frac{\mathbf{S}^{RG}}{|\mathbf{S}^{RG}|} a \sin\phi, \tag{6}$$

where $\mathbf{S}^{RG}(U^{RG}, V^{RG}) = \mathcal{R}^{RG}\mathbf{S}$ is the vector $\mathbf{S}(U, V)$ moved from R to G by corresponding rotation $\mathcal{R}^{RG}$ (Yessad 1995).

Expressing the vectorial relation (6) in the earth reference frame, we obtain

$$\left.\begin{aligned} \sin\theta_R &= \sin\theta_G \cos\phi - \frac{V^{RG}}{|\mathbf{S}^{RG}|} \cos\theta_G \sin\phi \\ \cos\theta_R \cos(\lambda_R - \lambda_G) &= \cos\theta_G \cos\phi + \frac{V^{RG}}{|\mathbf{S}^{RG}|} \sin\theta_G \sin\phi \\ \cos\theta_R \sin(\lambda_R - \lambda_G) &= \frac{-U^{RG}}{|\mathbf{S}^{RG}|} \sin\phi. \end{aligned}\right\} \tag{7}$$

The factor f in the denominator of the shift central angle ϕ makes a difficulty in applying these relations in the equatorial regions. Since the transformation is defined to be applicable mainly in midlatitudes areas, we introduce in section 5 an empirical relaxation of the factor f^{-1} when reaching the equator.

4. Iterative building and filtering of the transformation

The relations (7), defining the coordinate change, are nonlinear in (λ_R, μ_R), with $\mu = \sin\theta$, since the components (U^{RG}, V^{RG}) of the shift vector moved from R to G, as defined in the previous section, depend on the components (u_g, v_g) of the geostrophic wind at R. In the appendix, we propose an iterative procedure to solve these relations.

Because some small-scale phenomena cannot be taken into account by the semigeostrophic formalism, the given representation of the wind components cannot be abruptly used. The consequence is that they require filtering, but not in real space before the transformation, because this would cause the strong cyclonic vorticity values associated with sharp cold fronts to vanish where the transformation is expected to be the most efficient.

Thus, we propose to iteratively filter the wind components, along the iterations used to solve (7) (see the appendix), with the following modified expression of the basic coordinate transformation (6):

$$\mathbf{R} = \mathbf{G} \cos\phi - \frac{\mathcal{A}(S^{RG})}{|\mathcal{A}(S^{RG})|} a \sin\phi,$$

where $\mathcal{A}$ stands for a filtering operator of the components of the positional shift vector moved from R to G. This filtering operator is more precisely described in the next section.

Note that at the end of the iterations, S^{RG} defines a

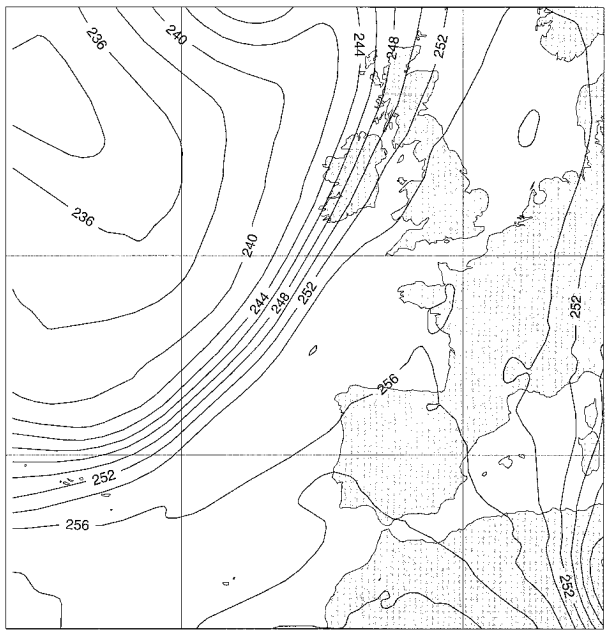

FIG. 2. Six-hour ARPEGE forecast of the temperature field at 500 hPa, valid for 1200 UTC 9 January 1988. The contour interval is 2°C.

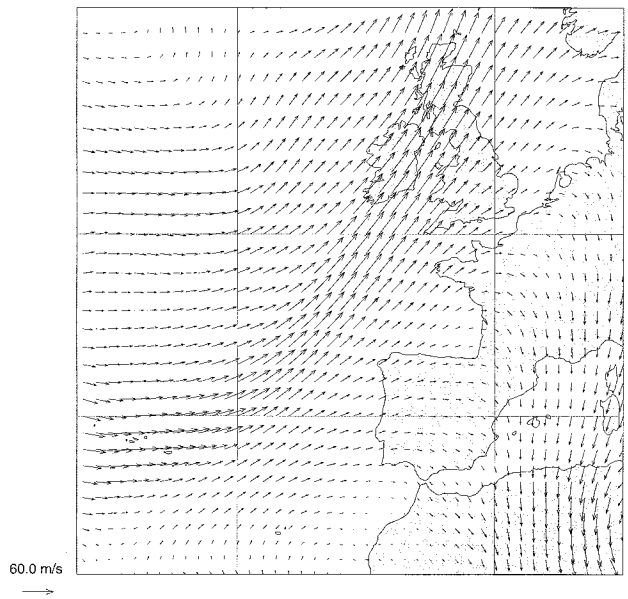

FIG. 3. Six-hour ARPEGE forecast of the wind field at level 10 of the model with a 500-hPa mean pressure over the sea.

representation of the resulting filtered wind field in transformed space.

5. Application to a real meteorological situation

a. Presentation of the meteorological situation and construction of the coordinate change

In this section, we present an application of the transformation to a simulation of a real meteorological case by the ARPEGE system developed by Météo-France for its short range forecasts.

The ARPEGE system (Courtier et al. 1991) includes a global spectral model with pressure-based hybrid coordinates η following the orography in the vertical. In the horizontal the prognostic variables and the orography are represented by truncated series of spherical harmonics. On the other hand, a conformal transform of the η surfaces is introduced, based on the Schmidt transformation (1977).

The situation presented here is the IOP 7 of the European experiment FRONTS 87 organized to study mesoscale cold-front dynamics. This situation has been thoroughly described by Thorpe and Clough (1991) and was essentially characterized by a two-dimensional front with a sharp temperature gradient at level 500 hPa (Fig. 2). The wind field at the closest η level shows a strong jet reaching 50 m s^{-1} over Ireland, with a strong cyclonic gradient on its westerly side (Figs. 3 and 4). Both temperature and cyclonic wind gradients are almost collocated. On the contrary, the warm-air region is characterized by the absence of strong variations in the temperature field and negative values in the vorticity field.

In the perspective of an application in a data assimilation problem and following the approach described in section 2, the transformation is built here with a wind field issued from a 6-h forecast. The initial

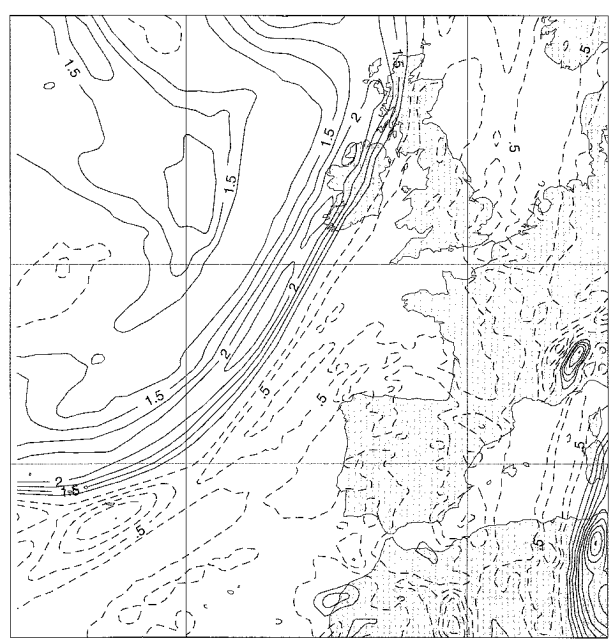

FIG. 4. Absolute vorticity field, normalized by $f/\alpha(\theta)$, corresponding to the initial ARPEGE wind field shown in Fig. 3. Contour interval is 0.25. Values more than one are drawn with solid lines and values less than one with dashed lines.

"geostrophic" wind components are approximated by the components of the rotational part of the real wind.

On the other hand, the transformation is gradually relaxed when reaching the equatorial regions by defining a weighted value $\alpha(\theta)f^{-1}$ of f^{-1}, with $\alpha(\theta)$ an empirical function such as $\alpha(\theta)f^{-1} \simeq f^{-1}$ in the extratropical regions and $\alpha(\theta)f^{-1} \simeq 0$ near the equator [here, $\alpha(\theta)$ is defined by $\alpha(\theta) = 1 - e^{-\theta^2/2\theta_0^2}$, with $\theta_0 = 15°$].

The transformation is directly applied on the η surfaces following orography. The construction of the coordinate change on each one of these surfaces requires the definition of the complete set of coordinates (λ_R, μ_R) corresponding to the coordinates (λ_G, μ_G) of the Gauss grid points of the fictitious transformed space.

These coordinates (λ_R, μ_R) are obtained by the iterative procedure presented in section 4 and in the appendix.

The definition of the filtering operator $\mathcal{A}$ introduced in section 4 is done as follows. At each iteration of the procedure, the streamfunction and velocity potential, associated with the new wind field in transformed space, are deduced; a spectral filtering of the streamfunction is then applied (with an equivalent cutoff wavenumber $n_f = 64$ on the real geographical sphere in the ARPEGE high resolution zone); before the inverse spectral transform to the corresponding wind components, the spectral coefficients of the velocity potential are also reset to zero.

About five iterations are needed to obtain a satisfactory convergence in most regions but some more iterations are required to get this convergence in strong vorticity areas.

As described above, this procedure iteratively determines a wind field in transformed space that should be filtered from the scales that are not relevant with the semigeostrophic formalism. From the representation of this wind field in transformed space, we can derive the corresponding representation in physical space by performing the inverse transform. The vorticity of this retrieved wind field (Fig. 5) can be compared to the vorticity of the initial wind representation given by the model (Fig. 4). It clearly appears that some small-scale structures of this field, due in part to mountainous effects, are removed in the filtered wind representation. However, one can notice that the values in the bands of strong cyclonic vorticity are nearly kept unchanged by this filtering, which is performed in transformed space.

The value of the absolute vorticity, normalized by the local value of the Coriolis factor, is close to the Jacobian of the transformation and indicates the regions where the coordinate change will have a stretching effect: this is the case in the cyclonic areas where the Jacobian is greater than one (solid lines in Fig. 5). On the contrary, values of the Jacobian less than one (dashed lines in Fig. 5) correspond to a contraction effect of the transformation: this effect particularly appears in the warm air region with low field gradients. Note that some values of this Jacobian far less than one in the initial wind field on the south and east part of the jet, which would result in too strong contractions, have also been successfully removed by the iterative procedure.

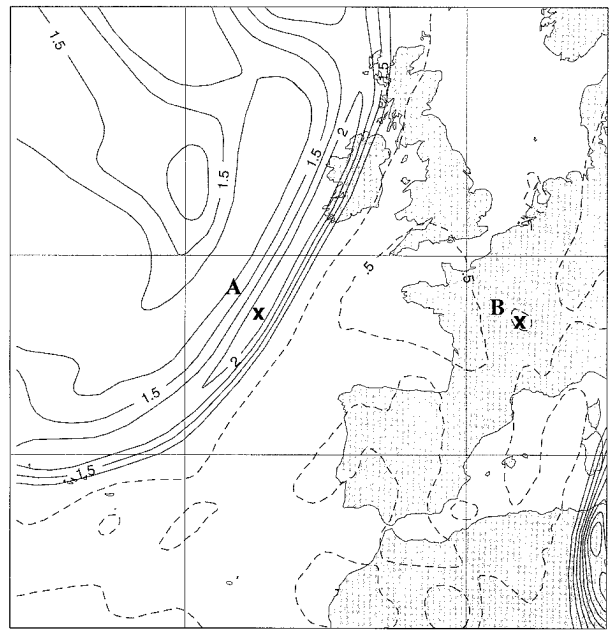

FIG. 5. As in Fig. 4 but for the filtered wind field used to build the coordinate change and issued from the iterative procedure.

b. *Implicit anisotropic correlation functions produced by the transformation*

As the way to integrate the transformation in a data assimilation scheme has been discussed in section 2, this implementation has not been yet performed. However, the discussion of the interest of such an implementation can be started from the observation of the deformation that this transformation would implicitly introduce in the forecast error structure functions used for the analysis.

For the sake of simplicity we assume that we have at our disposal a continuous model for the correlation function. Furthermore, if we postulate that this correlation ρ^g is homogeneous in transformed space, then, according to the relation (4) found in section 2, the form of the corresponding function ρ^r in physical space at a given point $R_0(\lambda_{R_0}, \mu_{R_0})$ can be simply built by writing for any point $R(\lambda_R, \mu_R)$ surrounding R_0

$$\rho^r[(\lambda_{R_0}, \mu_{R_0}); (\lambda_R, \mu_R)] = \rho^g[(\lambda_{G_0}, \mu_{G_0}); (\lambda_G, \mu_G)], \quad (8)$$

where $(\lambda_{G_0}, \mu_{G_0})$ and (λ_G, μ_G) are the transformed coordinates corresponding, respectively, to the real space coordinates $(\lambda_{R_0}, \mu_{R_0})$ and (λ_R, μ_R). As pointed out in section 2, these coordinates can be built explicitly by applying the expression (5), once the representation in

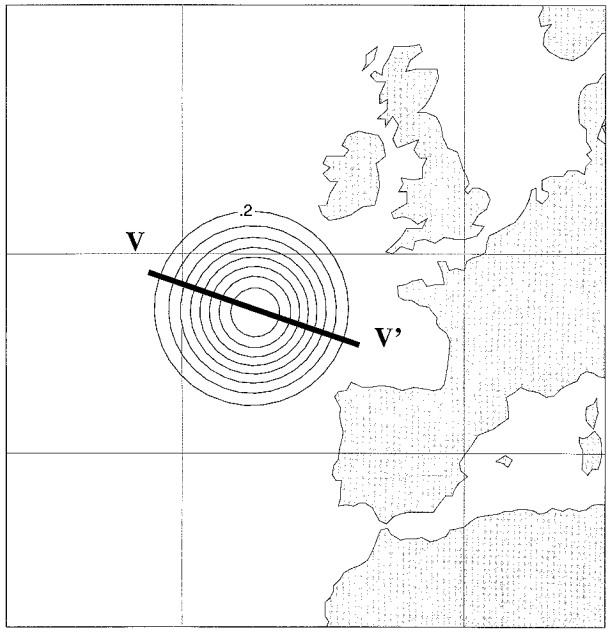

FIG. 6. Horizontal representation of an isotropic temperature error correlation function, based at point *A* defined in Fig. 5. The contour interval is 0.1.

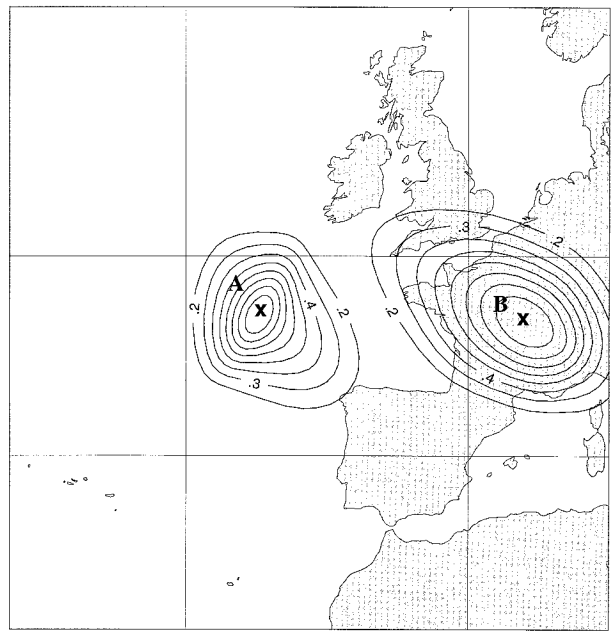

FIG. 7. As in Fig. 6 but for the implicit anisotropic correlation function yielded by the transformation at points *A* and *B* defined in Fig. 5.

physical space of the filtered wind field, determined by the iterative procedure, is known.

As an example, we can study the corresponding deformation of a correlation function for the temperature errors. This correlation function is herein defined by a 300-km horizontal length scale (Fig. 6) and an about 50-hPa vertical length scale and also a slight anticorrelation (Fig. 8) at long-range distance.

For a point located in the cold front, the implicit structure function induced by the transformation and defined by (8) shows a typical anisotropic behavior with a contraction across the discontinuity and an anisotropy ratio locally reaching two (Fig. 7). If we focus now at a point located in the warm-air sector (point *B* in Fig. 7), the transformation has, on the contrary, a stretching effect with an anisotropy ratio close to 2.

The impact of the transformation is also noticeable in the vertical: if we compare the vertical structure of the specified homogeneous correlation function (Fig. 8) to the retrieved correlation function in physical space (Fig. 9), it clearly appears that the correlation is tilted with the transformation. The slope of this tilt is about 2/100 and is roughly equal to the slope of the front as observed with the dense dropsonde network imple-

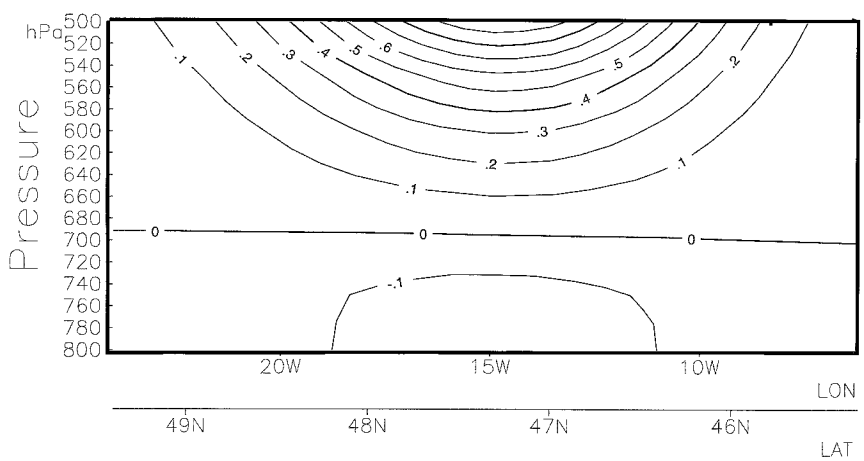

FIG. 8. Vertical cross section, as defined by line *VV'* in Fig. 6, of the isotropic correlation function for the temperature forecast errors, based at point *A*. The contour interval is 0.1. The horizontal extension of the cross section is about 1500 km.

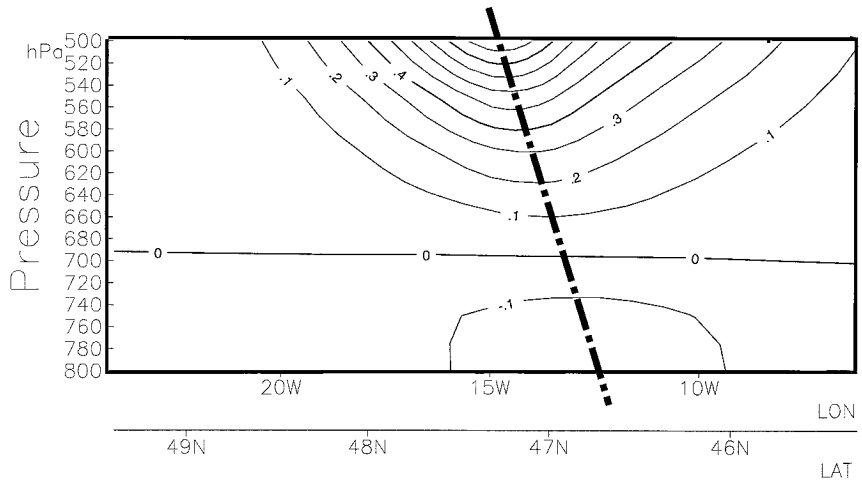

FIG. 9. As in Fig. 8, but for the implicit anisotropic correlation function yielded by the transformation. The dashed line is drawn in order to indicate the vertical tilt induced by the coordinate change.

mented during the FRONTS 87 experiment (Thorpe and Clough 1991).

This is another important and positive result: as vertical structures in transformed space tend to orient themselves along a front, a vertical correlation will also be tilted with the same slope. Thus, the assimilation of data with such a transformation should provide a better dynamical coherence of fronts: one can in particular expect that the vertical resolution of a profile crossing a front could implicitly improve the horizontal resolution of the retrieved analysis.

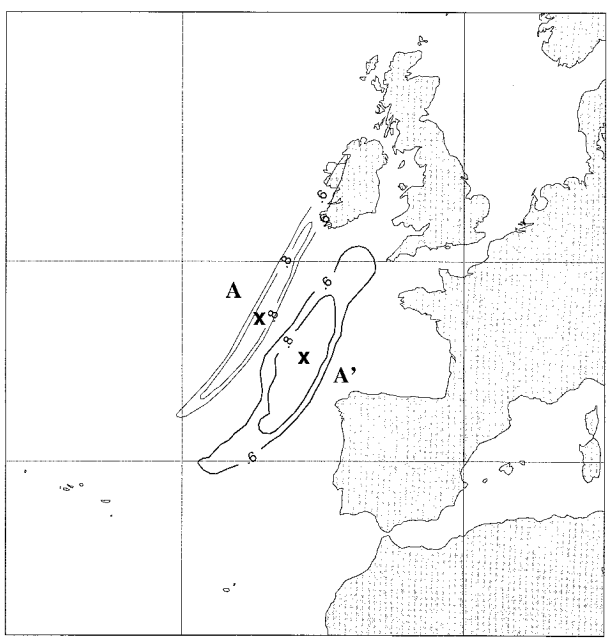

FIG. 10. Variation $1 - |\epsilon\zeta/\epsilon\zeta_A|$ of the vorticity differences $\epsilon\zeta$ between a 6-h forecast and a 30-h forecast valid at the same time (1200 UTC 9 January 1988). Here, $\epsilon\zeta$ is normalized by its value at point A inside the cold front (thin solid lines). The thick solid lines are for the same function but computed for the differences between the transformed vorticity fields, normalized by the Jacobian of the transformation. The point A' is the image of A in transformed space (the shaded map is valid only for the function in physical space). A spatial windowing of the functions is applied around points A and A' and only values more than 0.6 are plotted. The contour interval is 0.2.

Figure 10 gives an insight into the validity of the hypothesis of isotropy of the forecast error correlations in transformed space. It shows the variation of the differences between two forecasts of the vorticity field valid at the same time but with different ranges separated by 24 h and normalized by the forecast differences at point A inside the cold front. The strong anisotropy of the function in physical space is clearly reduced in transformed space.

6. Conclusions

In this note we have shown how a coordinate transformation inspired by the semigeostrophic formalism could be implemented in a variational assimilation scheme.

We have proposed a formulation of the coordinate change on the sphere, defined to be applicable mainly in midlatitude regions.

To build the transformation from real to transformed space, an iterative procedure has also been defined. This procedure encloses an iterative filtering of the wind field used in the transformation: it is shown that this formalism allows an efficient filtering of small-scale structures that cannot be taken into account by the semigeostrophic theory, but keeps nearly unchanged the strong cyclonic vorticity values associated with a frontal structure.

Starting from a continuous model of forecast error correlation, we show how the hypothesis of isotropy in

transformed space implicitly yields anisotropic error correlations in physical space.

The next logical step will be to introduce the transformation in a 3D-Var scheme and to perform assimilation tests. However, some points need further discussion. The amount of filtering in the coordinate change, which has been empirically determined, would require a better theoretical or statistical justification. The definition of the set of variables to which this transformation could be applied in a data assimilation procedure needs also to be investigated. New statistics of the forecast error correlations in transformed space should also been computed and compared to the corresponding statistics in real space.

Acknowledgments. I would like to thank P. Arbogast, P. Bernardet, A. Joly, and J.-N. Thépaut for their constructive comments on this note.

APPENDIX

Direct Transformation With the Secant Method

Relations (7), introduced in section 3 and defining the coordinate change, can be written under the form $(\mu_R, \gamma_R, \nu_R) = [\mathcal{F}(\mu_R, \gamma_R, \nu_R), \mathcal{G}(\mu_R, \gamma_R, \nu_R), \mathcal{H}(\mu_R, \gamma_R, \nu_R)]$, with $\gamma_R = \cos\theta_R \cos(\lambda_R - \lambda_G)$ and $\nu_R = \cos\theta_R \sin(\lambda_R - \lambda_G)$. Then, a simple way to solve this set of equations could be to iterate with $(\mu_R^{k+1}, \gamma_R^{k+1}, \nu_R^{k+1}) = (\mathcal{F}^k, \mathcal{G}^k, \mathcal{H}^k)$, where k stands for the iteration index.

This is a transposition of the fixed-point method used to solve a unidimensional equation $x = \mathcal{F}(x)$, where $\mathcal{F}(x)$ is a nonlinear expression of x. This method has been used by Schär and Davies (1990) to compute the real space representations of their semigeostrophic simulations on an f-plane. They have demonstrated why this method is likely to work in most cases when performing this inverse transformation. However, for direct transformation from real space to geostrophic space, it can be checked that this procedure may on the contrary fail in strong cyclonic area [the fixed-point procedure converges provided that the mapping $(\lambda_R^k, \mu_R^k) \rightarrow (\lambda_R^{k+1}, \mu_R^{k+1})$ is contractive in the area surrounding the fixed point with coordinates (λ_G, μ_G); this condition may break down when the geostrophic vorticity in real space reaches f]. Thus, in spite of its nice simplicity, the fixed-point method cannot be applied to the direct transformation from real space to geostrophic space. We have developed another procedure based on the secant method. For a unidimensional equation $\mathcal{F}_s(x) = 0$ to solve, this method is based on the following iteration:

$$x^{k+1} = x^k + \frac{x^k - x^{k-1}}{\mathcal{F}_s(x^k) - \mathcal{F}_s(x^{k-1})}\mathcal{F}_s(x^k).$$

This iteration can be deduced from the development of $\mathcal{F}_s$ into a Taylor series about x^k and with the condition $\mathcal{F}_s(x^{k+1}) = 0$.

Here, we define $\mathcal{F}_s$ and $\mathcal{G}_s$ by

$$\mathcal{F}_s(\lambda_R, \mu_R) = \mu_R - \widetilde{\mu_R}$$

$$\mathcal{G}_s(\lambda_R, \mu_R) = \lambda_R - \widetilde{\lambda_R},$$

where $\widetilde{\mu_R}$, $\widetilde{\lambda_R}$ are the values deduced from the evaluation of the right-hand members of Eq. (7) with μ_R and λ_R. Then, with these definitions, the secant method can be generalized in dimension two by writing

$$\left.\begin{aligned} \mathcal{F}_s^{k+1} &= \mathcal{F}_s^k + \delta\lambda_R^{k+1}\frac{\mathcal{F}_s(\lambda_R^k, \mu_R^{k-1}) - \mathcal{F}_s^{k-1}}{\delta\lambda_R^k} \\ &\quad + \delta\mu_R^{k+1}\frac{\mathcal{F}_s(\lambda_R^{k-1}, \mu_R^k) - \mathcal{F}_s^{k-1}}{\delta\mu_R^k} \\ \mathcal{G}_s^{k+1} &= \mathcal{G}_s^k + \delta\lambda_R^{k+1}\frac{\mathcal{G}_s(\lambda_R^k, \mu_R^{k-1}) - \mathcal{G}_s^{k-1}}{\delta\lambda_R^k} \\ &\quad + \delta\mu_R^{k+1}\frac{\mathcal{G}_s(\lambda_R^{k-1}, \mu_R^k) - \mathcal{G}_s^{k-1}}{\delta\mu_R^k} \end{aligned}\right\} \quad \text{(A1)}$$

where $\delta\lambda_R^k = \lambda_R^k - \lambda_R^{k-1}$, $\delta\mu_R^k = \mu_R^k - \mu_R^{k-1}$ and $\mathcal{F}_s^k$, $\mathcal{G}_s^k$ stand, respectively, for $\mathcal{F}_s(\lambda_R^k, \mu_R^k)$, $\mathcal{G}_s(\lambda_R^k, \mu_R^k)$.

At each iteration k, the computation of the right-hand members of the previous expressions requires the evaluation of U^{RG} and V^{RG} at different locations: this is performed by an interpolation in real space of the components of the initial wind v_g, as defined in section 5, and by the application of the rotation $\mathcal{R}^{RG}$ introduced in section 3.

Setting $\mathcal{F}_s^{k+1} = 0$ and $\mathcal{G}_s^{k+1} = 0$, the inversion of the previous linear system for λ_R^{k+1} and μ_R^{k+1} allows the updating of λ_R and μ_R.

The starting point of the minimization is set to $(\lambda_R^0, \mu_R^0) = (\lambda_G, \mu_G)$ and the coordinates (λ_R^1, μ_R^1) of R are deduced from a single iteration of the above-described fixed-point method.

With another fixed-point iteration added after each secant iteration, in order to suppress some local bad convergence, this procedure allows to build the coordinate change.

REFERENCES

Benjamin, S., K. Brewster, R. Brummer, B. Jewett, and T. Schlatter, 1991: An isentropic three-hourly data assimilation system using ACARS aircraft observations. *Mon. Wea. Rev.,* **119,** 888–906.

Courtier, P., C. Freydier, J.-F. Geleyn, F. Rabier, and M. Rochas, 1991: The ARPEGE project at Météo-France. *Proc. ECMWF Workshop on Numerical Methods in Atmospheric Models,* Reading, United Kingdom, ECMWF, 193–231.

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

Desroziers, G., and J. P. Lafore, 1993: A coordinate transformation for objective frontal analysis. *Mon. Wea. Rev.,* **121,** 1531–1553.

Gandin, L. S., 1965: *Objective Analysis of Meteorological Fields.* Israel Program for Scientific Translations, 242 pp.

Hoskins, B., 1975: The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.,* **32,** 233–242.

——, and F. Bretherton, 1972: Atmospheric frontogenesis models: Mathematical formulation and solution. *J. Atmos. Sci.,* **29,** 11–27.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 97–110.

Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus,* **37A,** 309–322.

Magnusdottir, G., and W. H. Schubert, 1991: Semigeostrophic theory on the hemisphere. *J. Atmos. Sci.,* **48,** 1449–1456.

Schär, C., and H. Davies, 1990: An instability of mature cold fronts. *J. Atmos. Sci.,* **47,** 929–950.

Schmidt, F., 1977: Variable fine mesh in spectral global model. *Beitr. Phys. Atmos.,* **50,** 211–217.

Thépaut, J.-N., R. Hoffman, and P. Courtier, 1993: Interaction of dynamics and observations in a four-dimensional variational assimilation. *Mon. Wea. Rev.,* **121,** 3393–3413.

——, P. Courtier, G. Belaud, and G. Lemaître, 1996: Dynamical structure functions in a four-dimensional variational assimilation: A case study. *Quart. J. Roy. Meteor. Soc.,* **122,** 535–561.

Thorpe, A., and S. Clough, 1991: Mesoscale dynamics of cold fronts: Structures described by dropsoundings in FRONTS 87. *Quart. J. Roy. Meteor. Soc.,* **117,** 903–941.

Yessad, K., 1995: Semi-Lagrangian computations in the cycle cy13r3 of ARPEGE/IFS. Météo-France/CNRM Technical Rep., 70 pp. [Available from Météo-France, CNRM/GMAP, 42 Av. G. Goriolis, 31057 Toulouse Cedex, France.]

Q. J. R. Meteorol. Soc. (2001), **127**, pp. 1433–1452

Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation

By GERALD DESROZIERS* and SERGUEI IVANOV
Météo-France, France

(Received 27 April 2000; revised 27 November 2000)

SUMMARY

Following the *a posteriori diagnosis* approach proposed by some authors, a practical computation of the expectation of sub-parts of the value of a cost function at the minimum is shown to be feasible by using a randomization technique based on a perturbation of observations or background fields. These computations allow the tuning of observation-error weighting parameters by applying a simple iterative fixed-point procedure. The procedure is first tested in a simplified variational scheme on a circular domain and then in a similar scheme but with the addition of the vertical coordinate. The relationship between the proposed approach and the Generalized Cross Validation is also shown. A test in the French Action de Recherche Petite Echelle Grande Echelle (ARPEGE) three-dimensional variational framework with both simulated observations and background fields is finally performed. It shows that a complete description of observation-error parameters can be retrieved with only a few iterations and, thus, at a reasonable cost.

KEYWORDS: Parameter estimation Variational assimilation

1. INTRODUCTION

Most modern assimilation schemes basically rely on linear estimation theory, or on an extension of this theory. In such an approach, each observation is given a weight that is proportional to the inverse of its specified error variance, measuring the confidence or the precision given to this particular observation. Because of the poor accuracy of certain observations or their sparse density in some areas, practical implementations of operational analysis schemes are based on the use not only of proper observation sets but also of background fields, given by a short-range forecast. In fact, these background fields can be seen as another source of observations (Talagrand 1997) with a given confidence that corresponds to the forecast-error covariances. Because the final analysis is very dependent on the specification of the relative weights given to each source of observations, through the error covariances, and because these errors are not perfectly known, a large potential for improvement on analyses is offered by methods producing a posteriori diagnoses of a mis-specification of a priori errors, or by procedures allowing an adaptive tuning of these parameters.

On the other hand, large operational centres are now using, or have planned to use, assimilation schemes based on a three-dimensional or four-dimensional variational (3D-Var or 4D-Var, respectively) approach, that especially allows the use of a wider range of observations (Lewis and Derber 1985; Courtier and Talagrand 1987). Diagnoses based on statistics of innovations (the differences between observations and background) (Bennett *et al.* 1993), or of departures between observations (including the background) and the minimizing solution (Talagrand 1999) that can be applied in a variational framework, have been proposed. In particular, it has been shown that a simple diagnosis is the value at the minimum of the cost function that measures the distance between observations and analysis.

In that case, the idea is to perform a posteriori statistics of such a criterion on a sufficiently large number of realizations of the analysis process in order to verify the a priori setting of certain parameters.

* Corresponding author: Météo-France, CNRM/GMME/RECYF, 42 av. G. Coriolis, 31057 Toulouse Cedex, France.

© Royal Meteorological Society, 2001.

A different kind of method tries to recover an appropriate tuning of error parameters from a single set of observations used in one analysis. In particular, methods based on the maximum-likelihood approach (Dee 1995) have been proposed. A second type of procedure inspired by the Generalized Cross Validation (GCV) approach was also introduced (Wahba and Wendelberger 1980) and proved to be feasible in a variational context (Wahba *et al.* 1995). Both types of methods have been applied to estimate a limited number of parameters, simultaneously with the estimation of the analysis. In that sense, these methods differ from others retrieving a specification of observation or background errors from a comparison over a period of time between observations and short-range forecasts (Hollingsworth and Lönnberg 1986; Lönnberg and Hollingsworth 1986).

This paper presents a method based on diagnoses of *observations-minus-analysis* differences but which aims to perform an adaptive tuning of observation-error parameters from a single batch of observations and background fields.

In section 2 the general variational framework for the implementation of such an approach is presented, as well as the *observations-minus-analysis* diagnosis approach; a practical way for computing these diagnoses is also described. Section 3 introduces the proposed procedure in a simple case, where one tries to tune the ratio between background and observation terms; the link with GCV for this simple case is also shown; an application in a simplified one-dimensional (1D) case is given. Then, it is shown how the procedure can be generalized to the determination of weights of parts of the background and observation terms, and an application to the tuning of the vertical profile of observation or background errors is performed still with a simple toy variational problem, yet including a second vertical dimension. Section 4 presents an application of the procedure, used to optimize radiosonde errors, in a more realistic framework given by the operational 3D-Var scheme used at Météo-France. Conclusions and perspectives are formulated in section 5.

2. Diagnoses in a variational scheme

(*a*) *Variational formulation*

The principle of the incremental formulation of 3D/4D-Var algorithms (Courtier *et al.* 1994) is to seek the increment $\delta\mathbf{x}$ to be added to the background $\mathbf{x}^{\mathrm{b}}$—so that the analysis is given by $\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \delta\mathbf{x}$—which minimizes the cost function

$$\begin{aligned} \mathrm{J}(\delta\mathbf{x}) &= \mathrm{J}^{\mathrm{b}}(\delta\mathbf{x}) + \mathrm{J}^{\mathrm{o}}(\delta\mathbf{x}) \\ &= \tfrac{1}{2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + \tfrac{1}{2}(\mathbf{d} - \mathrm{H}\delta\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{d} - \mathrm{H}\delta\mathbf{x}). \end{aligned}$$

The background term J^{b} measures the distance between the analysis $\mathbf{x}^{\mathrm{a}}$ and the short-range forecast $\mathbf{x}^{\mathrm{b}}$, with $\mathbf{B}$ the forecast-error covariance matrix. In the observation term J^{o}, H is the linearized observation operator, $\mathbf{R}$ stands for the observation-error covariance matrix, including representativeness errors (Lorenc 1986), and $\mathbf{d} = \mathbf{y}^{\mathrm{o}} - \mathrm{H}(\mathbf{x}^{\mathrm{b}})$ is the innovation vector, with H the observation operator that allows the computation of the model equivalents in the space of the observations.

The solution of the minimization of J is given by

$$\delta\mathbf{x}^{\mathrm{a}} = \mathbf{K}\mathbf{d} = \mathbf{K}\{\mathbf{y}^{\mathrm{o}} - \mathrm{H}(\mathbf{x}^{\mathrm{b}})\}, \tag{1}$$

where $\mathbf{K} = \mathbf{B}\mathrm{H}^{\mathrm{T}}(\mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}} + \mathbf{R})^{-1}$ is the gain matrix.

The previous formalism corresponds to the 3D-Var case; however, the extension to the temporal dimension (4D-Var) can be easily introduced: then H includes the integration of the forecast model, $\mathbf{x}^{\mathrm{a}}$ stands for the initial conditions at the beginning of the assimilation period, and an outer loop in the minimization process is introduced in order to take into account some parts of the nonlinearities contained in H. Although the following developments stand in a 4D-Var framework, this paper focuses on the 3D-Var case for the sake of simplicity.

(*b*) *Statistical average of the terms in* J

Following Talagrand (1997), the complete vector of observations, $\mathbf{z}^{\mathrm{o}}$, can be seen as the two-component vector of proper observations, $\mathbf{y}^{\mathrm{o}}$, with dimension p, and of the background estimate $\mathbf{x}^{\mathrm{b}}$, with the same dimension n as the true state $\mathbf{x}^{\mathrm{t}}$: $\mathbf{z}^{\mathrm{o}} = \{(\mathbf{x}^{\mathrm{b}})^{\mathrm{T}}(\mathbf{y}^{\mathrm{o}})^{\mathrm{T}}\}^{\mathrm{T}}$ and

$$\mathbf{x}^{\mathrm{b}} = \mathbf{x}^{\mathrm{t}} + \boldsymbol{\epsilon}^{\mathrm{b}},$$

where $\boldsymbol{\epsilon}^{\mathrm{b}}$ is the vector of unknown forecast errors with covariance matrix $\mathbf{B}$,

$$\mathbf{y}^{\mathrm{o}} = \mathrm{H}(\mathbf{x}^{\mathrm{t}}) + \boldsymbol{\epsilon}^{\mathrm{o}},$$

where $\boldsymbol{\epsilon}^{\mathrm{o}}$ is the vector of unknown observation errors with covariance matrix $\mathbf{R}$. Thus, $\mathbf{z}^{\mathrm{o}}$ could also be written

$$\mathbf{z}^{\mathrm{o}} = \Gamma(\mathbf{x}^{\mathrm{t}}) + \boldsymbol{\epsilon},$$

with Γ the complete *observation operator* and $\boldsymbol{\epsilon}$ the vector of forecast and observation errors, with dimension $n + p$.

An important result pointed out by Talagrand (1999) is that, if J_j stands for a term of J, which is the sum of m_j elements, then the expectation of J_j at the minimum is

$$\mathrm{E}\{\mathrm{J}_j(\mathbf{x}^{\mathrm{a}})\} = \tfrac{1}{2}\{m_j - \mathrm{Tr}(\Gamma_j^{\mathrm{T}}\mathbf{S}_j^{-1}\boldsymbol{\Gamma}_j\mathbf{P}^{\mathrm{a}})\}, \tag{2}$$

where Γ_j and $\mathbf{S}_j$ are, respectively, the linearized observation operator and the observation-error covariance matrix associated with these m_j elements (a proof of this relation is given in appendix A). Tr is the trace of the matrix to which it is applied. Here, $\mathbf{P}^{\mathrm{a}}$ stands for the estimation analysis-error covariance matrix resulting from the analysis with the whole set of m pieces of observations ($m = \sum_j m_j$ and $\mathrm{J} = \sum_j \mathrm{J}_j$).

A first application of the previous expression is for the complete observation term J^{o}, with, in this case, $\Gamma_j = \mathrm{H}$ and $\mathbf{S}_j = \mathbf{R}$:

$$\begin{aligned}\mathrm{E}(\mathrm{J}^{\mathrm{o}}) &= \tfrac{1}{2}\{p - \mathrm{Tr}(\mathrm{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathrm{H}\mathbf{P}^{\mathrm{a}})\}\\ &= \tfrac{1}{2}\{p - \mathrm{Tr}(\mathrm{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}})\}\\ &= \tfrac{1}{2}\,\mathrm{Tr}(\mathbf{I}_p - \mathrm{H}\mathbf{K}),\end{aligned}$$

where $\mathbf{I}_p$ stands for the identity matrix with order p, and using the fact that $\mathbf{K} = \mathbf{P}^{\mathrm{a}}\mathrm{H}^{\mathrm{T}}\mathbf{R}^{-1}$. Similarly, the application for the complete background term, with $\boldsymbol{\Gamma}_j = \mathbf{I}_n$ and $\mathbf{S}_j = \mathbf{B}$ gives

$$\begin{aligned}\mathrm{E}(\mathrm{J}^{\mathrm{b}}) &= \tfrac{1}{2}[n - \mathrm{Tr}\{\mathbf{B}^{-1}(\mathbf{B} - \mathbf{K}\mathrm{H}\mathbf{B})\}]\\ &= \tfrac{1}{2}\{n - \mathrm{Tr}(\mathbf{I}_n - \mathbf{B}^{-1}\mathbf{K}\mathrm{H}\mathbf{B})\}\\ &= \tfrac{1}{2}\,\mathrm{Tr}(\mathbf{K}\mathrm{H}),\end{aligned}$$

where $\mathbf{I}_n$ is the identity matrix with order n and using the relation $\mathbf{P}^{\mathrm{a}} = \mathbf{B} - \mathbf{KHB}$.

Thus, since $\mathrm{Tr}(\mathbf{HK}) = \mathrm{Tr}(\mathbf{KH})$, it follows that

$$\mathrm{E}(\mathrm{J}) = \mathrm{E}(\mathrm{J}^{\mathrm{o}}) + \mathrm{E}(\mathrm{J}^{\mathrm{b}}) = 1/2\,\mathrm{Tr}(\mathbf{I}_p) = p/2.$$

This means that if the observation-error covariances are properly specified, then the expectation of the value of the cost function at the minimum is simply proportional to the number of proper observations, p, which also corresponds to the number of degrees of freedom in the analysis problem. As shown by some authors (Bennett *et al.* 1993; Talagrand 1999), this provides a simple a posteriori diagnosis that requires no extra cost: a deviation of the expectation of J from $p/2$ will indicate a mis-specification of the statistics of the observation errors. It can also be shown that if observation errors are normally distributed then J has a χ^2-distribution with p degrees of freedom. This suggests that each part, J_j, of J should also have a χ^2-distribution but with a rough count of degrees of freedom given by $m_j - \mathrm{Tr}(\Gamma_j^{\mathrm{T}}\mathbf{S}_j^{-1}\Gamma_j\mathbf{P}^{\mathrm{a}})$.

Statistics of parts of J through Eq. (2) can be powerful diagnosis tools: however, there is a difficulty in their estimation coming from the computation of $\mathrm{Tr}(\Gamma_j^{\mathrm{T}}\mathbf{S}_j^{-1}\Gamma_j\mathbf{P}^{\mathrm{a}})$. The following section introduces a practical way for computing such an expression even when m_j is large.

(c) *Randomized estimation of* Tr(KH) *and* Tr(HK)

The computation of $\mathrm{E}(\mathrm{J}^{\mathrm{o}})$ is examined first: from the relation given in the previous section, it requires the computation of $\mathrm{Tr}(\mathbf{HK})$. The idea is to estimate this trace by a randomized method, using a method proposed by Girard (1987) and applied to the GCV computation (Girard 1991; Wahba *et al.* 1995), which implies the computation of the same expression (see section 3): it can be shown that a randomized estimation Rand Tr(HK) of $\mathrm{Tr}(\mathbf{HK})$ is given by

$$\mathrm{Rand\,Tr}(\mathbf{HK}) = \mathrm{Rand\,Tr}(\mathbf{R}^{-1/2}\mathbf{HKR}^{1/2}) = (\mathbf{R}^{-1/2}\boldsymbol{\xi})^{\mathrm{T}}\mathbf{HKR}^{1/2}\boldsymbol{\xi},$$

where $\boldsymbol{\xi}$ is a p-dimensional vector of numbers with a standard Gaussian distribution (mean zero and variance one).

In practice this is not sufficient, since $\mathbf{K}$ is not explicitly computed in an operational variational scheme. However, Eq. (1) yields

$$\mathrm{H}\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}}+\delta\mathbf{y}^{\mathrm{o}})} - \mathrm{H}\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}})} \simeq \mathbf{HK}\delta\mathbf{y}^{\mathrm{o}}$$

where $\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}}+\delta\mathbf{y}^{\mathrm{o}})}$ and $\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}})}$ are the analysis increments obtained with perturbed and unperturbed observations, respectively.

Combining both previous expressions, with $\delta\mathbf{y}^{\mathrm{o}} = \mathbf{R}^{1/2}\boldsymbol{\xi}$, a randomized estimate of $\mathrm{Tr}(\mathbf{HK})$ is given by

$$\mathrm{Rand\,Tr}(\mathbf{HK}) = (\mathbf{R}^{-1/2}\boldsymbol{\xi})^{\mathrm{T}}(\mathrm{H}\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}}+\mathbf{R}^{1/2}\boldsymbol{\xi})} - \mathrm{H}\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}})}).$$

The accuracy of this estimate depends not only on the value of Tr(Hk) itself, but also on the number of observations (Girard 1991; Wahba *et al.* 1995).

Since $\mathrm{Tr}(\mathbf{HK}) = \mathrm{Tr}(\mathbf{KH})$, this randomized estimation can also be used to compute $\mathrm{E}(\mathrm{J}^{\mathrm{b}})$. However, a similar procedure can be employed to determine $\mathrm{Tr}(\mathbf{KH})$ separately (this independent computation is also given here to introduce the computation of the expectation of parts of J^{b} addressed in a following section): using Eq. (1) again, it

follows that

$$\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{x}^{\mathrm{b}}+\delta\mathbf{x}^{\mathrm{b}})} - \delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{x}^{\mathrm{b}})} \simeq -\mathbf{K}\mathrm{H}\delta\mathbf{x}^{\mathrm{b}},$$

and then a randomized estimate of Tr(**K**H) can also be given by

$$\mathrm{Rand\ Tr}(\mathbf{K}\mathrm{H}) = -(\mathbf{B}^{-1/2}\boldsymbol{\xi})^{\mathrm{T}}(\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{x}^{\mathrm{b}}+\mathbf{B}^{1/2}\boldsymbol{\xi})} - \delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{x}^{\mathrm{b}})}),$$

where $\boldsymbol{\xi}$ is a vector of random numbers with dimension n in this case. Both previous expressions for Tr(H**K**) and Tr(**K**H) provide a practical means to compute $\mathrm{E}(\mathrm{J}^{\mathrm{o}})$ and $\mathrm{E}(\mathrm{J}^{\mathrm{b}})$: each computation requires two analysis steps, with unperturbed observations ($\mathbf{y}^{\mathrm{o}}$ and $\mathbf{x}^{\mathrm{b}}$, respectively) and perturbed observations ($\mathbf{y}^{\mathrm{o}} + \mathbf{R}^{1/2}\boldsymbol{\xi}$ and $\mathbf{x}^{\mathrm{b}} + \mathbf{B}^{1/2}\boldsymbol{\xi}$, respectively).

Note that Fisher (1999) proposed a different randomized estimation of Tr(**K**H) to compute GCV using the relation $\mathbf{K} = \mathbf{P}^{\mathrm{a}}\mathrm{H}^{\mathrm{T}}\mathbf{R}^{-1}$ and applying a Lanczos algorithm to the Hessian matrix $\mathrm{J}'' = (\mathbf{P}^{\mathrm{a}})^{-1}$ of the cost function.

3. Adaptive tuning procedure

(*a*) *On-line tuning of the ratio between background and observation errors*

The problem of tuning the weighting parameters of the background and observation terms is addressed first in this section. If $s^{\mathrm{b}2}$ and $s^{\mathrm{o}2}$ denote these two parameters, then the cost function introduced in the previous section can be rewritten as

$$\mathrm{J}(\delta\mathbf{x}) = \frac{1}{s^{\mathrm{b}2}}\mathrm{J}^{\mathrm{b}}(\delta\mathbf{x}) + \frac{1}{s^{\mathrm{o}2}}\mathrm{J}^{\mathrm{o}}(\delta\mathbf{x}). \tag{3}$$

As in some applications of the GCV or in the maximum-likelihood methods, the procedure that is proposed here tries to optimize this couple of parameters, or simply their ratio, $\lambda = s^{\mathrm{o}2}/s^{\mathrm{b}2}$, from a single batch of observations. The rationale behind this is that, if s^{b} and s^{o} are the proper weights to introduce, then the values of $(1/s^{\mathrm{b}2})\mathrm{J}^{\mathrm{b}}(\delta\mathbf{x})$ and $(1/s^{\mathrm{o}2})\mathrm{J}^{\mathrm{o}}(\delta\mathbf{x})$ should be close to their expected values, that is to say $(1/2)\ \mathrm{Tr}\{\mathbf{K}\mathrm{H}(s^{\mathrm{b}}, s^{\mathrm{o}})\}$ and $(1/2)\ \mathrm{Tr}\{\mathbf{I}_p - \mathrm{H}\mathbf{K}(s^{\mathrm{b}}, s^{\mathrm{o}})\}$, respectively.

Thus, defining $S^{\mathrm{b}} = 2\mathrm{J}^{\mathrm{b}}/\mathrm{Tr}(\mathbf{K}\mathrm{H})$ and $S^{\mathrm{o}} = 2\mathrm{J}^{\mathrm{o}}/\mathrm{Tr}(\mathbf{I}_p - \mathrm{H}\mathbf{K})$, the idea is to find s^{b} and s^{o} such that

$$\begin{cases} s^{\mathrm{b}2} = S^{\mathrm{b}}(s^{\mathrm{b}}, s^{\mathrm{o}}) \\ s^{\mathrm{o}2} = S^{\mathrm{o}}(s^{\mathrm{b}}, s^{\mathrm{o}}). \end{cases} \tag{4}$$

At this stage, it is important to point out the similarity between this simple approach and GCV: the GCV criterion, which is usually minimized with respect to the ratio λ, is defined by

$$V(\lambda) = \frac{(\mathbf{d} - \mathrm{H}\delta\mathbf{x}^{\mathrm{a}})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{d} - \mathrm{H}\delta\mathbf{x}^{\mathrm{a}})}{\{\mathrm{Tr}(\mathbf{I}_p - \mathrm{H}\mathbf{K})\}^2} \tag{5}$$

$$= \frac{S^{\mathrm{o}}(\delta\mathbf{x}^{\mathrm{a}})}{\mathrm{Tr}(\mathbf{I}_p - \mathrm{H}\mathbf{K})}. \tag{6}$$

A randomized estimation Rand V of $V(\lambda)$ can also be computed using the randomization procedure previously described.

Therefore, the GCV criterion is the ratio between S^{o} and $\mathrm{Tr}(\mathbf{I}_p - \mathbf{HK})$, this last expression appearing itself in the computation of S^{o}. In appendix B, it is shown that if s^{b} and s^{o} are the optimal values of these parameters, then V reaches a minimum for their ratio λ and also $\lambda = s^{\mathrm{o}2}/s^{\mathrm{b}2} = S^{\mathrm{o}}(s^{\mathrm{b}}, s^{\mathrm{o}})$, so that both approaches are equivalent in this case.

The form of the nonlinear pair of equations (4) with respect to s^{b} and s^{o}, suggests the use of an iterative fixed-point method with the sequence

$$\begin{cases} s^{\mathrm{b}}_{k+1} = \{S^{\mathrm{b}}(s^{\mathrm{b}}_k, s^{\mathrm{o}}_k)\}^{1/2} \\ s^{\mathrm{o}}_{k+1} = \{S^{\mathrm{o}}(s^{\mathrm{b}}_k, s^{\mathrm{o}}_k)\}^{1/2} \end{cases} \tag{7}$$

at iteration k.

(*b*) *Test in a simplified one-dimensional variational formulation*

A first toy problem was constructed to test the iterative tuning procedure introduced in the previous section. The framework is given by an analysis problem on a circular domain (say on an earth meridian). A spectral Fourier decomposition of the signal (assumed here to be a temperature increment) is used. The variational problem is formulated as in Eq. (3), where $\delta\mathbf{x}$ stands for the vector of spectral coefficients.

The length of the domain is arbitrarily set to 40 000 km and the truncation to 100, which gives $n = 201$ spectral coefficients. To define the background-error covariance matrix $\mathbf{B}$, a Gaussian shape structure function is first defined in physical space with a 300 km length-scale. Assuming homogeneity on the domain, it can be shown that matrix $\mathbf{B}$ is diagonal and that its diagonal is obtained by applying the Fourier transform to the Gaussian correlation in physical space. Here, it is first assumed that the square root of the background error is also uniform on the domain and equal to 1 K, so that matrix $\mathbf{B}$ also corresponds to a correlation matrix.

The observation operator H is obtained thanks to an inverse Fourier transform and a simple linear interpolation at the observation location: thus, H is linear in this particular case. We also consider that the p observations have the same error variances and, therefore, matrix $\mathbf{R}$ is simply defined as the $p \times p$ identity matrix, the normalization by the variance being introduced through coefficient $s^{\mathrm{o}2}$ in Eq. (3). The number of observations, p, is set to 100.

In order to precondition the problem, a change of variable is performed with $\delta\mathbf{x} = \mathbf{B}^{1/2}\boldsymbol{\chi}$, where $\boldsymbol{\chi}$ is the new control variable, so that the cost function to minimize becomes $\mathrm{J}(\boldsymbol{\chi}) = 1/2\boldsymbol{\chi}^{\mathrm{T}}\boldsymbol{\chi} + \mathrm{J}^{\mathrm{o}}(\mathbf{B}^{1/2}\boldsymbol{\chi})$.

The minimization is performed with a Quasi-Newton algorithm (Gilbert and Lemaréchal 1989) involving a maximum of 70 iterations. The solution of this simple linear problem is also explicitly computed using Eq. (1).

To test the tuning procedure introduced in the previous section, both background and observation errors are simulated. If $\mathbf{x}^{\mathrm{t}}$ denotes the true signal, then $\mathbf{x}^{\mathrm{t}} = \mathbf{x}^{\mathrm{b}} - \boldsymbol{\epsilon}^{\mathrm{b}}$. We will further simplify the problem by assuming that $\mathbf{x}^{\mathrm{b}} = 0$ (so that $\mathbf{x}^{\mathrm{t}} = -\boldsymbol{\epsilon}^{\mathrm{b}}$). If $\boldsymbol{\zeta}^{\mathrm{b}}$ is an n-dimensional vector of random numbers with a Gaussian distribution and variance 1, then writing $\boldsymbol{\epsilon}^{\mathrm{b}} = \mathbf{B}^{1/2}\boldsymbol{\zeta}^{\mathrm{b}}$ is a way of simulating background errors consistent with the matrix $\mathbf{B}$ since $\mathrm{E}(\boldsymbol{\epsilon}^{\mathrm{b}}\boldsymbol{\epsilon}^{\mathrm{b}^{\mathrm{T}}}) = \mathbf{B}^{1/2}\mathrm{E}(\boldsymbol{\zeta}^{\mathrm{b}}\boldsymbol{\zeta}^{\mathrm{b}^{\mathrm{T}}})\mathbf{B}^{1/2} = \mathbf{B}$ (using the fact that $\mathrm{E}(\boldsymbol{\zeta}^{\mathrm{b}}\boldsymbol{\zeta}^{\mathrm{b}^{\mathrm{T}}})$ is the identity matrix since $\boldsymbol{\zeta}^{\mathrm{b}}$ is a vector of random values with variance 1).

Similarly, a p-dimensional vector of observation errors can be obtained with $\mathbf{y}^{\mathrm{o}} = \mathrm{H}(\mathbf{x}^{\mathrm{t}}) + \mathbf{R}^{1/2}\boldsymbol{\zeta}^{\mathrm{o}}$, where $\boldsymbol{\zeta}^{\mathrm{o}}$ is a vector of random numbers with a Gaussian distribution. Figure 1 shows the result of such a definition of $\mathbf{x}^{\mathrm{t}}$ and $\mathbf{y}^{\mathrm{o}}$: here $s^{\mathrm{o}} = 0.1$.

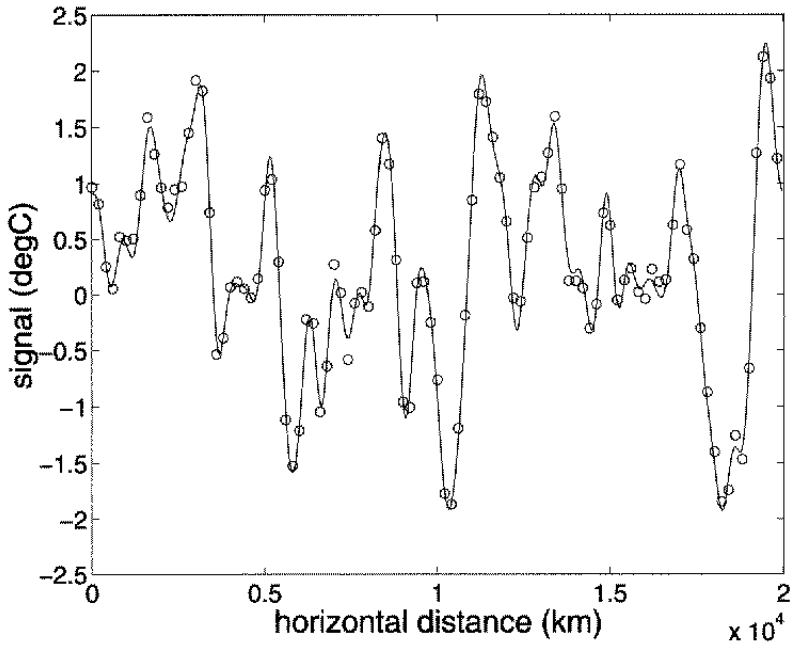

Figure 1. True signal $\mathbf{x}^t$ (full line) and observations $\mathbf{y}^o$ (circles) simulated with the procedure described in section 3(b); the length-scale of the structure function is 300 km and $\lambda = s^{o^2}/s^{b^2} = 0.01$. See text for further details.

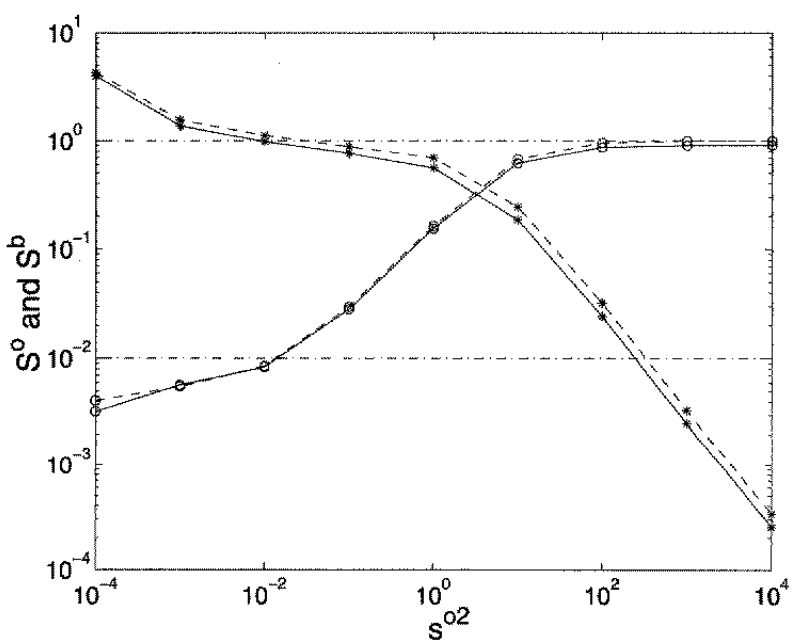

Figure 2. Exact criterion S^o (full line with circles) as a function of s^{o^2} computed with the explicit solution, and randomized estimation of S^o (dashed line with circles) computed with the two solutions $\mathbf{x}^a_{(\mathbf{y}^o)}$ and $\mathbf{x}^a_{(\mathbf{y}^o+\delta\mathbf{y}^o)}$ resulting from the minimization of the cost function J with the Quasi-Newton descent algorithm; solid line with stars and dashed line with stars: same as previously but for the criterion S^b. See text for further details.

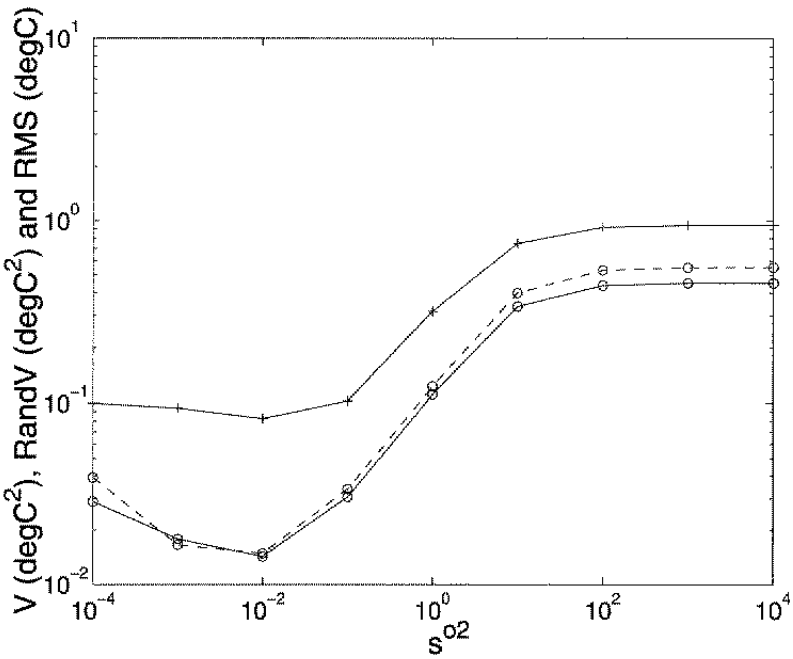

Figure 3. Exact Generalized Cross Validation criterion, V, as a function of s^{o^2} computed with the explicit solution (full line with circles) and randomized estimation Rand V of V (dashed line with circles) computed with the two solutions $\mathbf{x}^a_{(\mathbf{y}^o+\delta\mathbf{y}^o)}$ and $\mathbf{x}^a_{(\mathbf{y}^o)}$; the root mean square error (r.m.s.) between $\mathbf{x}^a(s^{o^2})$ and $\mathbf{x}^T$ is also plotted (solid line with crosses). See text for further details.

The first step of this test addresses the problem of recovering the correct value of s^{o2} or $\lambda = s^{o2}/s^{b2}$ since here, the parameter s^{b2} is assumed to be known ($s^{b2} = 1$). Figure 2 shows the variation of S^o with respect to s^{o2}, with two computations: one with the explicit solution $\mathbf{x}^a$ and the exact computation of $\mathrm{Tr}(\mathbf{I}_p - \mathbf{KH})$ and another evaluation of S^o obtained with the pair of solutions $\mathbf{x}^a_{(\mathbf{y}^o)}$ and $\mathbf{x}^a_{(\mathbf{y}^o+\mathbf{R}^{1/2}\xi)}$, produced by minimizing the cost function J with the Quasi-Newton descent algorithm. Accordingly, with the approach introduced in section 3(a), it appears that both exact and randomized computations of S^o are very similar, and that the relation $s^{o2} = S^o$ is fulfilled for a value of s^{o2} very close to the true value (0.01) used for simulating the observations. In this case, both exact and randomized computations of the GCV criterion V (Fig. 3) also reach a minimum for this optimal value of s^{o2} (see appendix B for a theoretical justification of such a result) and the root mean square (r.m.s.) error between $\mathbf{x}^a$ and $\mathbf{x}^t$ is logically minimum for this optimal value. On the other hand, the exact and randomized computations of S^b show that the relation $s^{b2} = S^b$ is true for the optimal value $s^{b2} = 1$ used for simulating the vector of background errors (Fig. 2), which is again in agreement with the pair of equations (4).

Then, the simple iterative method proposed in section 3(a) with Eq. (7) is applied, but for the determination of s^o only (s^b is kept equal to the optimal value). Table 1 shows that the optimal value is retrieved with this algorithm, with only a few iterations, whatever the initial value of the parameter is (no modification of s^o appears after these few iterations and, therefore, Table 1 displays results until iteration 5 only). The value of the GCV criterion V is also reduced in agreement with Fig. 3 and with the theoretical justification given in appendix B.

TABLE 1. VALUES OF OBSERVATION-ERROR WEIGHTING PARAMETERS (s^o) ALONG THE FIXED-POINT ITERATIONS

s^o it. 0	0.01	0.10	1.00
r.m.s.	0.1	0.08	0.32
V	0.4×10^{-3}	0.1×10^{-3}	1.2×10^{-3}
s^o it. 1	0.06	0.09	0.40
2	0.08	0.09	0.20
3	0.08	0.09	0.13
4	0.09	0.09	0.10
5	0.09	0.09	0.09
r.m.s.	0.08	0.08	0.08
V	0.1×10^{-3}	0.1×10^{-3}	0.1×10^{-3}
$(s^o)^s$		0.1	
r.m.s.		0.08	
V		0.1×10^{-3}	

The three right-hand columns of the table correspond to three different starting points; the values of the root mean square (r.m.s.) error and of the Generalized Cross Validation (GCV) criterion (V) are also given at the beginning and at the end of the iterations (it. 1–it. 5); the last three rows give the reference value of s^o used for the simulation of observations $(s^o)^s$ and the corresponding r.m.s. error and GCV criterion obtained with the use of this value in the analysis.

Indeed, this first test problem is rather simple due to the choice of homogeneous error covariances: the accuracy of such an estimation of the parameter with more realistic forms of covariances will have to be further investigated.

(c) *Tuning the weights given to parts of the background or observation terms*

In the previous section, it was shown that a simple iterative algorithm can be used for retrieving the global ratio between the background and the observation terms in a variational formulation. The possibility of tuning the weights given to parts of these two terms is now investigated.

First, if J^o_j stands for a part of J^o with p_j observations, and using again Eq. (2) introduced in section 2(b), then the following sequence of equalities can be written:

$$\begin{aligned}
\mathrm{E}(\mathrm{J}^o_j) &= \tfrac{1}{2}\{p_j - \mathrm{Tr}(\mathrm{H}_j^{\mathrm{T}}\mathbf{R}_j^{-1}\mathrm{H}_j\mathbf{P}^{\mathrm{a}})\} \\
&= \tfrac{1}{2}\{p_j - \mathrm{Tr}(\mathrm{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathrm{P}_j^{\mathrm{T}}\mathrm{P}_j\mathrm{H}\mathbf{P}^{\mathrm{a}})\} \\
&= \tfrac{1}{2}\{p_j - \mathrm{Tr}(\mathrm{P}_j\mathrm{H}\mathbf{P}^{\mathrm{a}}\mathrm{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathrm{P}_j^{\mathrm{T}})\} \\
&= \tfrac{1}{2}\{p_j - \mathrm{Tr}(\mathrm{P}_j(\mathrm{HK})\mathrm{P}_j^{\mathrm{T}})\},
\end{aligned}$$

where P_j is the projection operator that allows the passage from the complete set of observations to the subset of p_j observations, and $\mathbf{R}_j$ is the observation-error covariance matrix associated with this subset.

This means that once matrix **HK** is known, then the expectation of a part of J^o can be computed by applying the proper projection operator.

Similarly, the expectation of a part, $\mathrm{J}_j^{\mathrm{b}}$, of the background term can be computed using the following relations:

$$\begin{aligned}\mathrm{E}(\mathrm{J}_j^{\mathrm{b}}) &= \tfrac{1}{2}\{n_j - \mathrm{Tr}(\mathrm{P}_j^{\mathrm{T}}\mathbf{B}_j^{-1}\mathrm{P}_j\mathbf{P}^{\mathrm{a}})\}\\ &= \tfrac{1}{2}\{n_j - \mathrm{Tr}(\mathbf{B}_j^{-1}\mathrm{P}_j^{\mathrm{T}}\mathrm{P}_j\mathbf{P}^{\mathrm{a}})\}\\ &= \tfrac{1}{2}\{n_j - \mathrm{Tr}(\mathrm{P}_j\mathbf{P}^{\mathrm{a}}\mathbf{B}^{-1}\mathrm{P}_j^{\mathrm{T}})\}\\ &= \tfrac{1}{2}(n_j - \mathrm{Tr}[\mathrm{P}_j\{\mathbf{I}_n - (\mathbf{K}\mathrm{H})\}\mathrm{P}_j^{\mathrm{T}}]),\end{aligned}$$

where P_j is now the projection operator that allows the passage from the complete background vector $\mathbf{x}^{\mathrm{b}}$ to a subset with n_j elements, and $\mathbf{B}_j$ is the background-error covariance matrix associated with this subset.

Once more, these relations do not suffice, since neither $\mathrm{H}\mathbf{K}$ nor $\mathbf{K}\mathrm{H}$ are explicitly computed in a variational framework. A randomized estimation of $\mathrm{E}(\mathrm{J}_j^{\mathrm{o}})$ and $\mathrm{E}(\mathrm{J}_j^{\mathrm{b}})$, can however, be proposed in the form of

$$\mathrm{Rand}\ \mathrm{Tr}\{\mathrm{P}_j(\mathrm{H}\mathbf{K})\mathrm{P}_j^{\mathrm{T}}\} = (\mathrm{P}_j\mathbf{R}^{-1/2}\boldsymbol{\xi})^{\mathrm{T}}\mathrm{P}_j(\mathrm{H}\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}}+\mathbf{R}^{1/2}\boldsymbol{\xi})} - \mathrm{H}\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{y}^{\mathrm{o}})}), \tag{8}$$

and

$$\mathrm{Rand}\ \mathrm{Tr}\{\mathrm{P}_j(\mathbf{K}\mathrm{H})\mathrm{P}_j^{\mathrm{T}}\} = -(\mathrm{P}_j\mathbf{B}^{-1/2}\boldsymbol{\xi})^{\mathrm{T}}\mathrm{P}_j(\delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{x}^{\mathrm{b}}+\mathbf{B}^{1/2}\boldsymbol{\xi})} - \delta\mathbf{x}^{\mathrm{a}}_{(\mathbf{x}^{\mathrm{b}})}). \tag{9}$$

Equation (8) allows the computation of the expectation of parts of J^{o}, from only two analyses, one with the complete set of true values $\mathbf{y}^{\mathrm{o}}$ of the observations, and the other one with the whole set of perturbed observations, $\mathbf{y}^{\mathrm{o}} + \mathbf{R}^{1/2}\boldsymbol{\xi}$, followed by the application of P_j to the observation perturbation set $\mathbf{R}^{1/2}\boldsymbol{\xi}$ and to the difference between the pair of perturbed and unperturbed analyses.

Similarly, Eq. (9) proves that the expectation of parts of J^{b}, can be determined from two other analyses, one with the true values $\mathbf{x}^{\mathrm{b}}$ of the background field, and the other one with a perturbed background $\mathbf{x}^{\mathrm{b}} + \mathbf{B}^{1/2}\boldsymbol{\xi}$, and then by using the corresponding projection operator P_j.

The cost function defined in Eq. (3) can again be rewritten with

$$\mathrm{J}(\delta\mathbf{x}) = \sum_j \frac{1}{s_j^{\mathrm{b}2}}\mathrm{J}_j^{\mathrm{b}}(\delta\mathbf{x}) + \sum_j \frac{1}{s_j^{\mathrm{o}2}}\mathrm{J}_j^{\mathrm{o}}(\delta\mathbf{x}), \tag{10}$$

where $s_j^{\mathrm{b}2}$ and $s_j^{\mathrm{o}2}$ are, respectively, the background and observation-error weighting parameters supposed to be homogeneous for a given subset of observations, j.

Thus, following the same idea as in section 3(a), and defining $S_j^{\mathrm{b}} = 2\mathrm{J}_j^{\mathrm{b}}/(n_j - \mathrm{Tr}[\mathrm{P}_j\{\mathbf{I}_n - (\mathbf{K}\mathrm{H})\}\mathrm{P}_j^{\mathrm{T}}])$ and $S_j^{\mathrm{o}} = 2\mathrm{J}_j^{\mathrm{o}}/[p_j - \mathrm{Tr}\{\mathrm{P}_j(\mathrm{H}\mathbf{K})\mathrm{P}_j^{\mathrm{T}}\}]$, a procedure for finding the optimal values of parameters $s_j^{\mathrm{b}2}$ and $s_j^{\mathrm{o}2}$ is to solve the following set of equations

$$\begin{cases} s_j^{\mathrm{b}2} = S_j^{\mathrm{b}} \\ s_j^{\mathrm{o}2} = S_j^{\mathrm{o}}. \end{cases} \tag{11}$$

(*d*) *Test in a simplified two-dimensional variational formulation*

The possibility of recovering error variances with such a procedure is first tested in a simple problem similar to the previous case, but with the addition of the vertical

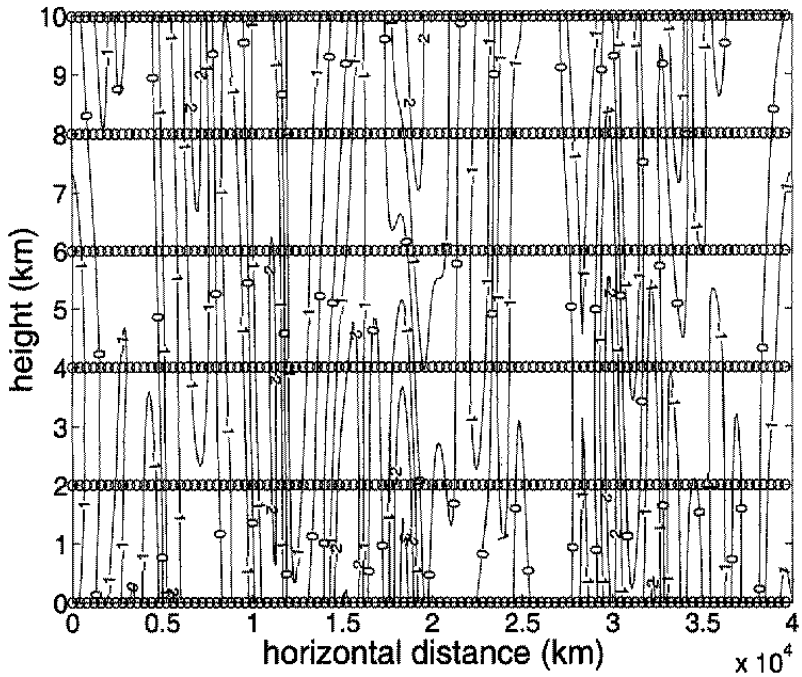

Figure 4. Two-dimensional simulated true signal $\mathbf{x}^{\mathrm{t}}$; the length-scale of the structure function is $\mathcal{L}_{\mathrm{h}} = 300$ km in the horizontal and $\mathcal{L}_{\mathrm{v}} = 1/200\ \mathcal{L}_{\mathrm{h}}$ in the vertical; isotherms are every 1 K; circles show the location of the observations. See text for further details.

coordinate. Such a framework allows the opportunity to specify different background- or observation-error variances at each level. The same spectral resolution as in the 1D case is used, but now with six vertical levels from 0 to 10 000 m. If m stands for the number of spectral coefficients in the horizontal and l is the number of levels, then $\mathbf{B}$ is defined as a block-diagonal matrix, each $l \times l$ block corresponding to the background correlations for a given horizontal wave number. Although such a formulation allows the full specification of non-separable correlation structure functions with a different vertical correlation for each wave number (Rabier *et al.* 1998), the correlations are simply assumed to be all the same here, and defined by a Gaussian function with length-scale $\mathcal{L}_{\mathrm{v}} = 1/200\ \mathcal{L}_{\mathrm{h}}$, where $\mathcal{L}_{\mathrm{h}}$ is the length-scale of the horizontal correlation (300 km as in the 1D case). As in the 1D case, a fictitious true signal $\mathbf{x}^{\mathrm{t}}$ is simulated by $\mathbf{x}^{\mathrm{t}} = -\boldsymbol{\epsilon}^{\mathrm{b}} = -\mathbf{B}^{1/2}\boldsymbol{\zeta}^{\mathrm{b}}$, where $\boldsymbol{\zeta}^{\mathrm{b}}$ is a vector of $m \times l$ random values with a Gaussian distribution.

• Firstly, the background-error weighting parameters, s_j^{b}, are assumed to be equal to 1 at each level. Figure 4 shows the field obtained with such a constant background-error variance. On the contrary, the vector of observations is simulated with $\mathbf{y}^{\mathrm{o}} = \mathrm{H}(\mathbf{x}^{\mathrm{t}}) + \mathbf{R}^{1/2}\boldsymbol{\zeta}^{\mathrm{o}}$, where $\boldsymbol{\zeta}^{\mathrm{o}}$ is a vector of random numbers with a Gaussian distribution, but with $\mathbf{R}$ containing values of the standard deviations of observation errors that increase with altitude: $(s_j^{\mathrm{o}},\ j = 1, l) = (1.0, 1.5, 2.0, 2.5, 3.0, 3.5)$ (here, the observations are placed at field levels for simplicity). The values of the background-error parameters, s_j^{b}, being kept constant, the optimal values of the observation-error parameters, s_j^{o}, are searched for by the same iterative fixed-point procedure as in the 1D case, but for all l levels simultaneously: $({s_{j,k+1}^{\mathrm{o}}}^2 = S_{j,k}^{\mathrm{o}},\ j = 1, l)$, where k stands for the iteration index. The results of such an iterative procedure are presented in Table 2: they show that a vertical profile of observation-error weighting parameters can be obtained after a few

TABLE 2. VALUES OF OBSERVATION-ERROR WEIGHTING PARAMETERS (s_j^o) AT DIFFERENT LEVELS ($j = 1, 6$) ALONG THE FIXED-POINT ITERATIONS

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
s_j^o it. 0	1.00	1.00	1.00	1.00	1.00	1.00
r.m.s.	0.64	0.56	0.56	0.68	0.91	1.17
s_j^o it. 1	1.02	1.49	1.89	2.13	2.72	2.85
2	0.97	1.52	1.97	2.14	2.67	3.03
3	0.96	1.52	1.97	2.14	2.67	3.04
4	0.95	1.52	1.98	2.14	2.66	3.04
5	0.95	1.52	1.98	2.14	2.66	3.04
r.m.s.	0.55	0.45	0.49	0.61	0.74	0.83
$(s^o)^s$	1.00	1.50	2.00	2.50	3.00	3.50
r.m.s.	0.56	0.46	0.49	0.61	0.74	0.83

The first row gives the values at the initial point of the iterative procedure (it. 0) (constant values in this case) and the second row shows the corresponding root mean square (r.m.s.) errors of the analyses performed with these values; the r.m.s. error at iteration 5 (it. 5) is also given; the last two rows give the reference values of s_j^o used for the simulation of observations and the corresponding r.m.s. errors obtained with the use of these values in the analyses.

TABLE 3.

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
s_j^b it. 0	1.00	1.00	1.00	1.00	1.00	1.00
r.m.s.	0.52	0.52	0.59	0.60	0.74	1.27
s_j^b it. 1	1.36	1.61	1.93	2.19	2.32	2.35
2	1.35	1.63	2.11	2.47	2.65	2.76
3	1.33	1.69	2.14	2.50	2.69	2.84
4	1.33	1.69	2.15	2.50	2.69	2.85
5	1.33	1.69	2.15	2.50	2.69	2.86
r.m.s.	0.47	0.45	0.41	0.41	0.45	0.68
$(s^b)^s$	1.00	1.50	2.00	2.50	3.00	3.50
r.m.s.	0.46	0.45	0.41	0.41	0.44	0.65

Same as Table 2, but for the background-error weighting parameters (s_j^b) (the values of s_j^o are kept constant during the iterations).

iterations only. Note that, as in the 1D case, each iteration requires two minimizations of the cost function J with unperturbed and perturbed observations.

• Secondly, the background-error weighting parameters, s_j^b, are now specified as varying in the vertical (s_j^b, $j = 1, l$) = (1.0, 1.5, 2.0, 2.5, 3.0, 3.5) in the background-error simulations and, on the contrary, the observation-error weighting parameters, s_j^o, are kept constant with a value of 1 in the simulation of observations. This time, the iterative fixed-point procedure is used for optimizing the s_j^b profile but the s_j^o profile is unchanged during the iterations: each iteration now requires two computations of the analysis with unperturbed and perturbed backgrounds. Table 3 shows that, again, this iterative method allows the recovery of a profile that is very close to the optimal one: although the values are not exactly the same as those used in the simulation of the background errors, the corresponding analysis errors are very close.

TABLE 4.

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
s_i^b it. 0	1.00	1.00	1.00	1.00	1.00	1.00
s_i^o it. 0	1.00	1.50	2.00	2.50	3.00	3.50
r.m.s.	0.50	0.64	0.85	1.20	1.85	2.63
s_i^b it. 1	1.18	1.33	1.52	1.67	1.69	1.62
s_i^o it. 1	0.90	1.17	1.37	1.39	2.06	2.48
s_i^b it. 2	1.26	1.53	1.90	2.19	2.25	2.17
s_i^o it. 2	0.89	1.06	1.07	0.96	1.27	1.76
s_i^b it. 3	1.30	1.63	2.07	2.43	2.53	2.51
s_i^o it. 3	0.89	1.03	0.98	0.88	1.00	1.32
s_i^b it. 4	1.32	1.67	2.14	2.51	2.64	2.70
s_i^o it. 4	0.89	1.02	0.95	0.88	0.95	1.10
s_i^b it. 5	1.32	1.69	2.17	2.53	2.68	2.80
s_i^o it. 5	0.89	1.02	0.93	0.88	0.94	1.00
r.m.s.	0.46	0.44	0.41	0.41	0.45	0.68
$(s^b)^s$	1.00	1.50	2.00	2.50	3.00	3.50
$(s^o)^s$	1.00	1.00	1.00	1.00	1.00	1.00
r.m.s.	0.46	0.45	0.41	0.41	0.44	0.65

Same as Table 2, but both s_j^b and s_j^o profiles are optimized along the iterations.

• Finally, the values of s_j^b and s_j^o used for background and observation errors are kept the same as in the previous case, but now a simultaneous iterative tuning of both sets of error parameters is performed: in this case each fixed-point iteration requires a pair of minimizations of J with unperturbed and perturbed observations for computing the $S_{j,k}^o$ expressions at iteration k, and another couple of minimizations with unperturbed and perturbed backgrounds to compute the $S_{j,k}^b$ expressions. Again, it can be checked (Table 4) that this iterative procedure allows the production of a very good approximation of the optimal values of the parameters, although the convergence is slightly slower here for the determination of the s_j^b weights only.

Again, these first good results are certainly related not only to the simplicity of the assumptions made regarding the form of the error covariances, but also to the vertical length-scale of the correlation that is short enough to leave uncoupled the information coming from the different levels: some tests have shown that the estimation becomes less accurate when this length-scale increases. This means that more work is needed to investigate the scope of validity of the procedure in a more realistic framework. Such a preliminary test is described in next section, but for the tuning of the s_j^o parameters only.

4. Application in a three-dimensional variational framework

A test of the parameter estimation procedure is performed here in the framework provided by the French Action de Recherche Petite Echelle Grande Echelle (ARPEGE) 3D-Var assimilation scheme. This scheme is based on an incremental formulation (Thépaut *et al.* 1998), with a global spectral model (Courtier *et al.* 1991). Although a particular feature of the ARPEGE system is the use of a stretched geometry that operationally provides a locally very high resolution for the trajectory, the spectral

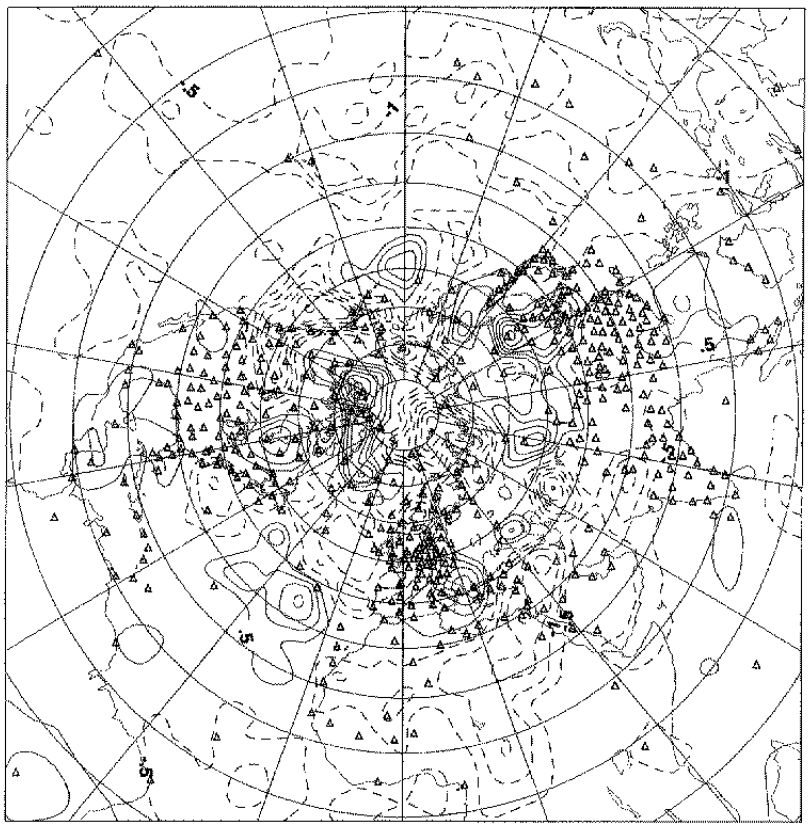

Figure 5. Field of temperature errors at 500 hPa produced by the procedure described in section 4; isotherms are every 0.5 K; triangles indicate the location of radio-sounding measurements.

resolutions for both the trajectory and the increments are set identical (T63 with no stretching), in order to simplify the framework of this test.

Once more, the idea is to simulate both background and observations corresponding to a given set of background and observation errors and then to recover the optimal specification of the variances in the analysis step, that should also give the best analysis. It is assumed that the true values $\mathbf{x}^t$ of the ARPEGE variables (temperature, vorticity, divergence, and specific humidity at 31 levels, and surface pressure) are given by a particular 6 h forecast (valid at 12 UTC 17 February 1997 here). Then, a vector $\boldsymbol{\epsilon}^b$ of forecast errors is obtained by applying the square root of the background-error covariance matrix $\mathbf{B}^{1/2}$ to a vector of random numbers (this can be done in practice by applying the change-of-variable operator, $L = \mathbf{B}^{1/2}$, introduced for pre-conditioning the problem as in section 3(b)). Given that the J^b term formulation is multivariate in the ARPEGE 3D-Var method, following the approach developed by Bouttier *et al.* (1997), this operation produces a vector $\boldsymbol{\epsilon}^b$ in which the temperature and the vorticity fields are, for example, correlated. Figure 5 presents the corresponding simulated background-error field for the temperature at 500 hPa. The simulated observations are given by a set of radio-sounding measurements with true locations (see Fig. 5): the observation errors are produced by using the operational error variances for geopotential, temperature, wind, and specific humidity (relative-humidity measurements are converted to specific-humidity observations before being used in the ARPEGE 3D-Var model). The variation of the standard deviation for these parameters is given by the full lines in Figs. 6 to 9. If $\boldsymbol{\epsilon}^o$ stands for the vector of observation errors, then the observations $\mathbf{y}^o$ are produced

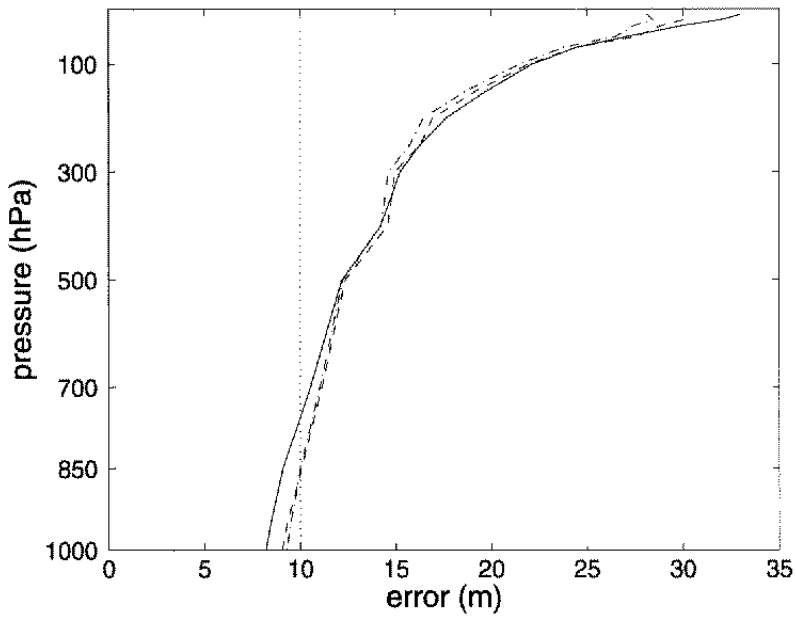

Figure 6. Geopotential-height error profiles used for the simulation of observations (full line), imposed at the beginning of the iterative procedure (dotted line), after 1 iteration (dash–dot line) and after five iterations (dashed line).

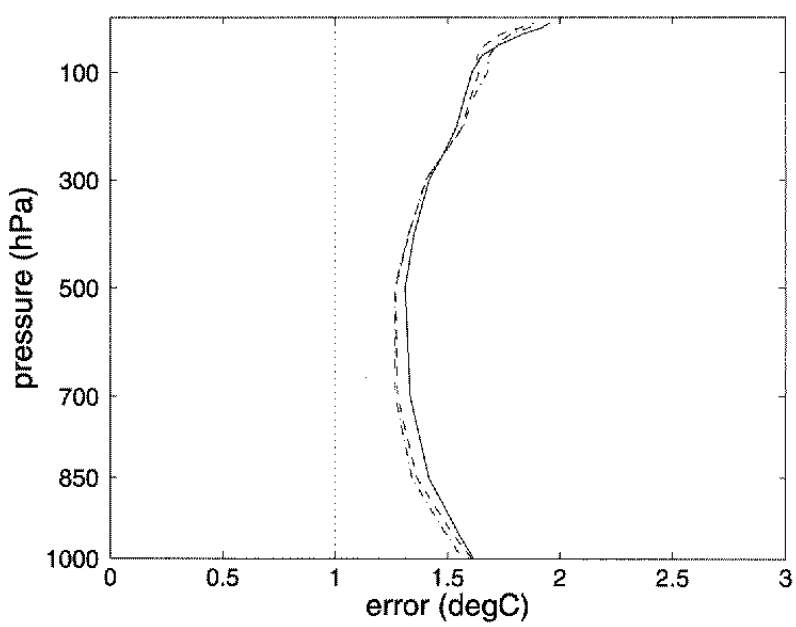

Figure 7. Same as Fig. 6 but for temperature.

with $\mathbf{y}^o = \mathrm{H}(\mathbf{x}^t) + \boldsymbol{\epsilon}^o$, with $\mathbf{x}^t$ the fictitious true representation of the atmospheric flow given by the ARPEGE forecast, and H is the observation operator associated with the radio-sounding observations.

The procedure described in previous sections was implemented in the ARPEGE 3D-Var model to optimize the values of the error variances associated with radiosondes. In this test, no simultaneous search of background-error variances is performed: their values are kept constant along the iterations, and set equal to the optimal values used

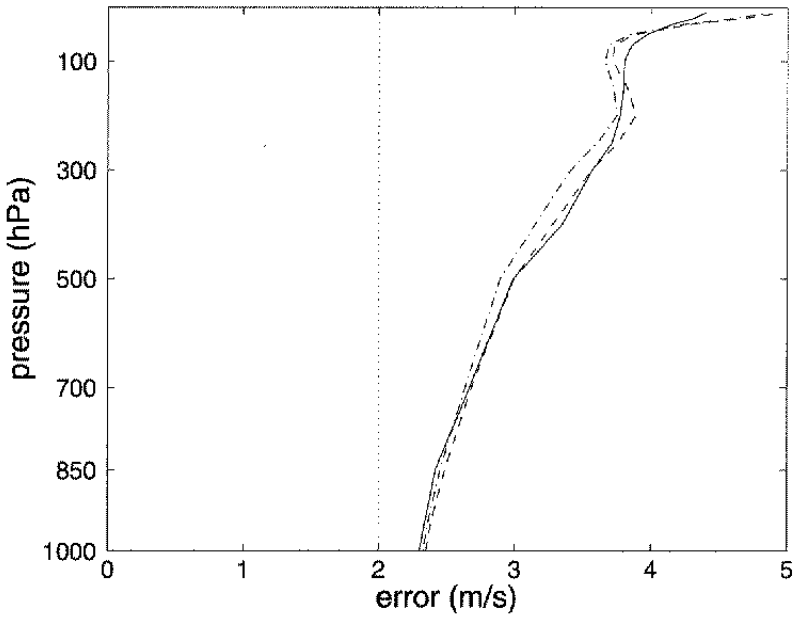

Figure 8. Same as Fig. 6 but for wind speed.

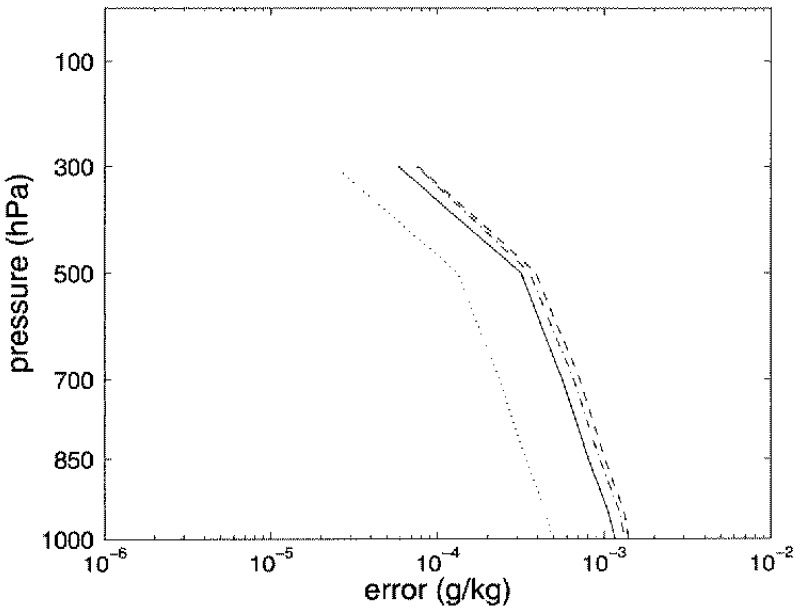

Figure 9. Same as Fig. 6 but for specific humidity.

for the background errors. Thus, each iteration of the tuning procedure requires a pair of 3D-Var analyses with unperturbed and perturbed observations. At each iteration, the r.m.s. error between the analysis obtained with the corresponding set of parameters and the fictitious truth is computed for the same variables as in the measurements: in order to identify the impact of tuning the parameters more clearly, this r.m.s. error is shown for a sub-domain roughly corresponding to North America, where the radio-sounding coverage is the most homogeneous over the globe (see Fig. 5). The initial vertical profiles of observation-error parameters (dotted lines in Figs. 6 to 9) were

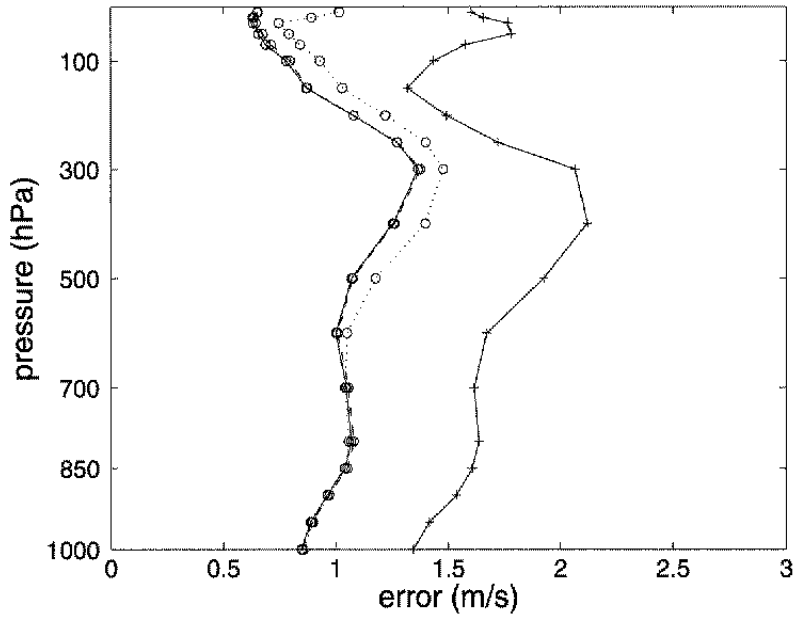

Figure 10. Wind-speed root mean square (r.m.s.) errors over a domain covering North America for: the background (full line and crosses), the analysis with wrong initial error profile (dotted line and circles), the analysis with true error profile (full line and circles), and the analysis with optimized profile at iteration five (dashed line and circles); the last two curves are nearly superimposed.

deliberately chosen to be very different from the true ones: they are constant in the vertical for geopotential, temperature and wind speed; the initial vertical profile for specific humidity is obtained by imposing a constant profile for relative-humidity and temperature errors and, then, going to the specific-humidity errors (that explains why the initial profile is not constant in the vertical for specific humidity).

Figures 6 to 9 show that the convergence of the iterative algorithm is extremely fast: the profiles obtained after a single iteration (dash–dot lines) are already very close to those obtained at iteration 5 (dashed lines), although this convergence is slightly slower for the wind speed. Moreover, the figures also show that these optimized profiles are very similar to the operational ones used for the simulation of observations (full lines). Finally, it can be verified that the optimized profiles give an r.m.s. error that almost shows no difference relative to the r.m.s. error resulting from an analysis obtained with the reference parameters (see Fig. 10 for the wind speed r.m.s. error; r.m.s. errors for other parameters are not presented here but show the same behaviour).

5. Conclusion

In this paper it was shown that following the a posteriori *diagnosis* approach proposed by some authors, a practical computation of the expectation of sub-parts of the value of a cost function at the minimum can be achieved by using a randomization technique based on a perturbation of either observations or background fields. Furthermore, these computations allow the tuning of observation-error parameters by applying a simple iterative fixed-point procedure. This procedure was successfully tested in two simplified cases: specifically, the possibility of tuning both observation- and background-error parameters was demonstrated. The link with GCV in the simple problem of determining the ratio between observation and background terms was also pointed out. Finally, a test

in the ARPEGE 3D-Var framework showed that a complete description of observation-error parameters can be retrieved with only a few iterations and, thus, at a reasonable cost (this cost is of the order of some 3D-Var analyses). Although feasible, the determination of background-error parameters in the ARPEGE 3D-Var model was not attempted here, since it is more difficult to implement.

Such promising results with simulated observations and background fields suggest that the procedure should now be applied, in a second step, to real observation sets (however, more tests with such an idealized problem are also planned in order to have a better idea of the accuracy and the sensitivity of the parameter estimates to the number of observations and the background-error covariances). The computation of partial diagnoses for subsets of observations could be, for example, applied occasionally to verify or to adjust the specification of observation errors for particular subsets, as was done here for radio soundings. In particular, the determination of satellite observation-error statistics, such as radiance errors, is of great interest since these errors, including the part associated with the observation operator, are not perfectly known. Another point of interest is the capability of using such a method for tuning the level of representativeness error due to the resolution of the model or of the analysis increment, or to the observation operator, as just mentioned for the use of radiances.

Because background-error covariances should evolve in time, as the observation error statistics are less dependent on the temporal dimension, there should still be a greater advantage to using the proposed approach for performing an adaptive determination of background errors with, for example, a different specification for predefined sub-domains. This was shown to be feasible in the simple two-dimensional variational analysis problem, and the application to the 3D- or 4D-Var ARPEGE analysis scheme will be addressed in future work.

Finally, although this study focused on the optimization of the variance of observation errors, the possibility of adapting the procedure to the determination of correlation-error parameters, such as observation- or background-error length-scales, will also be investigated.

Acknowledgements

The authors would like to thank Thierry Bergot, Gwenaëlle Hello, Alain Joly, Philippe Lopez and Florence Rabier for helpful discussion and comments on this paper. We also wish to thank the two referees for their constructive comments on the first version of this paper.

Appendix A

Proof of Eq. (2)

For any part of J, the following sequence of equalities can be written:

$$\begin{aligned}
\mathrm{E}(\mathrm{J}_j) &= \tfrac{1}{2}\mathrm{E}\{(\Gamma_j\mathbf{x}^{\mathrm{a}} - \mathbf{y}_j^{\mathrm{o}})^{\mathrm{T}}\mathbf{R}_j^{-1}(\Gamma_j\mathbf{x}^{\mathrm{a}} - \mathbf{y}_j^{\mathrm{o}})\} \\
&= \tfrac{1}{2}\mathrm{E}\{\mathrm{Tr}(\mathbf{R}_j^{-1}(\Gamma_j\mathbf{x}^{\mathrm{a}} - \mathbf{y}_j^{\mathrm{o}})(\Gamma_j\mathbf{x}^{\mathrm{a}} - \mathbf{y}_j^{\mathrm{o}})^{\mathrm{T}})\} \\
&= \tfrac{1}{2}\mathrm{E}\{\mathrm{Tr}(\mathbf{R}_j^{-1}(\Gamma_j\boldsymbol{\epsilon}^{\mathrm{a}} - \boldsymbol{\epsilon}_j^{\mathrm{o}})(\Gamma_j\boldsymbol{\epsilon}^{\mathrm{a}} - \boldsymbol{\epsilon}_j^{\mathrm{o}})^{\mathrm{T}})\} \\
&= \tfrac{1}{2}\mathrm{Tr}(\mathbf{R}_j^{-1}[\Gamma_j\mathrm{E}(\boldsymbol{\epsilon}^{\mathrm{a}}\boldsymbol{\epsilon}_j^{\mathrm{o}\mathrm{T}})\Gamma_j^{\mathrm{T}} + \mathrm{E}(\boldsymbol{\epsilon}_j^{\mathrm{o}}\boldsymbol{\epsilon}_j^{\mathrm{o}})^{\mathrm{T}} - \mathrm{E}\{(\Gamma_j\mathbf{x}^{\mathrm{a}})\boldsymbol{\epsilon}_j^{\mathrm{o}\mathrm{T}}\} - \mathrm{E}\{\boldsymbol{\epsilon}_j^{\mathrm{o}}(\Gamma_j\mathbf{x}^{\mathrm{a}})^{\mathrm{T}}\}]),
\end{aligned}$$

since $\Gamma_j\mathbf{x}^{\mathrm{a}} - \mathbf{y}_j^{\mathrm{o}} = \Gamma_j\mathbf{x}^{\mathrm{a}} + \Gamma_j\mathbf{x}^{\mathrm{T}} - \Gamma_j\mathbf{x}^{\mathrm{T}} - \mathbf{y}_j^{\mathrm{o}} = \Gamma_j\boldsymbol{\epsilon}^{\mathrm{a}} - \boldsymbol{\epsilon}_j^{\mathrm{o}}$, and using the fact that $\mathrm{E}\{\mathrm{Tr}(.)\} = \mathrm{Tr}\{\mathrm{E}(.)\}$ and the linearity properties of the expectation operator E.

On the other hand, Eq. (1) yields $\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}} = \mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}} + \mathbf{K}\{\mathbf{y}^{\mathrm{o}} - \mathrm{H}(\mathbf{x}^{\mathrm{b}})\}$ or $\boldsymbol{\epsilon}^{\mathrm{a}} = \boldsymbol{\epsilon}^{\mathrm{b}} + \mathbf{K}\{\mathrm{H}(\mathbf{x}^{\mathrm{t}}) + \boldsymbol{\epsilon}^{\mathrm{o}} - \mathrm{H}(\mathbf{x}^{\mathrm{b}})\}$. Then $\mathrm{E}(\boldsymbol{\epsilon}^{\mathrm{a}}\boldsymbol{\epsilon}^{\mathrm{o}\mathrm{T}}) = \mathbf{K}\mathrm{E}(\boldsymbol{\epsilon}^{\mathrm{o}}\boldsymbol{\epsilon}^{\mathrm{o}\mathrm{T}}) = \mathbf{P}^{\mathrm{a}}\mathrm{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{R} = \mathbf{P}^{\mathrm{a}}\mathrm{H}^{\mathrm{T}}$. This last equality gives $\mathrm{E}\{(\Gamma_j\mathbf{x}^{\mathrm{a}})\boldsymbol{\epsilon}_j^{\mathrm{o}\mathrm{T}}\} = \Gamma_j\mathrm{E}(\mathbf{x}^{\mathrm{a}}\boldsymbol{\epsilon}^{\mathrm{o}\mathrm{T}})\mathrm{P}_j^{\mathrm{T}} = \Gamma_j\mathbf{P}^{\mathrm{a}}\mathrm{H}^{\mathrm{T}}\mathrm{P}_j^{\mathrm{T}} = \Gamma_j\mathbf{P}^{\mathrm{a}}\Gamma_j^{\mathrm{T}}$, where P_j is the projection operator such that $\boldsymbol{\epsilon}_j^{\mathrm{o}} = \mathrm{P}_j\boldsymbol{\epsilon}^{\mathrm{o}}$ (as also introduced in section 3(c)). It is also easy to check that $\mathrm{Tr}[\mathrm{E}\{\boldsymbol{\epsilon}_j^{\mathrm{o}}(\Gamma_j\mathbf{x}^{\mathrm{a}})^{\mathrm{T}}\}] = \mathrm{Tr}[\mathrm{E}\{(\Gamma_j\mathbf{x}^{\mathrm{a}})\boldsymbol{\epsilon}_j^{\mathrm{o}\mathrm{T}}\}] = \mathrm{Tr}(\Gamma_j\mathbf{P}^{\mathrm{a}}\Gamma_j^{\mathrm{T}})$. Since $\Gamma_j\mathrm{E}(\boldsymbol{\epsilon}^{\mathrm{a}}\boldsymbol{\epsilon}_j^{\mathrm{o}\mathrm{T}})\Gamma_j^{\mathrm{T}} = \Gamma_j\mathbf{P}^{\mathrm{a}}\boldsymbol{\Gamma}_j{}^{\mathrm{T}}$ and $\mathrm{E}(\boldsymbol{\epsilon}_j^{\mathrm{o}}\boldsymbol{\epsilon}_j^{\mathrm{o}})^{\mathrm{T}} = \mathbf{R}_j$, all those relations yield

$$\begin{aligned}\mathrm{E}(\mathrm{J}_j) &= \tfrac{1}{2}\,\mathrm{Tr}\{\mathbf{R}_j^{-1}(\mathbf{R}_j - \Gamma_j\mathbf{P}^{\mathrm{a}}\Gamma_j^{\mathrm{T}})\}\\ &= \tfrac{1}{2}\{m_j - \mathrm{Tr}(\mathbf{R}_j^{-1}\Gamma_j\mathbf{P}^{\mathrm{a}}\Gamma_j^{\mathrm{T}})\}\\ &= \tfrac{1}{2}\{m_j - \mathrm{Tr}(\Gamma_j^{\mathrm{T}}\mathbf{R}_j^{-1}\Gamma_j\mathbf{P}^{\mathrm{a}})\}.\end{aligned}$$

APPENDIX B

Relationship with Generalized Cross Validation minimum

Equation (5) in section 3(a) can also be written

$$\begin{aligned}V(\lambda) &= \frac{\mathrm{Tr}\{(\mathbf{I}_p - \mathrm{H}\mathbf{K})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{I}_p - \mathrm{H}\mathbf{K})\mathbf{d}\mathbf{d}^{\mathrm{T}}\}}{\{\mathrm{Tr}(\mathbf{I}_p - \mathrm{H}\mathbf{K})\}^2}\\ &= \frac{\mathrm{Tr}\{\mathbf{R}^{-1}(\mathbf{I}_p - \mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}}(\mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}} + \lambda\mathbf{R})^{-1})^2\mathbf{d}\mathbf{d}^{\mathrm{T}}\}}{\{\mathrm{Tr}(\mathbf{I}_p - \mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}}(\mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}} + \lambda\mathbf{R})^{-1})\}^2}\\ &= \frac{\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-2}\mathbf{d}\mathbf{d}^{\mathrm{T}})}{\{\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-1})\}^2},\end{aligned}$$

with $\mathbf{D} = \mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}} + \lambda\mathbf{R}$ and $\lambda = s^{\mathrm{o}2}/s^{\mathrm{b}2}$ as defined in section 3(a).

Computing the derivative of this expression with respect to the parameter λ, the numerator $\mathrm{N}(V_\lambda')$ of this derivative is

$$\mathrm{N}(V_\lambda') = (-2\mathbf{R}^{\mathrm{T}}\mathbf{R}\mathbf{D}^{-3}\mathbf{d}\mathbf{d}^{\mathrm{T}})\{\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-1})\}^2 - \{\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-2}\mathbf{d}\mathbf{d}^{\mathrm{T}})\}\{-2\,\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-1})\mathbf{R}^{\mathrm{T}}\mathbf{R}\mathbf{D}^{-2}\}.$$

This last expression vanishes if $\mathbf{d}\mathbf{d}^{\mathrm{T}} = \mathbf{D} = \mathrm{H}\mathbf{B}\mathrm{H}^{\mathrm{T}} + \lambda\mathbf{R}$, that is to say, if the value of λ agrees with the statistics $\mathbf{d}\mathbf{d}^{\mathrm{T}}$ of innovations computed with a single batch of observations and background.

On the other hand, the expression S^{o} also defined in section 3(a) can be rewritten

$$S^{\mathrm{o}} = \frac{\mathrm{Tr}(\lambda^2\mathbf{R}\mathbf{D}^{-2}\mathbf{d}\mathbf{d}^{\mathrm{T}})}{\mathrm{Tr}(\lambda\mathbf{R}\mathbf{D}^{-1})} = \lambda\frac{\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-2}\mathbf{d}\mathbf{d}^{\mathrm{T}})}{\mathrm{Tr}(\mathbf{R}\mathbf{D}^{-1})},$$

so that $S^{\mathrm{o}} = \lambda = s^{\mathrm{o}2}/s^{\mathrm{b}2}$ if $\mathbf{d}\mathbf{d}^{\mathrm{T}} = \mathbf{D}$.

Thus, searching for λ, such that $\lambda = S^{\mathrm{o}}(\lambda)$ or as the minimum of $V(\lambda)$, are two equivalent approaches in this simple case where one simply aims to optimize the ratio between background and observation terms.

REFERENCES

Bennett, A. F., Leslie, L. M., Hagelberg, C. R. and Powers, P. E. 1993 Tropical cyclone prediction using a barotropic model initialized by a generalized inverse method. *Mon. Weather Rev.*, **121**, 1714–1729

Bouttier, F., Derber J. and Fisher, M. 1997 'The 1997 revision of the J_b term in 3D/4D-Var'. Technical Report 192, ECMWF, Reading, UK

Courtier, P., Freydier, C., Geleyn, J.-F., Rabier, F. and Rochas, M. 1991 'The Arpege project at Météo-France'. Pp. 193–231 in Proceedings of the ECMWF workshop on numerical methods in atmospheric models, 9–13 September 1991, Reading, UK

Courtier, P. and Talagrand, O. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Numerical results. *Q. J. R. Meteorol. Soc.*, **113,** 1329–1347

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994 A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120,** 1367–1387

Dee, D. P. 1995 On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Weather Rev.*, **123,** 1128–1145

Fisher, M. 1999 'Implementing GCV in the IFS'. Technical Report, ECMWF, Reading, UK

Gilbert, J. Ch. and Lemaréchal, C. 1989 Some numerical experiments with variable storage Quasi-Newton algorithms. *Mathematical Programming*, **B25,** 407–435

Girard, D. 1987 'A fast Monte Carlo cross-validation procedure for large least squares problems with noisy data'. Technical Report 687-M, IMAG, Grenoble, France

1991 Asymptotic optimality of the fast randomized versions of GCV and C_L in ridge regression and regularization. *Ann. Stat.*, **19,** 1950–1963

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

Lewis, J. M. and Derber, J. C. 1985 The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A,** 309–322

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

Lorenc, A. C. 1986 Analysis methods for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **112,** 1177–1194

Rabier, F., McNally, A., Anderson, E., Courtier, P., Undén, P., Eyre, J., Hollingsworth, A. and Bouttier, F. 1998 The ECMWF implementation of three-dimensional variational assimilation (3D-Var). II: Structure functions. *Q. J. R. Meteorol. Soc.*, **124,** 1809–1829

Talagrand, O. 1997 Assimilation of observations, an introduction. *J. Meteorol. Soc. Jpn.*, **75,** 191–209

1999 'A posteriori verification of analysis and assimilation algorithms'. Pp. 17–28 in Proceedings of the ECMWF Workshop on Diagnosis of data assimilation systems, 2–4 November 1999, Reading, UK

Thépaut, J.-N., Alary, P., Caille, P., Cassé, V., Geleyn, J.-F., Moll, P., Pailleux, J., Piriou, J.-M., Puech, D. and Taillefer, F. 1998 'The operational global data assimilation system at Météo-France'. Pp. 25–31 in Proceedings of HIRLAM 4 workshop on variational analysis in limited area models, Toulouse, France

Wahba, G., Johnson, D. R., Gao, F. and Gong, J. 1995 Adaptative tuning of numerical weather prediction models: Randomized GCV in three- and four-dimensional data assimilation. *Mon. Weather Rev.*, **123,** 3358–3369

Wahba, G. and Wendelberger, J. 1980 Some new mathematical methods for variational objective analysis using splines and cross validation. *Mon. Weather Rev.*, **108,** 1122–1143

Q. J. R. Meteorol. Soc. (2003), **129**, pp. 1301–1315 doi: 10.1256/qj.01.182

A 4D-Var re-analysis of FASTEX

By GÉRALD DESROZIERS*, GWENAÉLLE HELLO and JEAN-NOËL THÉPAUT
Météo-France, France

(Received 29 October 2001; revised 3 July 2002)

SUMMARY

A complete re-analysis of the Fronts and Atlantic Storm-Track EXperiment (5 January–27 February 1997) is performed. It provides a reference description of the life cycle of an ensemble of cyclogeneses over the north Atlantic. Many aspects make this re-analysis noteworthy. In particular, the FASTEX period is characterized by an exceptional increase in the number of dropsoundings or radiosoundings concentrated on the storm track. Another striking point of the re-analysis is the use of a four-dimensional variational (4D-Var) assimilation scheme on such a long period associated with the French Arpège (Action des Recherche Petite Echelle et Grand Echelle) global stretched model, that allows a locally high resolution in the area where dense networks of dropsondes were deployed. A twin re-analysis with the same characteristics but without any FASTEX data is also performed. It shows that the impact of the special FASTEX data is quite noticeable in the resulting analyses and is lower, but still perceptible, in the associated short-range forecasts (6 h–36 h). The impact of FASTEX observations on a particular FASTEX case is also shown.

KEYWORDS: Variational assimilation

1. INTRODUCTION

The field phase of the Fronts and Atlantic Storm-Track EXperiment (FASTEX) took place from 5 January to 27 February 1997 with 19 Intensive Observing Periods (IOPs). It produced a unique dataset of observations of a number of midlatitude cyclones (Joly *et al.* 1999). The main objectives of FASTEX (Joly *et al.* 1997) were to verify or to complete cyclogenesis theories, but also to address the predictability problem of such cyclones and to improve their operational forecast, which is still a difficult and important task. The cases of the Christmas 1999 cyclones which had such devastating impacts over western and central Europe are more recent examples of such phenomena and of the difficulty of forecasting them at short range.

In particular, FASTEX was the first opportunity to test the concept of adaptive observations concentrated in areas where small analysis errors will amplify rapidly. These observations were provided during FASTEX by dropsondes launched by aircraft. Other dropsonde samplings were also deployed in the FASTEX upstream or downstream areas in order to observe cyclones from genesis to maturity. A large number of additional vertical profiles were also provided by land-based sounding sites or ships.

The impact of FASTEX adaptive or 'targeted' observations on forecasts has been investigated by Bergot (1999), Montani *et al.* (1999) and Szunyogh *et al.* (1999). Amstrup and Huang (1999) have also investigated the impact of all additional FASTEX radiosonde observations. These studies were based on the use of measurements through analysis schemes associated with research or operational weather-prediction systems.

The optimal way to fulfil the FASTEX scientific objective of improving the description of the life cycle of the cyclones observed during the experiment was also clearly identified as performing a complete re-analysis of the FASTEX period with such an assimilation scheme.

A pilot study performed by Desroziers *et al.* (1999) on a single IOP (IOP 17) showed that the quality of the re-analyses was rather dependent on the assimilation

* Corresponding author: Météo-France, CNRM/GMME/RECYF, 42 av. G. Coriolis, F-31057 Toulouse Cedex 1, France. e-mail: gerald.desroziers@meteo.fr
© Royal Meteorological Society, 2003.

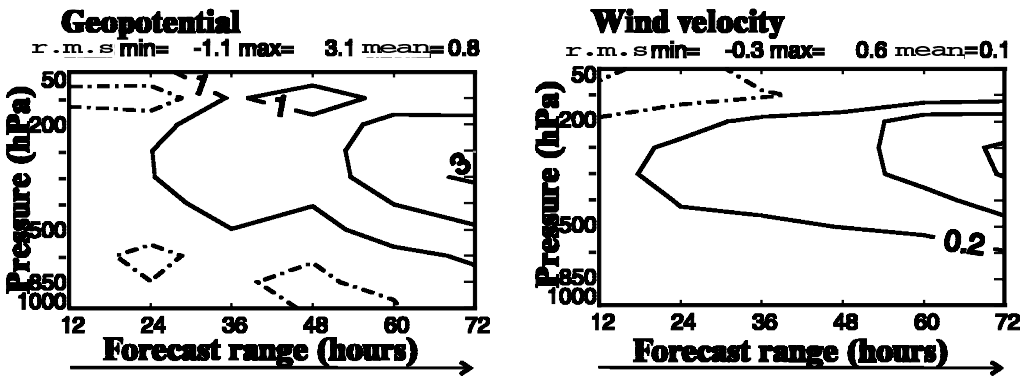

Figure 1. Difference in root mean square (r.m.s) errors in terms of geopotential and wind velocity between the four-dimensional variational (4D-Var) and three-dimensional variational re-analyses with Fronts and Atlantic Storm-Track EXperiment (FASTEX) observations for the period from 3 January to 28 February 1997. Verification is done against ECMWF analyses and for a domain centred over Europe (20.25°N–71.25°N, 32°W–36°E). Positive and zero impacts of 4D-Var are plotted with full and dashed–dotted lines, respectively. Values are every 1 m for geopotential and 0.2 m s^{-1} for wind velocity.

scheme itself and that a four-dimensional variational (4D-Var) assimilation scheme significantly improved the use of the FASTEX data which were especially characterized by their irregular distribution in time. These results led to the project to perform a complete re-analysis of the FASTEX period with the 4D-Var assimilation scheme developed at Météo-France. However, the re-analysis was only performed after recovering the validated and corrected form of the FASTEX soundings that were made available in the FASTEX data base (Jaubert *et al.* 1999), and when a version of the pre-operational 4D-Var assimilation scheme was made available.

4D-Var has been operational at Météo-France since June 2000 following ECMWF's implementation of a similar assimilation scheme in November 1997. The implementation of such a sophisticated scheme was made possible at Météo-France by applying a multi-incremental formalism. Thus, the FASTEX re-analysis presented in this paper is the first re-analysis of such a long period in a 4D-Var mode.

The use of 4D-Var has made possible the assimilation of all the FASTEX soundings, as only a part of them were assimilated in a parallel 3D-Var re-analysis that was also performed as a test of the improvement brought by 4D-Var. The comparison between the scores of the forecasts issued from FASTEX 4D-Var and 3D-Var re-analyses (Fig. 1) was an additional argument to make 4D-Var operational in June 2000. To measure the weights of the FASTEX data in the re-analyses it was also decided to perform an additional complete 4D-Var re-analysis with no FASTEX observations.

The aim of this paper is to present the results obtained from these two parallel suites with and without FASTEX observations. A description of the two re-analyses is given in section 2. The impact of FASTEX data on analyses and forecasts is discussed in section 3, and an example of the improvement brought by FASTEX observations is presented in section 4. Conclusions derived from the comparison between the two re-analyses are formulated in section 5.

2. Description of the FASTEX re-analysis

(*a*) *Configuration of the re-analysis*

The assimilation scheme used for the FASTEX re-analysis is based on a variational formulation. This formulation, which is also related to the general framework of estimation theory, especially allows the use of observations that are not directly linked to the atmospheric fields described by the forecast model. In particular, such a formulation permits a better assimilation of satellite data. A presentation of the variational approach can be found, for example, in Courtier (1997) and a description of the 4D-Var developed at Météo-France in Janisková *et al.* (1999) and Geleyn *et al.* (2001). This formulation is based on the minimization of a cost function that measures the distance between the evolution of the state variable during the assimilation period, given by the forecast model, and the observations. This cost function also includes a term that measures the distance between the state variable at the beginning of the period and a priori or 'background' information given by a very short-range forecast. On the other hand, the 4D-Var formulation developed at Météo-France includes an innovative and efficient additional constraint term based on digital filtering to damp spurious oscillations associated with gravity-inertia waves (Gauthier and Thépaut 2001). The forecast model is the French Arpège (Action des Recherche Petite Echelle et Grand Echelle) operational model (Courtier *et al.* 1991), a spectral global model based on a stretched geometry. The re-analysis was performed with a T199 spectral truncation on the stretched sphere and a stretching factor of 3.5, roughly corresponding to a 90 km resolution near Newfoundland, in the FASTEX upstream area, and a 30 km resolution near Ireland in the FASTEX downstream area.

To reduce the cost of the 4D-Var scheme, requiring about 70 direct integrations of the model and of its adjoint, the formalism applied for the FASTEX re-analysis is based on a so-called 'multi-incremental' formulation (Veersé and Thépaut 1998) where the resolution of the correction or 'increment' to the initial background is progressively increased, with the final corrections performed at a uniform T95 spectral resolution. The re-analysis was performed from 0000 UTC 1 January 1997 until 1800 UTC 28 February 1997 with a 6-hour assimilation period.

(*b*) *Observations*

As mentioned above, two 4D-Var re-analyses of the whole FASTEX period were made: a reference analysis (hereafter OPER) with only the operational set of observations and the proper FASTEX re-analysis (hereafter FULL), with the addition of FASTEX observations. The operational observations include all conventional data (SYNOP, AIREP, SATOB, TEMP, PILOT) but also Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) cloud-cleared radiances whose direct use is allowed by the variational formalism. The set of FASTEX observations used in FULL correspond to the SYNOP-SHIP data given by the FASTEX ships, the data issued from the supplementary buoys deployed during the experiment, and the TEMP profiles provided by dropsondes launched by the FASTEX aircraft or by additional radiosoundings performed by land-based upper-air stations, by the four FASTEX ships or by Automated Shipboard Aerological Programme (ASAP) ships. Figure 2 shows the locations of the set of FASTEX soundings performed during the experiment. Note that the FASTEX ships were not stationary during the experiment but moved in order to stay close to the main baroclinic zone. The advantage of such a procedure was confirmed by an idealized observing-system simulation study (Fischer *et al.* 1998). All the FASTEX

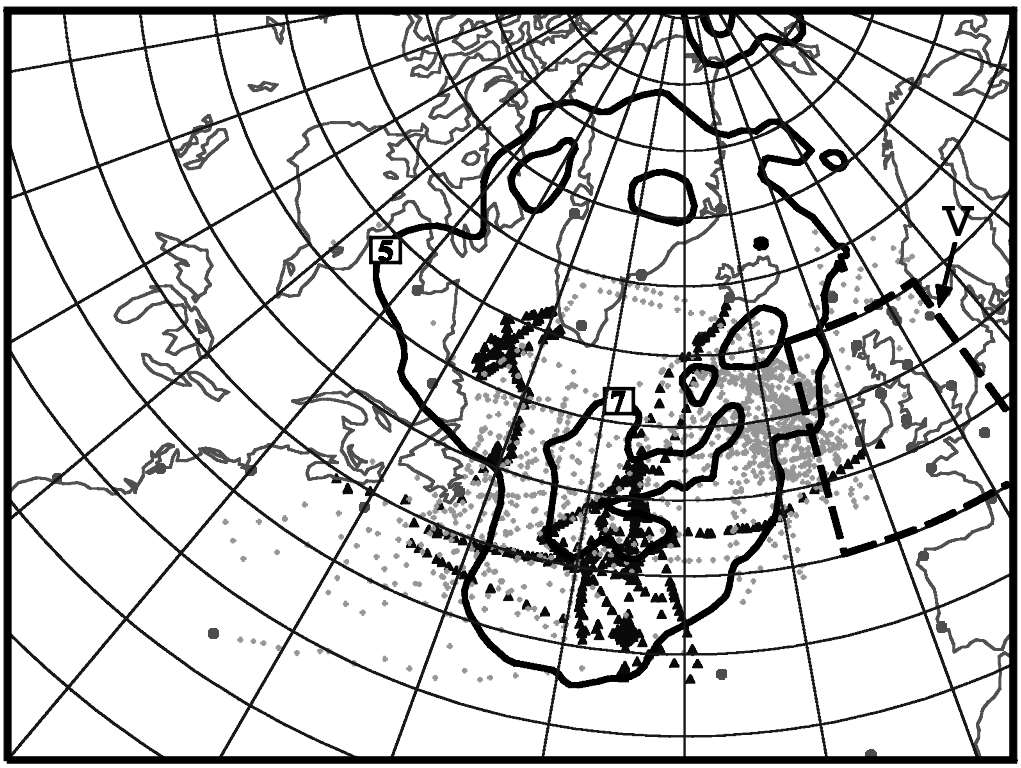

Figure 2. Standard deviation of the difference OPER−FULL (see text) for geopotential at 500 hPa. Statistics are computed over the two months of the field phase. Isolines 5 m and 7 m are plotted. Additional Fronts and Atlantic Storm-Track EXperiment (FASTEX) soundings are superimposed: ships (dark triangles), dropsondes (light crosses) and land-based stations (dark bullets). The verification domain V (see section 3(b)) is indicated.

soundings assimilated in FULL were taken from the FASTEX data archive in their corrected high-resolution form and converted back to the conventional TEMP format in order to be compatible with the resolution of the final analysis. All these profiles were checked very carefully before the re-analysis (additional quality control integrated in the assimilation scheme was also performed during the course of the re-analysis).

3. Statistical impact of FASTEX observations

(*a*) *Difference between the two analyses*

Figure 2 presents the standard deviation of the difference between the two re-analyses for the 500 hPa geopotential. This gives an indication of the expected reduction of the analysis error due to FASTEX observations although, from a strict point of view, these two quantities are different. This difference can also be compared with the TEMP observation error for the same parameter and level. The same order of difference is observed for other parameters and levels (not shown here). Similar differences are obtained for 0600 UTC and 1800 UTC (not shown) but with an extension to the east coast of Canada and the USA due to the addition of radiosoundings for those times and areas. It is also remarkable that the structure of the standard-deviation field is positioned along the maximum density of cyclone tracks during FASTEX (see Fig. 3 in Baehr *et al.* 1999).

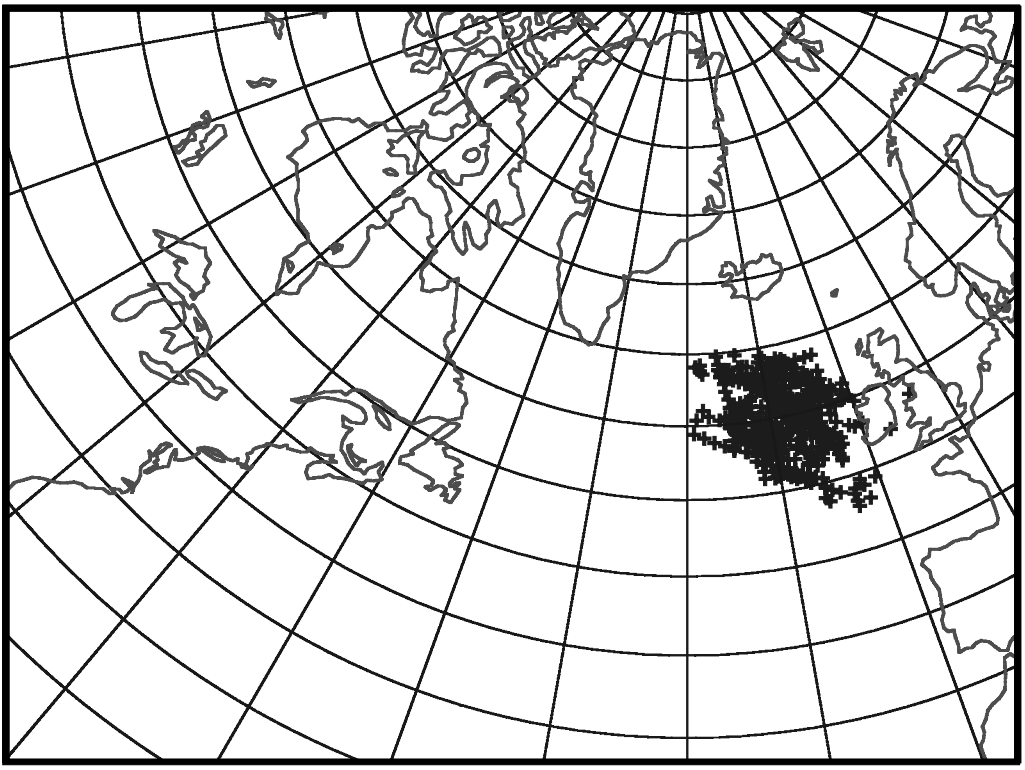

Figure 3. Positions of the dropsondes launched by the UK C130.

(*b*) *Distance to observations*

The variable resolution of the Arpège model is rather well adapted to the design of the FASTEX experiment: the corresponding resolution is sufficient ($\simeq$90 km) to describe the rather large-scale cyclone precursors in the upstream zone near Newfoundland and also high enough to describe mesoscale structures of mature cyclones in the downstream area close to Ireland ($\simeq$30 km). However, it should be noted that the analysis is not performed at this full resolution. The UK C130 aircraft performed a set of flights providing arrays of dropsondes with a typical separation close to the Arpège resolution in this downstream area (see Fig. 3). In this section the ability of the FULL re-analysis to adjust the UK C130 dropsonde data is examined and compared with the OPER reference analysis. In a second step, forecasts run from FULL and OPER analyses are also compared.

Root mean square (r.m.s.) differences between analyses and dropsonde observations were computed over ten UK C130 flights performed during the experiment and, respectively, corresponding to systems observed during IOPs 2, 7, 10, 11, 12, 15, 16, 17 and 19. A description of these flights can be, for example, found in Joly *et al.* (1999). Because some of the flights had a long duration, their observations were assimilated in two adjacent 6-hour assimilation periods. This is the case for IOPs 2, 16 and 19. It clearly appears that FULL is much closer to UK C130 dropsonde observations than OPER (Fig. 4). This indicates that these observations were correctly assimilated by the 4D-Var scheme. The reduction of the departure between analyses and temperature or wind observations is rather homogeneous within the vertical. Its amplitude is about

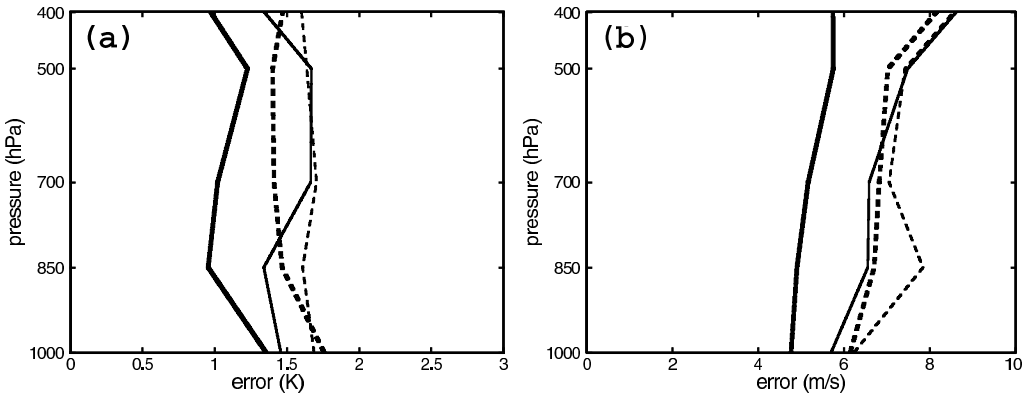

Figure 4. (a) Root mean square (r.m.s.) departures between analyses (bold lines) or backgrounds (thin lines) and temperature observations given by UK C130 dropsondes; (b) same as (a) but for wind velocity. In each panel, full lines correspond to FULL re-analyses and dashed lines to OPER re-analyses (see text).

30% which is quite substantial. The better fit of FULL to observations is in fact already present in the corresponding background fields (thin lines in Fig. 4), showing that there is a cumulative impact in time of the FASTEX observations in the assimilation cycle. A synthetic criterion to evaluate the impact of FASTEX observations is given by the computation of the mean square departure between both analyses and UK C130 observations of temperature, relative humidity and wind velocity, normalized by their error variances. This criterion corresponds to the evaluation of the normalized observation cost function minimized in the variational assimilation scheme. The normalized observation cost function is always lower for FULL than for OPER (Fig. 5(a)) and the mean reduction over all flights is about 30%, in agreement with the reduction seen on vertical profiles (Fig. 4). The same cost function can also be evaluated for the background fields corresponding, respectively, to FULL and OPER. Figure 5(b) shows that, in most cases, FULL backgrounds are closer to observations than OPER backgrounds with a mean reduction of about 10%.

This distance to UK C130 observations was also computed for 12-hour forecasts valid at the time of these observations (Fig. 6(a)). In this case, the aim was to evaluate the impact of FASTEX observations on the forecasts in the FASTEX downstream area, west of Ireland, where the UK C130 dropsondes were launched. It shows that most of these forecasts had a better score for FULL than for OPER with some large improvements (IOP 19 in particular). This leads to a noticeable mean impact comparable to the one found for backgrounds (corresponding to 6-hour forecasts). However, the mean impact of FASTEX data on the quality of the forecasts becomes difficult to distinguish at the 24-hour range (Fig. 6(b)). For this range there is roughly the same number of better forecasts for FULL than for OPER, and mean scores become the same for the whole set of flights. However, it is probable that the validation area based on UK C130 dropsondes is not well adapted to evaluate the scores of 24 h forecasts. Although these UK C130 flights were designed to study cyclones at their mature stage and to verify the impact of FASTEX upstream observations, it is likely that they were located too far west to identify a downstream forecast error at that range. This is also clear from Fig. 2 which shows that the area where FULL and OPER forecasts should differ is certainly downstream of the zone where the corresponding analyses are different.

Thus, a comparison of the forecasts against observations located inside an extended and more downstream domain was also performed. The extended verification domain

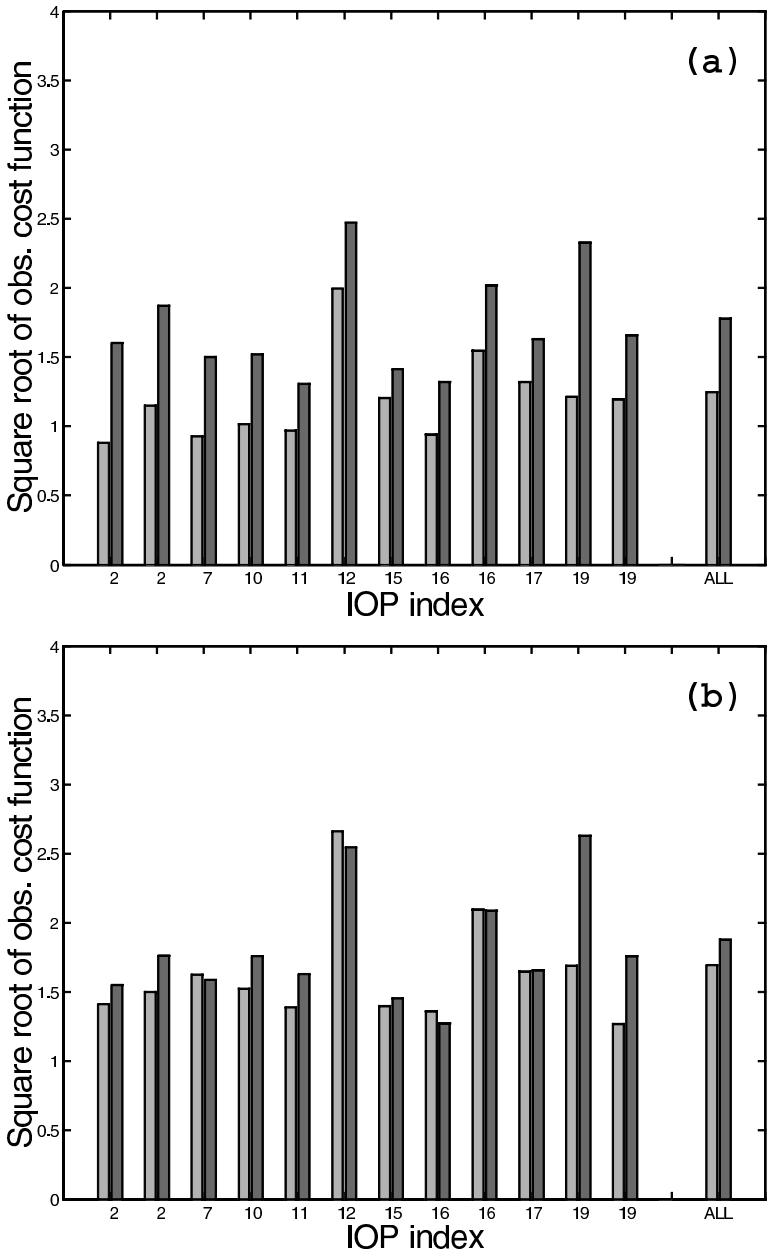

Figure 5. (a) Square root of observation cost function measuring the distance between OPER (dark bars) and FULL (light bars) analyses (see text) and UK C130 observations of temperature, relative humidity and wind velocities; bars labelled 'ALL' correspond to all flights; (b) same as (a) but for corresponding background fields. Intensive Observing Periods (IOPs) 2, 16 and 19 are split into two assimilation 6-hour windows because of the long duration of the corresponding flights.

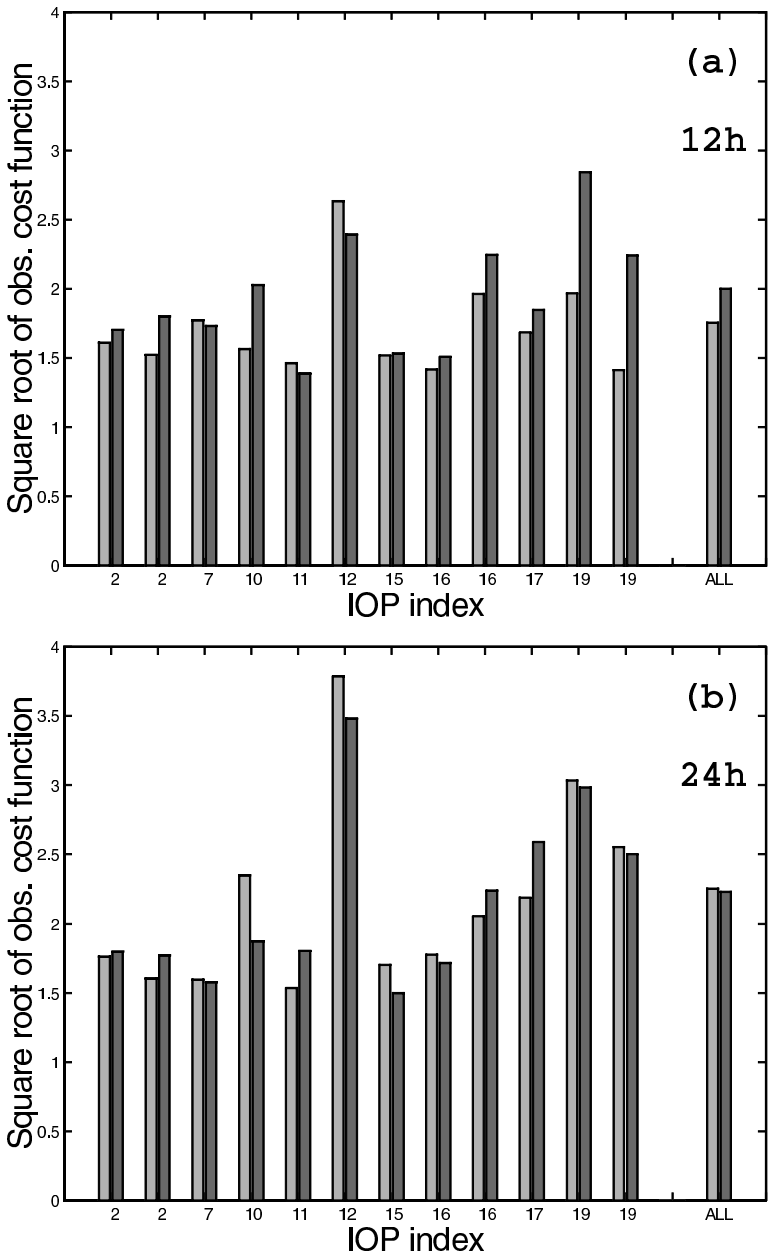

Figure 6. (a) Square root of observation cost function measuring the distance between OPER (dark bars) and FULL 12 h (light bars) forecasts (see text) and UK C130 observations of temperature, relative humidity and wind velocity; (b) same as (a) but for 24 h forecasts.

(denoted V) is shown in Fig. 2. This domain (45°N–60°N, 30°W–10°E) contains some of the previous UK C130 dropsoundings but also all the radiosoundings performed over the UK and northern France during FASTEX IOPs (the number of these radiosoundings was increased during the IOPs). The comparison between forecasts and those observations was performed during the same periods as previously when UK C130 flights occurred. Indeed, the arbitrary choice of domain V and of these periods is a necessary

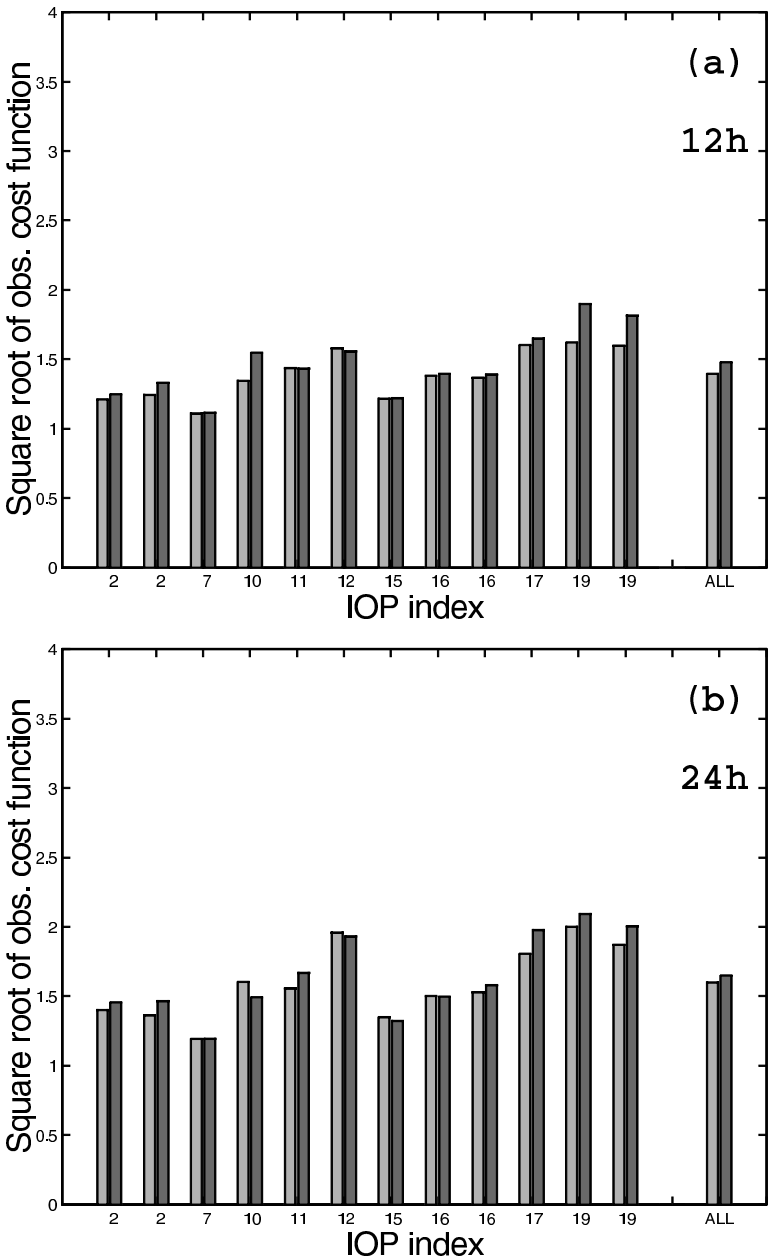

Figure 7. (a) Square root of observation cost function measuring the distance between OPER (dark bars) and FULL (light bars) 12 h forecasts (see text) and observations of temperature, relative humidity and wind velocity given by dropsondes and radiosondes inside domain V (see Fig. 2); (b) same as (a) but for 24 h forecasts.

compromise to catch a signal in the comparison of FULL and OPER forecasts (a too large domain or period makes this signal vanish).

The behaviour of 12 h forecast errors found with this new criterion (Fig. 7(a)) is different from the one obtained in the comparison against UK C130 dropsondes (Fig. 6(a)): forecast errors are lower and also appear smoother in time and closer between FULL and OPER. This is due to the averaging effect of a larger verification

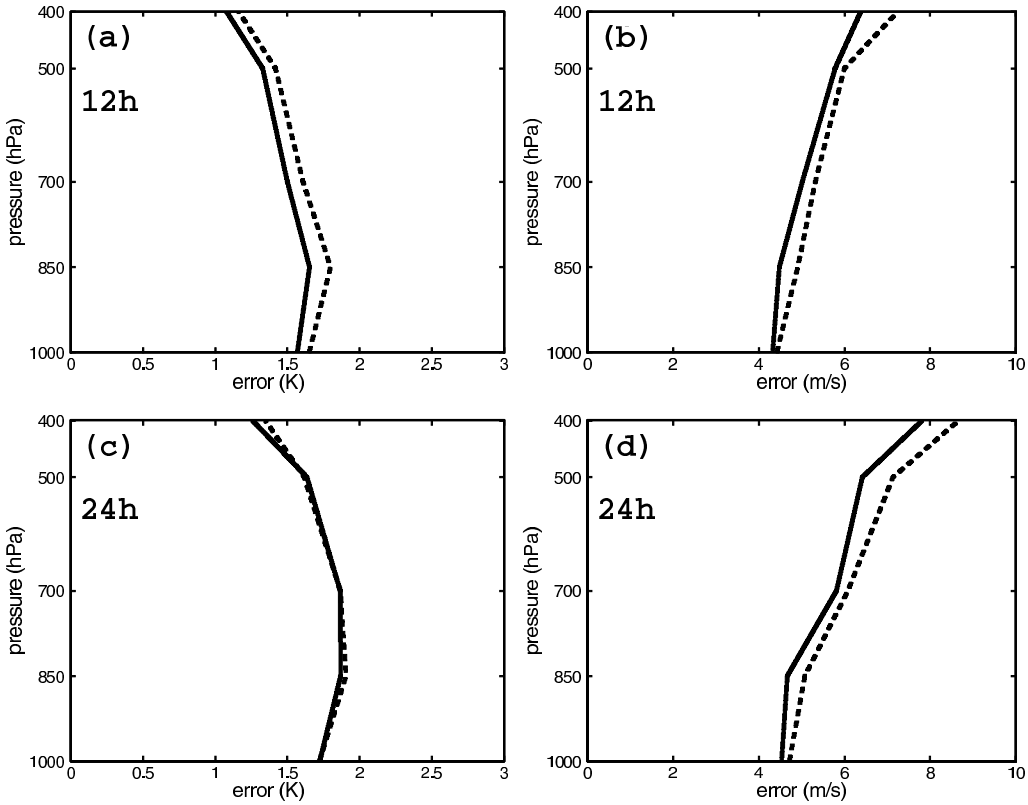

Figure 8. (a) Root mean square (r.m.s.) departure between FULL (solid line) and OPER (dashed line) (see text) 12 h forecasts and temperature observations given by UK C130 dropsondes and radiosondes inside domain V (see Fig. 2); (b) same as (a) but for wind velocity; (c) same as (a) but for 24 h forecasts; (d) same as (c) but for wind velocity.

domain. But the main point is that the mean impact of FASTEX observations now appears positive in the comparison of 24 h forecasts with radiosondes distributed inside domain V (Fig. 7(b)). Moreover, this positive impact is still present in 36 h forecasts with the same domain of validation (not shown here). Looking at the comparison of forecasts for each parameter, it appears that the mean positive improvement of FULL forecasts can be seen for both temperature and wind at 12 h (Figs. 8(a) and (b)), but is only noticeable for wind at ranges of 24 h and 36 h (Figs. 8(c) and (d), not shown for 36 h).

4. An example of the impact of FASTEX observations: IOP 19

The last IOP of the FASTEX field phase, IOP 19, was conducted on FASTEX low 46B (L46B hereafter). One characteristic of IOP 19 is that the low of interest started to show up in the satellite imagery on the day before the last flights, but then these signatures weakened. The forecast of the development was uncertain, and the *in situ* data collected during the flights are critical in finding out what actually occurred. Some characteristics of the IOP 19 life cycle can be found in Joly *et al.* (1999) and from the FASTEX website (http://www.cnrm.meteo.fr/fastex/, December 2002). L46B developed in the wake of L45. The respective locations of these two lows at 0600 UTC 27 February

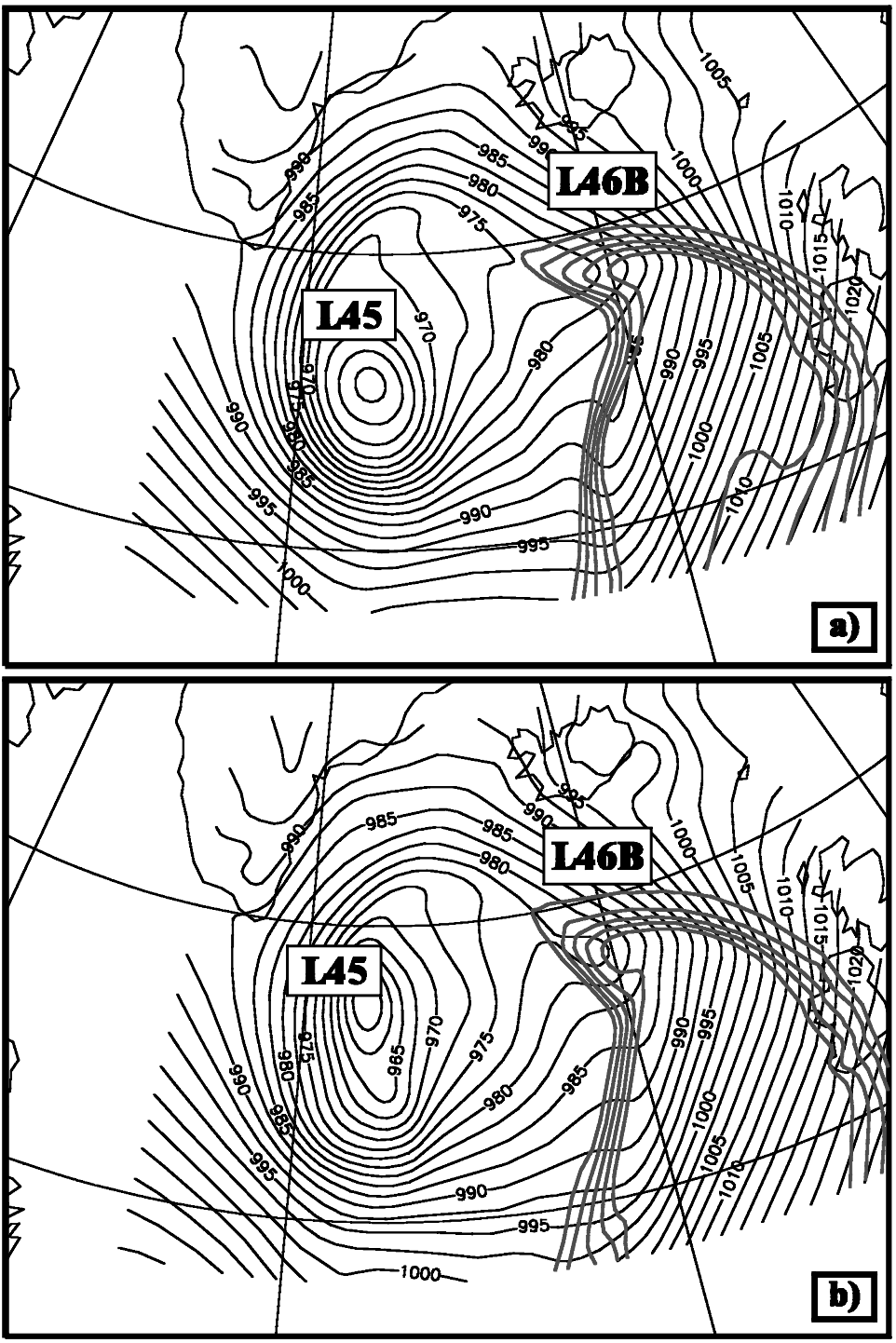

Figure 9. Mean-sea-level pressure for 0600 UTC 27 February 1997 (thin contours every 2.5 hPa) superimposed on wet-bulb temperature at 850 hPa (bold contours between 6 °C and 10 °C every 1 °C). Fields are drawn for FULL (a) and OPER (b) analyses (see text). L45 and L46B stand for the lows numbered 45 and 46B (Intensive Observing Period 19) during the Fronts and Atlantic Storm-Track EXperiment (FASTEX) campaign.

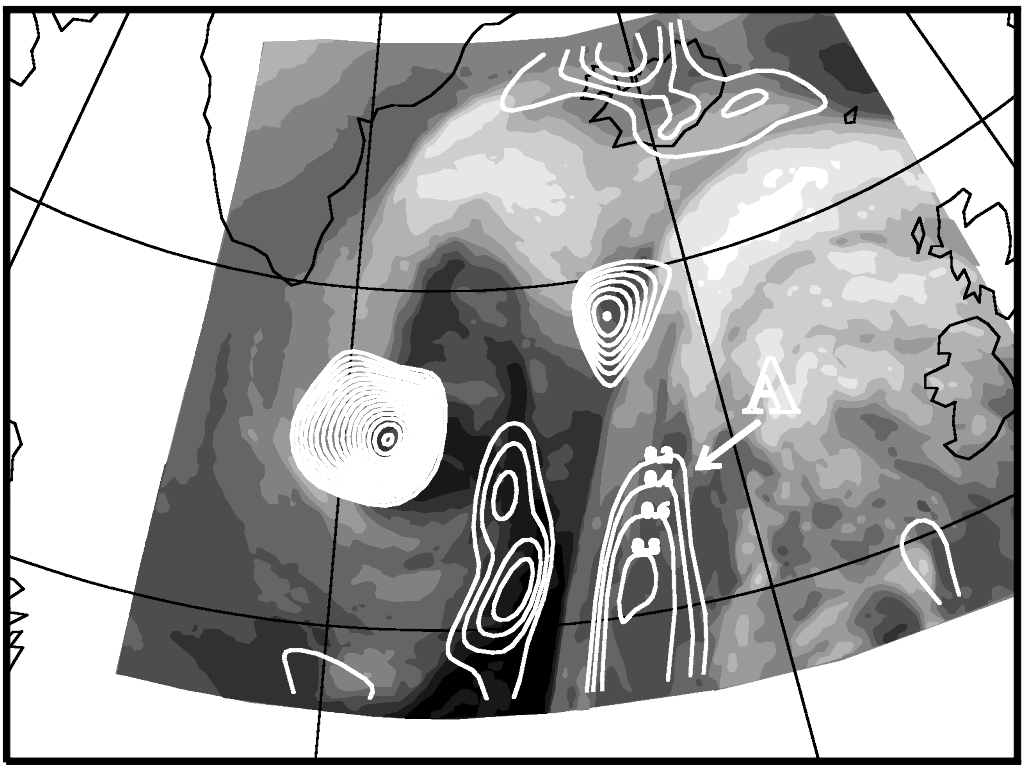

Figure 10. Water-vapour Meteosat brightness temperatures for 0600 UTC 27 February. Values are plotted from −60 °C (clear areas) to −30 °C (dark areas) every 3 °C. The differences between FULL and OPER (see text) re-analyses for temperature at 300 hPa are superimposed (contours every 0.2 °C). A is a positive anomaly of this last field above the cold front.

1997 are shown in Fig. 9. The tracking of cyclone L46B began on 26 February 1997 and ended on 28 February 1997. The low showed a maximum deepening rate for 0600 UTC 27 February, as noted by Joly *et al.* (1999). Figure 9 shows some features of L46B with the associated fronts in terms of wet-bulb temperature and mean-sea-level pressure fields, as represented by FULL and OPER analyses. It can be noted in Fig. 9 that the shape of the cold front associated with L46B is different in the two analyses. In FULL (panel (a)), a small wave develops along the undulating cold front of L46B. In the OPER reference analysis (panel (b)), the shape of the cold front is more regular and does not show any wave.

The differences at upper levels between FULL and OPER analyses are shown in Fig. 10. In the area above the cold front, a positive anomaly of the upper-level temperature difference FULL−OPER (labelled A in Fig. 10) indicates that FULL presents a lower tropopause than OPER. It should be underlined that this anomaly is well connected to an observed dark dry area in the satellite imagery (see Fig. 10). This last point gives some confidence to the FULL upper-level features and consequently to the shape of the cold front drawn in the FULL analysis. Indeed, the dark area of the picture is associated with a tropopause anomaly (high potential vorticity (PV) values) and also with a balanced vortex at the same level. The wind field associated with this PV coherent structure, which could be deduced from a PV inversion (Chaigne and Arbogast 2000), decreases vertically and horizontally as the distance to the PV source decreases.

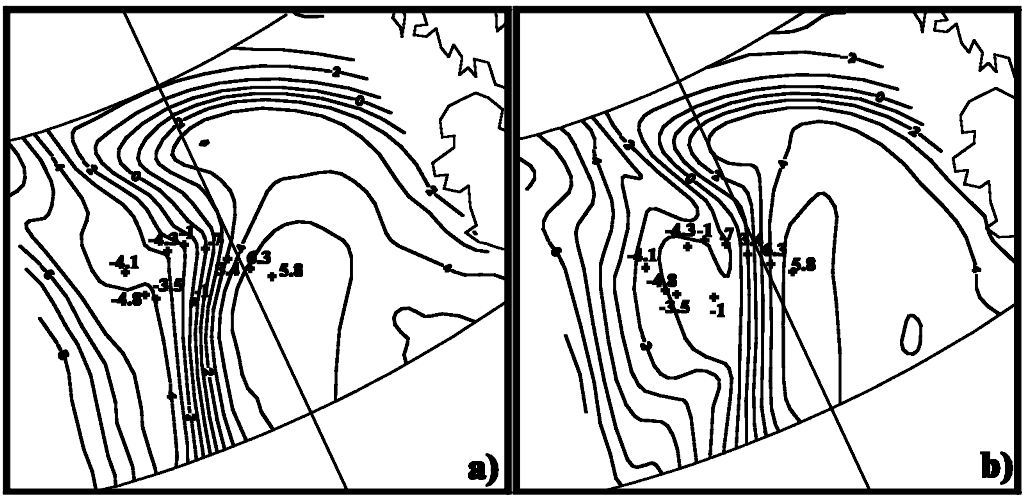

Figure 11. Temperature at 850 hPa for 0600 UTC 27 February (contours every 1 °C) drawn for FULL (a) and OPER (b) (see text) analyses. Temperature observations at 850 hPa retrieved from the dropsondes launched by the UK C130 between 0500 UTC and 0700 UTC are also plotted.

But, it would still not be zero at the boundary layer where the surface front is present. This attributed wind should be able to advect cyclonic flow associated with the surface front, leading to a frontal wave or, at least, to some modifications of its shape as the FULL analysis shows.

The comparison of both analyses with the observations coming from the dropsondes launched by the UK C130 between 0500 UTC and 0700 UTC shows that the description of the shape of the cold front of L46B and the frontal wave is closer to observations for the FULL re-analysis than for the OPER reference analysis (Fig. 11). This is also the case for the warm core associated with L46B and in the area of the strong gradient of temperature. Finally, there is an area of cold air behind the front and north-west of the frontal wave that exists in the observations and in the FULL analysis, but that is not present in the OPER reference analysis.

Note that this frontal wave did not finally develop although this non-development was questionable from the operational forecasts in real time during the field phase. However, the reason why this wave did not develop is still an interesting predictability and dynamical problem, and from the appreciable differences between the two re-analyses it should be investigated on the basis of the FULL re-analysis including FASTEX observations.

5. Conclusion

The FASTEX re-analysis was a long-term project initiated to derive the maximum benefit from the additional observations collected during the experiment, in subsequent case-studies. It has reached its objective by providing a complete set of meteorological fields over the north Atlantic domain and during the nearly two-month FASTEX period. Many aspects make this re-analysis rather unique. First, a large effort was devoted to handling and checking the quality of a very large number of pseudo-conventional observations, such as dropsondes or radiosoundings, performed by the FASTEX ships or land-based sites. Second, the use of a variational analysis scheme has also allowed better benefit to be taken of the global coverage of observations given by TOVS

radiances. But one of the most striking characteristics of this re-analysis is indeed the use of a 4D-Var formulation over such a long period. This permits the use of asynchronous FASTEX observations that would have been lost or not correctly handled by an intermittent assimilation scheme such as 3D-Var. 4D-Var is also known to implicitly spread corrections by the observations in a way more consistent with the evolution of forecast errors.

On the other hand, the variable resolution of the global French Arpège model used in the 4D-Var formulation was rather well adapted to the needs of the re-analysis, avoiding the boundary problem of a limited-area model, and providing a sufficient resolution to describe incipient precursors, and a high resolution near Ireland allowing a mesoscale description of FASTEX mature systems. In particular there is a clear agreement between the FASTEX re-analyses in the downstream area and the very dense networks of dropsondes launched by the UK C130. It has been shown in the IOP 19 case that this assimilation of FASTEX data was making a noticeable difference to the final description of structures such as frontal waves.

The consistency in time of the important dynamical structures (vorticity maxima, thermal gradient areas) in terms of shape and intensity has also been noted by the group of scientists that has already used these re-analyses (Donnadille *et al.* 2001; P. Arbogast, personal communication).

This consistency is probably linked to the improvement of the very short-range forecasts (6 h/12 h) issued from the FASTEX analyses. However, this mean positive impact is difficult to distinguish at a 24 h range on the basis of the same comparison to UK C130 dropsondes, but is noticeable when an extended and more downstream verification domain is used. The impact of FASTEX observations is particularly noticeable for IOPs 17 and 19. However, even for those cases, the impacts are relatively small. This unexpected result could be due to (at least) two reasons: the rather high predictability of developing systems during FASTEX and the use of a 4D-Var assimilation scheme allowing better benefit to be taken of both conventional and non-conventional (TOVS radiances in particular) observations. This result requires further research which will be important in perspective of THe Observing-system Research and Predictability EXperiment (THORPEX) project (however, due to the north Pacific's larger size, the error-reducing impact of targeted observations may be larger over the north Pacific than the north Atlantic).

FASTEX re-analyses are up to now archived internally but can be provided upon request to any interested scientist. There is also a plan to make them available from the on-line FASTEX data archive.

Acknowledgements

We are grateful to Béatrice Pouponneau for her help in preparing this FASTEX re-analysis and Genevive Jaubert and Catherine Piriou for making the FASTEX observations readily available. We also thank Francis Pouponneau who prepared the scores of the two analyses against ECMWF analyses. Finally we wish to thank Philippe Arbogast and Alain Joly for helpful discussions and comments on this work. We also thank the two referees for their constructive comments on the first version of this paper.

References

Amstrup, B. and Huang, X.-Y. 1999 Impact of the additional FASTEX radiosonde observations on the High-Resolution Limited-Area Model (HIRLAM) data-assimilation and forecasting system. *Q. J. R. Meteorol. Soc.*, **125,** 3359–3374

Baehr, C., Pouponneau, B., Ayrault, F. and Joly, A. 1999 Dynamical characterization of the FASTEX cyclogenesis cases. *Q. J. R. Meteorol. Soc.*, **125,** 3469–3494

Bergot, T. 1999 Adaptive observations during FASTEX: A systematic survey of upstream flights. *Q. J. R. Meteorol. Soc.*, **125,** 3271–3298

Chaigne, E. and Arbogast, P. 2000 Multiple potential-vorticity inversions in two FASTEX cyclones. *Q. J. R. Meteorol. Soc.*, **126,** 1711–1734

Courtier, C. 1997 Variational methods. *J. Meteorol. Soc. Jpn*, **75,** 211–218

Courtier, P., Freydier, C., Geleyn, J.-F., Rabier, F. and Rochas, M. 1991 'The Arpège project at Météo-France'. Pp. 193–231 in proceedings of the ECMWF workshop on numerical methods in atmospheric models, 9–13 September, Reading, UK

Desroziers, G., Pouponneau, B., Thépaut, J.-N., Janisková, M. and Veersé, F. 1999 4D-Var analyses of FASTEX situations using special observations. *Q. J. R. Meteorol. Soc.*, **125,** 3393–3414

Donnadille, J., Cammas, J.-P., Mascart, P., Lambert, D. and Gall, R. 2001 FASTEX IOP18: A very deep tropopause fold. I: Synoptic description and modelling. *Q. J. R. Meteorol. Soc.*, **127,** 2247–2268

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

Gauthier, P. and Thépaut, J.-N 2001 Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Weather Rev.*, **129,** 2089–2102

Geleyn, J.-F., Banciu, D., Bellus, M., El Khatib, R., Moll, P., Saez, P. and Thépaut, J.-N. 2001 'The operational 4D-Var data assimilation system of Météo-France: Characteristics and behaviour in the special case of the 99 Xmas storms over France'. Pp. 9–12 in proceedings of the 14th AMS conference on numerical weather prediction, 30 July–2 August, Fort Lauderdale, USA

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

Jaubert, G., Piriou, C., Loehrer, S. M., Petitpa, A. and Moore, J. A. 1999 Development and quality control of the FASTEX data archive. *Q. J. R. Meteorol. Soc.*, **125,** 3165–3188

Joly, A., Jorgensen, D., Shapiro, M. A., Thorpe, A., Bessemoulin, P., Browning, K. A., Cammas, J.-P., Chalon, J.-P., Clough, S. A., Emmanuel, K. A., Eymard, L., Gall, R., Hildebrand, P. H., Langlang, R. H., Lemaitre, Y., Lynch, P., Moore, J. A., Persson, P. O. G., Snyder, C. and Wakimoto, R. M. 1997 Definition of the Fronts and Atlantic Storm-Track EXperiment (FASTEX). *Bull. Am. Meteorol. Soc.*, **78,** 1917–1940

Joly, A., Browning, K. A., Bessemoulin, P., Cammas, J. P., Caniaux, G., Chalon, J. P., Clough, S. A., Dirks, R., Emanuel, K. A., Eymard, L., Lalaurette, F., Gall, R., Hewson, T. D., Hildebrand, P. H., Jorgensen, D., Langland, R. H., Lemaitre, Y., Mascart, P., Moore, J. A., Persson, P. O. G., Roux, F., Shapiro, M. A., Snyder, C., Toth, Z. and Wakimoto, R. M. 1999 Overview of the field phase of the Fronts and Atlantic Storm-Track EXperiment (FASTEX) project. *Q. J. R. Meteorol. Soc.*, **125,** 3131–3163

Montani, A., Thorpe, A. J., Buizza, R. and Undén, P. 1999 Forecast skill of the ECMWF model using targeted observations during FASTEX. *Q. J. R. Meteorol. Soc.*, **125,** 3219–3240

Szunyogh, I., Toth, Z., Emanuel, K. E., Bishop, C. H., Snyder, C., Morss, R. E., Woolen, J. and Marchok, T. 1999 Ensemble-based targeting experiments during FASTEX: The effect of dropsonde data from the Lear jet. *Q. J. R. Meteorol. Soc.*, **125,** 3189–3217

Veersé, F. and Thépaut, J.-N. 1998 Multiple-truncation incremental approach for four-dimensional variational data assimilation. *Q. J. R. Meteorol. Soc.*, **124,** 1889–1908

Q. J. R. Meteorol. Soc. (2005), **131**, pp. 2821–2837 doi: 10.1256/qj.04.151

Use of randomization to diagnose the impact of observations on analyses and forecasts

By GÉRALD DESROZIERS*, PIERRE BROUSSEAU and BERNARD CHAPNIK
Météo-France, Centre National de Recherches Météorologiques, Toulouse, France

(Received 15 October 2004; revised 1 July 2005)

SUMMARY

A method is proposed to diagnose the impact of a complete set, or subsets, of observations on the reduction of error variance in an analysis and in the subsequent forecasts run from this analysis. A practical method to estimate the error reduction, based on a randomization procedure, is also introduced and investigated in a simple framework given by the analysis and forecast of wind on a circular domain using the nonlinear Burger's equation. The randomization procedure is also applied and tested in the French ARPEGE 4D-Var assimilation. The first results in a real-size data assimilation system are realistic and provide useful information on the use of observations in an operational analysis.

KEYWORDS: Analysis and forecast errors Data assimilation

1. INTRODUCTION

The role of the analysis step in an operational prediction system is to reduce the error in the initial conditions in order to diminish, in turn, the amplitude of forecast errors. In a typical assimilation cycle, the analysis can be seen as the result of a correction applied to a background given by a very short-term forecast. When solving the corresponding least-square error problem, specific error statistics need to be prescribed for background and observations. As a result, the analysis is associated with a decreased error variance when compared to background error, in model space, and measurement errors, in observation space. It can also be shown that each piece of observation should in particular reduce the initial background error variance.

The variational formalism (Lewis and Derber 1985; Courtier and Talagrand 1987), which is nowadays operational in many numerical weather-prediction (NWP) centres (Rabier *et al.* 2000), allows the use of various types of observations, in particular satellite data, that are not directly linked with model variables.

The evaluation of the impact of observations on forecasts is typically done by performing observing system experiments (OSEs). Such experiments consist in performing continuous assimilation/forecast cycles with various observation deployment scenarios and by comparing the forecast scores of the different experiments (Dumelow 2003). OSEs are simple to implement but require a large amount of computational resources. However, diagnostics of the impact of observations on the analyses themselves have been barely implemented. One difficulty is that the true state is never known, preventing computation of analysis errors, whereas forecasts can be compared to observations or analyses considered as an approximation of the truth. Another difficulty is that the statistical variance reduction induced by the use of observations is not easily available in a variational scheme. The variational analysis at the European Centre for Medium-Range Weather Forecasts (ECMWF) uses a procedure to estimate analysis error variance as a by-product of the minimization, by a combined Lanczos–conjugate gradient method (Fisher and Courtier 1995). More recently, Fisher (2003) and Cardinali *et al.* (2004) have used the same Lanczos–conjugate gradient method to measure the degree of freedom for signal (DFS) given by the complete set or subsets of observations.

* Corresponding author: Météo-France, CNRM, 42 av. G. Coriolis, 31057 Toulouse Cedex, France.
e-mail: Gerald.Desroziers@meteo.fr
© Royal Meteorological Society, 2005.

On the other hand, Talagrand (1999) has shown that the values at the minimum of subparts of the cost function associated with subsets of observations also give information on the contribution of these observations to the final estimation covariance, corresponding to the same partial DFSs. Desroziers and Ivanov (2001) showed that such quantities can be computed by a randomization procedure consisting of a perturbation of observations. Chapnik *et al.* (2004) have applied such a procedure to compute the DFSs associated with subsets of observations in a quasi-operational setting of the French ARPEGE* four-dimensional variational (4D-Var) assimilation scheme. In particular, they compared the relative contribution of conventional and satellite data in the analysis. Rabier and Courtier (1992) also produced randomized estimations of the diagonal elements of the Hessian matrix of a 4D-Var assimilation problem.

While the previous studies focused on the impact of observations in an analysis, other studies have investigated the possibility of evaluating the impact of observations on subsequent short-term forecasts, especially in the framework of adaptive observations. In that case, the idea is to choose the location of additional observations in sensitive areas that will produce the largest error decrease in an area of interest. In particular, Bergot and Doerenbecher (2002) have developed a method to predict the reduction of error variance of a forecast aspect measured by a specific cost function. Leutbecher (2003) has proposed to use a reduced rank estimate of forecast error variance to evaluate the impact of additional data. Both methods rely on the use of the approximation of the analysis covariance matrix provided by the Lanczos–conjugate gradient algorithm proposed by Fisher and Courtier (1995). Langland and Baker (2004) also proposed an adjoint-based procedure for estimating the impact of observations on short-range forecast errors. From another point of view, Bishop *et al.* (2001) introduced a method based on the ensemble Kalman filter to optimize the sampling of adaptive observations.

The aim of this paper is to propose a method to estimate the error variance reduction due to particular subsets of observations in the analysis and also in the subsequent forecasts run from this analysis. Unlike the previously cited studies, the method proposed in this paper is an extension of the randomization procedure described by Desroziers and Ivanov (2001) and Chapnik *et al.* (2004).

In section 2, a general overview of the variational formalism is given and the opportunity is introduced to compute analysis error variance reduction associated with subsets of observations. Then it is shown in section 3 how the same quantities can be computed for forecast error variance reduction. The randomization procedure used to compute the error variance reductions is introduced in section 4. The application of the method to Burger's equation on a circle is presented in section 5. First results obtained from the application of the randomization procedure to the French ARPEGE 4D-Var assimilation are presented in section 6. Final discussion and conclusions are given in section 7.

2. Diagnosis on analyses

This section shows how, starting from the general formalism of variational analysis, it is possible to produce an expression that allows the computation of the reduction of error variance due to a subset of observations.

(*a*) *General formulation*

The principle of the incremental formulation of a variational algorithm (Courtier *et al.* 1994) is to seek the increment $\delta\mathbf{x}$ to add to the background $\mathbf{x}^{\mathrm{b}}$ which minimizes

* Action de Recherche Petite Echelle Grande Echelle.

the cost function

$$J(\delta\mathbf{x}) = \tfrac{1}{2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{B}^{-1}\delta\mathbf{x} + \tfrac{1}{2}(\mathbf{d} - \mathbf{H}\delta\mathbf{x})^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{d} - \mathbf{H}\delta\mathbf{x}),$$

where $\mathbf{B}$ and $\mathbf{R}$ respectively stand for the assumed background and observation error covariance matrices, $\mathbf{H}$ is the linearized observation operator and $\mathbf{d} = \mathbf{y}^{\mathrm{o}} - H(\boldsymbol{x}^{\mathrm{b}})$ is the innovation vector, with H the observation operator that allows the computation of the model equivalents in observation space. The solution of the minimization of J gives

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}\mathbf{d},$$

where $\mathbf{K}$ is the gain matrix. It can be shown (see for example Cohn 1997) that, if background and observation error covariance matrices are well specified, then the analysis error covariance matrix is given by

$$\mathbf{A} = \mathbf{B} - \mathbf{K}\mathbf{H}\mathbf{B} = \mathbf{B} - \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{B}. \tag{1}$$

As a consequence, the total variance reduction given by the assimilation of the observations is

$$r = \mathrm{Tr}(\mathbf{B}) - \mathrm{Tr}(\mathbf{A}) = \mathrm{Tr}(\mathbf{K}\mathbf{H}\mathbf{B}) = \mathrm{Tr}(\mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H}\mathbf{B}), \tag{2}$$

where Tr stands for the trace of a matrix.

(*b*) *Impact of a subset of observations*

It is now shown how to express the reduction of error variance due to subsets of observations with independent errors.

First, if the complete set of observations can be split between observation subsets i with independent errors, then matrix $\mathbf{R}$ is block-diagonal and it is easy to check that the expression (1) for matrix $\mathbf{A}$ can be rewritten

$$\mathbf{A} = \mathbf{B} - \sum_i \mathbf{A}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{B}, \tag{3}$$

where $\mathbf{R}_i$ and $\mathbf{H}_i$ respectively are the corresponding error covariance matrices and observation operators associated with subset i.

The last expression shows that the reduction of error variance due to subset i is

$$r_i = \mathrm{Tr}(\mathbf{A}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{B}), \tag{4}$$

which can be compared to the total error reduction (2) brought by the whole set of observations.

However, Eq. (4) cannot be used in practice since matrix $\mathbf{A}$ is not explicitly determined in a real-size assimilation system. To obtain another expression for r_i, each term under the sum of (3) can be first rewritten as

$$\mathbf{A}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{B} = \mathbf{A}(\boldsymbol{\Pi}_i\mathbf{H})^{\mathrm{T}}\boldsymbol{\Pi}_i\mathbf{R}^{-1}\boldsymbol{\Pi}_i{}^{\mathrm{T}}(\boldsymbol{\Pi}_i\mathbf{H})\mathbf{B} = \mathbf{A}\mathbf{H}^{\mathrm{T}}\boldsymbol{\Pi}_i{}^{\mathrm{T}}\boldsymbol{\Pi}_i\mathbf{R}^{-1}\boldsymbol{\Pi}_i{}^{\mathrm{T}}\boldsymbol{\Pi}_i\mathbf{H}\mathbf{B}, \tag{5}$$

where $\boldsymbol{\Pi}_i$ is the projection operator that allows us to pass from the complete set of observations to the subset i of p_i observations. Matrix $\mathbf{S}_i = \boldsymbol{\Pi}_i{}^{\mathrm{T}}\boldsymbol{\Pi}_i$ acts as an operator which sets to zero all components which do not correspond to subset i, and keeps all the others equal. It satisfies the relation $\mathbf{S}_i\mathbf{S}_i = \mathbf{S}_i$. Since $\mathbf{R}^{-1}$ is assumed to be block-diagonal, $\mathbf{S}_i$ also commutes with $\mathbf{R}^{-1}$. Then, expression (5) can be rewritten

$$\begin{aligned}\mathbf{A}\mathbf{H}_i^{\mathrm{T}}\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{B} &= \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{S}_i\mathbf{S}_i\mathbf{H}\mathbf{B}\\ &= \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{S}_i\mathbf{H}\mathbf{B}\\ &= \mathbf{K}\mathbf{S}_i\mathbf{H}\mathbf{B},\end{aligned}$$

and thus, the error variance reduction due to an independent subset of observations can take the following alternate expression:

$$r_i = \mathrm{Tr}(\mathbf{KS}_i\mathbf{HB}).$$

This last expression for r_i is similar to expression (2), except that operator $\mathbf{S}_i$ is first applied to $\mathbf{HB}$ in order to select observations associated with subset i.

It will be shown, in section 4, how these expressions can be computed with a randomization procedure.

3. Transformation of covariance matrices

(*a*) *General formulation*

Different aspects of the error variance reduction can be investigated by using a transformation $\mathbf{L}$ of the estimation error $\boldsymbol{\epsilon}$ on the state vector. If $\boldsymbol{\epsilon}$ is transformed in $\mathbf{L}\boldsymbol{\epsilon}$, then the covariance of the transformed vector becomes $E(\mathbf{L}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}) = \mathbf{L}E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}})\mathbf{L}^{\mathrm{T}}$.

In particular, $\mathbf{B}$ and $\mathbf{A}$ respectively become $\mathbf{LBL}^{\mathrm{T}}$ and $\mathbf{LAL}^{\mathrm{T}}$.

Similarly, the variance reductions for the complete set and for a subset of the observations are

$$r = \mathrm{Tr}(\mathbf{LKHBL}^{\mathrm{T}})$$

and

$$r_i = \mathrm{Tr}(\mathbf{LKS}_i\mathbf{HBL}^{\mathrm{T}}). \tag{6}$$

(*b*) *Link with degrees of freedom for signal*

Many expressions for the transformation $\mathbf{L}$ can be imagined. In particular, $\mathbf{L}$ can include a projection or correspond to a specific norm.

A special case is the use of $\mathbf{L} = \mathbf{B}^{-1/2}$, for which $\mathbf{B}$ is reduced to the identity matrix in the transformed space, $\mathbf{A}$ to $\mathbf{B}^{-1/2}\mathbf{AB}^{-1/2}$ and the variance reduction to $r = \mathrm{Tr}(\mathbf{KH})$. This former expression is known as the number of degrees of freedom for signal (Rodgers 2000). In that case, it can also be shown that

$$\begin{aligned} r_i &= \mathrm{Tr}(\mathbf{LKS}_i\mathbf{HBL}^{\mathrm{T}}) \\ &= \mathrm{Tr}(\mathbf{KS}_i\mathbf{H}) \\ &= \mathrm{Tr}(\mathbf{AH}_i{}^{\mathrm{T}}\mathbf{R}_i{}^{-1}\mathbf{H}_i) \end{aligned}$$

is equal to the expectation of the subpart $J_i^{\mathrm{o}}(\mathbf{x}^{\mathrm{a}})$ of the cost function associated with observation subset i, if prior statistics matrices $\mathbf{B}$ and $\mathbf{R}$ are well specified (Desroziers and Ivanov 2001). These quantities are also those computed by Fischer (2003) and Cardinali *et al.* (2004).

(*c*) *Diagnostics on forecasts*

The transformation $\mathbf{L}$ can also include a forecast. In that case the diagnostics on the background and analysis proposed in section 2 can be extended to diagnostics on forecasts. A projection operator $\mathbf{P}$ onto a verification region can also be introduced, as well as a particular metric $\mathbf{N}$ to normalize forecast errors in this region. In this case, the variance of forecast error becomes

$$E(\|\mathbf{PM}\boldsymbol{\epsilon}\|_{\mathbf{N}}^2) = \mathrm{Tr}\{\mathbf{N}^{1/2}\mathbf{PM}E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}})\mathbf{M}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{N}^{1/2}\},$$

where $\boldsymbol{\epsilon}$ corresponds to a vector of initial errors and $\mathbf{M}$ is the tangent linear model associated with the forecast model.

4. Randomized estimates

(*a*) *Principle*

It is now shown how to compute the partial (or total) reduction of error on forecasts (or on analyses) given by expression (6).

First, the expression of r_i can be again rewritten

$$\begin{aligned} r_i &= \mathrm{Tr}(\mathbf{LKS}_i\mathbf{HBL}^{\mathrm{T}}) \\ &= \mathrm{Tr}(\mathbf{S}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{LK}) \\ &= \mathrm{Tr}(\mathbf{R}^{-1/2}\mathbf{S}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{LKR}^{1/2}), \end{aligned}$$

using the fact that for any matrices $\mathbf{C}$ and $\mathbf{D}$, $\mathrm{Tr}(\mathbf{CD}) = \mathrm{Tr}(\mathbf{DC})$.

Neither $\mathbf{B}$ nor $\mathbf{K}$ are explicitly known in an operational variational analysis scheme; rather, background covariances are specified through the operator corresponding to the change of variables $\boldsymbol{\chi} = \mathbf{B}^{-1/2}\delta\mathbf{x}$, and matrix $\mathbf{K}$ is not computed. However, it is now shown that this variance reduction can be approximated by a randomization method. First, it can shown (Girard 1989) that for any matrix $\mathbf{C}$, an estimation of the trace of this matrix is given by $\mathcal{T} = \boldsymbol{\eta}^{\mathrm{T}}\mathbf{C}\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a vector of numbers with a standard Gaussian distribution. It can also be shown that the standard deviation of $\mathcal{T}$ is $\mathrm{Tr}(\mathbf{C}^2)$, that can indicate the number of repeated estimates, with different random vectors $\boldsymbol{\eta}$, in order to reach a given accuracy on the estimation. Thus, a randomized estimation of r_i is given by

$$\begin{aligned} r_i &\simeq \boldsymbol{\eta}^{\mathrm{o}\mathrm{T}}\mathbf{R}^{-1/2}\mathbf{S}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{LKR}^{1/2}\boldsymbol{\eta}^{\mathrm{o}} \\ &= \delta\mathbf{y}^{\mathrm{o}\mathrm{T}}\mathbf{R}^{-1}\mathbf{S}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{LK}\delta\mathbf{y}^{\mathrm{o}}, \end{aligned}$$

where $\delta\mathbf{y}^{\mathrm{o}} = \mathbf{R}^{1/2}\boldsymbol{\eta}^{\mathrm{o}}$ is a vector of perturbations on the whole set of observations. This can be rewritten

$$\begin{aligned} r_i &= \delta\mathbf{y}^{\mathrm{o}\mathrm{T}}\mathbf{R}^{-1}\boldsymbol{\Pi}_i^{\mathrm{T}}\boldsymbol{\Pi}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{LK}\delta\mathbf{y}^{\mathrm{o}} \\ &= \delta\mathbf{y}_i^{\mathrm{o}\mathrm{T}}\mathbf{R}_i^{-1}\boldsymbol{\Pi}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{LK}\delta\mathbf{y}^{\mathrm{o}}, \end{aligned}$$

where $\delta\mathbf{y}_i^{\mathrm{o}} = \boldsymbol{\Pi}_i\delta\mathbf{y}^{\mathrm{o}}$ is the projection of $\delta\mathbf{y}^{\mathrm{o}}$ onto the particular subset i of observations.

Now, the key point is that $\mathbf{K}\delta\mathbf{y}^{\mathrm{o}}$ is nothing other than the perturbation $\delta\mathbf{x}^{\mathrm{a}}(\delta\mathbf{y}^{\mathrm{o}})$ on the analysis produced by a perturbation on the observations, since $\delta\mathbf{x}^{\mathrm{a}}(\delta\mathbf{y}^{\mathrm{o}}) = \mathbf{K}\delta\mathbf{y}^{\mathrm{o}}$ in the linear case. Then a randomized and practical estimation of r_i is

$$r_i \simeq \delta\mathbf{y}_i^{\mathrm{o}\mathrm{T}}\mathbf{R}_i^{-1}\boldsymbol{\Pi}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{L}\delta\mathbf{x}^{\mathrm{a}}(\delta\mathbf{y}^{\mathrm{o}}). \tag{7}$$

Thus, the following chain of operations is necessary to compute such an estimation for a particular subset of observations:

(i) perform two analyses, one with the complete set of unperturbed observations and another one with the same complete set of perturbed observations,
(ii) compute the difference between the two analyses,
(iii) apply the linear operator $\mathbf{L}$ to the difference,
(iv) apply the adjoint of $\mathbf{L}$,
(v) multiply the result by $\mathbf{B}$, that is to say apply the operator $\mathbf{B}^{\mathrm{T}/2}$ followed by the operator $\mathbf{B}^{1/2}$,
(vi) evaluate the resulting field at the locations and for the variables contained in observations subset i,

(vii) normalize the evaluations by the corresponding observation error,
(viii) sum the products of the normalized evaluations by the corresponding perturbations on observations.

Note that, for a given operator $\mathbf{L}$, the first chain of operations $\mathbf{HBL}^{\mathrm{T}}\mathbf{L}\delta\mathbf{x}^{\mathrm{a}}$, which is the most expensive, has only to be performed once. In particular the perturbation is performed on the whole set of observations, but the evaluation of r_i for any subset is performed as a last and very cheap computation.

If $\mathbf{L}$ involves a model integration, then $\mathbf{L}^{\mathrm{T}}\mathbf{L}$ includes the forward integration of the tangent linear model $\mathbf{M}$ and a backward integration of the corresponding adjoint model $\mathbf{M}^{\mathrm{T}}$.

(*b*) *Extension to the nonlinear case*

Although the previous expression has been derived in the case where the analysis increment is a linear combination of innovations, it can be easily extended to a slightly nonlinear analysis with $\delta\mathbf{x}^{\mathrm{a}} \simeq \mathbf{x}^{\mathrm{a}}(\mathbf{y}^{\mathrm{o}} + \delta\mathbf{y}^{\mathrm{o}}) - \mathbf{x}^{\mathrm{a}}(\mathbf{y}^{\mathrm{o}})$, where $\mathbf{x}^{\mathrm{a}}(\mathbf{y}^{\mathrm{o}} + \delta\mathbf{y}^{\mathrm{o}})$ and $\mathbf{x}^{\mathrm{a}}(\mathbf{y}^{\mathrm{o}})$ are the analyses with perturbed and unperturbed observations, respectively. An example of such a slightly nonlinear analysis is given by the incremental 4D-Var where the nonlinearities appear through the periodic updates of the full model trajectory (Courtier *et al.* 1994; Rabier *et al.* 2000). The opportunity of performing such computations on a slightly nonlinear scheme such as 4D-Var, that is progressively implemented in main operational centres, appears as an advantage of the randomized method proposed here.

5. Application to the analysis and forecast on a circle

(*a*) *Analysis of wind on a circle and Burger's forecast model*

The opportunity of applying such a procedure to compute error reductions due to subsets of observations has been first tested in a simple framework given by the analysis and the evolution of wind on a circular domain. The forecast equation for the evolution of the wind u is given by the nonlinear Burger's equation with diffusion:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial s} = \nu\frac{\partial^2 u}{\partial s^2},$$

where ν is the diffusion coefficient and t and s respectively stand for the time and the spatial coordinate along the circle.

The representation of u is spectral on a 40 000 km long domain (say an earth meridian). A spectral resolution is used of $k = 100$ modes, corresponding to $n = 2k + 1 = 201$ spectral coefficients and grid points.

The Burger model is an adaptation of the model developed by Liu (2002). The integration is performed with a semi-implicit leapfrog scheme initialized by a forward Euler step. Three forms of the model are used here: the nonlinear, tangent-linear and associated adjoint models.

The true initial condition (dashed line in Fig. 1) has a simple sinusoidal form

$$u(s) = -U\sin\left(\frac{s}{a}\right),$$

with $U = 20$ m s^{-1} and a is the length of the domain. The value of the diffusion coefficient is $\nu = 1000$, and with such a value a localized sharp gradient is obtained after a 72 h integration (dashed line in Fig. 2).

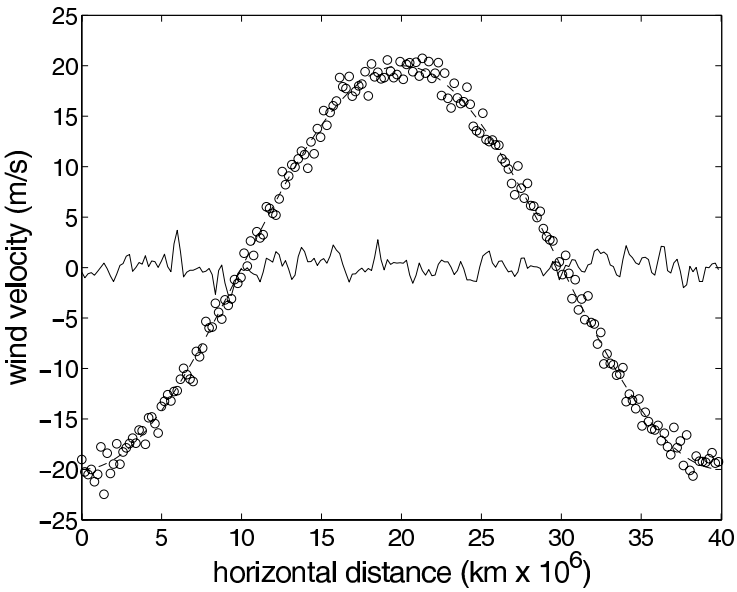

Figure 1. True signal (dashed line) at time 0 h, corresponding simulated observations (circles) and difference between simulated background and true signal (solid line); the length-scale of background errors is 200 km, background error σ^{b} and observation error σ^{o} are both homogeneous with $\sigma^{\mathrm{b}} = \sigma^{\mathrm{o}} = 1$ m s^{-1}.

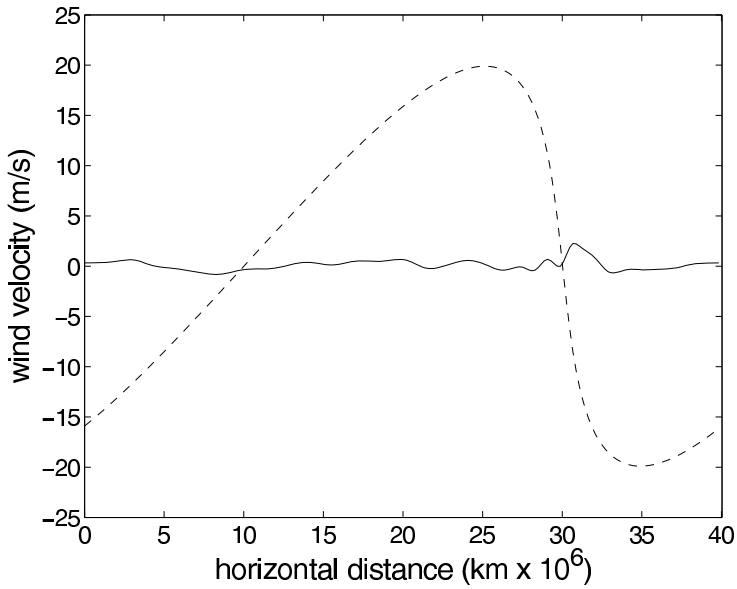

Figure 2. As Fig. 1 but at time 72 h.

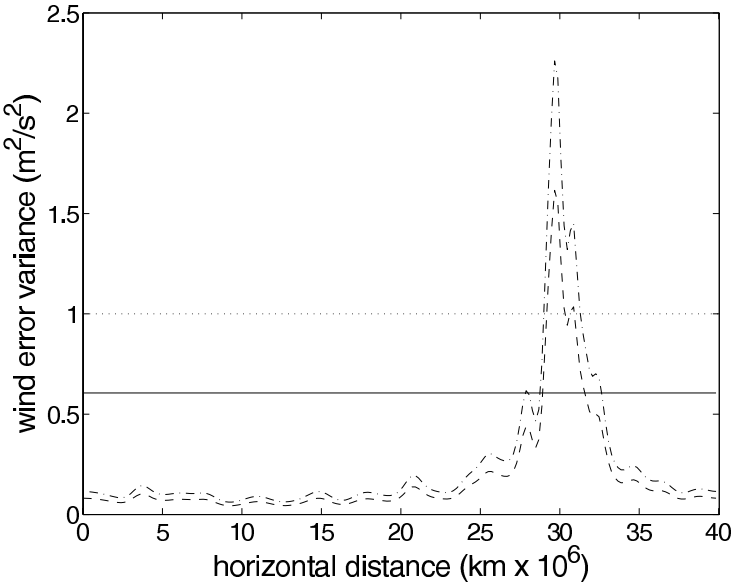

Figure 3. Background error variance at time 0 h (dotted line) and time 72 h (dash-dot line), total error reduction r at time 0 h (solid line) and time 72 h (dashed line) given by observations. There are $p = 201$ observations and $\sigma^{\mathrm{b}} = \sigma^{\mathrm{o}} = 1\ \mathrm{m\ s^{-1}}$. Corresponding exact and randomized error reductions are given in Table 1.

Simulated background and observations are produced at initial time (Fig. 1). The background is given by a perturbation of the true signal by an error vector $\boldsymbol{\epsilon}^{\mathrm{b}} = \mathbf{B}^{1/2}\boldsymbol{\eta}^{\mathrm{b}}$, where $\boldsymbol{\eta}^{\mathrm{b}}$ is a random vector with Gaussian distribution. Matrix $\mathbf{B}$ is defined as a diagonal matrix in spectral space whose diagonal is given by the Fourier transform of a Gaussian covariance function, with a length-scale equal to 200 km. Observations are simulated as the evaluation of the true signal perturbed by random noise with variance σ^{o^2}. They are regularly placed along the domain (circles in Fig. 1). Both background error variance σ^{b^2} and observation error variance σ^{o^2} are set equal to $1\ \mathrm{m^2 s^{-2}}$.

(*b*) *Computation of error reductions in analysis and forecast*

Different experiments have been made in order to demonstrate the validity of the present approach. The first experiment shown in Fig. 3 corresponds to an assimilation with 201 observations regularly spaced along the domain (in this case, there is exactly one observation per grid point since the number of grid points is also equal to 201). The analysis is an incremental analysis scheme like 3D-Var. The spectral analysis increment $\delta\mathbf{x}$ has the same resolution as the nonlinear model used to compute the innovation vector (i.e. a spectral resolution with $k = 100$ modes).

Although the wind signal is represented and evolved in spectral space, all the error variances or error reductions are determined in grid-point space, that is with the application of a linear operator corresponding to the inverse Fourier transform $\mathbf{F}^{-1}$. Hence, the dotted line in Fig. 3 corresponds to the diagonal of matrix $\mathbf{F}^{-1}\mathbf{B}\mathbf{F}^{-\mathrm{T}}$ and the solid line to the total error reduction given by the diagonal of matrix $\mathbf{F}^{-1}\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{F}^{-\mathrm{T}}$. Here, matrix $\mathbf{K}$ is exactly and explicitly used in the formula, since the small size of the problem allows its direct computation. Note that, in that case, the error reduction due to the assimilation of observations is homogeneous over the domain. This is due to

TABLE 1. ERROR VARIANCE REDUCTION (m^2s^{-2}) ON ANALYSIS (r^a) AND FORECAST (r^f) DUE TO THE COMPLETE SET OF 201 OBSERVATIONS, WITH EXACT REDUCTION AND RANDOMIZED ESTIMATES WITH DIFFERENT NUMBERS OF PERTURBATIONS

	Exact values	1 perturbation	5 perturbations	10 perturbations
r^a	0.61	0.59	0.60	0.61
r^f	0.18	0.21	0.16	0.18

The error reductions correspond to Fig. 3.

the choice of a 200 km error length-scale for background errors that is larger than the separation between observations, roughly equal to 100 km (40 000 km/201). The dash-dot line in Fig. 3 shows the field of background error variance evolved after 72 h, that is when the localized sharp gradient has occurred (see also Fig. 2 for a representation of the wind signal at this time). This evolved background error variance is given by the diagonal of matrix $\mathbf{F}^{-1}\mathbf{M}\mathbf{B}\mathbf{M}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}}$, where $\mathbf{M}$ is the matrix associated with the integration of the tangent-linear model, linearized around the background trajectory. Again, thanks to the small size of the problem, matrix $\mathbf{M}$ is explicitly built by applying the spectral Burger tangent linear model to matrix identity. In agreement with the physics described by Burger's equation, the evolved background error variance shows a decrease outside the area of sharp wind gradient, due to the dissipation induced by the diffusion term, and on the contrary a very localized increase in the discontinuity zone, where the variance is roughly multiplied by a factor of two. Finally, the dashed line in Fig. 3 corresponds to the error variance reduction at 72 h given by the assimilation of observations at initial time. This reduction is given by the diagonal of the matrix involved in formula (2), with $\mathbf{L} = \mathbf{F}^{-1}\mathbf{M}$, that is $r = \mathbf{F}^{-1}\mathbf{M}\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{M}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}}$. The evolved error variance reduction is no longer homogeneous and has rather the same structure as the evolved background error variance. It is smaller than the initial reduction outside the discontinuity zone and larger in the discontinuity area.

At initial time, the mean error reduction over the whole domain is given by $r^a = \mathrm{Tr}(\mathbf{F}^{-1}\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{F}^{-\mathrm{T}})$ and, at final time, by $r^f = \mathrm{Tr}(\mathbf{F}^{-1}\mathbf{M}\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{M}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}})$. These exact values, normalized by the number of grid points, are given in the first column of Table 1.

The error variance is globally decreased since the discontinuity area has a small extent and outside this zone the variance decreases. Table 1 also shows the computation of the randomized estimates of these same quantities. They are given for different numbers (1, 5 and 10) of perturbations. Note that, in that case, the randomized estimate is computed according to expression (7). In particular, the analysis is produced as the result of a minimization procedure with a quasi-Newton algorithm, which is an adaptation of the M1QN3 minimizer (Gilbert and Lemaréchal 1989). This is done to mimic a real-size assimilation scheme. Similarly the term $\mathbf{L}^{\mathrm{T}}\mathbf{L}\delta\mathbf{x}^{\mathrm{a}}$ results from the application of tangent-linear and adjoint model operators to the perturbation of the analysis resulting from the difference between unperturbed and perturbed analyses. The values in Table 1 show that the randomized estimates are close to the true values, both at initial and final times, even with a single perturbation of observations, that is with a single additional analysis. The values improve with the number of perturbations and are in perfect agreement with the exact computation with 10 perturbations.

Figure 4 corresponds to the same case with 201 observations assimilated, but now the reductions shown are associated with a subset of a quarter of the observations, i.e. one observation out of four along the domain. Note that this variance reduction

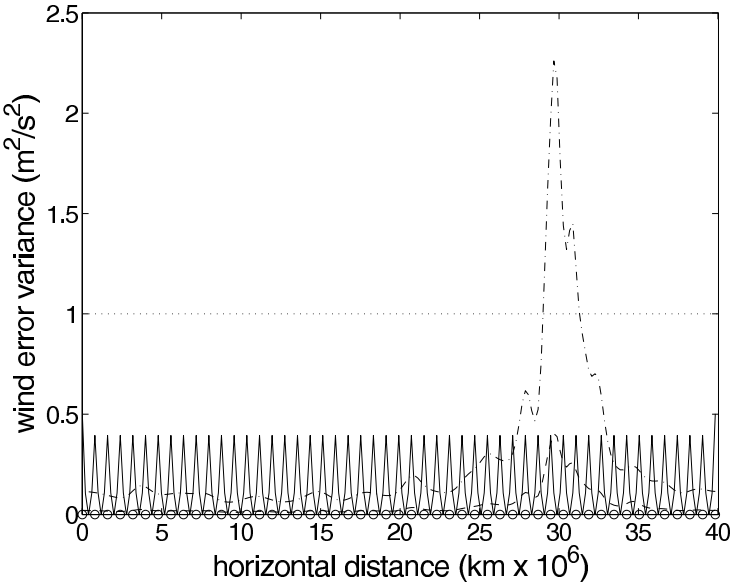

Figure 4. As Fig. 3, but the reductions correspond to a subset of $p = 51$ observations whose locations are shown as circles. Corresponding exact and randomized error reductions are given in Table 2.

TABLE 2. AS TABLE 1, BUT WITH ERROR REDUCTIONS CORRESPONDING TO A SUBSET OF OBSERVATIONS SHOWN IN FIG. 4

	Exact values	1 perturbation	5 perturbations	10 perturbations
r^{a}	0.15	0.21	0.17	0.17
r^{f}	0.05	0.07	0.05	0.05

is different from the reduction that would give such a set of observations without the presence of other observations. The reduction brought by one observation out of four but in the presence of other observations (Fig. 4) shows a typical pattern centred at each location of the subset of data. It evolves toward a smoother structure at 72 h (dashed line), with again decreased variance outside the discontinuity zone and a peak in the discontinuity area.

Logically, the exact mean error reduction brought by the subset of observations (Table 2) is roughly equal to a quarter of the reduction brought by the complete set of observations at initial time. At final time, the ratio is slightly different (0.05 $\mathrm{m^2s^{-2}}$ in comparison with 0.18 $\mathrm{m^2s^{-2}}$). In that case, because the computation is made with fewer observations ($\approx$50 data), the agreement between randomized and exact values is slightly worse than with the complete set, but becomes quite good with only 5 perturbed analyses.

Figure 5 shows the same assimilation experiment as Fig. 3, but with a geographical projection $\mathbf{P}$ added at initial and final times. In particular, the error reduction covariance matrices respectively becomes $\mathbf{PF}^{-1}\mathbf{KHBF}^{-\mathrm{T}}\mathbf{P}^{\mathrm{T}}$ and $\mathbf{PMF}^{-1}\mathbf{KHBM}^{\mathrm{T}}\mathbf{F}^{-\mathrm{T}}\mathbf{P}^{\mathrm{T}}$. Here, the area of interest is chosen to be the second part of the domain where the sharp wind gradient occurs. Quite logically, the exact mean error reduction at initial time (Table 3) is equal to half the value computed without the geographical projector (0.30 $\mathrm{m^2s^{-2}}$ compared with 0.61 $\mathrm{m^2s^{-2}}$) and because all the variance is concentrated in the discontinuity

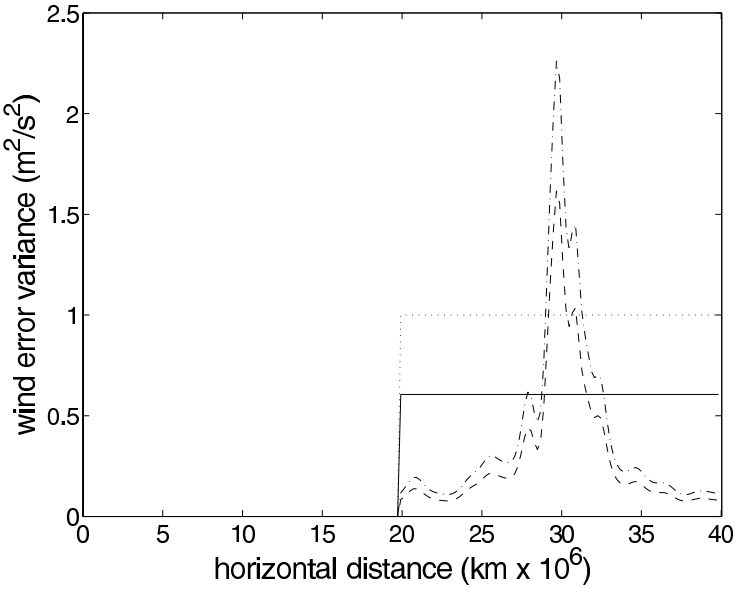

Figure 5. As Fig. 3, but the background errors and the error reduction are restricted to the second part of the domain. Corresponding exact and randomized error reductions are given in Table 3.

TABLE 3. AS TABLE 1, BUT WITH ERROR REDUCTIONS RESTRICTED TO THE SECOND PART OF THE DOMAIN AS SHOWN IN FIG. 5

	Exact values	1 perturbation	5 perturbations	10 perturbations
r^{a}	0.30	0.31	0.32	0.31
r^{f}	0.15	0.16	0.17	0.15

zone at final time, the value at 72 h for the area of interest is not very different from the total mean error reduction. Note that to obtain the randomized estimates of the error reduction, the operator **P** has to be introduced in the chain of operators. In particular, the reduction at final time becomes $r = \delta\mathbf{y}^{\mathrm{o}\mathrm{T}}\mathbf{R}^{-1}\mathbf{\Pi}_i\mathbf{HBL}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{PL}\delta\mathbf{x}^{\mathrm{a}}$, where $\mathbf{P}^{\mathrm{T}}\mathbf{P}$ simply corresponds to a zeroing outside the domain of interest. Again, the agreement between exact and randomized estimates is acceptable with only 1 observation perturbation and becomes quite good with 10 perturbations.

Finally, Fig. 6 shows the error reductions with the same geographical projector, but only for a subset of observations that are outside the domain of interest. As expected, the error reduction is nearly equal to zero. (There might be a slight contribution on the error reduction of observations placed on the border of the domain, due the width of the covariance function specified in matrix **B**.) However, the contribution of these observations is no longer negligible at final time (dashed line in Fig. 6); there is a clear upward and backward advection of information into the area of interest. Table 4 shows that the nearly zero mean error reduction at initial time is well estimated by the randomization method. Again, the small exact reduction is also well captured at final time with a single perturbation of the analysis and perfectly recovered with 10 perturbations.

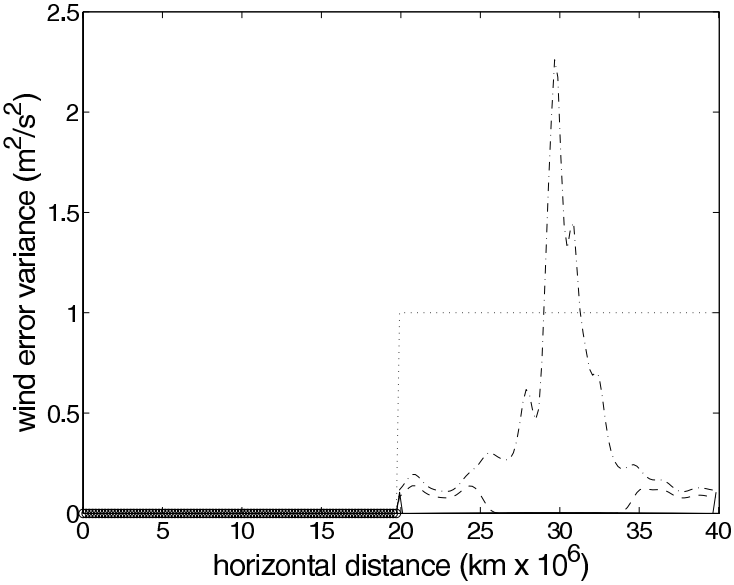

Figure 6. As Fig. 5, but the error reductions are determined with respect to observations placed in the first part of the domain (observation locations are shown as circles, but because of the high density of the observations, they appear as a dark bar in the left half of the figure). Corresponding exact and randomized error reductions are given in Table 4.

TABLE 4. AS TABLE 3, BUT WITH ERROR REDUCTIONS COMPUTED WITH RESPECT TO OBSERVATIONS PLACED IN THE FIRST PART OF THE DOMAIN AS SHOWN IN FIG. 6

	Exact values	1 perturbation	5 perturbations	10 perturbations
r^{a}	$\simeq$0.00	$\simeq$0.00	$\simeq$0.00	$\simeq$0.00
r^{f}	0.03	0.02	0.02	0.03

6. APPLICATION TO THE FRENCH ARPEGE 4D-VAR SCHEME

The computation of the error reduction brought by observations has also been implemented in the French ARPEGE 4D-Var scheme. Until now, only the reduction on analysis error has been investigated. The French 4D-Var assimilation is based on an incremental formulation (Courtier *et al.* 1994) and shares many aspects with the ECMWF 4D-Var analysis (Rabier *et al.* 2000). It is a global assimilation scheme which uses a large number ($\approx$300 000) different observations in each 4D-Var analysis.

The randomization procedure described in section 4 has been applied in order to compute total or partial reductions of error variance brought by observations. In order to make the results easier to interpret, the opportunity of performing a projection $\mathbf{P}$ of the model fields onto a particular field and level has been implemented. This corresponds to the particular case where $\mathbf{L} = \mathbf{P}$ in expression (7) of the randomized estimate of error variance reduction.

In a first step, tests have been made in order to evaluate the number of perturbations necessary to obtain a realistic estimation of the error variance reduction brought by the assimilation of observations. Figure 7 shows how total error reduction evolves as a function of the number of perturbed analyses for a particular date (5 January 2005). The figure shows that about five perturbed analyses are necessary in order to obtain a

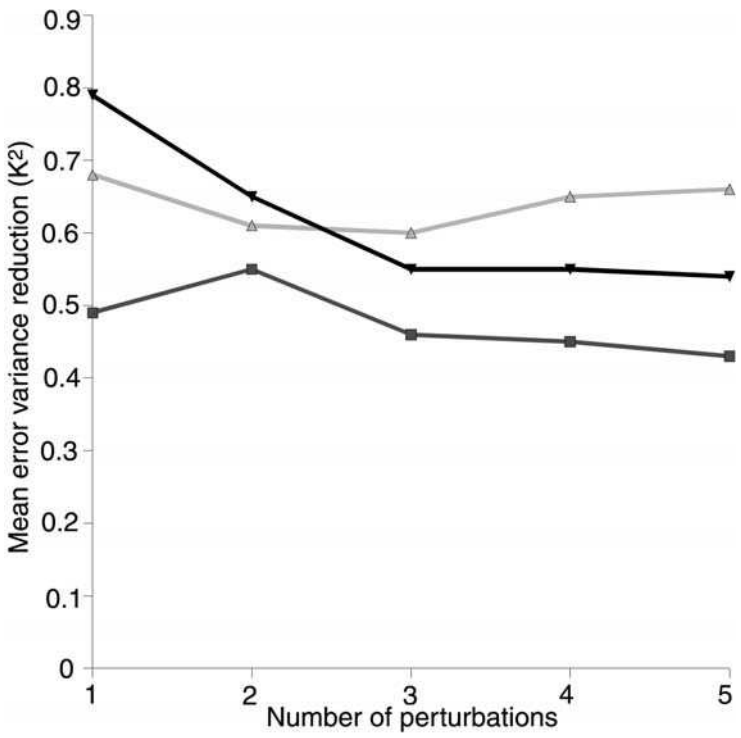

Figure 7. Mean error variance reduction per observation for the temperature field at 850 hPa, as a function of the number of perturbed analyses and for three different areas. Analysis date is 5 January 2005. The three areas are: northern hemisphere (grey line with triangles), southern hemisphere (black line with triangles), and tropics (line with squares).

stabilized estimation of the error variance reduction. This was also found sufficient for the computation of the error reduction brought by a particular subset of observations. As a consequence, all error variance reductions were computed with five perturbations of observations.

Figure 7 also shows that the mean error variance reduction (the total error variance reduction divided by the number of observations) for the particular temperature field at 850 hPa depends on the area. This can be mainly explained by the fact that the mean background variance error, and the error reduction, are largest for the northern hemisphere in winter, and smallest for the tropics.

Another validation of the retrieved error variance reductions has been performed by comparing them to the total background error variance. This total background error variance has been obtained by another Monte Carlo procedure (Fisher and Courtier 1995; Andersson *et al.* 2000). In this procedure, an ensemble of perturbations, $\boldsymbol{\epsilon}^{\mathrm{b}} = \mathbf{B}^{1/2}\boldsymbol{\eta}^{b}$, is built, where $\mathbf{B}^{1/2}$ is the operator corresponding to the square root of the background error covariance matrix, and $\boldsymbol{\eta}^{\mathrm{b}}$ is a vector of random numbers. The background error variances are deduced from the statistics of the squared background perturbations $\boldsymbol{\epsilon}^{\mathrm{b}}$. Table 5 gives such total background error variances for three different fields, and the corresponding reductions brought by the whole set of observations. The percentages of error variance reductions are similar for temperature at the two levels (850 and 250 hPa) and look sensible (around 30%). The lower percentage of error reduction for specific

TABLE 5. TOTAL BACKGROUND ERROR VARIANCE AND TOTAL ERROR VARIANCE REDUCTION FOR TEMPERATURE AT 850 AND 250 hPa AND SPECIFIC HUMIDITY AT 850 hPa

	Background error variance	Error variance reduction	Percentage reduction
T at 850 hPa	20864 (K^2)	6558 (K^2)	31
T at 250 hPa	18373 (K^2)	6819 (K^2)	39
q at 850 hPa	0.0263 (kg^2kg^{-2})	0.0038 (kg^2kg^{-2})	14

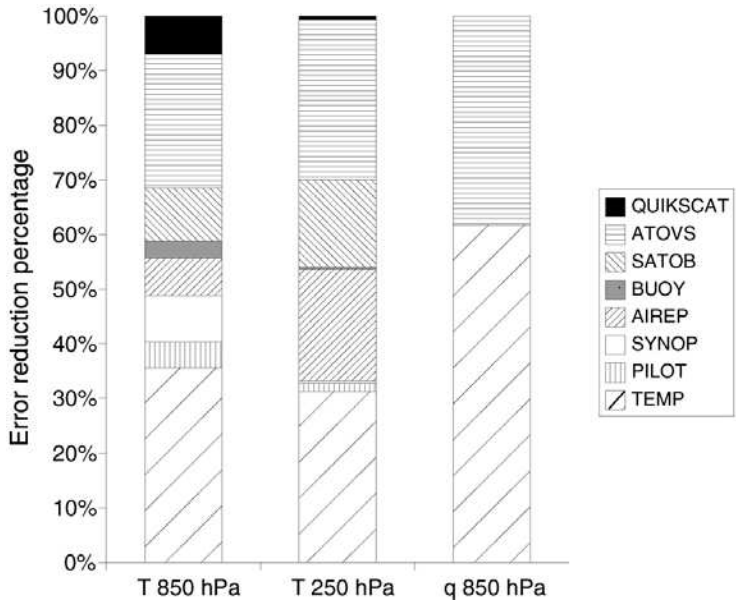

Figure 8. Percentage of error variance reduction due to different sets of observations and for three different fields (temperature at 850 hPa and 250 hPa, and specific humidity at 850 hPa).

humidity at 850 hPa is likely due to the fact that the analysis of humidity field is mainly univariate and then less efficient in terms of the reduction of the estimation error.

Once the randomization procedure is verified, it becomes possible to investigate the error variance reductions brought by different sets of observations. This has been done for three particular fields: temperature at 850 and 250 hPa, and specific humidity at 850 hPa. The observations used in the ARPEGE assimilation are:

- QUIKSCAT: surface wind observations from satellite scatterometers,
- ATOVS: radiances from polar-orbiting satellites,
- SATOB: velocity data inferred from cloud motions detected by geostationary satellite imagery,
- BUOY: ocean surface observations provided by buoys,
- AIREP: aircraft observations,
- SYNOP: land and ocean surface observations,
- PILOT: wind observations from manual tracking of pilot balloons,
- TEMP: upper-air observations of temperature, wind velocity and humidity by radiosondes.

Figure 8 shows the contribution of these different sets of observations to the reduction of the error for the selected fields. For the 850 hPa temperature field, the main contributions are provided by TEMP and ATOVS. This is logical since both temperature

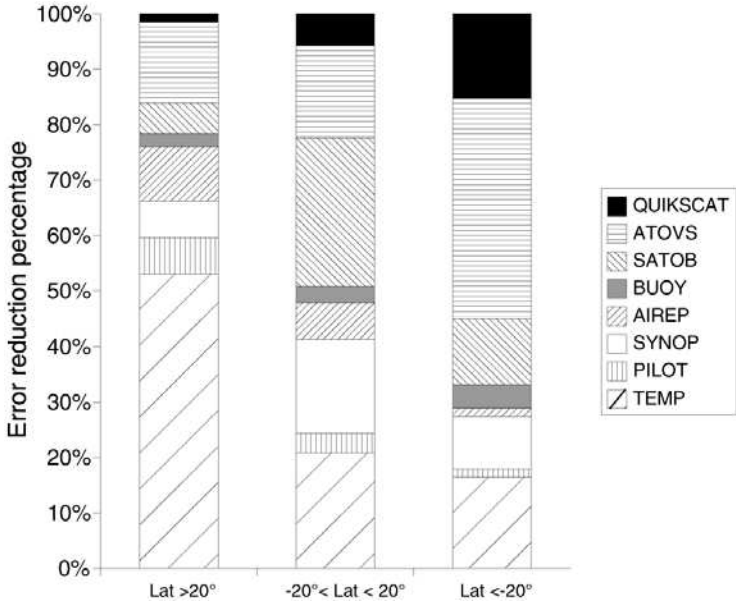

Figure 9. Percentage of error variance reduction due to different sets of observations for temperature at 850 hPa and for three different areas.

and radiance measurements are expected to have an impact on the analysis of temperature. Also, SATOB, AIREP and PILOT appear to bring a significant contribution to the analysis of temperature at 850 hPa, although they provide only wind velocity data. This is evidence of the impact of the multivariate formulation of the analysis through the description of covariances between parameters in the background error covariance matrix **B**.

Figure 8 also shows that the contributions of BUOY and SYNOP surface observations vanish for temperature at 250 hPa, which is to be expected.

Finally, it can be observed that the analysis of humidity at 850 hPa is only influenced by TEMP and ATOVS observations. This is also logical since there are no covariances between humidity and other fields in the ARPEGE covariance matrix for background errors and since TEMP and ATOVS are the only source of humidity information in the whole set of observations.

Figure 9 gives another view on the contribution of the subsets of observations for different areas. The contributions again appear consistent since, as expected, satellite data have proportionally more impact in the southern hemisphere than in the northern hemisphere, which is simply due to the fact that *in situ* data are less numerous in the southern hemisphere.

7. Conclusion

A method has been proposed to diagnose the impact of the complete set of observations or particular subsets of observations on an analysis itself and also on the forecast that can be run from this analysis. It consists of investigating the error variance reduction and thus is directly linked with the formalism of linear estimation. However, it has been shown that it could also be applied to a slightly nonlinear analysis method such as the incremental 4D-Var. Of course, the error reduction produced by such a

method assumes that the specified error covariances for background and observations are reasonably close enough to the actual ones.

It has also been shown with the simple Burger's equation that a realistic randomized estimate can be obtained even with a modest number of perturbations.

The first application of the randomization procedure to the French ARPEGE 4D-Var assimilation seems to confirm that reliable estimations of the error variance reduction due to different subsets of observations can be obtained with a limited number of perturbations on the analyses. Furthermore, the first results obtained in this operational framework show that useful and new information on the impact of observations on the analysis can be obtained from the application of such a method. Other results, not presented here, especially show that the randomization procedure has, for example, the capability to compare the impact of two different channels of the same satellite on the analysis. Another potential interest of the procedure is that it allows a better understanding of the impact of observations in a complex analysis scheme such as 4D-Var, and especially to diagnose the impact of observations along the 6-hour assimilation period. All these points are still under investigation and results should be presented in a further paper.

The first application of the randomization procedure in the ARPEGE 4D-Var operational framework is promising but also shows the difficulty of validating and interpreting the large number of results concerning the impact of observations on such an analysis. The application of the method to measure the impact of observations on forecasts will represent a next step, from which new and informative results are expected. However, this next step will require additional validation and understanding of the limitations of the approach proposed in this paper. In particular, an obvious limitation of the estimation of the impact of a subset of observations on the reduction of forecast error is imposed by the tangent linear hypothesis and by the fact that model error is not taken into account in the present approach. On another hand, such an approach is not intended to replace OSEs but rather to complement them.

Nevertheless, it is expected that the proposed approach could be useful in the field of numerical weather prediction and in particular in the framework of the international THORPEX project (Shapiro and Thorpe 2004), one of whose objectives is to optimize the design of observational networks.

Finally, the fact that the proposed procedure is based on perturbations of observations—and thus analyses—also suggests that it should be closely related to ensembles of analyses (Evensen 1994; Houtekamer and Mitchell 1998), which are more and more intended to be implemented in large NWP systems. It is felt that the proposed approach has to be seen from that emerging point of view.

Acknowledgements

We would like to thank C. Fischer and F. Rabier for fruitful discussions and their careful reading of the manuscript.

References

Andersson, E., Fisher, M., Munro, R. and McNally, A. 2000 Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc.*, **126,** 1455–1472

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

Bishop, C. H., Etherton, B. J. and Majumdar, S. J. 2001 Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspect. *Mon. Weather Rev.*, **129,** 420–436

Cardinali, C., Pezzuli, S. and Andersson, E. 2004 Influence matrix diagnostic of a data assimilation system. *Q. J. R. Meteorol. Soc.*, **130,** 2767–2786

Chapnik, B., Desroziers, G., Rabier, F. and Talagrand, O. 2004 Properties and first application of an error statistics tuning method in a variational assimilation. *Q. J. R. Meteorol. Soc.*, **130,** 2253–2275

Cohn, S. E. 1997 An introduction to estimation theory. *J. Meteorol. Soc. Japan*, **75,** 257–288

Courtier, P. and Talagrand, O. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Numerical results. *Q. J. R. Meteorol. Soc.,* **113,** 1329–1347

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994 A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120,** 1367–1387

Desroziers, G. and Ivanov, S. 2001 Diagnosis and adaptive tuning of information error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.*, **127,** 1433–1452

Dumelow, R. 2003 'Overview of observing system experiments'. Pp. 97–123 in Proceedings of seminar on recent developments in data assimilation for atmosphere and ocean, 8–12 September 2003. ECMWF, Reading, UK

Evensen, G. 1994 Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**(C5), 3104–3127

Fisher, M. 2003 'Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems'. Technical Memo. 397, ECMWF, Reading, UK

Fisher, M. and Courtier, P. 1995 'Estimating the covariance matrices of analysis and forecast error in variational data assimilation'. Technical Memo. 220, ECMWF, Reading, UK

Gilbert, J.-C. and Lemaréchal, C. 1989 Some numerical experiments with variable storage quasi-Newton algorithms. *Mathematical Programming*, **B25,** 407–435

Girard, D. 1989 A fast Monte Carlo cross-validation procedure for large least-squares problems with noisy data. *Numer. Math.*, **56,** 1–23

Houtekamer, P. L. and Mitchell, H. L. 1998 Data assimilation using an ensemble Kalman filter technique. *Mon. Weather Rev.*, **126,** 796–811

Langland, R. H. and Baker, N. L. 2004 Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A,** 189–201

Leutbecher, M. 2003 A reduced rank estimate of forecast error variance changes due to intermittent modifications of the observing network. *J. Atmos. Sci.*, **60,** 729–742

Lewis, J. M. and Derber, J. C. 1985 The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A,** 309–322

Liu, Z. 2002 'Influence de la résolution d'observations sur l'assimilation de données'. Doctoral thesis, Université Paul Sabatier, Toulouse, France

Rabier, F. and Courtier, P. 1992 Four-dimensional assimilation in the presence of baroclinic instability. *Q. J. R. Meteorol. Soc.*, **118,** 649–672

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F. and Simmons, A. 2000 The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126,** 1143–1170

Rodgers, C. D. 2000 *Inverse methods for atmospheric sounding; theory and practice.* World Scientific Publishing, London

Shapiro, M. A. and Thorpe, A. J. 2004 'THORPEX: A global atmospheric research programme for the beginning of the 21st century'. *WMO Bulletin*, **54,** 222–226. Also available at http://www.wmo.int/thorpex/ publications.html

Talagrand, O. 1999 'A posteriori verification of analysis and assimilation algorithms'. Pp. 17–28 in Proceedings of Workshop on diagnosis of data assimilation systems, 2–4 Nov. 1998, ECMWF, Reading, UK

Q. J. R. Meteorol. Soc. (2005), **131**, pp. 3385–3396 doi: 10.1256/qj.05.108

Diagnosis of observation, background and analysis-error statistics in observation space

By G. DESROZIERS†, L. BERRE, B. CHAPNIK and P. POLI
Météo-France, Centre National de Recherches Météorologiques, Toulouse, France

(Received 31 May 2005; revised 31 October 2005)

SUMMARY

Most operational assimilation schemes rely on linear estimation theory. Under this assumption, it is shown how simple consistency diagnostics can be obtained for the covariances of observation, background and estimation errors in observation space. Those diagnostics are shown to be nearly cost-free since they only combine quantities available after the analysis, i.e. observed values and their background and analysis counterparts in observation space. A first application of such diagnostics is presented on analyses provided by the French 4D-Var assimilation. A procedure to refine background and observation-error variances is also proposed and tested in a simple toy analysis problem. The possibility to diagnose cross-correlations between observation errors is also investigated in this same simple framework. A spectral interpretation of the diagnosed covariances is finally presented, which allows us to highlight the role of the scale separation between background and observation errors.

KEYWORDS: Estimation theory Optimality criterion Parameter estimation

1. INTRODUCTION

Most main operational assimilation systems are now based on the variational formalism (Lewis and Derber 1985; Courtier and Talagrand 1987, Rabier *et al.* 2000). Such a formalism allows the use of a large spectrum of observations and in particular satellite data that are not directly and linearly linked with model variables. However, those variational algorithms still rely on the theory of least-variance linear statistical estimation (Talagrand 1997). In the linear estimation theory, each set of information is given a weight proportional to the inverse of its specified error covariance. The pieces of information are classically given by observations and a background estimate of the state of the atmospheric flow. Analysis systems are then dependent on appropriate statistics for observation and background errors. Unfortunately those statistics are not perfectly known and their determination remains a major challenge in assimilation systems. One source of information on the observation and background errors is contained in the statistics of the innovations, that is the differences between observations and their background counterparts. Those statistics have for example been used by Hollingsworth and Lönnberg (1986), assuming that background errors carry cross-correlations while observation errors do not. From a slightly different point of view, Dee and da Silva (1999) have used a maximum likelihood method to estimate the information error statistics. Desroziers and Ivanov (2001) have proposed an approach based on a consistency criterion of the analysis relying on statistics of *observation-minus-analysis* differences to adapt observation-error statistics. The consistency criterion used in this method was defined by Talagrand (1999). Chapnik *et al.* (2004) investigated the properties of the algorithm and especially showed that it was equivalent to a maximum likelihood method, though less expensive to implement. Chapnik *et al.* (2006) also applied the same algorithm in an operational framework to tune observation-error variances.

This paper presents a set of diagnostics based on combinations of *observation-minus-background* (O–B), *observation-minus-analysis* (O–A) and *background-minus-analysis* (B–A) differences, which provide an additional consistency check of an analysis scheme.

† Corresponding author: Météo-France, CNRM, 42 av. G. Coriolis, 31057 Toulouse Cedex, France.
e-mail: gerald.desroziers@meteo.fr
© Royal Meteorological Society, 2005.

In section 2 the general least-variance statistical estimation framework is introduced, as well as the consistency diagnostics. A geometrical interpretation of those diagnostics is given in section 3. An application of the computation of the diagnostics on operational analyses given by a four-dimensional variational (4D-Var) assimilation scheme is presented in section 4. Then, it is shown in section 5 how the diagnostics can be used to optimize observation and background errors, and an application of such a tuning algorithm in a simple assimilation toy problem is presented. It is shown in section 6 how such a method can also be used to determine cross-correlations between the errors corresponding to different observations. A spectral interpretation of the diagnosed covariances is proposed in section 7. Conclusions and perspectives are given in section 8.

2. Diagnostics in observation space

(*a*) *Consistency diagnostic on innovations*

In statistical linear estimation theory, the expression of the analysed state $\mathbf{x}^{\mathrm{a}}$ is given by

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \delta\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}},$$

where $\mathbf{x}^{\mathrm{b}}$ is the background state, $\delta\mathbf{x}^{\mathrm{a}}$ the analysis increment,

$$\mathbf{K} = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$$

the gain matrix in the analysis process and $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$ the innovation vector (Talagrand 1997). The vector $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$ is the difference between observations $\mathbf{y}^{\mathrm{o}}$ and their background counterparts $H(\mathbf{x}^{\mathrm{b}})$, where H is the possibly nonlinear observation operator and $\mathbf{H}$ the matrix corresponding to the linearized version of H. $\mathbf{B}$ is the background-error covariance matrix.

From the definition of the innovation vector, the following sequence of relations can be derived:

$$\mathbf{d}_{\mathrm{b}}^{\mathrm{o}} = \mathbf{y}^{\mathrm{o}} - H(\mathbf{x}^{\mathrm{b}}) = \mathbf{y}^{\mathrm{o}} - H(\mathbf{x}^{\mathrm{t}}) + H(\mathbf{x}^{\mathrm{t}}) - H(\mathbf{x}^{\mathrm{b}}) \simeq \boldsymbol{\epsilon}^{\mathrm{o}} - \mathbf{H}\boldsymbol{\epsilon}^{\mathrm{b}},$$

where $\mathbf{x}^{\mathrm{t}}$ is the unknown true state, $\boldsymbol{\epsilon}^{\mathrm{o}}$ the vector of observation errors and $\boldsymbol{\epsilon}^{\mathrm{b}}$ the vector of background errors. Then, the covariance of innovations is

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] = E[\boldsymbol{\epsilon}^{\mathrm{o}}(\boldsymbol{\epsilon}^{\mathrm{o}})^{\mathrm{T}}] + \mathbf{H}E[\boldsymbol{\epsilon}^{\mathrm{b}}(\boldsymbol{\epsilon}^{\mathrm{b}})^{\mathrm{T}}]\mathbf{H}^{\mathrm{T}},$$

using the linearity of the statistical expectation operator E, and assuming that observation errors $\boldsymbol{\epsilon}^{\mathrm{o}}$ and background errors $\boldsymbol{\epsilon}^{\mathrm{b}}$ are uncorrelated.

As a consequence, it is easy to check that the relation

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} \tag{1}$$

should be fulfilled, if the covariance of observation errors, $\mathbf{R}$, and the covariance of background errors in observation space, $\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}}$, are correctly specified in the analysis. This is a classical result that provides a global check on the specification of those covariances (Andersson 2003).

It is shown below how additional relations can be obtained, which provide separate diagnostics on the background-, observation- and analysis-error statistics.

(*b*) *Consistency diagnostic on background errors*

From the previous expression for $\mathbf{x}^{\mathrm{a}}$, the $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}$ (A–B) differences in observation space can be written

$$\mathbf{d}_{\mathrm{b}}^{\mathrm{a}} = H(\mathbf{x}^{\mathrm{a}}) - H(\mathbf{x}^{\mathrm{b}}) \simeq \mathbf{H}\delta\mathbf{x}^{\mathrm{a}} = \mathbf{HKd}_{\mathrm{b}}^{\mathrm{o}}.$$

As a consequence, the cross-product between the $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}$ (A–B) differences in observation space and the $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$ (O–B) differences is

$$\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}} = \mathbf{HKd}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}.$$

Matrix $\mathbf{HK}$ is given by $\mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}$. Thus, the statistical expectation of this expression is given by

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}E[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}],$$

using the linearity of the statistical expectation operator E. It is easy to check that this whole expression simplifies to

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{HBH}^{\mathrm{T}}, \tag{2}$$

if matrix $\mathbf{HK} = \mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}$ is in agreement with the true covariances for background and observation errors.

This is a first additional diagnostic to the diagnostic on innovations. It provides a separate consistency check on background-error covariances in observation space.

(*c*) *Consistency diagnostic on observation errors*

Similarly, the $\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}$ (O–A) differences are given by

$$\begin{aligned}\mathbf{d}_{\mathrm{a}}^{\mathrm{o}} &= \mathbf{y}^{\mathrm{o}} - H(\mathbf{x}^{\mathrm{b}} + \delta\mathbf{x}^{\mathrm{a}})\\ &\simeq \mathbf{y}^{\mathrm{o}} - H(\mathbf{x}^{\mathrm{b}}) - \mathbf{HKd}_{\mathrm{b}}^{\mathrm{o}}\\ &= (\mathbf{I} - \mathbf{HK})\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}\\ &= \mathbf{R}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}},\end{aligned}$$

then the statistical expectation of the cross-product between the $\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}$ (O–A) differences and the $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$ (O–B) differences is

$$E[\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{R}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}E[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}],$$

which simplifies to

$$E[\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{R}, \tag{3}$$

if matrix $\mathbf{HK} = \mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}$ agrees with exact covariances for background and observation errors.

This is a second additional diagnostic providing a separate consistency check on observation-error covariances.

(*d*) *Diagnosis of analysis errors*

Finally, the cross-product between the $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}$ (A–B) differences in observation space and the $\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}$ (O–A) differences can also be derived:

$$\begin{aligned}\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{a}}^{\mathrm{o}})^{\mathrm{T}} &= \mathbf{HKd}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}(\mathbf{I} - \mathbf{HK})^{\mathrm{T}}\\ &= \mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{R}.\end{aligned}$$

Again, the statistical expectation of this expression simplifies to

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{a}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{R},$$

if matrix $\mathbf{HK} = \mathbf{HBH}^{\mathrm{T}}(\mathbf{HBH}^{\mathrm{T}} + \mathbf{R})^{-1}$ is in agreement with the true covariances for background and observation errors.

It is easy to show that the right-hand side of the last relation is an expression for $\mathbf{HAH}^{\mathrm{T}}$ (where $\mathbf{A}$ is the analysis-error covariance matrix in model space), if $\mathbf{HBH}^{\mathrm{T}}$ and $\mathbf{R}$ are correctly specified in the analysis. Thus, in that case, the following relation should hold:

$$E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{a}}^{\mathrm{o}})^{\mathrm{T}}] = \mathbf{HAH}^{\mathrm{T}}. \tag{4}$$

This is a third additional diagnostic providing information on analysis errors in observation space.

3. A GEOMETRICAL INTERPRETATION

From what has been shown in the previous section, relations (1) to (4) should be fulfilled in an optimal linear analysis, with linearized observation operators. They can be summarized as follows:

$$\begin{aligned} E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] &= \mathbf{HBH}^{\mathrm{T}}, & E[\mathbf{d}_{\mathrm{a}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] &= \mathbf{R}, \\ E[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})^{\mathrm{T}}] &= \mathbf{HBH}^{\mathrm{T}} + \mathbf{R}, & E[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}(\mathbf{d}_{\mathrm{a}}^{\mathrm{o}})^{\mathrm{T}}] &= \mathbf{HAH}^{\mathrm{T}}. \end{aligned}$$

These relations are matricial, that is to say that they should be also true for cross-covariances between differences associated with different observations.

A geometrical interpretation of these relations can be proposed in the space of the eigenvectors $\mathbf{V}$ of matrix $\mathbf{HK}$, such as $\mathbf{HK} = \mathbf{V\Lambda V}^{\mathrm{T}}$, where $\mathbf{\Lambda}$ is the diagonal matrix of the eigenvalues of $\mathbf{HK}$. With this decomposition of matrix $\mathbf{HK}$, the vector $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}$ of the differences between analysis and background in observation space can be rewritten: $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}} = \mathbf{H}\delta\mathbf{x}^{\mathrm{a}} = \mathbf{HKd}_{\mathrm{b}}^{\mathrm{o}} = \mathbf{V\Lambda V}^{\mathrm{T}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$. Then, the projection of $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}$ onto the eigenvectors of $\mathbf{HK}$ is given by $\mathbf{V}^{\mathrm{T}}\mathbf{d}_{\mathrm{b}}^{\mathrm{a}} = \mathbf{V}^{\mathrm{T}}\mathbf{V\Lambda V}^{\mathrm{T}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}} = \mathbf{\Lambda V}^{\mathrm{T}}\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$. If $[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}]_i$ and $[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}]_i$ stand respectively for the projections of $\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}$ and $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$ onto a particular eigenvector $\mathbf{v}_i$ of $\mathbf{HK}$, it follows that $[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}]_i = \lambda_i[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}]_i$, where λ_i is the corresponding eigenvalue of $\mathbf{HK}$. (The notations $[\mathbf{d}_{\mathrm{b}}^{\mathrm{a}}]_i$, $[\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}]_i$ are used here to specify that these two vectors are unidimensional vectors.)

In Figure 1, $H(\mathbf{x}^{\mathrm{t}})_i$, $\mathbf{y}^{\mathrm{o}}{}_i$ and $H(\mathbf{x}^{\mathrm{b}})_i$ respectively stand for the projections onto a particular eigenvector $\mathbf{v}_i$ of the true, observed and background equivalents of $\mathbf{x}$ at observation locations. The triangle $H(\mathbf{x}^{\mathrm{b}})_i$, $H(\mathbf{x}^{\mathrm{t}})_i$, $\mathbf{y}_i^{\mathrm{o}}$ is right in $H(\mathbf{x}^{\mathrm{t}})_i$ since the projections $[\boldsymbol{\epsilon}^{\mathrm{o}}]_i$ and $[\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{b}}]_i$ of observation errors and background errors onto vector $\mathbf{v}_i$ are assumed to be uncorrelated and hence orthogonal. This orthogonality is defined for the particular scalar product $\langle[\boldsymbol{\epsilon}_1], [\boldsymbol{\epsilon}_2]\rangle = E(\epsilon_1\epsilon_2)$, where $[\boldsymbol{\epsilon}_1]$ and $[\boldsymbol{\epsilon}_2]$ are two vectors of random errors, with respective components ϵ_1 and ϵ_2 being two random variables. It means that the orthogonality of $[\boldsymbol{\epsilon}^{\mathrm{o}}]_i$ and $[\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{b}}]_i$ is only true from a statistical point of view. With this definition of the scalar product, the angle between two random error vectors is also directly linked with the correlation of the errors; a zero angle and a right angle respectively correspond to perfect correlation and decorrelation and an angle between 0 and $\pi/2$ to an intermediate correlation.

It is a classical result that the analysis error $\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{a}}$ is also orthogonal to the innovation vector $\mathbf{d}_{\mathrm{b}}^{\mathrm{o}}$, again from a statistical point of view. Since Fig. 1 corresponds to a projection

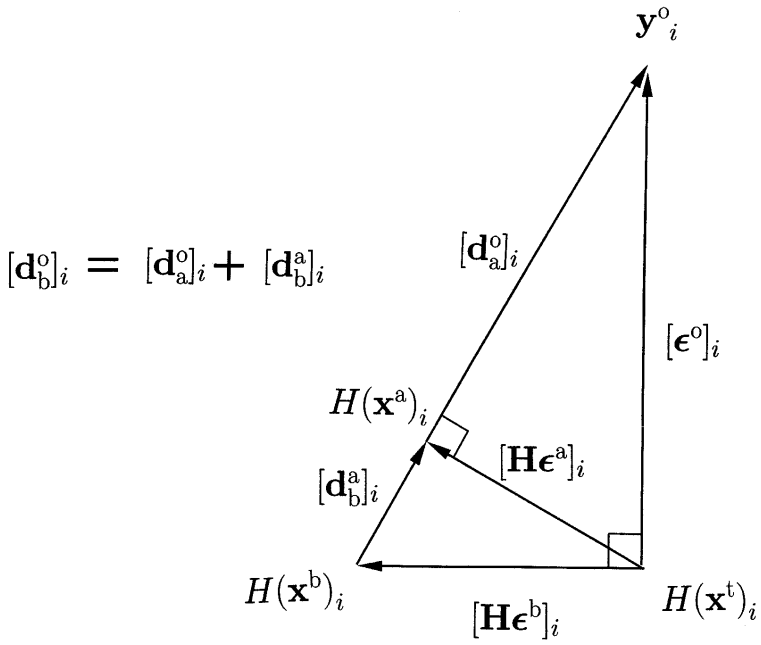

Figure 1. Geometrical representation of the analysis projected onto a particular eigenvector $\mathbf{v}_i$ of matrix **HK**.

onto a particular eigenvector $\mathbf{v}_i$, it follows that $[H(\mathbf{x}^{\mathrm{b}})_i H(\mathbf{x}^{\mathrm{a}})_i] = [\mathbf{d}^{\mathrm{a}}_{\mathrm{b}}]_i = \lambda_i [\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i$ and then $H(\mathbf{x}^{\mathrm{a}})_i$ is on the line defined by the point $H(\mathbf{x}^{\mathrm{b}})_i$ and the vector $[\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i$. The vector $[\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{a}}]_i$ is thus orthogonal to $[\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i$ in $H(\mathbf{x}^{\mathrm{a}})_i$.

Hence, the application of the Pythagoras theorem to this triangle implies that $\|[\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i\|^2 = \|[\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{b}}]_i\|^2 + \|[\boldsymbol{\epsilon}^{\mathrm{o}}]_i\|^2$, where the norm $\|\ \|$ is associated with the previous scalar product. This relation corresponds to the first classical diagnostic on innovation covariances.

From the application of Euclid's theorems in a right triangle, the three additional relations can be written:

$$\langle [\mathbf{d}^{\mathrm{a}}_{\mathrm{b}}]_i, [\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i \rangle = \|[\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{b}}]_i\|^2,$$
$$\langle [\mathbf{d}^{\mathrm{o}}_{\mathrm{a}}]_i, [\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i \rangle = \|[\boldsymbol{\epsilon}^{\mathrm{o}}]_i\|^2,$$
$$\langle [\mathbf{d}^{\mathrm{a}}_{\mathrm{b}}]_i, [\mathbf{d}^{\mathrm{o}}_{\mathrm{b}}]_i \rangle = \|[\mathbf{H}\boldsymbol{\epsilon}^{\mathrm{a}}]_i\|^2,$$

corresponding to the diagnostics of background, observation and analysis-error variances, respectively.

In the case where background and observation-error covariances are homogeneous, and the data density is uniform, the eigenvectors of matrix **HK** are the spectral harmonics (Fourier on the plane, spherical harmonics on the sphere). Then, Fig. 1 shows a projection onto a particular spectral component. This case will be further developed in section 7.

4. Application of the diagnostics to ARPEGE 4D-Var analyses

The diagnostics shown in section 2 potentially provide information on the full covariances of observation, background and analysis errors in observation space.

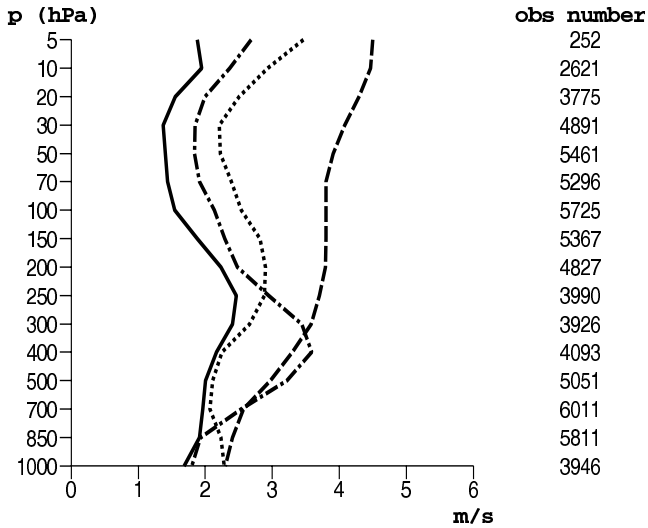

Figure 2. Vertical profiles of diagnosed square roots of background (solid) and observation (dotted) error variances for radiosonde wind observations in the northern hemisphere, compared with profiles for corresponding background (dash-dotted) and observation errors (dashed). All values are in m s^{-1}. The numbers of observations used to compute statistics are shown on the right.

One first application of these diagnostics is to diagnose observation and background-error variances. Thus for any subset of observations i with p_i observations, it is possible simply to compute the quantities

$$(\widetilde{\sigma_i^{\mathrm{b}}})^2 = (\mathbf{d}_{\mathrm{b}}^{\mathrm{a}})_i^{\mathrm{T}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})_i/p_i = \sum_{j=1}^{p_i}(y_j^{\mathrm{a}} - y_j^{\mathrm{b}})(y_j^{\mathrm{o}} - y_j^{\mathrm{b}})/p_i$$

$$(\widetilde{\sigma_i^{\mathrm{o}}})^2 = (\mathbf{d}_{\mathrm{a}}^{\mathrm{o}})_i^{\mathrm{T}}(\mathbf{d}_{\mathrm{b}}^{\mathrm{o}})_i/p_i = \sum_{j=1}^{p_i}(y_j^{\mathrm{o}} - y_j^{\mathrm{a}})(y_j^{\mathrm{o}} - y_j^{\mathrm{b}})/p_i,$$

where y_j^{o} is the value of observation j and y_j^{b}, y_j^{a} respectively their background and analysis counterparts. The quantities $(\widetilde{\sigma_i^{\mathrm{b}}})^2$ and $(\widetilde{\sigma_i^{\mathrm{o}}})^2$ are the diagnosed values of background and observation errors that can be different from the specified values in the analysis. They correspond to the use of relations (2) and (3) respectively, but for mean diagonal elements of those matrices only. These computations are nearly cost-free and can be performed, a posteriori, using one or several analyses.

Such computations have been performed on analyses produced by the French operational ARPEGE† 4D-Var assimilation. This 4D-Var assimilation is based on an incremental formulation (Courtier *et al.* 1994) and shares many aspects with the 4D-Var analysis of the European Centre for Medium-range Weather Forecasts (Rabier *et al.* 2000). Figure 2 shows an example of diagnosed background and observation variances for wind observations given by radiosondes in the northern hemisphere. These statistics have been computed after ten 4D-Var analyses. They show that both background and observation errors seem to be overestimated in the analysis. This is consistent for observation errors with what has been found by Chapnik *et al.* (2006) using a different diagnostic. On the other hand, similar diagnosed profiles have been retrieved for the

† Action de Recherche Petite Echelle Grande Echelle.

TABLE 1. TRUE AND TUNED VALUES OF σ_o AND σ_b AFTER FIXED-POINT ITERATIONS

	True value	Initial value	Iteration 1	2	3	4	5
σ_o	2	1	1.73	1.89	1.95	1.97	1.98
σ_b	1	2	1.41	1.19	1.10	1.07	1.03

southern hemisphere (not shown), thus giving some confidence in those diagnostics. The same kind of diagnostics have been produced for all observations taken into account in the ARPEGE 4D-Var. Results (not presented here) confirm the overestimation of background and observation error for most of the observations.

5. TUNING OF ERROR VARIANCES

Since background and observation errors are expected to be incorrectly specified in an operational analysis, a method can be envisaged to tune them. The rationale of such a tuning procedure is to find the values of the σ_i^b and σ_i^o, for the different subsets i of observations such that those values fulfil the relations $(\sigma_i^b)^2 = (\mathbf{d}_b^a)_i^T(\mathbf{d}_b^o)_i/p_i$ and $(\sigma_i^o)^2 = (\mathbf{d}_a^o)_i^T(\mathbf{d}_b^o)_i/p_i$. This is a nonlinear problem since the $(\mathbf{d}_b^a)_i$ and $(\mathbf{d}_a^o)_i$ depend themselves on the σ_i^b and σ_i^o values. However the form of those nonlinear equations suggests the use of an iterative fixed-point method to solve this tuning problem.

This iterative procedure is similar to the procedure proposed by Desroziers and Ivanov (2001) to solve the same kind of problem but with a different optimality criterion.

A preliminary test of the previous tuning algorithm has been made in a toy problem given by a spectral analysis on a circular domain (say on an earth meridian) and also used by Desroziers and Ivanov (2001). The length of the domain is set to 40 000 km and the truncation to 200 corresponding to $n = 401$ spectral coefficients. The background-error covariance matrix $\mathbf{B}$ is built in spectral space from a Gaussian structure function in physical space. Assuming homogeneity on the domain makes the $\mathbf{B}$ matrix diagonal, so that the diagonal of $\mathbf{B}$ is given by the Fourier transform of the Gaussian correlation in physical space. Here the length-scale of the Gaussian correlation is set to 300 km. The analysis problem is solved with $p = 401$ observations and a diagonal observation-error covariance matrix $\mathbf{R}$, assuming that error observations are uncorrelated. Observations have the same nature as the state variable $\mathbf{x}$. Thus the observation operator $\mathbf{H}$ only involves interpolation of $\mathbf{x}$ at observation locations. Both background and observations are simulated in agreement with the corresponding $\mathbf{B}$ and $\mathbf{R}$ covariance matrices with homogeneous values $\sigma^b = 1$ and $\sigma^o = 2$.

Table 1 shows that, starting from erroneous values $\sigma^b = 2$ and $\sigma^o = 1$, the fixed-point iterative algorithm allows us to recover a good approximation of the true values with only a few iterations.

6. DIAGNOSIS OF OBSERVATION ERROR CROSS-COVARIANCES

The diagnostic $E[\mathbf{d}_a^o(\mathbf{d}_b^o)^T] = \mathbf{R}$ on observation errors, introduced in section 2, can potentially provide some information on cross-correlations between two different sets of observations, for instance observations issued from different channels of the same satellite instrument. The capability of such a diagnostic to recover observation error cross-correlations is investigated here with the same previous analysis toy problem but

TABLE 2. TRUE, INITIAL AND DIAGNOSED VALUES OF OBSERVATION ERROR CROSS-COVARIANCES FOR THREE SIMULATED OBSERVATION SETS

	1			2			3		
Set	True	Init.	Diag.	True	Init.	Diag.	True	Init.	Diag.
1	1.00	1.00	1.00	1.00	0.00	1.01	0.10	0.00	0.10
2				4.00	1.00	3.73	−0.20	0.00	−0.18
3							0.25	1.00	0.25

with the addition of a vertical dimension. The analysis domain is now a periodic vertical plane with three levels with 5 km separation. All parameters are the same as previously, and the additional vertical length-scale of background errors is set to 3 km. At each horizontal observation location, a set of three simulated observations y_1^o, y_2^o and y_3^o are defined as simple combinations of the state variable at the three levels:

$$y_1^o = 0.75x_1 + 0.25x_2, \quad y_2^o = 0.10x_1 + 0.80x_2 + 0.10x_3, \quad y_3^o = 0.30x_2 + 0.70x_3,$$

where x_1, x_2 and x_3 are the state variable values at the three levels. This is done to mimic, in a rough way, satellite observations integrated within the vertical. These observations are simulated with correlated observation errors as given in Table 2.

The diagnosed values of observation error cross-covariances (Table 2) show that such a simple diagnostic is able to recover most of the information on observation-error covariances starting from a mis-specified representation of those covariances (with no cross-correlations and homogeneous error variance).

7. A SPECTRAL INTERPRETATION OF THE DIAGNOSED COVARIANCES

(*a*) *The adjustment equations of the spectral variances*

The diagnosed covariance matrices in Eqs. (2) and (3) may be seen as some adjusted covariance estimates, noted $\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^T$ and $\widetilde{\mathbf{R}}$ respectively:

$$\begin{aligned} E[\mathbf{d}_b^a(\mathbf{d}_b^o)^T] &= \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^T = \mathbf{H}\mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{H}\mathbf{B}^*\mathbf{H}^T + \mathbf{R}^*) \\ E[\mathbf{d}_a^o(d_b^o)^T] &= \widetilde{\mathbf{R}} \qquad = \mathbf{R}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{H}\mathbf{B}^*\mathbf{H}^T + \mathbf{R}^*), \end{aligned}$$

where $\mathbf{B}^*$, $\mathbf{R}^*$ are the true covariance matrices.

Following Hollingsworth (1987) and Daley (1991, pp. 125–128), one may consider for example the 1D or 2D case where the background and observation-error covariances are homogeneous, and the data density is uniform.

In this case, the different covariance matrices have common eigenvectors, which are the spectral modes: $\mathbf{H}\mathbf{B}\mathbf{H}^T = \mathbf{S}\Lambda^b\mathbf{S}^T$, $\mathbf{R} = \mathbf{S}\Lambda^o\mathbf{S}^T$, where $\mathbf{S}$ is an orthogonal matrix whose columns are the common eigenvectors of $\mathbf{H}\mathbf{B}\mathbf{H}^T$ and $\mathbf{R}$, and Λ^b, Λ^o are diagonal matrices which contain the corresponding eigenvalues. The eigenvectors are the spectral harmonics (Fourier on the plane, spherical harmonics on the sphere), and $\mathbf{S}$ is the inverse spectral transform. Moreover, the eigenvalues correspond to the variances of the spectral components.

Similar equations hold for the diagnosed and true covariance matrices respectively. The equation of the diagnosed background-error covariance matrix can thus be written as follows:

$$\begin{aligned} \mathbf{S}\widetilde{\boldsymbol{\Lambda}}^b\mathbf{S}^T &= \mathbf{S}\boldsymbol{\Lambda}^b\mathbf{S}^T(\mathbf{S}\boldsymbol{\Lambda}^b\mathbf{S}^T + \mathbf{S}\boldsymbol{\Lambda}^o\mathbf{S}^T)^{-1}(\mathbf{S}\boldsymbol{\Lambda}^{b*}\mathbf{S}^T + \mathbf{S}\boldsymbol{\Lambda}^{o*}\mathbf{S}^T) \\ &= \mathbf{S}\boldsymbol{\Lambda}^b\mathbf{S}^T\mathbf{S}^{T-1}(\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o)^{-1}\mathbf{S}^{-1}\mathbf{S}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*})\mathbf{S}^T \\ &= \mathbf{S}\boldsymbol{\Lambda}^b(\boldsymbol{\Lambda}^b + \boldsymbol{\Lambda}^o)^{-1}(\boldsymbol{\Lambda}^{b*} + \boldsymbol{\Lambda}^{o*})\mathbf{S}^T. \end{aligned}$$

The same kind of relations can be written for the diagnosed observation-error covariance matrix. This provides the final eigenvalue relations:

$$\begin{aligned}\widetilde{\boldsymbol{\Lambda}}^{\mathrm{b}} &= \boldsymbol{\Lambda}^{\mathrm{b}}(\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}})^{-1}(\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*})\\ \widetilde{\boldsymbol{\Lambda}}^{\mathrm{o}} &= \boldsymbol{\Lambda}^{\mathrm{o}}(\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}})^{-1}(\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}),\end{aligned}$$

where $\{\boldsymbol{\Lambda}^{\mathrm{b}*}, \boldsymbol{\Lambda}^{\mathrm{o}*}\}$, $\{\boldsymbol{\Lambda}^{\mathrm{b}}, \boldsymbol{\Lambda}^{\mathrm{o}}\}$ and $\{\widetilde{\boldsymbol{\Lambda}}^{\mathrm{b}}, \widetilde{\boldsymbol{\Lambda}}^{\mathrm{o}}\}$ respectively stand for the eigenvalues of the true, specified and diagnosed covariance matrices for background and observation errors.

(*b*) *Two visions of the adjustment equations*

A first insight into the previous relations is to note that the true innovation variance spectrum $\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}$ is multiplied by two filtering ratios

$$\mathbf{F}^{\mathrm{b}} = \frac{\boldsymbol{\Lambda}^{\mathrm{b}}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}}, \quad \mathbf{F}^{\mathrm{o}} = \frac{\boldsymbol{\Lambda}^{\mathrm{o}}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}}$$

to provide the respective (adjusted) variance spectra of $\boldsymbol{\epsilon}^{\mathrm{b}}$ and $\boldsymbol{\epsilon}^{\mathrm{o}}$:

$$\widetilde{\boldsymbol{\Lambda}}^{\mathrm{b}} = \frac{\boldsymbol{\Lambda}^{\mathrm{b}}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}}(\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}), \quad \widetilde{\boldsymbol{\Lambda}}^{\mathrm{o}} = \frac{\boldsymbol{\Lambda}^{\mathrm{o}}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}}(\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}).$$

These two filters are represented in Fig. 3(a) for the example evoked in section 5. Due to the shape of the two kinds of error correlation, $\mathbf{F}^{\mathrm{b}}$ and $\mathbf{F}^{\mathrm{o}}$ are respectively low-pass and high-pass filters. Figures 3(c) and (d) illustrate that applying $\mathbf{F}^{\mathrm{b}}$, $\mathbf{F}^{\mathrm{o}}$ to $(\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*})$ allows us to extract its large-scale and small-scale components respectively. They correspond to the (estimated) contributions of the background and observation errors, respectively. These provide some adjusted variance spectra, which are closer to the true spectra than the specified ones. This is also consistent with the fact that the equations indicate, in particular, that if the analysis is optimal, then these two filtering steps will provide the two exact error variance spectra.

Another complementary insight is to notice that the adjustments amount to scaling each specified information error spectral variance by the ratio between the exact and specified innovation variances:

$$\widetilde{\boldsymbol{\Lambda}}^{\mathrm{b}} = \boldsymbol{\Lambda}^{\mathrm{b}}\frac{\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}}, \quad \widetilde{\boldsymbol{\Lambda}}^{\mathrm{o}} = \boldsymbol{\Lambda}^{\mathrm{o}}\frac{\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}}.$$

The perfect adjustment ratios would instead be $(\boldsymbol{\Lambda}^{\mathrm{b}*}/\boldsymbol{\Lambda}^{\mathrm{b}})$ and $(\boldsymbol{\Lambda}^{\mathrm{o}*}/\boldsymbol{\Lambda}^{\mathrm{o}})$ respectively. However, as illustrated in Fig. 3(b), the innovation misfit ratio $\boldsymbol{N} = (\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*})/(\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}})$ is close to $(\boldsymbol{\Lambda}^{\mathrm{b}*}/\boldsymbol{\Lambda}^{\mathrm{b}})$ at the large scales (i.e. where the background errors have their largest amplitudes), and it is close to $(\boldsymbol{\Lambda}^{\mathrm{o}*}/\boldsymbol{\Lambda}^{\mathrm{o}})$ over a wide range of small scales (i.e. where the observation errors predominate). In other words, the close links between $\boldsymbol{N}$ and $(\boldsymbol{\Lambda}^{\mathrm{b}*}/\boldsymbol{\Lambda}^{\mathrm{b}})$ or $(\boldsymbol{\Lambda}^{\mathrm{o}*}/\boldsymbol{\Lambda}^{\mathrm{o}})$, as a function of scale, indicate that the adjustments are done in a relevant scale-dependent way.

The two visions are consistent, in the sense that they both indicate that the adjustments are particularly relevant when the background errors and the observation errors tend to predominate in different scales. This is summarized by the fact that

$$\frac{\boldsymbol{\Lambda}^{\mathrm{b}*} + \boldsymbol{\Lambda}^{\mathrm{o}*}}{\boldsymbol{\Lambda}^{\mathrm{b}} + \boldsymbol{\Lambda}^{\mathrm{o}}} \simeq \frac{\boldsymbol{\Lambda}^{\mathrm{b}*}}{\boldsymbol{\Lambda}^{\mathrm{b}}}$$

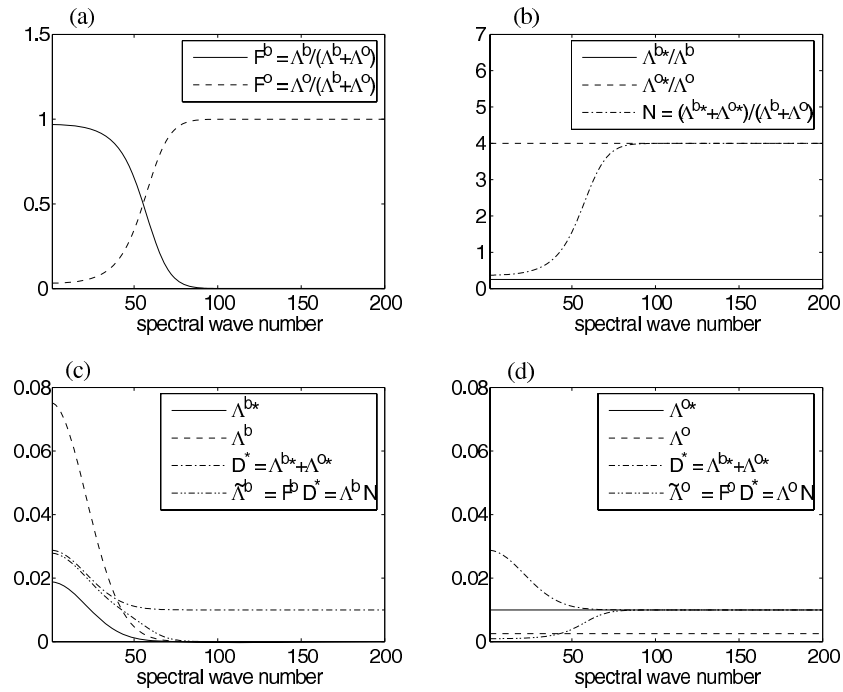

Figure 3. Spectral interpretation of diagnosed covariances: (a) filtering functions, (b) misfit ratios, (c) background error, and (d) observation error. See text for definitions.

in the large scales if both $\mathbf{\Lambda}^{b*} \gg \mathbf{\Lambda}^{o*}$ and $\mathbf{\Lambda}^{b} \gg \mathbf{\Lambda}^{o}$ and

$$\frac{\mathbf{\Lambda}^{b*} + \mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{b} + \mathbf{\Lambda}^{o}} \simeq \frac{\mathbf{\Lambda}^{o*}}{\mathbf{\Lambda}^{o}}$$

in the small scales if both $\mathbf{\Lambda}^{b*} \ll \mathbf{\Lambda}^{o*}$ and $\mathbf{\Lambda}^{b} \ll \mathbf{\Lambda}^{o}$.

The adjusted standard deviations that were presented in section 5 are simply the square root of the sum of the different (adjusted) spectral variances. The relevance of the standard deviation adjustments is therefore consistent with the adjustments of the spectral variances.

From the previous discussion, it also appears that the adjustment of background and observation-error variances is only relevant if those errors have different structures. Hence, the application of such an adjustment will not work in the case where the two spectra (either $\mathbf{\Lambda}^{b*}$ and $\mathbf{\Lambda}^{o*}$ or $\mathbf{\Lambda}^{b}$ and $\mathbf{\Lambda}^{o}$) are proportional.

However, if the background and observation-error spectra are sufficiently different (both in the exact and specified statistics), the adjustment will be able to modify the variances in a correct way even if the correlations are not perfectly specified.

8. Conclusion

On the basis of linear estimation theory, simple consistency diagnostics should be fulfilled in an optimal analysis. These diagnostics can potentially provide information on imperfectly known observation and background-error statistics. Another advantage of these diagnostics is that they are nearly cost-free and can be applied to any analysis scheme.

The application of the computation of the diagnostics to analyses issued from the operational French 4D-Var system shows likely diagnosed values for observation and background errors. Even if the values of background errors cannot be directly used in a model-space assimilation scheme, the study of these errors can be quite useful to understand the relative impact in the analysis of observations that are not directly related to the state variables. This is particularly the case for satellite data, for which the diagnosed errors can also be compared to randomized estimates of $\mathbf{HBH}^{\mathrm{T}}$, where $\mathbf{B}$ is the specified covariance background-error matrix in the analysis, as proposed by Andersson *et al.* (2000). Since the observation operator $\mathbf{H}$ includes the model integration in a 4D-Var scheme, the proposed diagnostic can be similarly used to diagnose the implicit evolution of background errors in 4D-Var.

Furthermore, it has been shown that it is possible to adapt the values specified in an analysis scheme by an iterative method. This can be useful to adapt observation errors but also background errors in observation space that are classically used in first-guess quality control of observations.

The use of such consistency diagnostics also seems to be a promising way to tackle the problem of the estimation of correlation between observation errors.

Another domain of interest is the use of the diagnostic of estimation error in observation space, which will have to be investigated in the future.

Nevertheless, such diagnostics also have to be understood from theoretical and practical points of view. In particular it has been shown that a spectral vision can be helpful in this perspective, by highlighting the role of scale separation between background and observation errors.

REFERENCES

Andersson, E. 2003 'Modelling of innovation statistics'. Pp. 153–164 of Proceedings of Workshop on recent developments in data assimilation for atmosphere and ocean, ECMWF, Reading, UK

Andersson, E., Fisher, M., Munro, R. and McNally, A. 2000 Diagnosis of background errors for radiances and other observable quantities in a variational data assimilation scheme, and the explanation of a case of poor convergence. *Q. J. R. Meteorol. Soc.*, **126,** 1455–1472

Chapnik, B., Desroziers, G., Rabier, F. and Talagrand, O. 2004 Properties and first application of an error statistics tuning method in variational assimilation. *Q. J. R. Meteorol. Soc.*, **130,** 2253–2275

2006 Diagnosis and tuning of observational error statistics in a quasi-operational data assimilation setting. *Q. J. R. Meteorol. Soc.*, **132,** 543–565

Courtier, P. and Talagrand, O. 1987 Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Q. J. R. Meteorol. Soc.*, **113,** 1329–1347

Courtier, P., Thépaut, J.-N. and Hollingsworth, A. 1994 A strategy for operational implementation of 4D-Var, using an incremental approach. *Q. J. R. Meteorol. Soc.*, **120,** 1367–1387

Daley, R. 1991 *Atmospheric data analysis*. Cambridge atmospheric and space series. Cambridge University Press, UK

Dee, D. P. and da Silva, A. 1999 Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Weather Rev.*, **127,** 1822–1834

Desroziers, G. and Ivanov, S. 2001 Diagnosis and adaptive tuning of observation-error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.*, **127,** 1433–1452

Hollingsworth, A. 1987 'Objective analysis for numerical weather prediction'. Pp. 11–59 in Proceedings of WMO/IUGG Symposium on short- and medium-range numerical weather prediction, Tokyo, 4–8 August 1986, Ed. T. Matsuno. *J. Meteorol. Soc. Japan*, special volume

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

Lewis, J. M. and Derber, J. C. 1985 The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A,** 309–322

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F. and Simmons, A. 2000 The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q. J. R. Meteorol. Soc.*, **126,** 1143–1170

Talagrand, O. 1997 Assimilation of observations; an introduction. *J. Meteorol. Soc. Japan*, **75,** 191–209

1999 'A posteriori verification of analysis and assimilation algorithms'. Pp. 17–28 in Proceedings of Workshop on diagnosis of data assimilation systems, 2–4 November 1998, ECMWF, Reading, UK