Apport des sondeurs hyperspectraux au sein du système global d'observation pour l'étude des dépressions météorologiques

Nadia Fourrié

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

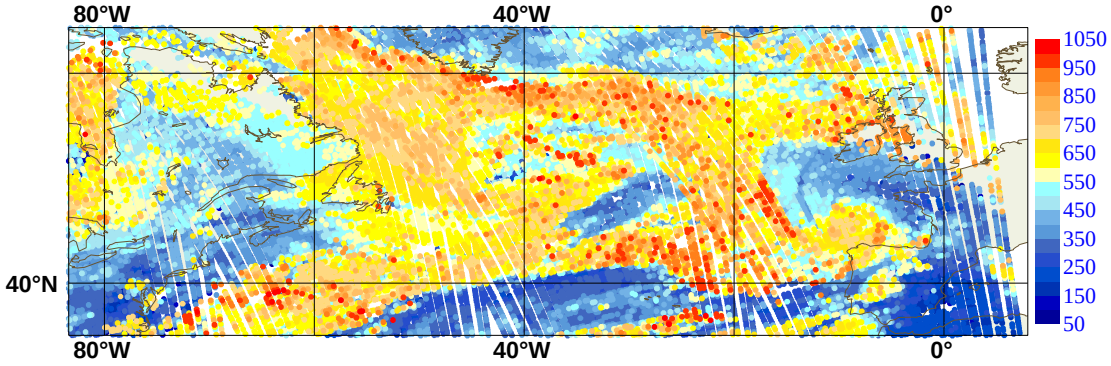
Présenté à L'UNIVERSITÉ TOULOUSE III
Spécialité : Sciences de l'Univers, de l'Environnement et de l'Espace

par

Nadia FOURRIÉ

le 13 décembre 2010, à Toulouse

Apport des sondeurs hyperspectraux au sein du système global d'observation pour l'étude des dépressions météorologiques



Pression de sommet des nuages (hPa) dérivée de IASI le 23 janvier 2009 vers 21h30 UTC.

JURY

Florence Rabier	**Directeur de recherche**
Pierre Brasseur	**Rapporteur**
Pierre Gauthier	**Rapporteur**
Philippe Drobinski	**Rapporteur**
Philippe Bougeault	**Examinateur**
Yvon Lemaître	**Examinateur**
Sylvain Coquillat	**Président du jury**

CNRM-GAME/GMAP
MÉTÉO-FRANCE et CNRS
42, Avenue G. Coriolis - 31057 Toulouse Cedex 1

Remerciements

Je tiens à remercier Pierre Brasseur, Philippe Drobinski et Pierre Gauthier d'avoir accepté d'être rapporteurs de mon Habilitation à Diriger des Recherches. Je suis également reconnaissante à Pierre Gauthier d'avoir traversé l'Océan Atlantique pour assister à la soutenance. J'adresse mes remerciements aux autres membres du jury, Philippe Bougeault, Sylvain Coquillat et Yvon Lemaître d'avoir lu mon mémoire et participé au jury de mon Habilitation à Diriger des Recherches.

Je souhaite également remercier Alain Joly qui a été mon encadrant de thèse au CNRM et qui m'a presenté à Florence Rabier pour la suite de mon doctorat. Un grand merci à Florence qui m'a prise sous son aile et m'a permis de mener ces travaux de recherche dans son équipe, d'abord en tant que post-doc puis en tant que chargée de recherche au CNRS dans son équipe. C'est tout naturellement que je lui ai demandé d'être ma directrice de recherche pour mon habilitation et je la remercie d'avoir accepté cette tâche.

Merci aux stagiaires et doctorants qui m'ont beaucoup appris et permis de présenter ces travaux de recherche, notamment David Marchal, Mohamed Dahoui, Thomas Pangaud et Fanny Duffourg. Merci à Vincent Guidard dont l'aide et les discussions sont si enrichissantes et à Fatima Karbou pour ses conseils avisés pour la rédaction. Merci à Laurent Jacquin, Jean-Daniel Gril et Vincent Guidard d'avoir assuré la partie administrative du dossier de HDR et les allers-retour à l'Université Paul Sabatier. Merci à Cécile Loo et Jean Maziejewski pour avoir pris en charge la partie restauration/détente de l'habilitation.

Merci aux membres du GMAP que je côtoie quotidiennement et qui font que l'ambiance de travail est si agréable. Je les remercie egalement pour leur aide en toute occasion.

Enfin un grand merci à Jérôme pour avoir "assuré" à la maison pendant la préparation de l'habilitation, son soutien et sa patience, ainsi que d'avoir relu mon manuscrit pour les ultimes corrections.

Table des matières

Tables des matières **1**

Glossaire **3**

1 Curriculum Vitae **5**
1.1 Position Actuelle et expérience professionnelle 5
1.2 Publications essentielles pour la soutenance de l'habilitation à diriger des recherches 6
1.3 Activités 6
1.3.1 Encadrement 6
1.3.2 Enseignement 7
1.3.3 Implication dans les projets nationaux et internationaux 7
1.3.4 Autres activités 8
1.4 Publications 9
1.4.1 Publications dans des revues à comité de lecture 9
1.4.2 Publications dans des revues sans comité 10
1.4.3 Communications à des congrès, Conférences avec acte de lecture 10
1.4.4 Communications à des Conférences sans acte de lecture 12
1.4.5 Rapports 13

2 Synthèse des travaux de recherche **15**
2.1 Introduction 15
2.2 Le sondage satellitaire 16
2.3 L'assimilation de données 18
2.3.1 L'opérateur d'observation 18
2.3.2 L'assimilation variationnelle 19
2.3.3 L'assimilation des radiances satellitaires 21
2.3.4 Le contenu en information 23
2.4 Le ciblage des observations 24
2.5 Préparation de l'assimilation des sondeurs hyperspectraux 27
2.5.1 Evaluation d'un jeu de canaux AIRS pour la prévision numérique 27
2.5.2 Détermination d'un jeu optimal de canaux IASI pour l'étude des tempêtes 28
2.6 Amélioration de l'opérateur d'observation 33
2.6.1 Assimilation des radiances nuageuses 33
2.6.2 De nouveaux opérateurs d'observation pour l'assimilation à méso-échelle sur la Méditerranée 37
2.7 Impact des observations 42
2.7.1 Etude de l'impact des données d'une campagne de mesures 42

2.7.2 Etudes des transitions extratropicales de cyclones avec les observations IASI 45
2.8 Conclusions et perspectives 51

Références bibliographiques **55**

Publications essentielles **59**

Glossaire

1D-Var assimilation variationnelle unidimensionnelle
3D-Var assimilation variationnelle tridimensionnelle
4D-Var assimilation variationnelle quadridimensionnelle
AIRS Atmospheric InfraRed Sounder
AMMA Analyse Multidisciplinaire de la Mousson Africaine
ALADIN Aire Limitée Adaptation dynamique et Développement InterNational
AMS American Meteorological Society
AMSU-A Advanced Microwave Sounding Unit, module A
AMSU-B Advanced Microwave Sounding Unit, module B
AROME Application de la Recherche vers l'Opérationel à Méso-échellE
ARPEGE Action de Recherche Petite Echelle Grande Echelle
ATOVS Advanced TIROS Operational Vertical Sounder
AVHRR Advanced Very High Resolution Radiometer
CEPMMT Centre Européen de Prévisions Météorologiques à Moyen Terme
CIMSS Cooperative Institute For Meteorological Satellite Studies
CMC Service de Météorologie Canadienne
CMS Centre de Météorologie Spatiale
CNES Centre National d'Etudes Spatiales
CNRM Centre National de Recherches Météorologiques
CNRS Centre National de Recherche Scientifique
CYPRIM CYclogenèses et Précipitations Intenses en Méditerranée
DFS Degrees of Freedom for Signal
ECAM European Conference on Applications of Meteorology
ECMWF European Centre for Medium-range Weather Forecasts (voir CEPMMT)
EMS European Meteorological Society
ENM Ecole Nationale de la Météorologie
ER Entropy Reduction (Réduction d'Entropie)
ET Extra Tropical
EUMETSAT EUropean Organisation for the Exploitation of METeorological SATellites
FASTEX Fronts And Storm Track EXperiment

GAME Groupe d'etude de l'Atmosphère MétéorologiquE
GMAP Groupe de Modélisation pour l'Assimilation et la Prévision
GPS Global Positioning System
HIRLAM HIgh Resolution Limited Area Model
HIRS High resolution Infra-Red Sounder
HyMeX HYdrological cycle in Mediterranean EXperiment
IASI Infrared Atmospheric Sounding Interferometer
ICARE Interactions Clouds Aerosols Radiations Etc
IFS Integrated Forecasting System
IRS InfraRed Sounding
KFS Kalman Filter Sensitivity
MAIA Mask AVHRR for Inversion ATOVS
MetOP METeorological OPerational satellite
MHS Microwave Humidity Sounder
MODIS MODerate Imager Spectroradiometer
MSG Météosat Seconde Génération
NASA National Aeronautics and Space Administration
NA-TreC North Atlantic Thorpex REgional Campaign
NCAR National Center for Atmospheric Research
NESDIS National Environmental Satellite Data and Information Service
NOAA National Oceanic and Atmospheric Administration
NWP Numerical Weather Prediction
PNT Prévision Numérique
PNT Prévision Numérique du Temps
PTOP Pression de sommet des nuages (Pressure Top)
RTTOV Radiative Transfer for TIROS Operational Vertical Sounder
SAFNWP Satellite Application Facility on support to Numerical Weather Prediction
SEVIRI Spinning Enhanced Visible and Infrared Imager
THORPEX Programme international de l'Organisation Mondiale Météorologique
TOVS TIROS Operational Vertical Sounder
UKMO United Kingdom Meteorological Office
URA Unité de Recherche Associée
WMED West MEDiterranean

Chapitre 1

Curriculum Vitae

Nadia Fourrié
36 ans, mariée, 4 enfants
Chargée de recherche 1^{re} classe au CNRS

CNRM-GAME/GMAP/OBS
Météo France et CNRS
42 av Coriolis
31057 TOULOUSE Cedex 1
Nadia.Fourrie@meteo.fr

Diplômes

2000 : Thèse de l'Université Pierre et Marie Curie, Paris 6 soutenue le 30 novembre 2000 : "Intérêt des observations du sondeur satellitaire TOVS pour l'analyse et la prévision des dépressions pendant FASTEX". Mention très honorable avec les félicitations du jury.

1997 : DEA Méthodes Physiques en Télédétection, Université Paris 6. Mention Bien.

1996 : Maîtrise Physique et Applications, Université Paris 6. Mention Assez Bien.

1995 : License Physique et Applications, Université Paris 6. Mention Assez Bien.

1994 : DEUG , Université Paris 6. Mention Bien.

1.1 Position Actuelle et expérience professionnelle

- **En poste depuis le 1er Novembre 2003** : Chargée de Recherche au CNRS au GAME URA 1357.
- ***2002-2003*** : Post-doc EUMETSAT "Data assimilation of satellite data in regional models" au CNRM-GAME.
- ***2000-2002*** : Post-doc CNES "Sélection des canaux IASI pour l'assimilation dans un modèle de prévision numérique du temps" au CNRM-GAME
- ***Octobre-Décembre 2001*** : Visite scientifique du SAF NWP au Centre Européen pour les Prévisions Météorologiques à Moyen Terme (CEPMMT) à Reading "Validation of the NESDIS near real time AIRS channel selection".
- ***1997-2000*** : Travaux de thèse intitulée "Intérêt des observations du sondeur satellitaire TOVS pour l'analyse et la prévision des dépressions pendant FASTEX" au Laboratoire de Météorologie Dynamique à Palaiseau.

– ***Février-Juin 1997*** : Stage de DEA sous la direction de Chantal Claud au Laboratoire de Météorologie Dynamique à Palaiseau. Analyse Dynamique de 3 périodes d'observations Intensives de FASTEX

1.2 Publications essentielles pour la soutenance de l'habilitation à diriger des recherches

Fourrié N. and J.-N. Thépaut (2003) : Evaluation of the AIRS Near real Time channel selection for application to numerical weather prediction. *Quart. J. R. Met. Soc.*, 129, 2425-2439.

Fourrié and Rabier (2004) : Clouds Characteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Quart. J. R. Met. Soc.*, 130, 1839-1856.

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

Pangaud T., Fourrié N., Guidard V., Dahoui M. and Rabier F. (2009) : Assimilation of AIRS radiances affected by mid- to low-level clouds. *Monthly Weather Review*, 4276-4292 (17) DOI : 10.1175/2009MWR3020.1.

Duffourg, F. Ducrocq V., Fourrié N., Jaubert G. and Guidard V. (2010) : Simulation of satellite infrared radiances for convective-scale data assimilation over the Mediterranean. *Journal of Geophysical Research*.115, D15107, 14 pp., DOI :10.1029/2009JD012936

1.3 Activités

1.3.1 Encadrement

– Responsable de l'équipe OBS pendant 6 mois (mars-septembre 2006).

Mémoires d'étudiant ayant permis de soutenir un diplôme

– David Marchal (2004) : " Test d'impact de diverses données d'une campagne de mesures "ciblées" dans le 4D-Var ARPEGE" Stage de fin d'étude n^{o}941 de l'Ecole Nationale de la Météorologie et de DEA 'Océan Atmosphère et Environnement' de l'Université Paul Sabatier Toulouse 3.
– Dewi Daniel (2005) : "Les transitions de régimes sont-elles liées à des tempêtes caractéristiques", Stage de Master Océan Atmosphère surfaces continentales 2ème année de l'Ecole Nationale de la Météorologie et de l'Université Paul Sabatier, Toulouse 3.
– Pierre Brousseau (2005) : "Evaluation de l'impact des observations dans un système d'assimilation - prévision", co-encadrement avec G. Desroziers. Stage de fin d'étude n^{o}991 de l'Ecole Nationale de la Météorologie.
– Fanny Duffourg (2007) : "Assimilation à méso-échelle de radiances satellitaires - optimisation de la densité spatiale des observations assimilées" co-direction avec G. Jaubert et P. Brousseau, stage de fin d'étude n^{o}1059 de l'Ecole Nationale de la Météorologie.
– Thomas Pangaud (2009) : Assimilation des radiances des sondeurs infrarouges hyperspectraux en condition nuageuses : application à des cyclogenèses extratropicales. Thèse soutenue le 20 novembre 2009 à l'Universite Paul Sabatier Toulouse 3.

– Fanny Duffourg (2010) : Assimilation de données satellitaires infrarouges pour la prévision et la compréhension des pluies intenses en région méditerranéenne. Co-encadrement avec Véronique Ducrocq. Thèse préparée à l'Université Paul Sabatier Toulouse 3.

Autres stages accompagnés d'un mémoire

– Dahoui M., Rabier F., Lavanant L. and Fourrié N. (2005) : Use of AIRS cloudy radiances. NWP SAF visiting scientist report.
– Zahra Salahoui (2006) : Assimilation des radiances micro-ondes sur terre. Etude d'impact dans ALADIN. Stage effectué en novembre-décembre 2006.
– Abdelaziz Babqiqi, Fourrié N., Rabier F. et Karbou F. (2006) : "Détermination de l'émissivité de surface sur terre des données AIRS sur terre". Rapport de stage effectué du 1^{er} août au 2 novembre 2006.
– Hua Z., F. Rabier, M. Ssczech-Gajewska, N. Fourrié. and T. Auligné (2006) : Impact study on the assimilation of the Atmospheric Infrared Sounder radiances over land. Note de Centre n^{o}7 , Centre National de Recherches Météorologiques.
– Anwender D, Fourrié N., Rabier F. and Arbogast P. (2009) : "Data impact experiments using IASI observations during the ETs of Hurricanes Gustav, Hannah et Ike (2008)" Stage de 6 mois du 15 juin 2009 au 15 décembre 2009).

1.3.2 Enseignement

– Cours annuel (2h) sur l'assimilation des observations télédétectées donné aux Élèves Ingénieurs des Travaux de la Météorologie depuis 2006.
– Cours annuel (2h) de formation permanente "Les modèles : assimilation et impact des données", ENM/Formation Permanente stage "Techniques et organisation de la prévision à Météo-France" depuis 2009.
– Organisation de rencontres doctorants/docteurs au GAME pour sensibiliser les doctorants et les postdocs du GAME à la préparation de leur avenir professionnel les 4 avril 2005, 11 juin 2007 et le 18 mai 2009.

1.3.3 Implication dans les projets nationaux et internationaux

– FASTEX : étude de précurseurs d'altitude de cyclogenèse, étude de l'assimilation des sondeurs dans les zones sensibles pour la cyclogenèse.
– THORPEX : Etude d'impact de données ciblées dans le modèle global ARPEGE
– Cyprim : Cyclogenèses et précipitations intenses en Méditerranée. Assimilation des radiances infrarouges AIRS en présence de nuages.
– ENVISTORM : Rôle de l'environnement (humidité et déformation) sur le cycle de vie et la variabilité des tempêtes. Assimilation des canaux vapeur d'eau de IASI.
– AMMA : Etude d'impact des radiosondages supplémentaires de la campagne sur la prévision.
– HyMeX : Coordinatrice de la tâche de modélisation de l'assimilation de données atmosphériques pour HyMeX dont l'objectif est d'améliorer la caractérisation et la compréhension du cycle de l'eau sur le bassin Méditerranéen.

1.3.4 Autres activités

Activité éditoriale

Relecteur d'articles dans des revues scientifiques : Quarterly Journal of the Royal Meteorological Society et Monthly Weather Review.

Administration de la recherche

- Membre du Comité scientifique du Programme National de Télédétection Spatiale depuis 2006.
- Membre du Conseil De Laboratoire depuis 2007.
- Membre du comité utilisateurs ICARE (2004-2007).

Animation scientifique

- Coordinatrice de la tâche de modélisation de l'assimilation de données atmosphériques pour HyMeX depuis 2010.
- Organisation de la session « Observation targeting and observation impact studies" de la 9th EMS Annual Meeting/9th European Conference on Applications of Meteorology (ECAM) à Toulouse du 28 septembre au 2 octobre 2009.

1.4 Publications

1.4.1 Publications dans des revues à comité de lecture

1. Fink A.H., Agustí-Panareda A., Parker D. J., Lafore J.-P., Ngamini J.-B., Afiesimama E., Beljaars A., Bock O., Christoph M., Didé F., Faccani C. , **Fourrié N.**, Karbou F., Polcher J., Mumba Z., Nuret M., Pohle S., Rabier F., Tompkins A. M., and Wilson G. (2010) : Operational meteorology : observational networks, weather analysis and forecasting. Accepté à Atmospheric Science Letters.
2. Doerenbecher A., Argence S., Cassou C., Caumont O., Descamps L., Ducrocq V., **Fourrié N.**, Guidard V., Jaubert G., Joly A., Lambert D., Lebeaupin Brossier C., Pangaud T., Sanchez-Gomez E., Talagrand O., Terray L., Vincendon B. (2010) : Cyclogenèses et précipitations intenses : éléments de prévisibilité, besoins en observations, La Météorologie. N^{o}68 février 2010, p 23-34.
3. Duffourg F., Ducrocq V., **Fourrié N.**, Jaubert G. and Guidard V. (2010) : Convective scale data assimilation of satellite infrared radiances over the Mediterranean : sensitivity to the observation operator. Accepté au Journal of Geophysical Research.
4. Pangaud T., **Fourrié N.**, Guidard V., Dahoui M. and Rabier F. (2009) : Assimilation of AIRS radiances affected by mid to low-level clouds, Monthly Weather Review, 4276-4292 (17) DOI : 10.1175/2009MWR3020.1.
5. Faccani F., Rabier F. **Fourrié N.**, Agusti-Panareda A., Karbou F., Moll P., Lafore J.-P., Nuret M., Hdidou F. and Bock O. (2009) : The impact of the AMMA radiosonde data on the French global assimilation and forecast system, Weather and forecasting, 24,1268-1284.
6. **Fourrié N .**, Marchal D., Rabier F., Chapnik B. and Desroziers G. (2006) : Impact study of the 2003 North Atlantic THORPEX Regional Campaign, Quart. J. R. Met. Soc, 132, 275-295.
7. **Fourrié N .** and Rabier F. (2004) : Cloud Characteristics and channel selection for IASI radiances in meteorologically sensitive areas. Quart. J. R. Met. Soc., 130, 1839-1856.
8. **Fourrié N .** and Thépaut J.-N. (2003) : Evaluation of the AIRS Near Real Time channel selection for application to numerical weather prediction. Quart. J. R. Met. Soc., 129, 2425-2439.
9. **Fourrié, N .**, Claud C. and Chédin A. (2003) : On the depiction of upper-level precursors of the 1999 december storms from TOVS observations. Weather and Forecasting, 18, 417-430.
10. **Fourrié, N .**, Doerenbecher A., Bergot, T and Joly A (2002) : Adjoint Sensitivity of the forecast to TOVS observations. Quart. J. R. Met. Soc., 128, 2759-2777.
11. Rabier F., **Fourrié N.**, Chafaï D. and Prunet P. (2002) : Channel selection methods for infrared atmospheric sounding interferometer radiances. Quart. J. R. Met. Soc., 128, 1011-1027.
12. **Fourrié N .**, Claud C., Donnadille J., Cammas J.-P., Pouponneau B., et Scott N. A. (2000) : Use of TOVS observations for the identification of tropopause-level thermal anomalies. Quart. J. R. Met. Soc., 126, 1473-1494.
13. Jacquinet-Husson, N., Arié E., Ballard J., Barbe A., Bjoraker G., Bonnet B., Brown L., Camy-Peyret C., Champion J., Chédin A., Chursin A., Clerbaux C., Duxbury G., Flaud J.-M., **Fourrié N.**, Fayt A., Graner G., Gamache R., Goldman A., Golovko V.I., Guelachvili

G., Hartman J.M., Hilico J.C., Hillman J., Lefevre G., Lellouch E., Mikaïlenko S.N., Naumenko O.V., Nemtchinov V., Newnham D.A., Nikitin A., Orphal J., Perrin A., Reuter D.C., Rinsland C.P., Rosenmann L.S., Scott N.A., Selby J., Sinitsa L.N., Sirota J.M., Smith A.M., Smith K.M., Tyuterev V.G., Tipping R.H., Urban S., Varanasi P., et Weber M. (1998) : The 1997 spectroscopic GEISA databank . J. Quant. Spectrosc. Radiat. Transfer, 62, 205 –254.

1.4.2 Publications dans des revues sans comité

- **Fourrié N.**, Guidard V. et Rabier F. (2010) : Impact of hyperspectral infrared sounders in Numerical Weather Prediction models of Meteo-France. WMO CAS/JSC WGNE Blue Book 2010, Edited by J. Côté.
- Rabier F. with contributions from A. Bouchard, **N. Fourrié**, E. Gérard, S. Guedj, V. Guidard, F. Karbou, P. Moll, T. Pangaud and C. Payan (2009) : Update on satellite data assimilation system at Météo-France. JCSDA Quarterly, N 28 september 2009. Edited by George Ohring www.jcsda.noaa.gov
- Rabier F., Faccani C., **Fourrié N.**, Gérard E., Guidard V., Karbou F., Moll P., Payan C. (2009) : Current developments in global and regional data assimilation at Météo-France WMO CAS/ JSC WGNE Blue Book 2009, Edited by J. Côté.
- Doerenbecher A., **Fourrié N.** et Rabier F. (2005) : Etudes diagnostiques d'une campagne de ciblage, Atmosphériques, Avril 2005, 26.

1.4.3 Communications à des congrès, Conférences avec acte de lecture

- Anwender D., **Fourrié N.**, Rabier F. and Arbogast P. (2010) : Data impact experiments using IASI observations during the life cycle of Hurricanes Gustav, Hannah and Ike in the Atlantic (2008) : 29th AMS conference on Hurricanes and Tropical Meteorology, 10-14 May 2010, Tucson, Arizona, Etats Unis.
- **Fourrié N.**, Guidard V., Pangaud T. and Rabier F. (2010) : Assimilation of cloud affected radiances in the french global numerical weather prediction model. 14-20 avril 2010. 17th International TOVS Study Conference, Monterey Etats-Unis.
- Anwender D., **Fourrié N.**, Rabier F. and Arbogast P. (2010) : Data impact experiments using IASI observations during the extratropical transitions of Hurricanes Gustav, Hannah and Ike in the Atlantic. 14-20 avril 2010. 17th International TOVS Study Conference, Monterey Etats-Unis.
- Pangaud T., **Fourrié N.**, Guidard V. (2008) : Assimilation of cloudy AIRS observations in the French global atmospheric model ARPEGE. The 16th International TOVS Study Conference (ITSC-16), 7-13 May, 2008, Angra dos Reis, Brazil
- Bouchard A., Rabier, F. Guidard V., Karbou F., **Fourrié N.** and Pangaud T. (2008) Satellite Data Assimilation over Antarctica : The Concordiasi Field Experiment The 16th International TOVS Study Conference (ITSC-16), 7-13 May, 2008, Angra dos Reis, Brazil.
- Duffourg F., Ducrocq V., Jaubert G., **Fourrié N.** and Guidard V. (2008) : Convective-scale data assimilation of satellite infrared radiances over the Mediterranean : adaptation of the observation operator to the high-resolution. The 16th International TOVS Study Conference (ITSC-16), 7-13 May, 2008, Angra dos Reis, Brazil.
- Montmerle T., **Fourrié N.**, Gérard E., Guidard V., Karbou F., Moll P., Payan C., Poli P. and Rabier F. (2008) : Recent advances in the use of satellite data in the French NWP models. The 16th International TOVS Study Conference (ITSC-16), 7-13 May, 2008, Angra dos Reis, Brazil.

- **Fourrié N.**, Dahoui M., Guidard V., Pangaud T., Poli P. and Rabier F. (2007) : Assimilation of advanced sounder radiances in the French global numerical weather prediction ARPEGE model. Joint 2007 EUMETSAT Meteorological Satellite Conference and the 15th Satellite Meteorology & Oceanography Conference of the American Meteorological Society, 24-28 September 2007, Amsterdam, Netherland.
- Babqiqi A. **Fourrié N**, F. Rabier et Karbou F.(2007) : Détermination de l'émissivité de surface pour l'assimilation des données AIRS sur terre Journées Recherche et développement de Météo-France, Atelier de Modélisation de l'Atmosphère, 16-18 janvier 2007, Toulouse, France.
- **Fourrié N.**, Dahoui M. and Rabier F. (2006) : Towards assimilation of AIRS cloudy radiances, 15th International TOVS Study Conference, 4-10 octobre 2006, Maratea, Italie.
- **Fourrié N.**, Marchal, D., Chapnik B. and Rabier F. (2005) : Case studies of the 2003 Atlantic THORPEX Regional Campaign in the ARPEGE model, 4th WMO Data Assimilation Symposium (Prague, 18-22 April 2005)
- Moll P., T. Auligné, **N. Fourrié**, E. Gérard, D. Lacroix, T. Montmerle, C. Payan, F. Rabier and M. Dahoui (2005) : Recent improvements in the use of observations in the global Arpege NWP system at Météo-France, 4th WMO Data Assimilation Symposium (Prague, 18-22 April 2005)
- Desroziers G., P. Brousseau, B. Chapnik, **N. Fourrié** (2005) : Use of randomization to diagnose the impact of observations on analyses and forecasts, 4th WMO Data Assimilation Symposium, Prague, 18-22 April 2005.
- **Fourrié N.**, Marchal D. and Rabier F. (2004) : Impact of the TReC additional data in the ARPEGE model. First THORPEX Science Symposium, 6-10 Décembre 2004, Montréal Canada.
- Rabier F., **Fourrié N.** , Chapnik B. Desroziers G. and Marchal, D. (2004) : Use of information content measures for various applications. First THORPEX Science Symposium, 6-10 Décembre 2004, Montréal Canada.
- Murray J. J., Smith B. L., **Fourrié N.**, Zhou D. K., Larar A. M and Avery M. A. (2004) : Assimilation of NPOESS Atmospheric Sounder Testbed-Interferometer and dropsonde observations from the 2003 Atlantic THORPEX Regional Campaign. First THORPEX Science Symposium, 6 –10 Décembre 2004, Montréal, Canada.
- **Fourrié N.** and Rabier F. (2003) : Clouds characteristics and channel selection for IASI radiances in meteorologically sensitive areas. Thirteenth International TOVS Study Conference, Sainte Adele, Canada, 29 Octobre – 4 Novembre 2003.
- Rabier F., **Fourrié N.**, Auligné T. et Szczech-Gajewska M. (2003) : Preparation for the assimilation of advanced sounders at Meteo-France. AMS Conference, Long Beach, California, 9-13 février 2003, 11 pages.
- **Fourrié N.** et Rabier F. (2002) : Préparation de l'assimilation des observations IASI : étude de sélection dans les zones sensibles. Atelier de Modélisation de l'Atmosphère, Toulouse, 17-19 Décembre 2002. pages 61-64.
- **Fourrié N.** and Thépaut J.-N. (2002) : Channel selection for AIRS radiance assimilation. Twelfth International ATOVS Study Conference, Lorne, Australie, 27 février-5 mars, 8 pages.
- **Fourrié N.** et Rabier F. (2002) : Use of Advanced Infrared Sounders in cloudy conditions. Twelfth International ATOVS Study Conference, Lorne, Australie, 27 février-5 mars.
- Rabier F. and **Fourrié N.** (2002) : Channel selection for IASI in clear - sky conditions. Twelfth International ATOVS Study Conference, Lorne, Australie, 27 février-5 mars, 8 pages.

- McNally A., **Fourrié N.**, Matricardi M., Thépaut J.-N. and Watts P. (2002) : Progress towards an assimilation strategy for AIRS at ECMWF. Twelfth International ATOVS Study Conference, Lorne, Australie, 27 février-5 mars, 8 pages.
- **Fourrié, N.** et Rabier F. (2001) : Channel selection method for infrared atmopheric sounding interferometer radiances. Dans Proceedings of the Conference on satellite meteorology and oceanography, Madison, 15-18 octobre 2001, 4 pages.
- Doerenbecher, T. Bergot, F. Bouttier and N. Fourrié (2000) : Sensibilité aux observations et ciblage des observations. Dans Ateliers de Modélisation de l'Atmosphère, Météo-France, 29-30 novembre 2000, Toulouse, France, pages 157-161.
- **Fourrié N.** (1999) : Utilisation de l'assimilation des observations TOVS dans une optique de ciblage. Dans Atelier de Modélisation de l'Atmosphère, Météo-France, Toulouse, France, pages 227-230.
- **Fourrié N.**, Claud C., Armante R., Pouponneau B., et Cammas J.-P. (1998) : Use of TOVS observations for the identification of tropopause upper-level thermal anomalies. Dans Ninth Conference on Satellite Meteorology and Oceanography, AMS, 25-29 Mai 1998.

1.4.4 Communications à des Conférences sans acte de lecture

- **Fourrié N.**, Lavanant L., Arriaga A., August T., Gambacorta A., Heilliette S., Nishihata H., Pavelin E., Rabier F., Stubenrauch C. and Cros S. (2010) : Use of cloudy IASI radiances, an intercomparison over the Southern polar region. 2nd Concordiasi Workshop, 29-31 Mars Toulouse, France.
- **Fourrié N.**, Guidard V., Pangaud T. and Rabier F. (2010) : Assimilation of cloudy radiances in global numerical weather prediction model. Seconde conférence internationale IASI, 25-29 janvier 2010, Sévrier, France
- **Fourrié N.**, Guidard V., Pangaud T. (2009) : Update on IASI usage in Météo-France NWP models. ISSWG, 12-13 octobre 2009. CNES Toulouse France.
- **Fourrié N**, Pangaud T., Guidard V. and Rabier F. (2009) : Impact of advanced infrared sounders in the french global NWPARPEGE model. 9th EMS and 9th ECAM conference, Toulouse, 25-29 september 2009.
- **Fourrié N.**, Guidard V, Rabier F. and Bouchard A (2009) : Studies on the assimilation of AIRS/IASI near surface channels in the French global NWP model, 2nd Workshop on remote sensing and modelling of surface properties, Toulouse, 9-11 juin 2009.
- **Fourrié N.**, Dahoui M., Guidard V., Pangaud T., Poli P. and Rabier F. (2007) : Assimilation of advanced sounder radiances in the French global numerical weather prediction ARPEGE model. Joint 2007 EUMETSAT Meteorological Satellite Conference and the 15th Satellite Meteorology & Oceanography Conference of the American Meteorological Society, 24-28 September 2007, Amsterdam, Netherland.
- **Fourrié N.**, Dahoui M. et Rabier F. (2006) : Vers une assimilation des radiances nuageuses, Journées Recherche et développement de Météo-France, 29-30 novembre 2006.
- Babqiqi A. **Fourrié N.**, Karbou F. et F. Rabier (2006) : Détermination de l'émissivité de surface pour l'assimilation des données AIRS sur terre Journées Recherche et développement de Météo-France, 29-30 novembre 2006.
- **Fourrié N.**, Dahoui M. Rabier F. (2006) : Assimilation des radiances AIRS en ciel nuageux dans le modèle ARPEGE. Atelier Thématique sur l'assimilation de données, INSU, 9-10 mai 2006, Toulouse, France.
- **Fourrié N.**, Marchal D. et Rabier F. (2004) : Impact d'observations ciblées en automne 2003, Présentation à l'Atelier PREVI/GMAP, 7 Mai 2004, CIC Météo-France, Toulouse.

- **Fourrié N.** (2003) : Preparation of the assimilation of radiances from ATOVS satellites in global and regional models at Meteo-France. EUMETSAT Seminar. Darmstadt, 1rst December 2003.
- **Fourrié N.** et Rabier F. (2002) : Use of Advanced Infrared Sounders in the sensitive areas. 27ème Assemblée Générale de l'EGS, Nice, France, 21-26 Avril 2002.
- Rabier F. et **Fourrié N.** (2002) : Channel selection for IASI in clear-sky conditions. 27ème Assemblée Générale de l'EGS, Nice, France, 21-26 Avril 2002.
- **Fourrié N.**,Claud C. et Arbogast P. (2001) : Detection and monitoring of the upper- level precursor of the Martin storm. 26ème Assemblée Générale de l'EGS, Nice, France, 25-30 Mars 2001.
- **Fourrié N.** et Doerenbecher A. (2001) : Sensitivity of the forecast to TOVS observations. 26ème Assemblée Générale de l'EGS, Nice, France, 25-30 Mars 2001.
- Fourrié N., Claud C., Donnadille J., Joly A., et Armante A. (1999) : Determination of the thermal structure around the tropopause from TOVS observations for 3 FASTEX IOPs. Dans IUGG99, Birmingham, 19-24 Juillet 1999.
- **Fourrié N.**, Claud C., Armante R., Pouponneau B., et Cammas J.-P. (1998a) : Identification of tropopause upper-level thermal anomalies from TOVS observations. Dans FASTEX Workshop, Toulouse, 27-30 Avril 1998.
- **Fourrié N.**, Claud C., Joly A. et Cammas J.-P. (1998b) : Identification of tropopause upper-level thermal anomalies from TOVS observations. Dans 23 ème Assemblée Générale de l'EGS, Nice, 20-24 Avril 1998.

1.4.5 Rapports

- **Fourrié N.** (2003) : Assimilation of satellite data in a regional model. Part 1 : Preparation of the direct read-out radiances from the ATOVS satellites in global and regional weather prediction models. First year annual report of EUMETSAT Research Fellowship.
- **Fourrié N.** and Thépaut J.-N. (2002) : Validation of the NESDIS Near Real Time AIRS channel selection. Technical Memorandum du CEPMMT., n 390, 12 pages.
- **Fourrié, N.** (2000) : Intérêt des observations du sondeur satellitaire TOVS pour l'analyse et la prévision des dépressions pendant FASTEX. Thèse de doctorat de l'Université Paris VI, Palaiseau, France, 215 pages.
- **Fourrié N.** (1997) : Analyse dynamique de cyclogenèse sur l'Atlantique Nord au moyen d'observations satellitaires TOVS dans le cadre de la campagne FASTEX. Rapport de DEA Méthodes Physiques en Télédétection, 34 pages.
- **Fourrié N** (1996) : Contribution à une étude de l'impact de la qualité des données spectroscopiques infrarouges sur la modélisation du transfert radiatif atmosphérique. Application à la banque de données GEISA et au projet IASI, 39 pages.

Chapitre 2

Synthèse des travaux de recherche

2.1 Introduction

La prévision numérique du temps a connu dans les dernières décennies d'importants progrès. Ceux-ci sont dus aux développements de plus en plus poussés des simulations de l'évolution de l'état atmosphérique. Ainsi la résolution du modèle global français ARPEGE[1] est à ce jour de 10 km sur la France et le modèle à aire limitée sur la France AROME[2] de 2.5 km. Des progrès ont été également réalisés dans la définition de l'état initial de l'atmosphère utilisé comme point de départ des simulations numériques. Cet état, appelé analyse, est constitué par l'assimilation de données qui consiste à combiner les différentes sources d'information disponible. Celles-ci proviennent en grande partie des observations météorologiques et d'une ébauche de l'état atmosphérique, généralement une prévision à courte échéance.

Depuis la fin des années 70, les Etats-Unis, l'Europe et le Japon ont lancé des satellites météorologiques afin d'observer l'atmosphère depuis l'espace. Les premiers satellites dits opérationnels ont été lancés par la NOAA[3] aux Etats-Unis. Deux types de satellites existent : les satellites géostationnaires qui observent toujours la même partie du globe et les satellites défilants. Ce dernier type de satellites est en orbite basse (à environ 800 km) et tourne autour de la Terre en 1h40 environ. Les satellites d'observations de la Terre fournissent de façon continue des observations de l'atmosphère ainsi que de la surface terrestre. Celles-ci sont des mesures de rayonnement, généralement appelées *radiances*. Deux types d'utilisation des données satellites sont répandus en météorologie : l'utilisation qualitative où des images sont étudiées pour le suivi des structures météorologiques et l'utilisation quantitative de l'information contenue dans les observations. L'assimilation directe des observations satellitaires, généralement sous la forme de radiances, dans les modèles de prévision numérique fait partie de cette dernière approche.

Malgré les progrès de la prévision numérique et l'expertise des prévisionnistes, l'actualité récente nous rappelle sans cesse la fragilité de nos sociétés face aux menaces des tempêtes. Certaines restent encore difficiles à prévoir. De vastes projets internationaux comme FASTEX (Fronts And Storm Track EXperiment, (Joly *et al.*, 1999)) dont la campagne de mesure s'est déroulée en Janvier-Février 1997, ont pour vocation d'étudier ces phénomènes atmosphériques afin de mieux comprendre les mécanismes qui entrent en jeu et d'améliorer leur prévision. Cette campagne de mesures a constitué le cadre de mes premiers travaux de recherches en

[1] Action de Recherche Petite Echelle Grande Echelle
[2] Action de Recherche vers l'Opérationnel à Meso-Echelle
[3] National Oceanic and Atmospheric Administration

thèse puis en post-doctorat. Ma thèse portait sur l'utilisation d'une combinaison de radiances pour détecter les précurseurs d'altitude des tempêtes et l'apport de l'assimilation directe des radiances TOVS sur la prévision de la cyclogenèse. Les travaux de mon postdoctorat CNES [4] traitaient, quant à eux, de la préparation de l'assimilation des radiances des sondeurs avancés infrarouges dans le cadre de la prévision numérique.

Mon projet de recherche au CNRS porte sur l'étude du rail des dépressions au moyen de l'assimilation de données satellitaires. Au cours de mon cursus professionnel, je me suis intéressée à divers aspects de l'assimilation des observations satellitaires. Le cadre de l'assimilation offre des outils pour étudier le contenu en information des observations et je décris ces études dans la partie 2.5. Ces outils sont très intéressants dans le contexte des études exploratoires pour les futurs sondeurs spatiaux.

Un aspect important de l'assimilation de données concerne l'opérateur d'observation. Il s'agit de l'ensemble des transformations nécessaires pour simuler une observation à partir des champs du modèle de l'état atmosphérique. Au cours de mes travaux, je me suis intéressée à la prise en compte des radiances nuageuses car les zones où se développent les dépressions sont souvent recouvertes de nuages. Par ailleurs avec le développement de modèles de prévision à méso-échelle comme AROME de nouvelles problématiques apparaissent. L'observation satellitaire à une échelle bien plus grande que la maille du modèle en est une. Ces deux aspects de l'assimilation des observations satellitaires sont traités dans la partie 2.6 avec les modèles opérationnels à Météo-France.

Le but de l'assimilation de données est de fournir une analyse qui contient l'information provenant du modèle et des observations diverses et variées. Ainsi lors de campagnes de mesures, on peut tirer profit d'observations supplémentaires ou spécifiques en les assimilant dans un modèle de prévision numérique. Dans la partie 2.7, j'étudie l'impact des observations de diverses campagnes comme le North Atlantic THORPEX Regional Campaign, la campagne AMMA[5] ou l'impact de l'assimilation des observations IASI sensibles à la vapeur d'eau. La dernière partie de ce mémoire contient les principales conclusions et les perspectives de mon travail.

2.2 Le sondage satellitaire

Le sondage atmosphérique à partir des satellites repose sur l'interprétation des mesures radiométriques dans des intervalles spectraux spécifiques sensibles à certaines caractéristiques physiques de l'atmosphère. Le choix de plages fréquentielles utiles à chaque instrument dépend donc de l'usage auquel il est destiné. Ainsi, pour obtenir des informations sur le profil de température, il est intéressant de choisir des fréquences pour lesquelles l'absorption du rayonnement est due à des gaz dont la distribution est constante dans l'atmosphère tels que le gaz carbonique (régions infrarouges situées à 15 μm et 4,3 μm par exemple) ou le dioxygène (certaines régions micro-ondes situées entre 50- 70 GHz et à 118 GHz). Pour des mesures d'humidité ou d'ozone, les plages spectrales dans lesquelles ces gaz sont des absorbants potentiels doivent être choisies (bande spectrale située autour de 6 μm pour l'humidité et autour de 9,6 μm pour l'ozone). Pour des observations de surface ou de nuages, l'utilisation de fenêtres atmosphériques (régions dans lesquelles l'absorption du rayonnement par l'atmosphère est faible) doit être privilégiée.

[4] Centre National d'Etudes Spatiales

[5] Analyse Multidisciplinaire de la Mousson Africaine

L'équation du transfert radiatif

L'équation du transfert radiatif représente la relation existant entre l'état de l'atmosphère et la radiance atteignant le sommet de l'atmosphère mesurée par le sondeur satellitaire. Cette radiance est calculée à partir de la fonction de Planck et de la transmission des différentes couches atmosphériques pour chacun des gaz traversés. La radiance monochromatique de fréquence ν quittant l'atmosphère vers l'espace avec un angle d'incidence θ est obtenue grâce à l'équation du transfert radiatif :

$$R_\nu(\theta) = \epsilon_\nu(\theta) B_\nu[T(z_0)] \tau_\nu(z_0, \theta) + \int_{z_0}^{\infty} B_\nu[T(z)] \frac{d\tau_\nu(z,\theta)}{dz} + \Psi \tag{2.1}$$

où z_0 désigne l'altitude du sol, ϵ_ν l'émissivité de surface à la fréquence ν, $B_\nu[T(z)]$ la fonction de Planck, $\tau_\nu(z, \theta)$ la transmittance à la fréquence ν entre la couche z et l'espace. Ψ représente un ensemble de processus complexes comprenant la réflexion des rayonnements solaires et atmosphériques, l'atténuation par les nuages ou encore la diffusion. Ce terme est ignoré dans la plupart des cas. Le premier terme de l'équation 2.1 correspond au rayonnement de la surface transmis par l'atmosphère. Le second est l'intégration de l'émission des couches de l'atmosphère vers le satellite.

Les sondeurs hyperspectraux

Depuis 2006, l'Europe par le biais d'EUMETSAT a également à sa charge un satellite météorologique défilant opérationnel MetOp (Meteorological Operational). A son bord se trouvent des instruments que l'on trouve également sur la série de la NOAA tels que le sondeur infrarouge HIRS (High resolution Infra Red Sounder), le sondeur micro-onde AMSU-A (Advanced Microwave Sounding Unit A) et le sondeur d'humidité MHS (Microwave Humidity Sounder). Ces intruments fournissent entre 10 et 20 radiances par point d'observation. Les progrès récents en détection et optique ont permis d'améliorer la résolution spectrale et la précision radiométrique des sondeurs infrarouges. Ainsi le sondeur avancé IASI[6] (Cayla, 2001) développé conjointement par le CNES et EUMETSAT fait également partie de la charge utile de MetOp. Cet instrument représente une avancée technologique par le nombre de données fournies (plus de 8000 par point d'observation) tout comme le sondeur AIRS [7] de la NASA lancé en 2002 à bord du satellite Aqua (Aumann *et al.*, 2003).

Ces deux derniers instruments sont des sondeurs dits "hyperspectraux". AIRS dispose de 2378 canaux couvrant 3 bandes spectrales de 645 cm^{-1} à 2665 cm^{-1} avec une résolution spectrale variant entre 0.25 et 1 cm^{-1}. La résolution spatiale au nadir est d'environ 13.5 km et sur chaque ligne de mesure 90 pixels sont positionnés de façon à ce que 9 pixels AIRS correspondent à un pixel AMSU-A.

IASI est un interféromètre de Michelson qui présente 8461 canaux couvrant l'intervalle [645 – 2760 cm^{-1}] avec une résolution spectrale de 0.25 cm^{-1}. Chaque pixel AMSU-A est observé simultanément par 4 champs de vue IASI. La résolution horizontale est de 12 km au nadir.

La figure 2.1 illustre la région spectrale sondée par IASI et quelques molécules absorbantes permettant le sondage atmosphérique à partir de cette région spectrale. Les régions situées entre 800 et 1200cm^{-1} sont utilisées pour caractériser la surface ou les nuages suivant les cas.

La région spectrale couverte par les sondages infrarouges est sensible à la présence de nuages. La figure 2.2 illustre le principe du sondeur en cas de ciel clair (a) ou de ciel nuageux (b). Si l'atmosphère est libre de tout nuage, l'observation prise par le sondeur satellitaire

[6] Interféromètre Atmosphérique de Sondage dans l'Infrarouge
[7] Atmospheric InfraRed Sounder

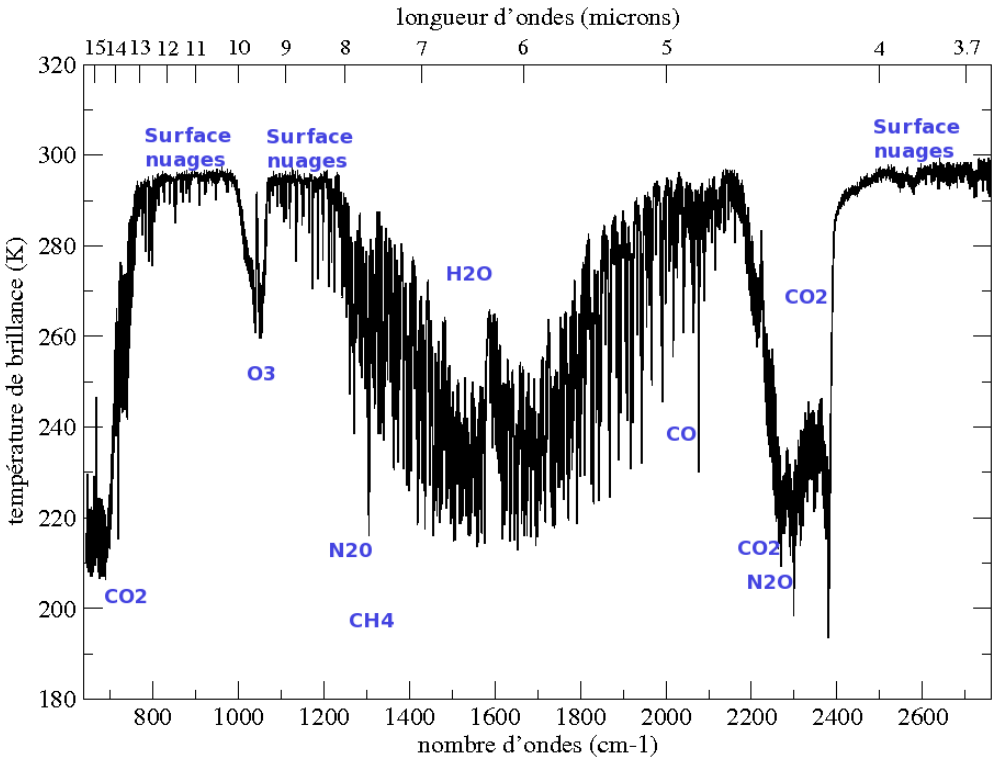


FIG. 2.1 – Spectre IASI et principales molécules absorbantes

permettra de réaliser un sondage de toute l'atmosphère et de recueillir des informations sur la surface terrestre. Si, au contraire, un nuage opaque est présent, seule la partie de l'atmosphère située au dessus du nuage et la surface du nuage sera sondée (Fig 2.2-b). Ainsi les températures de brillance observées seront plus faibles dans le cas d'un nuage (en raison de la température de l'atmosphère à ce niveau plus faible qu'à la surface) que dans le cas du ciel clair où la surface terrestre sera sondée (Fig. 2.3).

2.3 L'assimilation de données

2.3.1 L'opérateur d'observation

On appelle l'opérateur d'observation l'ensemble des transformations nécessaires pour simuler une observation à partir des champs du modèle atmosphérique. Ainsi pour les observations conventionnelles telles que les radiosondages, l'opérateur d'observation contient les interpolations verticales et horizontales pour se situer au point d'observation. Dans le cas des radiances satellitaires, l'opérateur comprend également un modèle de transfert radiatif rapide qui permet de simuler les températures de brillance. Dans notre cas il s'agit du modèle RTTOV développé par le SAF (Satellite Application Facilities) NWP (Numerical Weather Prediction) (Saunders *et al.*, 2006). Ce modèle est développé conjointement et utilisé par les centres de prévision numérique européens comme Météo-France, le UK Met Office et le Centre Européen.

On définit le jacobien comme étant la dérivée de la température de brillance par rapport aux variables du profil atmosphérique (température ou humidité) en entrée de RTTOV.

a) ciel clair b) ciel nuageux

FIG. 2.2 – Principe du sondage de l'atmosphère en cas de ciel clair ou de ciel nuageux.

2.3.2 L'assimilation variationnelle

Nous allons maintenant poser les bases mathématiques de l'assimilation de données. On appelle vecteur d'état le vecteur constitué des variables du modèle à ajuster. Il est noté $\mathbf{x}$ et contient généralement la température, l'humidité, les composantes du vent, la pression de surface. Les observations sont représentées par un vecteur $\mathbf{y}$. Les observations sont reliées au vecteur d'état atmosphérique par l'opérateur d'observation incluant l'interpolation verticale et horizontale et l'équation du transfert radiatif.

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \varepsilon_{\mathbf{O}} + \varepsilon_{\mathbf{F}} \tag{2.2}$$

où $\mathcal{H}$ est le modèle de transfert radiatif. Les erreurs d'observation $\varepsilon_{\mathbf{O}}$ et de modélisation $\varepsilon_{\mathbf{F}}$ sont supposées être des bruits gaussiens avec des matrices de covariances $\mathbf{O}$ et $\mathbf{F}$. Soit $\mathbf{R} = \mathbf{O} + \mathbf{F}$ la matrice de covariances d'erreurs résultante.

Notre connaissance a priori de l'état atmosphérique est représentée par le vecteur d'état de l'ébauche $\mathbf{x_b}$. Sa matrice de covariance d'erreurs associées est notée $\mathbf{B}$. Toutes les erreurs sont supposées non biaisées. L'équation du transfert radiatif est supposée être faiblement non-linéaire afin que l'approximation linéaire-tangente soit valable :

$$\mathcal{H}(\mathbf{x}) = \mathcal{H}(\mathbf{x_b}) + \mathbf{H}(\mathbf{x} - \mathbf{x_b}) \tag{2.3}$$

On appelle $\mathbf{H}$ la matrice linéaire tangente, dérivée de $\mathcal{H}$ au voisinage de $\mathbf{x_b}$.

L'état analysé optimal, au sens du maximum de vraisemblance, est donné par

$$\mathbf{x_a} = \mathbf{x_b} + \mathbf{K}(\mathbf{y} - \mathcal{H}(\mathbf{x_b})) \tag{2.4}$$

avec $\mathbf{K} = \mathbf{A}\mathbf{H}^T\mathbf{R}^{-1}$ et $\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H})^{-1}$.

$\mathbf{K}$ est la matrice de gain de Kalman et $\mathbf{A}$ est la matrice de covariances d'erreurs d'analyse. $\mathbf{K}$ peut être interprétée comme l'inverse généralisé de $\mathcal{H}$, permettant de reconstruire le profil atmosphérique analysé à partir des données.

Pour résoudre l'équation 2.4, l'approche variationnelle (Lorenc, 1986) et (Le Dimet et Talagrand, 1986) qui consiste à minimiser un critère objectif défini en fonction du problème posé

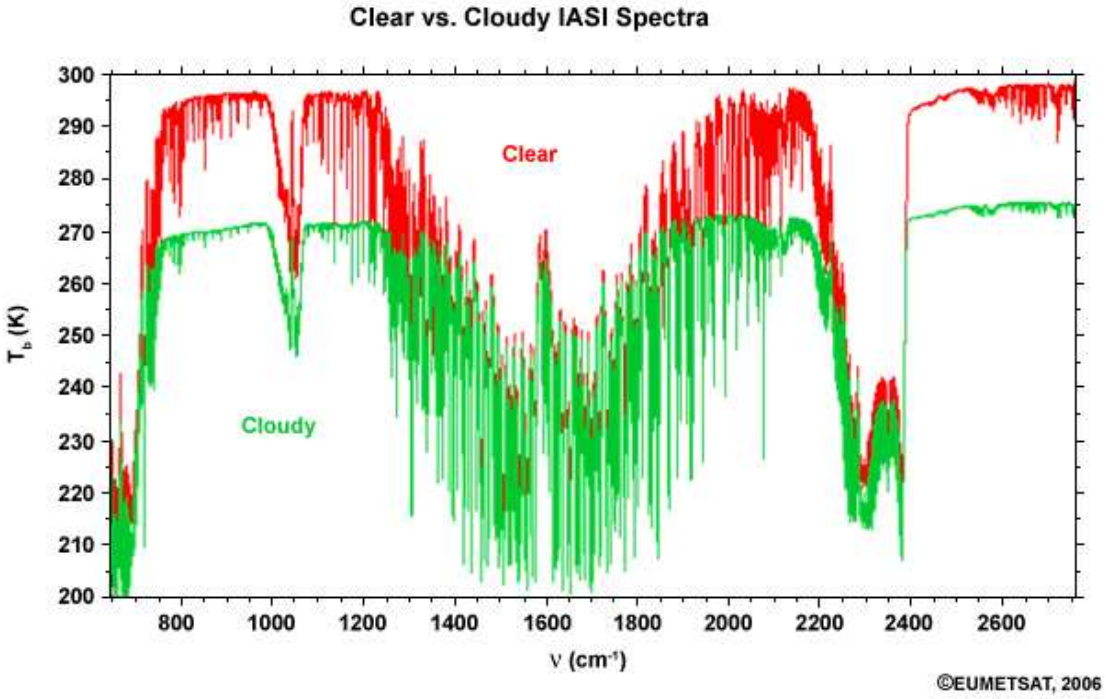


FIG. 2.3 – Spectre d'une observation claire (en rouge) et d'une observation nuageuse (en vert) de IASI

est utilisée depuis la fin des années 90 dans de nombreux centres de prévision numérique. On minimise la fonction quadratique $J(\mathbf{x})$ qui quantifie les écarts aux différentes sources d'information.

$$J(\mathbf{x}) = \overbrace{\frac{1}{2}(\mathbf{x}-\mathbf{x_b})^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x_b})}^{J_b} + \overbrace{\frac{1}{2}(\mathcal{H}(\mathbf{x})-\mathbf{y})^T\mathbf{R}^{-1}(\mathcal{H}(\mathbf{x})-\mathbf{y})}^{J_o} + J_c \tag{2.5}$$

Le premier terme de l'équation 2.5 J_b représente l'écart à l'ébauche et le second J_o l'écart aux observations. Le troisième terme J_c est un terme de contrainte, utilisé par exemple pour expliciter une condition d'équilibre dynamique imposée. Dans le cas où une seule dimension est étudiée, il s'agit du 1D-Var, l'analyse se fait au niveau du profil atmosphérique. L'espace des paramètres est ainsi le profil de température, d'humidité et d'ozone définis sur les 43 niveaux du modèle de transfert radiatif RTTOV. Si les trois dimensions de l'atmosphère sont concernées, il s'agit du 3D-Var. Ce système d'assimilation a été opérationnel à Météo-France de mai 1997 à juin 2000 pour le modèle ARPEGE. Depuis cette date, l'assimilation variationnelle est quadridimensionnelle (4D-Var). En 4D-Var, la notion temporelle intervient et $J(\mathbf{x})$ devient :

$$J(x) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=0}^{n}(\mathcal{H}_i[\mathbf{x}(t_i)]-\mathbf{y}_i)^T{\mathbf{R}_i}^{-1}(\mathcal{H}_i[\mathbf{x}(t_i)]-\mathbf{y}_i) + J_c \tag{2.6}$$

où $\mathcal{H}_i$ représente l'opérateur d'observation pour l'intervalle temporel i, $\mathbf{y}_i$ les observations pour l'intervalle i, $\mathbf{x}(t_i)$ l'état du modèle au temps i et $\mathbf{R}_i$ la matrice de covariances d'erreurs d'observations pour l'intervalle i.

L'état du modèle $\mathbf{x}(t_i)$, pour un instant donné i, résulte de la propagation de son état depuis l'instant initial t_0 jusqu'à t_i par l'opérateur *a priori* non linéaire $M_{t_0 \rightarrow t_i}$ tel que :

$$x(t_i) = M_{t_0 \rightarrow t_i}(x(t_0)) \tag{2.7}$$

Cependant pour faciliter la résolution du problème, on utilise la formulation incrémentale (Courtier *et al.*, 1994) qui consiste à linéariser le problème autour de $\mathbf{x_b}$ pour obtenir un incrément $\delta\mathbf{x} = \mathbf{x} - \mathbf{x_b}$. L'équation 2.6 devient alors :

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}\sum_{i=0}^{n}(\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i)^T\mathbf{R}_1^{-1}(\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i) + J_c \tag{2.8}$$

où $\delta\mathbf{x}(t_i)$ est l'incrément de l'intervalle temporelle i et $\mathbf{d}_i = \mathbf{y} - \mathcal{H}(\mathbf{x_b})$ est le vecteur d'innovation de l'intervalle i.

Au minimum de la fonction coût $J(\delta\mathbf{x})$, l'incrément d'analyse est ajouté à l'ébauche pour obtenir l'état analysé. L'importance pratique de la formulation incrémentale réside dans le fait qu'elle permet de linéariser et de réduire le coût de l'assimilation en évaluant l'incrément d'analyse dans un espace de dimension réduite par rapport à celui du modèle. Les calculs se font donc pour une variable $\delta\mathbf{x}$ dite *de contrôle* de dimension réduite par rapport à $\mathbf{x}$. La résolution du modèle d'assimilation dans la minimisation est ainsi également plus faible que celle du modèle de prévision.

Un des avantages de la méthode variationnelle est notamment la possibilité d'utiliser des observations qui ne sont pas reliées au vecteur d'état de manière simple. L'opérateur d'observation $\mathcal{H}$ peut être complexe et faiblement non linéaire, comme dans le cas des radiances.

2.3.3 L'assimilation des radiances satellitaires

Le système ARPEGE/IFS est développé en collaboration avec le CEPMMT. Le modèle ARPEGE dispose d'une assimilation 4D-Var avec une fenêtre de 6h depuis juin 2000 et les radiances "brutes" des sondeurs infrarouges et micro-ondes sont assimilées directement depuis 2002. Les observations assimilées dites conventionnelles sont les mesures issues des radiosondages, des stations sol, des bouées, des données des avions de lignes, des vents déduits des déplacements de nuages issus de l'imagerie satellitaire, des stations de réception du GPS. Les données issues des satellites sont les radiances des sondeurs micro-ondes et infrarouges (issues de satellites défilants ou géostationnaires), les occultations des satellites GPS et les vents de surface issus des diffusiomètres.

La figure 2.4 montre l'évolution du nombre d'observations assimilées dans les analyses du modèle global ARPEGE depuis 2002. On peut voir l'augmentation constante du nombre d'observations satellitaires assimilées. Les données AIRS ont été assimilées à partir de septembre 2006. Au départ il s'agissait des radiances sondant la basse stratosphère, non affectées par la présence de nuages. En février 2008, leur assimilation a été étendue aux radiances troposphériques simultanément à l'assimilation des radiances IASI. On voit ainsi que les radiances IASI et AIRS représentent une part très importante du volume d'observations assimilées. En avril 2010, le nombre d'observations satellitaires a presque quadruplé car la résolution des observations utilisées dans la minimisation est passée d'une observation par boîte de 250 km de côté à une observation par boîte de 125 km.

Deux modèles régionaux sont opérationnels à Météo-France : ALADIN et AROME. Le premier modèle a une résolution de 9.5 km au dessus de la France et une assimilation 3D-Var avec

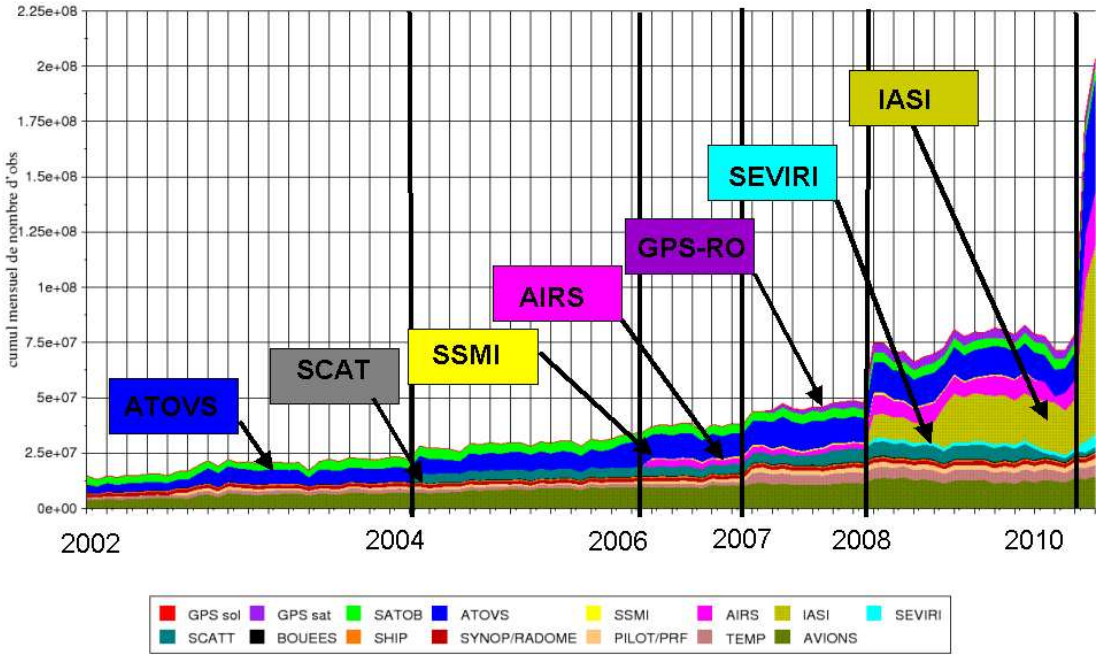
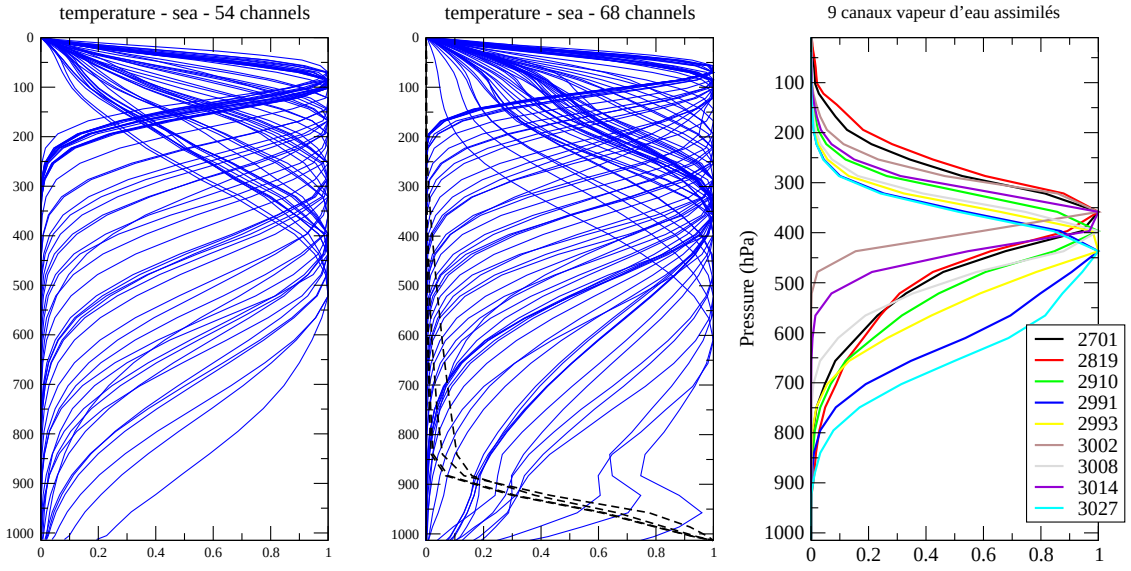


FIG. 2.4 – Evolution du nombre d'observations assimilées depuis 2002. Figure d'Hervé Bénichou.

une fenêtre temporelle de 6h. AROME est opérationnel depuis fin 2008 et dispose d'un cycle rapide d'assimilation en 3D-VAR de 3h. Ce modèle de méso-échelle centré sur la France assimile les mêmes observations que le modèle global ainsi que les données de vent de radar doppler et les réflectivités des radars météorologiques. IASI et AIRS ne couvrent le domaine d'AROME que deux fois par jour.

La figure 2.5 montre les fonctions de poids (dérivée de la transmittance par rapport à l'altitude, Kaplan, *et al*, 1977) des différents canaux AIRS et IASI assimilés dans les modèles ARPEGE, ALADIN et AROME. Le pic de sensibilité de la fonction poids est maximal à une certaine altitude qui correspond à la région de l'atmosphère principalement sondée. Si la majorité des radiances apportent de l'information sur la température de la moyenne et haute troposphère ainsi que de la stratosphère pour les deux instruments, dans le cas de IASI 4 canaux sensibles à la surface (en pointillé sur la figure 2.5 (b)) ainsi que 9 canaux vapeur d'eau sont également assimilés.

Pour assimiler les observations dans un schéma d'assimilation variationnelle, il est nécessaire de faire l'hypothèse que celles-ci sont non biaisées. Cependant des biais existent et peuvent être dus par exemple aux erreurs de calibration de l'instrument et au modèle de transfert radiatif utilisé pour simuler ces données. L'étude de ces biais montre qu'ils peuvent varier dans le temps, en fonction de la zone géographique et de l'angle de visée. Le biais est donc calculé à partir de prédicteurs dépendant de l'état de l'atmosphère, de l'angle de visée. Une constante est également appliquée à chaque canal. Après avoir utilisé pendant de nombreuses années une méthode de correction de biais (Harris et Kelly, 2001) calculée à partir des innovations en dehors du système d'assimilation, nous utilisons la correction de biais dite variationnelle (Dee, 2004) et (Auligné *et al.*, 2007) depuis février 2008. Cette technique ajuste les paramètres du biais calculé à partir de prédicteurs à chaque nouveau cycle d'analyse au moment de la minimisation.



a) AIRS b) IASI T c) IASI WV

FIG. 2.5 – Fonctions de poids des canaux AIRS (a) et IASI température (b) et vapeur d'eau (c) assimilés dans ARPEGE.

2.3.4 Le contenu en information

Des concepts de la théorie de l'information peuvent être utilisés pour quantifier le gain en information apporté par les observations. Ils sont expliqués dans Rodgers (2000) : ce sont principalement la réduction d'entropie (ER $= -\frac{1}{2}\log_2 det(\mathbf{AB}^{-1})$) et le nombre de degrés de liberté pour le signal (Degrees of Freedom for Signal, DFS $= Tr(\mathbf{I} - \mathbf{AB}^{-1})$, où $\mathbf{I}$ représente la matrice identité, $\mathbf{A}$ la matrice de covariances d'erreurs d'analyse et $\mathbf{B}$ la matrice de covariances d'erreurs d'ébauche. Si λ représente une valeur propre de $\mathbf{AB}^{-1}$ (une faible valeur de λ correspondant à une direction où les mesures ont réduit significativement la variance d'erreur d'analyse), il peut être montré que :

$$\mathrm{DFS} = n - \sum_{\lambda \in \sigma(\mathbf{AB}^{-1})} \lambda \tag{2.9}$$

$$\mathrm{ER} = -\frac{1}{2} \sum_{\lambda \in \sigma(\mathbf{AB}^{-1})} \log_2 \lambda \tag{2.10}$$

où n représente le nombre total de degrés de liberté du problème analysé.

Ces deux concepts sont très utiles car ils permettent de donner une mesure globale de la réduction de l'incertitude apportée par l'analyse. Le DFS permet de diagnostiquer quelles sont les observations les plus déterminantes dans l'analyse et de donner une mesure de l'impact des observations dans l'analyse.

Les travaux de la partie suivante reposent sur l'étude de ces deux concepts.

2.4 Le ciblage des observations

La prévision de certaines tempêtes de moyennes latitudes reste encore parfois difficile et une mauvaise prévision peut avoir de lourdes conséquences socio-économiques. On sait que ces erreurs de prévisions peuvent provenir d'erreurs dans les conditions initiales (Rabier *et al.*, 1996). Les zones dites "sensibles" sont des régions de l'atmosphère où de petites erreurs dans les conditions initiales peuvent croître très rapidement avec l'intégration temporelle du modèle atmosphérique et conduire à de fortes erreurs de prévisions numériques.

Un moyen de contrôler ces erreurs est de réduire ces incertitudes sur l'état du modèle dans ses conditions initiales et en particulier dans les zones sensibles. Pour cela, on cherche à collecter et assimiler plus d'observations. Le ciblage, ou observation adaptative, a été imaginé pour améliorer, à l'aide d'un nombre restreint d'observations supplémentaires, la prévision d'un phénomène météorologiques pour une date, une échéance et sur une zone géographique donnée. La difficulté est ici de déterminer où les observations supplémentaires doivent être rajoutées. Les premiers essais de ciblage ont eu lieu pendant la campagne FASTEX en 1997 et depuis cette technique est employée par les Etats-Unis chaque hiver sur l'Océan Pacifique. Elle est aussi communément utilisée pour la surveillance des cyclones tropicaux. Il existe diverses méthodes de calcul pour déterminer où déployer des observations supplémentaires.

L'état de l'atmosphère est caractérisé par le vecteur d'état $\mathbf{x}$, noté $\mathbf{x}_0$ à l'instant t_0. t_0 n'est pas forcément l'instant présent. Le modèle non linéaire $\mathcal{M}$ fait évoluer le vecteur d'état de l'instant initial t_0 à l'instant final t_1 de la prévision : $\mathbf{x}_1 = \mathcal{M}(\mathbf{x}_0)$. Une petite perturbation de l'état initial $\delta\mathbf{x}_0$ est transportée au premier ordre par le modèle linéaire-tangent $\mathbf{M}$ en une perturbation $\delta\mathbf{x}_1$ au voisinage de l'état $\mathbf{x}_1$:

$$\delta\mathbf{x}_1 = \mathbf{M}\delta\mathbf{x}_0$$

On définit à l'instant t_1 une fonction scalaire $\mathcal{J}$ de l'état $\mathbf{x}_1$. À l'aide du modèle adjoint (noté $\mathbf{M}^*$ ou $\mathbf{M}^T$), on peut alors relier une petite variation de la fonction $\mathcal{J}$ à l'instant t_1 à une petite variation de la variable $\mathbf{x}_0$.

$$\delta\mathcal{J} = \mathcal{J}(\mathbf{x}_1 + \delta\mathbf{x}_1) - \mathcal{J}(\mathbf{x}_1)$$

$$\delta\mathcal{J} = \langle \frac{\partial\mathcal{J}(\mathbf{x}_1)}{\partial\mathbf{x}_1}, \delta\mathbf{x}_1 \rangle + \mathcal{O}(\|\delta\mathbf{x}_1\|^2)$$

$$\delta\mathcal{J} = \langle \frac{\partial\mathcal{J}(\mathbf{x}_1)}{\partial\mathbf{x}_1}, \mathbf{M}\delta\mathbf{x}_0 \rangle = \langle \mathbf{M}^* \frac{\partial\mathcal{J}(\mathbf{x}_1)}{\partial\mathbf{x}_1}, \delta\mathbf{x}_0 \rangle$$

Mais $\mathcal{J}$ étant aussi une fonction explicite de $\delta\mathbf{x}_0$ on a aussi :

$$\delta\mathcal{J} = \langle \nabla_{\mathbf{x}_0}\mathcal{J}, \delta\mathbf{x}_0 \rangle$$

On appelle sensibilité aux conditions initiales ou encore gradient de la fonction $\mathcal{J}$ cette quantité $\nabla_{\mathbf{x}_0}\mathcal{J}$:

$$\nabla_{\mathbf{x}_0}\mathcal{J} = \mathbf{M}^* \frac{\partial\mathcal{J}(\mathbf{x}_1)}{\partial\mathbf{x}_1}$$

On obtient ce gradient en appliquant l'adjoint du modèle $\mathbf{M}^*$ au gradient de la fonction à l'instant t_1.

Si $\mathcal{J}$ est conçue comme l'erreur de prévision à t_1, on peut également calculer -a posteriori- les structures clés d'erreur d'analyse (Klinker *et al.*, 1998). Celles-ci représentent les perturbations optimales de $\mathbf{x}_0$ pour réduire l'erreur de prévision (différence entre la prévision à l'échéance $(t_1 - t_0)$ et l'analyse de vérification à t_1). Ces perturbations (à t_0) sont obtenues après quelques itérations de la minimisation de $\mathcal{J}$ à t_1.

Par ailleurs, l'approche des vecteurs singuliers permet par un calcul de vecteurs propres de l'opérateur $\mathbf{M}^*\mathbf{M}$ d'obtenir des zones sensibles assez similaires à celles obtenues avec $\nabla_{\mathbf{x}_0}\mathcal{J}$. En effet, la combinaison linéaire des vecteurs avec les plus forts taux de croissance est une approche très courante pour produire des zones sensibles. Les vecteurs singuliers de $\mathbf{M}$ sont les structures qui s'amplifient le plus entre les instants t_0 et t_1 (Palmer *et al.*, 1998) et (Bergot *et al.*, 1999).

L'ajout d'observations implique que $\delta\mathbf{x}_0$ ait nécessairement un forme d'incrément. De là vient l'idée de faire remontée le calcul de sensibilité jusqu'aux observations. La sensibilité aux observations (Baker et Daley, 2000 ; Doerenbecher et Bergot, 2001) fournit une mesure de l'influence et de l'importance de chaque observation en tenant compte de la dynamique du jour, des caractéristiques instrumentales ainsi que de la présence simultanée des autres observations.

$$\nabla_{\mathbf{y}_o}\mathcal{J} = \mathbf{R}^{-1}\mathbf{H}\mathbf{A}\mathbf{M}^*\nabla_{\mathbf{x}_0}\mathcal{J}$$

Pour obtenir $\nabla_{\mathbf{y}_o}\mathcal{J}$, on calcule d'abord la sensibilité aux conditions initiales puis on applique l'opérateur adjoint (pour résumer) du processus d'assimilation de données pour obtenir la sensibilité aux observations. Celle-ci est définie dans l'espace des observations, c'est-à-dire qu'aucune senibilité n'est définie en un point de l'espace qui ne serait pas observé.

À nouveau, si $\mathcal{J}$ est conçue comme l'erreur de prévision à t_1, la sensibilité aux observations permet de déboucher sur le diagnostic du FEC (Contribution à l'Erreur de Prévision, Cardinali, 2009) qui permet d'estimer - a posteriori - la contribution de chaque observation à la valeur $\mathcal{J}$ au cours du processus d'assimilation. Pour obtenir des résultats robustes de cette application, le calcul de sensibilité doit cependant être dérivé d'un développement de Taylor, non plus au premier, mais au second ordre.

Enfin, la *Sensibilité du Filtre de Kalman* (KFS) proposée par (Bergot et Doerenbecher, 2002) permet de prévoir le déploiement optimal des observations additionnelles en tenant compte des observations (routine et autres) dont on prévoit qu'elles seront déployées. Cette approche conçue pour le temps-réel fonctionne aussi en mode diagnostique.

La KFS est basée sur une estimation de la variance de l'erreur de prévision de la fonction $\mathcal{J}$. Il s'agit de réduire au maximum cette variance qui se calcule ainsi : $\sigma^2 = \nabla_{\mathbf{x}_0}\mathcal{J}^T\mathbf{A}\nabla_{\mathbf{x}_0}\mathcal{J}$, où $\mathbf{A}$ représente la matrice de covariance des erreurs d'analyse pour une configuration donnée du réseau d'observation. La KFS teste divers déploiements alternatifs et sélectionne celui qui produit la plus forte réduction de variance $\delta\sigma_i^2$ par rapport au réseau de routine.

Dans une formulation itérative, la KFS peut prendre en compte des choix successifs d'observations ciblées. La variance de l'erreur à l'étape de l'assimilation d'une observation $\mathbf{y}_i$ s'écrit alors ainsi :

$$\sigma_i^2 = \nabla_{\mathbf{x}_0}\mathcal{J}^T\mathbf{B}_{i-1}\mathbf{H}_i(\mathbf{R}_i + \mathbf{H}_i^T\mathbf{B}_{i-1}\mathbf{H}_i)^{-1}\mathbf{H}_i^T\mathbf{B}_{i-1}\nabla_{\mathbf{x}_0}\mathcal{J}$$

où $\mathbf{B}_{i-1}$ correspond à la matrice de covariances des erreurs d'analyse obtenue à l'étape précédente, $\mathbf{R}_i$ et $\mathbf{H}_i$ sont la matrice de covariances des erreurs et l'opérateur d'observation associés à l'observation $\mathbf{y}_i$.

2.5 Sélection de canaux pour la préparation de l'assimilation des sondeurs infrarouges hyperspectraux

Ces travaux ont été réalisés pendant mon post-doc CNES. Ils avaient pour but d'étudier l'apport d'information d'un jeu restreint de canaux. Avec le lancement d'AIRS et IASI, une nouvelle génération de sondeurs infrarouges a vu le jour. En effet, du fait de leur précision spectrale, ces deux sondeurs infrarouges avancés fournissent un très grand nombre de canaux. Cependant pour des raisons de coûts de transmission et de temps de calcul, l'ensemble des données ne peut pas être intégré dans les modèles de prévision numérique. Un moyen de réduire le coût consiste à réduire le nombre de canaux assimilés et on cherche alors à garder dans l'assimilation les canaux les plus informatifs.

J'ai tout d'abord participé à la finalisation de l'article de Rabier *et al* (2002). Il s'agissait de tester plusieurs méthodes pour choisir un jeu de canaux pour l'assimilation de données dans le cadre de la prévision numérique. Dans le cadre de la préparation de l'assimilation des radiances IASI simulées dans les modèles de prévision numérique, une méthode de sélection de canaux a été développée en se basant sur des études de contenu en information dans le contexte de la théorie de l'estimation linéaire (Rabier et al, 2002). Le cadre général de la sélection de canaux que j'ai étudiée est la théorie de l'estimation linéaire optimale et pour réaliser les analyses nous avons utilisé un 1D-Var couplé au modèle de transfert radiatif. Quatre méthodes de sélection de canaux ont été comparées sur une base de profils atmosphériques regroupant diverses stations terrestres. Deux d'entre elles, la méthode basée sur les jacobiens et l'autre sélectionnant de façon itérative les canaux avec le plus grand contenu en information par rapport à l'ébauche Rodgers (1996), donnent les sélections de canaux les plus performantes au sens du contenu en information et de l'erreur d'analyse. Une variante, qui est une sélection constante calculée à partir de la moyenne des jeux de canaux établis à partir de la méthode itérative sur toute la base de profils atmosphérique, a été mise au point afin de réduire le coût en temps de calcul. Cette variante a fourni des résultats de qualité intermédiaire entre les deux méthodes précédentes en terme d'erreur d'analyse et de résolution verticale.

2.5.1 Evaluation d'un jeu de canaux AIRS pour la prévision numérique

Cette partie concerne le papier Fourrié et Thépaut (2003) dans le Quarterly Journal of the Royal Meteorological Society. J'ai réalisé cette étude lors d'une visite scientifique au CEPMMT de deux mois à la fin de l'année 2001 financée par le SAF NWP. L'étude présentée par la suite consiste à reprendre la méthode de Rodgers (1996) afin d'établir une sélection de canaux du sondeurs AIRS avec le plus grand contenu en information et de la comparer au jeu de canaux établi par l'équipe scientifique AIRS de la NOAA/NESDIS (Susskind *et al.*, 2003) pour la transmission sur le Système Mondial de Télécommunications en temps réel. Ce dernier jeu de canaux comprend 324 canaux sélectionnés pour permettre l'inversion de la température de surface et l'émissivité de surface ainsi que les profils de température, humidité et ozone. La sélection AIRS de la NOAA/NESDIS prend les canaux sur les ailes des bandes d'absorption, ce qui permet d'avoir plus de canaux pointant au dessus des nuages. Un des objectifs de cette sélection est de bien échantillonner les régions fenêtres pour déterminer les variations spectrales de l'émissivité de surface. Un autre objectif de cette sélection est d'être optimisée pour toutes les conditions d'observations, en particulier en présence de nuages.

La sélection dite "optimale pour le contenu en information" dépend de l'état atmosphérique à travers sa dépendance à la matrice de covariances d'erreurs d'ébauche et au linéaire tangent du modèle de transfert radiatif. Cependant dans un contexte opérationnel il n'est pas

possible de faire une sélection atmosphère par atmosphère. La sélection a donc été construite à partir d'un ensemble global qui optimisait le contenu en information à partir de la méthode itérative de sélection de canaux proposée par Rodgers (1996) et utilisée par Rabier et al (2002). La méthode procède en effectuant des analyses successives qui n'utilisent qu'un seul canal supplémentaire à chaque fois. La matrice de covariances d'erreurs d'analyse est calculée à chaque fois et est utilisée comme matrice de covariances d'erreurs d'ébauche à la fois suivante. Ceci permet de prendre en compte toute l'information des canaux sélectionnés précédemment dans la sélection du nouveau canal. Ici le critère retenu est la maximisation de la réduction d'entropie ER. Ensuite comme dans le papier de Rabier et al (2002) la sélection de canaux globale a été calculée comme la moyenne des différentes sélections optimales sur une base de profils atmosphériques représentatifs. Cette base de données est un sous ensemble de la base de profils atmosphérique du CEPMMT (Chevallier, 1999) (Chevallier *et al.*, 2000) et forme un ensemble de 108 profils (75 profils des régions tempérées, 14 profils tropicaux et 19 profils des régions polaires). 63 canaux ont été toujours sélectionnés pour les 108 profils alors que 1504 canaux sur les 2378 n'ont jamais été choisis.

La comparaison entre ce jeu "optimal" et la sélection de la NOAA/NESDIS montre que 119 canaux sur 324 sont communs aux deux sélections. La sélection NOAA/NESDIS sélectionne des canaux qui sont mieux répartis sur le spectre infrarouge tandis que la sélection optimale privilégie certaines régions spectrales. Même si moins de la moitié des canaux sont communs aux deux sélections celles-ci couvrent de larges régions spectrales et sélectionnent des nombres d'ondes voisins. Les résultats en terme de contenu en information (DFS et ER) sont similaires pour les deux jeux de canaux ainsi qu'en terme d'erreurs d'analyses linéaires du 1D-Var pour la température (Fig. 2.6-a), l'humidité (Fig. 2.6-b), l'ozone et la température de surface. La figure 2.6 montre le gain apporté par le nombre important de canaux AIRS en comparaison de la quinzaine de canaux HIRS et si on utilise tous les canaux AIRS.

L'impact de la masse d'air sur la sélection de canaux a ensuite été étudié. La conclusion est qu'une sélection de canaux réalisée avec un ensemble de profils atmosphériques non représentatif peut avoir de mauvaises performances notamment pour la température et l'humidité de la basse troposphère. La performance de la sélection basée sur le contenu en information est sensible à la base de données d'apprentissage même si l'impact sur la qualité de l'inversion n'est pas très important. D'autre part la sélection pour la diffusion en temps réel semble capturer de façon satisfaisante la plupart de la variabilité des situations atmosphériques.

La robustesse de la sélection AIRS en temps réel pour l'inversion des structures clés des erreurs d'analyse a été validée en utilisant une matrice de covariances représentative des erreurs clés de l'analyse (Klinker *et al.*, 1998). Même avec ces conditions, la sélection AIRS pour le temps téel se comporte raisonnablement bien en comparaison d'une sélection dédiée à ce type d'erreurs.

2.5.2 Détermination d'un jeu optimal de canaux IASI pour l'étude des tempêtes

Après les aspects liés à la prévision numérique, je me suis intéressée à l'étude des canaux les plus informatifs pour la prévision des tempêtes dans l'article de Fourrié et Rabier (2004). Des études précédentes (Prunet *et al.*, 1998) et (Collard, 1998) avaient suggéré que IASI pouvait voir certaines structures baroclines de petite échelle, qui avaient été identifiées par des études de sensibilité comme étant cruciales dans le développement des erreurs de prévision (Rabier *et al.*, 1996). Néanmoins le principal obstacle pour utiliser les observations des sondeurs hyperspectraux est la présence de nuages, qui peut fortement limiter l'information des sondeurs infrarouges. Dans ce contexte, Mc Nally (2002) a étudié l'occurence des nuages dans

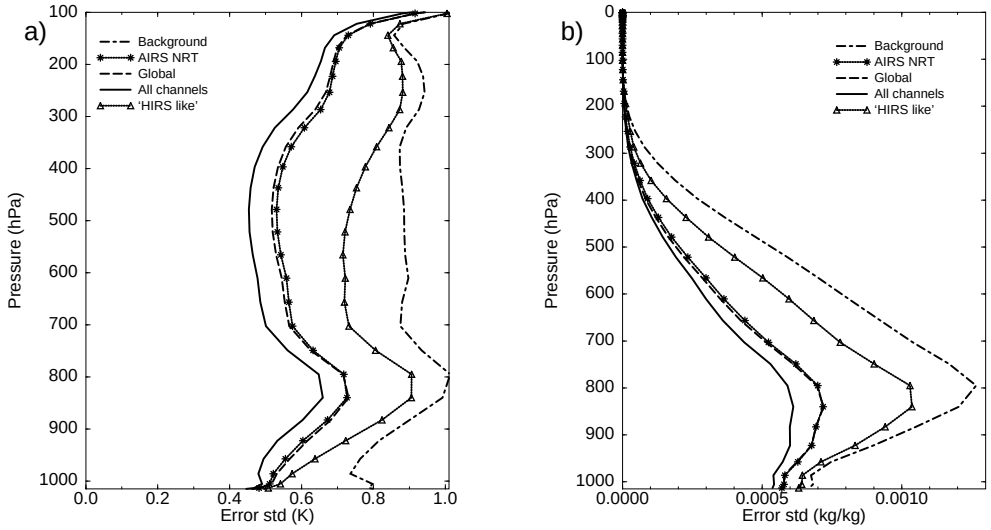


FIG. 2.6 – Ecart-types d'erreurs d'assimilation 1D-Var moyennés sur 108 profils atmosphériques pour la température (a) et l'humidité (b) pour différentes sélections de canaux. La courbe tiretée correspond à l'ébauche, la ligne continue avec des triangles à l'assimilation de canaux AIRS correspondant à ceux de HIRS, la ligne continue avec des étoiles à la sélection de la NOAA/NESDIS et tiretée à la sélection globale et la ligne continue à l'assimilation de l'ensemble des canaux AIRS.

les zones sensibles en utilisant les nuages du modèle du centre européen. Il a montré qu'il y avait une forte corrélation entre les zones météorologiquement sensibles et les nuages produits par le modèle du CEPMMT. Dans notre cas, nous avions pour but, d'une part, d'étudier la couverture nuageuse et la hauteur de sommet des nuages simulées à partir d'un imageur à bord d'un satellite défilant dans les pixels IASI. D'autre part, la robustesse de la sélection de canaux itérative de Rabier et al (2002) et de Fourrié et Thépaut (2003) a été étudiée pour ces régions et comparée à des sélections de canaux dédiées aux zones sensibles.

Le masque nuageux a été calculé à partir des observations de l'imageur AVHRR à bord des satellites défilants de la NOAA pour des cas de dépressions de la campagne FASTEX. Cet intrument a une résolution horizontale bien plus fine que le sondeur IASI et les données ont été traitées par le Centre de Météorologie Spatiale à Lannion à l'aide du masque nuageux MAIA (Mask AVHRR for Inversion ATOVS). Ce logiciel fournit différents paramètres nuageux qui sont le pourcentage de pixels AVHRR clairs ainsi que la température de sommet de nuages. La couverture nuageuse a été déterminée à partir du pourcentage de pixels clairs AVHRR dans le pixel simulé IASI. Un seuil arbitraire de 90% de pixels clairs a été choisi pour déclarer l'observation IASI claire. Pour les observations nuageuses, la pression de sommet des nuages est déduite de la température de sommet des nuages en comparant celle-ci à la température du profil atmosphérique de la réanalyse de la campagne FASTEX interpolé au point d'observation. La pression est trouvée quand la différence entre les deux températures est minimale.

Les zones sensibles pour la prévision ont été calculées à partir de la fonction objective de

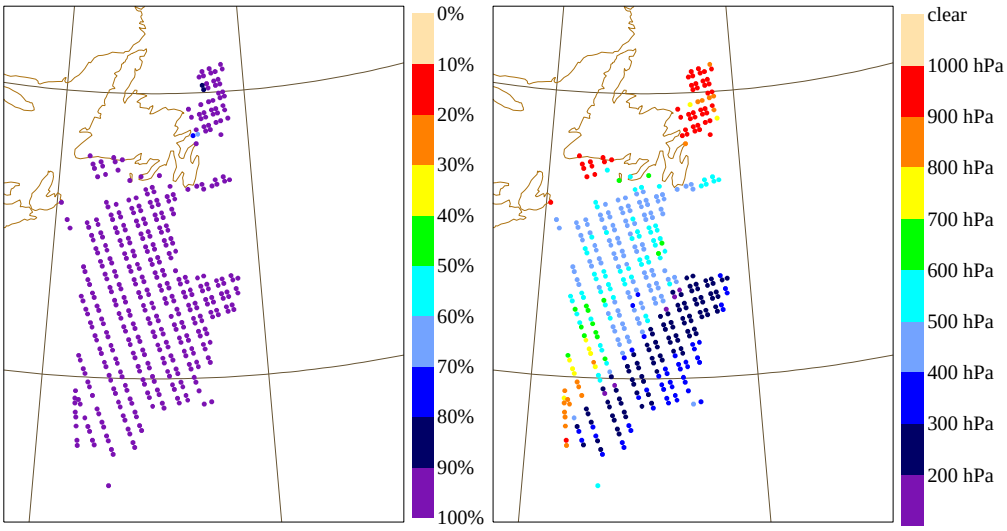


FIG. 2.7 – Pourcentage de sous-pixels AVHHR nuageux (à gauche) et pression de sommet des nuages (à droite) des pixels IASI pour la zone sensible du 18 février 1997 à 00 UTC.

Hello et al (2000) qui est la prévision de la pression au niveau de la mer sur la région des systèmes dépressionnaires de la campagne FASTEX. Huit systèmes du mois de février 1997 ont été étudiés. Les champs de température du gradient de cette fonction ont été élevés au carré et sommés pour définir une zone géographique. Ce traitement fait disparaître la nature penchée du gradient. Un seuil arbitraire de 30% de la valeur du maximum du champ de sensibilité a été fixé pour définir le masque horizontal des zones sensibles.

Ainsi la figure 2.7 montre la couverture nuageuse de la zone sensible quasi complète et la pression de sommet des nuages dans la zone zensible de la dépression du 18 Février 1997 à 00 UTC. La zone sensible pour la prévision de cette dépression est située au sud de Terreneuve. La partie sud de la zone sensible est recouverte de nuages élevés (entre 200 et 400 hPa) tandis que seule la partie située la plus au nord est recouverte de nuages bas (pression de sommet de nuages inférieure à 900 hPa).

Le résultat moyenné sur les cas de dépressions météorologiques de l'Atlantique Nord montre que les zones sensibles sont principalement localisées au large du continent Nord-Américain. Elles sont souvent situées à proximité de l'entrée sud du courant jet d'altitude mais leur position varie en fonction de la dépression étudiée. Plus la valeur de la sensibilité est faible, plus la zone est étendue. Les pixels clairs se trouvent majoritairement sur terre. Sur mer les pixels de la zones sensibles sont largement recouverts de nuages avec une couverture proche de 100%. Bien souvent, comme dans le cas de la figure 2.7, la partie sud de la zone sensible est recouverte de nuages élevés tandis que la partie nord est recouverte de nuages bas. Enfin, en moyenne il y a moins de pixels clairs dans les zones sensibles que sur l'ensemble de la région Atlantique Nord et le niveau des nuages y est plus élevé que pour l'ensemble de la région Atlantique Nord. Ces résultats sont en accord avec ceux trouvés par McNally (2002). Ceci est fort regrettable car l'information ne sera disponible qu'au dessus du nuage. L'échantillonnage de l'atmosphère et de la troposphère avec les observations des sondeurs hyperspectraux ne sera donc que partiel.

Nous avons cependant entrepris d'étudier une sélection de canaux adaptée aux régions sensibles avec des nuages très bas (1000 hPa). Même si ce cas est peu fréquent dans la réalité, il permet de maximiser l'information sondée par l'interféromètre. Nous avons repris la méthode

de Rabier et al (2002) basée sur la maximisation de la réduction d'entropie et nous l'avons comparée à d'autres méthodes issues du ciblage des observations. Les méthodes utilisées sont la sensibilité de la prévision aux observations (Baker et Daley, 2000) et (Doerenbecher et Bergot, 2001) et la sensibilité du filtre de Kalman (Bergot et Doerenbecher, 2002) qui permet de trouver le déploiement optimal des observations ciblées. La matrice $\mathbf{B}$ utilisée dans cette étude correspond aux erreurs clés de l'analyse, comme dans le cas de Fourrié et Thépaut (2003), qui a des erreurs plus faibles dans la stratosphère comparé à la matrice $\mathbf{B}$ climatologique.

En terme de contenu en information la méthode itérative basée sur l'ER est plus efficace que les méthodes issues du ciblage mais, en terme de qualité de l'analyse, les méthodes basées sur le gradient sont plus efficaces vers 850 hPa alors que dans la couche 450-100 hPa l'erreur d'analyse est plus grande que celle de l'ébauche. La sélection ER partage une centaine de canaux avec les méthodes de sélections basées sur l'adjoint du modèle de prévision, tandis que ces dernières ont 239 canaux en commun. Une étude sur un profil a ensuite été réalisée. L'étude des jacobiens en température (dérivée de la température de brillance selon le profil de température atmosphérique) montre que la sélection basée sur le contenu en information échantillonne l'atmosphère régulièrement tandis que les méthodes basées sur le gradient privilégient les régions vers 400, 500 et 800 hPa. Peu de canaux au dessus de 300 hPa sont sélectionnés par ces méthodes en raison des faibles valeurs dans la haute stratosphère de la matrice d'erreurs clés de l'analyse. En conclusion, la méthode basée sur le contenu en information semble assez robuste, même pour les zones sensible et ce, avec un coût de calcul moindre que les sélections basées sur le calcul du gradient.

Enfin nous avons montré l'importance de bien spécifier la matrice de covariances d'erreurs d'ébauche en utilisant pour les zones sensibles la matrice $\mathbf{B}$ dédiée ou la matrice $\mathbf{B}$ climatologique représentant les erreurs de prévision à court terme pour l'analyse. Cet aspect a beaucoup moins d'importance pour la sélection de canaux elle-même.

L'impact d'une sélection constante calculée à partir d'un ensemble de profils atmosphérique représentatif des régions des moyennes latitudes a été étudié comme pour les travaux de Rabier et al (2002) et de Fourrié et Thépaut (2003). Même si une légère dégradation par rapport à la méthode optimale est observée dans la troposphère et à la tropopause (Fig. 2.8), les erreurs obtenues avec la sélection constante sont du même ordre de grandeur que celles obtenues avec les autres méthodes. Le contenu en information de la sélection constante est plus grand que celui de la sélection basée sur la sensibilité du filtre de Kalman. La sélection constante apparaît donc comme un bon compromis entre la qualité de l'analyse obtenue et le coût de calcul.

Au moment du lancement de IASI en 2006, une sélection de canaux pour la transmission vers les centres de prévision numérique du temps opérationnels a été établie par Collard (2007). Elle a également été établie à partir d'études de contenu en information. Le critère choisi était le DFS. Elle comprend 300 canaux pour observer la température, l'humidité, l'ozone ainsi que le gaz carbonique. Après l'éviction de canaux IASI fortement sensibles aux erreurs de transfert radiatif, aux gaz trace dont la variabilité n'est pas considéré dans l'ébauche ou l'analyse, une sélection de canaux sensibles à la température a été réalisée. Ensuite les canaux sensibles à la vapeur d'eau ont été sélectionnés en considérant l'analyse de la température, puis les canaux sensibles à l'ozone, le gaz carbonique, le méthane, le monoxyde de carbone et le protoxyde d'azote et enfin les canaux affectés par le rayonnement solaire. Actuellement 14 canaux pour étudier les espèces chimiques ont été rajoutées à la sélection de 300 canaux et 314 canaux sont utilisés pour l'assimilation à Météo-France et au UK Met. Office. Ce jeu de canaux est plus grand au CEPMMT (366 canaux, Collard et McNally, 2009). Les 66 canaux supplémentaires se décomposent entre 43 canaux dans la bande à 15 μm sondant la basse stratosphère et 23 dédiés à l'intercomparaison des statistiques de monitoring entre les centres de PN et le CNES.

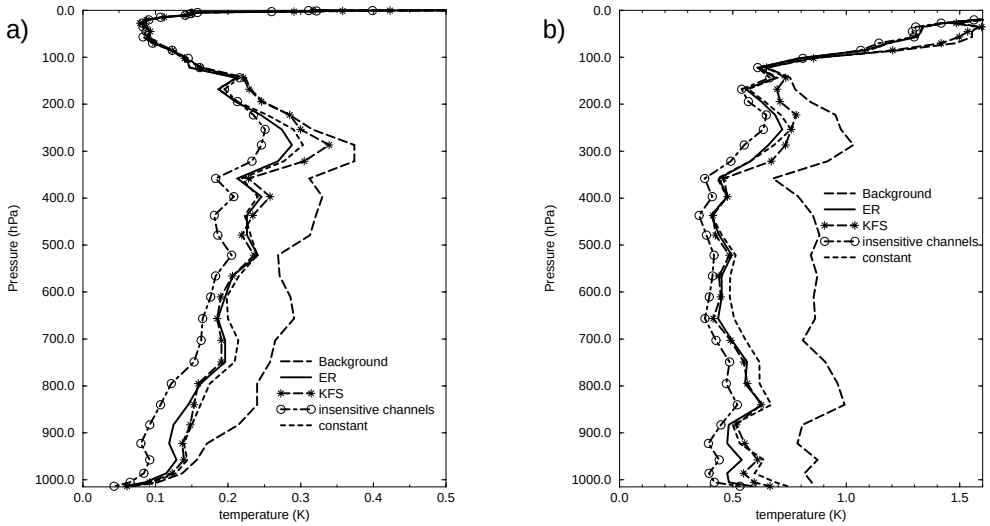


FIG. 2.8 – Erreur quadratique moyenne pour 4 méthodes de sélection de canaux (Réduction d'entropie ER, Sensibilité du filtre de Kalman KFS, canaux non contaminés par le nuage et sélection constante calculées sur 142 profils correspondants à un niveau de nuage situé à 1000 hPa pour une matrice $\mathbf{B}$ correspondant aux erreurs clés de l'analyse (a) et pour une matrice d'erreur d'ébauche climatologique (b).

L'approche proposée sera très utile pour d'autres futurs intruments. Cependant il est nécessaire d'avoir à disposition une base de données représentatives des profils atmosphériques pour réaliser une sélection adéquate. Par contre en ce qui concerne la sélection dans les zones sensibles, les résultats ne sont pas probants à ce stade. Il aurait peut-être fallu aller plus loin avec des données réelles en enchaînant les trois étapes : sélection, assimilation puis prévision à partir de l'analyse pour bien quantifier l'impact d'une sélection particulière dans les zones sensibles.

2.6 Amélioration de l'opérateur d'observation

Comme je l'ai expliqué dans la partie 2.3.1, l'opérateur d'observation est l'ensemble des opérations qui permettent de calculer la simulation de l'observation à partir des champs du modèle atmosphérique. Ainsi, dans les cas des radiances, il comprend les interpolation horizontales et verticales au point d'observations ainsi que le modèle de transfert radiatif RTTOV.

2.6.1 Assimilation des radiances nuageuses

Comme nous l'avons vu au chapitre précédent, les régions où les erreurs de prévision des tempêtes croissent le plus sont souvent recouvertes de nuages. Dès lors, je me suis intéressée à l'assimilation des radiances des sondeurs hyperspectraux en présence de nuages afin d'optimiser leur assimilation. Jusqu'il y a peu de temps, avant de procéder à l'assimilation, une vérification était faite pour déterminer si l'observation avait été réalisée en présence de nuages ou non. Si tel était le cas, elle était écartée du processus d'assimilation.

Ce travail s'inscrivait dans la continuité du travail de Mohamed Dahoui avec lequel j'ai travaillé lors de la fin de sa thèse. Dans un premier article (Dahoui *et al.*, 2005), Mohamed avait étudié différents schémas de détection et de caractérisation de nuages. Il a trouvé que la méthode dite du CO2-slicing (Chahine, 1974)(Menzel *et al.*, 1983) ainsi que le schéma de détection nuageuse de (McNally et Watts, 2003) avaient de bonnes performances et pouvaient permettre l'assimilation des radiances nuageuses sous certaines conditions pour le CO2-slicing et l'assimilation des radiances juste au dessus du nuage dans le deuxième cas.

Le schéma de détection nuageuse utilisée au CEPMMT (McNally et Watts, 2003) cherche quant à lui à déterminer à partir de quel canal on a une signature nuageuse et à distinguer les canaux non affectés par la présence d'un nuage. Elle est basée sur l'étude des innovations calculées en ciel clair et ordonnées verticalement en fonction de la sensibilité du canal à un nuage. En effet le spectre des innovations est plus ou moins nul sauf quand les canaux commencent à être affectés par la présence d'un nuage. On recherche ainsi le canal à partir duquel les innovations lissées par un filtre passe-bas croissent (ou décroissent) de façon monotone indiquant un effet de plus en plus marqué du nuage. Une fois ce canal trouvé, tous les canaux situés en dessous sont déclarés nuageux et ceux au dessus déclarés clairs.

Le CO2-slicing consiste à trouver pour certains canaux de la bande spectrale du CO_2 à 15 μm une pression p_k pour laquelle la fonction ci-dessous est minimale (Eq. 2.11) :

$$F_{k,p} = \frac{R^k_{clear} - R^k_{meas}}{R^{k_{ref}}_{clear} - R^{k_{ref}}_{meas}} - \frac{R^k_{clear} - R^{k,p}_{cld}}{R^{k_{ref}}_{clear} - R^{k_{ref},p}_{cld}} \tag{2.11}$$

avec :

- p : niveau de pression.
- k : numéro du canal dans la bande CO_2.
- k_{ref} : canal fenêtre de référence.
- R^k_{meas} : radiance mesurée pour le canal k.
- R^k_{clear} : radiance claire simulée pour le canal k.
- $R^{k,p}_{cld}$: radiance simulée d'un corps noir pour le canal k en considérant un nuage opaque au niveau p.

La pression de sommet effective du nuage est trouvée par une moyenne pondérée (Eq 2.12).

$$PTOP = \frac{\sum w_k^2 p_{c,k}}{\sum w_k^2} \tag{2.12}$$

où w_k représente la dérivée par rapport à la pression de la fonction $F_{k,p}$ ($w_k = \delta F_{k,p}/\delta lnp$).

Ensuite, la fraction nuageuse effective est trouvée à partir de cette pression de sommet des nuages avec un canal de surface :

$$N_e = \frac{R_{clear}^{k_{ref}} - R_{meas}^{k_{ref}}}{R_{clear}^{k_{ref}} - R_{cld}^{k_{ref},PTOP}} \tag{2.13}$$

Mohamed sur la fin de sa thèse avait testé l'assimilation des radiances AIRS nuageuses dans un contexte de recherche. En effet début 2006 les radiances AIRS n'étaient pas encore assimilées dans le modèle opérationnel. Il avait trouvé que l'on pouvait assimiler les radiances nuageuses avec un léger impact positif (Dahoui, 2006) pour des cas de nuages bas.

Avec Thomas Pangaud, doctorant que j'ai encadré, nous avons repris les travaux de Mohamed Dahoui, cette fois-ci dans un contexte opérationnel. En effet les radiances AIRS ont été assimilées opérationnellement en septembre 2006, d'abord des canaux stratosphériques puis des canaux sondant la haute troposphère. Les résultats que je vais présenter ici sont ceux de l'article de Pangaud et al (2009) publié dans Monthly Weather Review.

Dans un premier temps, les schémas de détection nuageuse (CO2-slicing et schéma de McNally& Watts (2003)) ont été évalués à l'aide de l'imageur MODIS (Moderate resolution Imaging Spectroradiometer) à bord du satellite Aqua sur le domaine Atlantique. Les données AIRS et MODIS ayant une résolution horizontale différente (5km pour MODIS au nadir), elles ont été colocalisées avec les pixels AIRS. Les deux types de détection donnent de bons résultats comparés à la nébulosité déduite de MODIS. Les performances de détection varient suivant la hauteur des nuages mais il y a une bonne qualité de détection pour les nuages de niveau moyen compris entre 400 et 800 hPa alors que les nuages bas et très bas ne sont pas bien détectés par les deux schémas. Ce dernier point a une moindre importance car ces nuages ont un faible impact radiatif et donc peu d'impact sur la simulation des observations nuageuses. Alors que la pression de sommet de nuages déduite de AIRS est bien corrélée avec celle déduite de MODIS, la fraction nuageuse l'est beaucoup moins. Ceci est en accord avec les résultats trouvés par Pavelin *et al* (2008).

A la vue de ces résultats, des expériences d'assimilation ont été menées sur le mois de septembre 2006 en utilisant les paramètres nuageux (pression de sommet des nuages et fraction nuageuse effective) en entrée du transfert radiatif pour la simulation des observations AIRS à partir de l'ébauche. Pour tester l'impact des radiances nuageuses, l'expérience a été comparée à une réference qui n'assimilait que les radiances AIRS claires et qui correspondait à ce qui était fait dans la version opérationnelle. L'option choisie combine le schéma de détection du CEPMMT et le CO2-slicing pour l'assimilation des radiances nuageuses. Le schéma de la figure 2.9 illustre l'assimilation des radiances nuageuses dans le modèle. La pression de sommet des nuages et la fraction nuageuse sont utilisées en entrée de RTTOV. Si la pression de sommet des nuages est comprise entre 600 et 950 hPa, alors les canaux déclarés nuageux par le schéma de McNally and Watts seront rachetés et assimilés. Si le pixel est considéré clair par le CO2-slicing, les canaux déclarés clairs par le schéma de détection seront assimilés.

La prise en compte des paramètres nuageux dans la simulation et l'assimilation permet d'accroître faiblement le nombre d'observations assimilées (3.5%) et de 10% le nombre de canaux assimilés. Les observations supplémentaires se situent principalement aux moyennes latitudes et dans les régions polaires. L'analyse est plutôt modifiée dans ces régions. Des scores globaux ont été calculés pour évaluer la qualité des simulations réalisées à partir de ces analyses. L'assimilation des radiances nuageuses AIRS a un impact neutre à légèrement positif sur les prévisions. Ceci peut s'expliquer par le fait que les canaux assimilés se situent principalement dans la haute troposphère et la stratosphère (voir Fig. 2.5 page 23) et les nuages bas

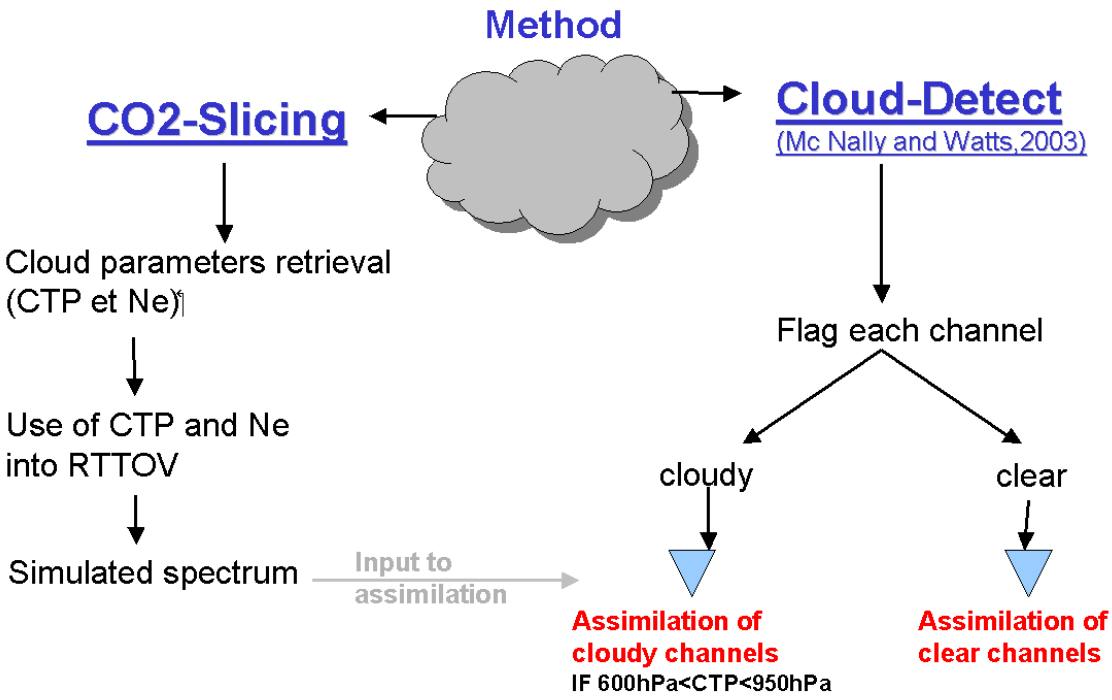
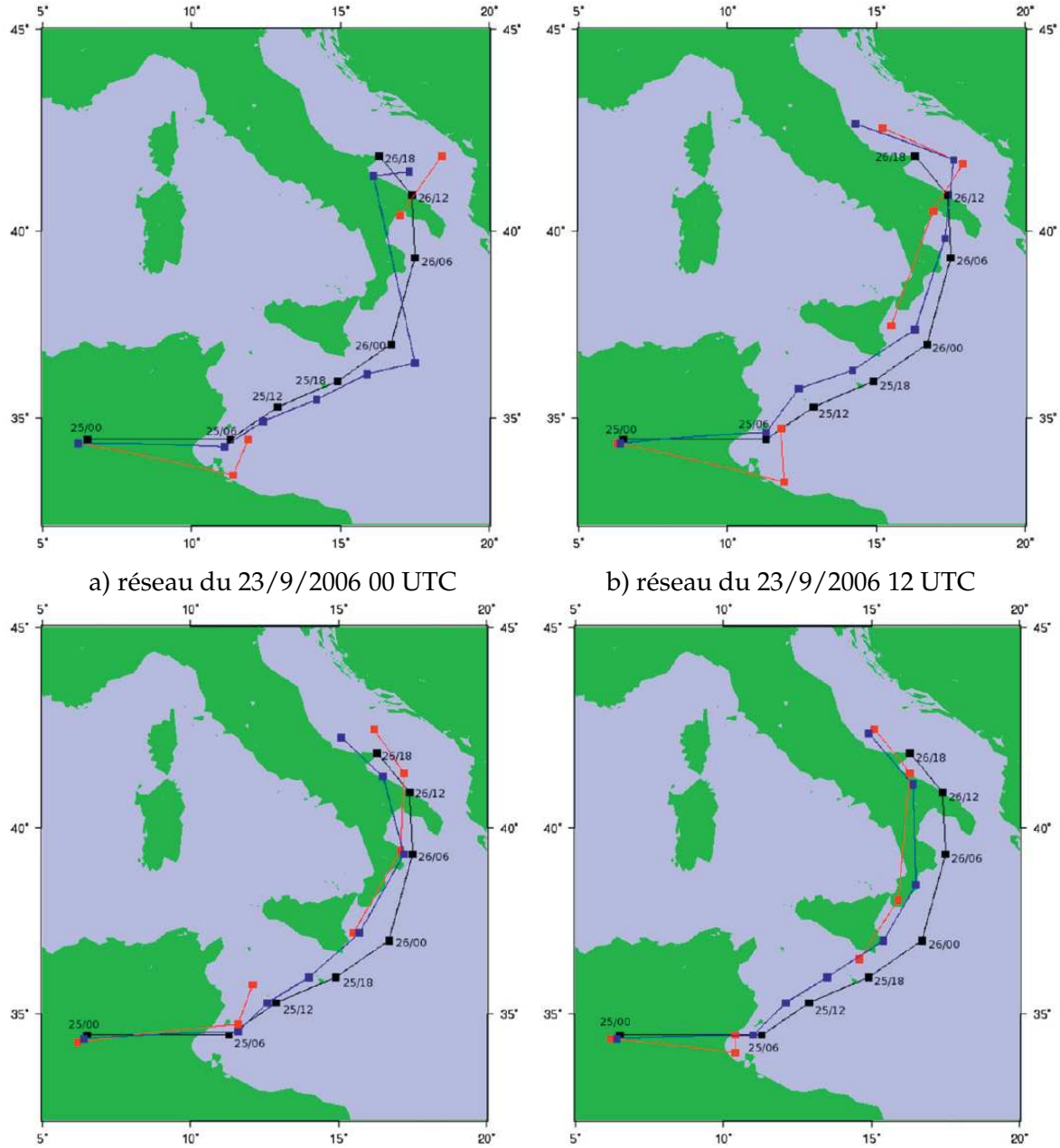


FIG. 2.9 – Schéma d'assimilation des radiances nuageuses dans ARPEGE

auront donc peu d'impact sur les observations. C'est d'ailleurs ce qui était souhaité dans notre démarche.

Une étude de cas a été également menée. Il s'agit du Medicane (Moscatello *et al.*, 2008), une dépression de petite échelle qui s'est développée au sud est de l'Atlas à la frontière de l'Algérie et de la Tunisie. Cette dépression a évolué en tempête en traversant la Méditerranée et a touché le sud-est de l'Italie avec de fortes précipitations. Des simulations ont été lancées à partir de 4 dates différentes et nous avons trouvé que l'assimilation des radiances nuageuses AIRS donnait de meilleures prévisions que la référence en terme de prévision (la dépression disparaît sur certains réseaux des simulations de la référence, Fig. 2.10), de localisation et d'intensité du système dépressionnaire. Enfin les précipitations engendrées par le système, même si elles restent encore sous-estimées (en raison du modèle utilisé qui est un modèle global peu adapté à la simulation de ce genre de paramètre), sont mieux simulées en assimilant les radiances nuageuses. Dans ce cas précis, il semble que ce soit l'amélioration de la qualité de l'ébauche par le cycle continu d'assimilation des radiances nuageuses qui soit responsable de l'amélioration des prévisions.

L'assimilation des radiances AIRS nuageuses a été introduite dans le modèle opérationnel en Février 2009. D'un point de vue international, des travaux de recherches similaires sur l'assimilation des radiances des sondeurs hyperspectraux ont été menés. Le UK Met Office (UKMO) (Pavelin *et al*, 2008), le Centre de Météorologie Canadien CMC (Heilliette *et al*, 2007) et le Centre Européen (McNally, 2009) ont également travaillé sur l'assimilation des radiances AIRS et IASI « nuageuses ». Ils ont adopté la même stratégie pour les assimiler dans un premier temps : avant de procéder à la minimisation, les paramètres nuageux nécessaires à une bonne simulation des observations en présence de nuages sont calculés puis fixés au cours du processus d'analyse pour le UKMO ou ajustés au cours des minimisations pour le Centre Européen et le Service Ca-



a) réseau du 23/9/2006 00 UTC

b) réseau du 23/9/2006 12 UTC

c) réseau du 24/9/2006 00 UTC

d) réseau du 24/9/2006 12 UTC

FIG. 2.10 – Trajectoire des dépressions prévues par l'expérience assimilant les radiances AIRS nuageuses (en bleu) et par la référence n'assimilant que les radiances claires (en rouge) en haut à partir de l'analyse du 23 septembre à 0 UTC (a) et à 12 UTC (b) et en bas à partir de l'analyse du 24 septembre 2009 à 0 UTC (c) et à 12 UTC (d). En noir est tracée la trajectoire de la dépression déduite des analyses opérationnelles du CEPMMT.

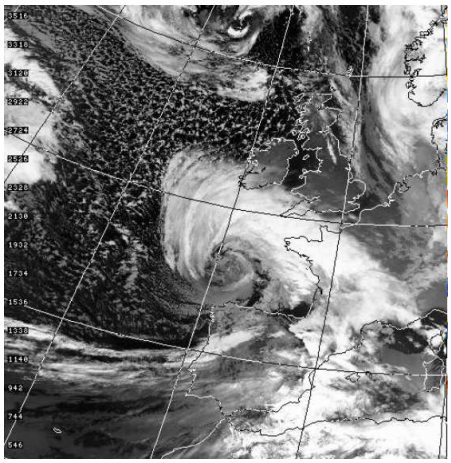


FIG. 2.11 – Pression de sommet des nuages (hPa) déduite des observations IASI pour la tempête "Klaus" du 24 janvier 2009 à 00 UTC et image AVHRR du canal 5 du 23 janvier 2009 vers 21h40.

nadien. L'assimilation des radiances AIRS nuageuses est opérationnelle depuis septembre 2008 au UKMO et depuis Mars 2010 pour IASI et celle de AIRS et IASI depuis septembre 2009 au Centre Européen. ARPEGE a donc été le deuxième modèle global à bénéficier de l'assimilation des radiances nuageuses.

Des développements sont en cours pour le sondeur IASI comme l'illustre la figure 2.11 avec la pression de sommet des nuages déduite des observations IASI pour la tempête "Klaus" qui a frappé le sud de la France le 24 janvier 2009. Ainsi, la forme en virgule de la dépression est bien retrouvée à partir des observations IASI ainsi que la bande de nuages associés au front le long de l'Océan Atlantique.

La méthode du CO2-slicing ne permet de retrouver qu'un seul niveau de nuages, ce qui est loin d'être la majorité des cas dans la réalité, où bien souvent plusieurs couches de nuages se recouvrent. Même si l'approche de l'assimilation des radiances nuageuses avec un nuage à une couche soit assez limitée, elle a permis d'avoir un gain rapide dans l'analyse et de me familiariser avec les radiances nuageuses.

2.6.2 De nouveaux opérateurs d'observation pour l'assimilation à méso-échelle sur la Méditerranée

Avec l'avènement des modèles régionaux tels que le modèle AROME, la problématique de la représentation des observations satellites dont la résolution est inférieure à celle du modèle se pose. C'est ce que nous avons voulu explorer dans la première partie du travail de thèse de Fanny Duffourg que je co-encadre avec Véronique Ducrocq. La thèse de Fanny Duffourg est intitulée « Assimilation des données satellitaires infrarouges pour la prévision et la compréhension des pluies intenses en région méditerranéenne » et se place dans le cadre de la préparation d'HyMeX dont l'objectif est d'améliorer la caractérisation et la compréhension du cycle de l'eau sur le bassin Méditerranéen.

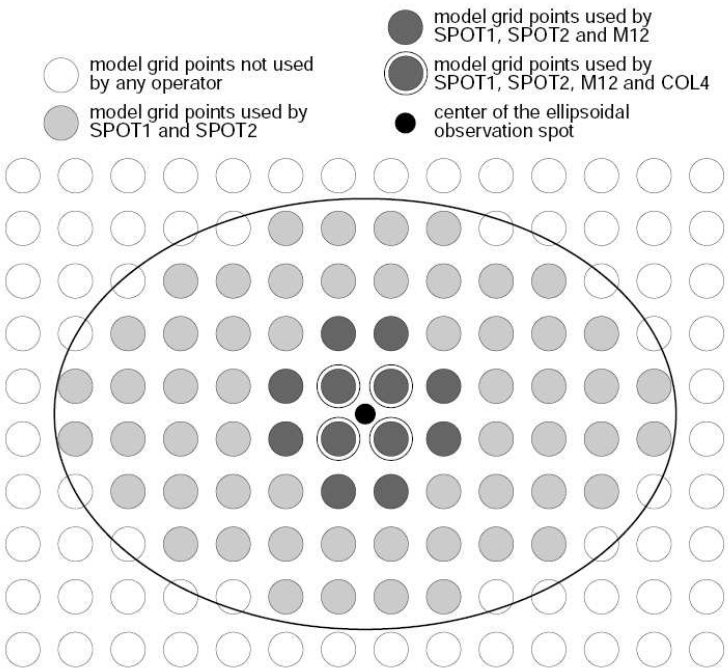


FIG. 2.12 – Points de grille du modéle agrégés par différents opérateurs d'observation pour simuler l'observation IASI représentée par l'ellipse.

Dans le cadre de la thèse de Fanny Duffourg, une approche multi-capteurs (AIRS/IASI) a été adoptée pour optimiser la présence d'observations satellites dans chaque fenêtre d'assimilation des systèmes d'assimilation 3D-Var ALADIN et AROME. On cherche notamment à exploiter toute la résolution spatiale (environ 12 km au nadir) des données satellitaires en utilisant un système d'assimilation à forte résolution spatiale. On a donc étudié la meilleure façon de simuler l'observation à partir du modèle. Jusqu'à présent, on a supposé que la maille du modèle est plus grande que le pixel observé, et l'on compare la radiance simulée en un point de grille à la radiance observée. Celle-ci est ainsi affectée au point de grille le plus proche (aux interpolations horizontales près). Actuellement, seuls 4 points sont utilisés pour calculer le profil interpolé utilisé pour simuler l'observation satellitaire. Ceci n'a plus de sens lorsque le pixel a une taille largement supérieure à la maille du modèle comme le montre la figure 2.12. Comme une observation satellitaire AIRS ou IASI couvre au minimum 12 points de grille au nadir, on a modifié l'opérateur d'observation pour agréger l'information du modèle dans le champ de vue de l'observation satellitaire afin de mieux représenter la sensibilité de la mesure à toute l'atmosphère comprise dans son champ de vue (Figure 2.12).

Plusieurs opérateurs d'observations ont été évalués pour les simulations de AIRS et IASI sur une période de 1 mois. Le premier appelé COL4 est celui qui utilise l'interpolation à partir de 4 points du modèle les plus proches pour simuler l'observation. C'est celui utilisé actuellement dans la version opérationnelle d'AROME (version juin 2010). Comme la fonction d'instrument est quasi-uniforme sur le champ de vue, le même poids est affecté à chaque profil du modèle dans les nouveaux schémas d'aggrégation de l'information du modèle. SPOT1 représente la moyenne de toutes les colonnes du modèle contenues dans le champ de vue. La colonne moyenne est fournie en entrée de RTTOV. Dans M12, le profil atmosphérique est estimée à partir des 12 points du modèle entourant le centre du pixel. Enfin un troisième opérateur

opérateur	méthode
COL4	interpolation 4 points
M12	moyenne sur 12 points avant le transfert radiatif
SPOT1	moyenne sur le spot avant le transfert radiatif
SPOT2	moyenne sur le spot après le transfert radiatif

TAB. 2.1 – Caractéristiques des différents opérateurs d'observation.

d'observation (SPOT2) où la moyenne est réalisée après le calcul du transfert radiatif a été testé. Ce dernier opérateur est très coûteux en temps calcul car pour chaque point du modèle contenu dans le champ de vue la simulation du transfert radiatif est réalisée. Le tableau 2.1 rappelle les différentes caractéristiques des opérateurs d'observations.

On s'est intéressé plus particulièrement aux observations en présence d'un gradient important sur le champ d'humidité. Les nouveaux opérateurs d'observation améliorent la simulation des canaux vapeur d'eau. Les simulations sont surtout modifiées pour les canaux vapeur d'eau dont les fonctions de poids sont maximales au dessus de 340hPa pour IASI et comprises entre 340 et 800 hPa pour AIRS (Figure 2.13). Les différences les plus importantes apparaissent pour les gradients d'humidité de fine échelle dans des couches atmosphériques sèches.

Agréger l'information du modèle contenue dans le pixel du satellite avant ou après le transfert radiatif ne conduit pas à des différences significatives dans la plupart des cas. La moyenne sur 12 points fournit des résultats intermédiaires entre l'ancien opérateur d'observation COL4 et SPOT1.

Pour la plupart des canaux de la bande vapeur d'eau (Figure 2.14) et de la bande 1, l'écart-type des innovations (observations – ébauche) est réduit et le biais est inchangé. Les nouveaux opérateurs d'observation améliorent les simulations à partir de l'ébauche en éliminant les structures de fine échelle non détectées par les instruments mais présentes dans les champs du modèle.

Les nouveaux opérateurs d'observations pourraient éviter le rejet de certaines observations par les procédures de contrôle de qualité durant l'assimilation de données et réduire les mauvais incréments d'analyse dus aux erreurs de réprésentativité. Des expériences d'assimilation à une observation ont ensuite été menées en utilisant les différents opérateurs d'observations. Elles montrent que malheureusement les modifications dans les simulations des observations n'ont que très peu d'impact sur l'analyse finale.

Ces résultats ont été rassemblés dans un article publié au JGR (Duffourg *et al.*, 2010). Il reste cependant à tester ces nouveaux opérateurs d'observations dans un cycle continu d'assimilation. Ainsi on pourra tester la rétroaction du nouvel opérateur d'observation sur l'analyse mais également sur la correction de biais et la détection nuageuse.

Joiner et Poli (2005) avaient étudié l'impact sur la simulation de la prise en compte des gradients horizontaux en ne considérant plus la température de brillance le long d'un profil vertical mais le long d'un profil incliné. Même si en général, l'effet de ces gradients est en dessous du bruit de l'instrument, il y a certains cas où la prise en compte de tels gradients dans l'opérateur d'observation serait bénéfique comme les gradients d'humidité de la haute troposphère. L'article décrit dans ce paragraphe pourrait être ainsi complété dans cette direction.

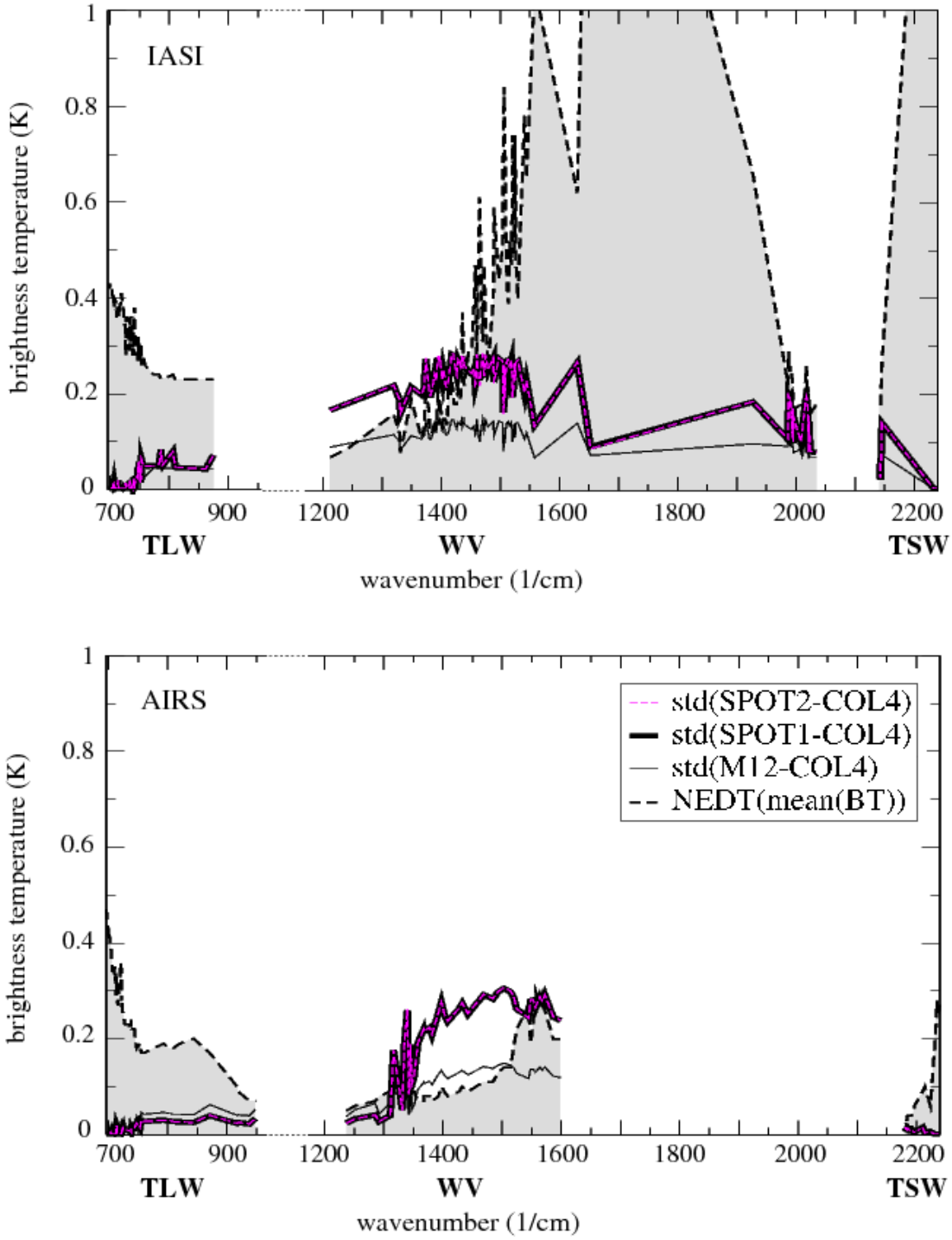


FIG. 2.13 – Ecart-type (en K) des différences entre les températures de brillance simulées entre l'ancien opérateur d'observation (COL4) et un nouvel opérateur d'observations (M12 en ligne fine, SPOT1 en ligne épaisse noire et SPOT2 en ligne tiretée rose pour les canaux IASI (en haut) et les canaux AIRS (en bas) en fonction du nombre d'ondes. Le bruit estimé de chaque instrument à la température moyenne mensuelle est représenté par la ligne tiretée noire avec l'aire sous la ligne colorée en gris.

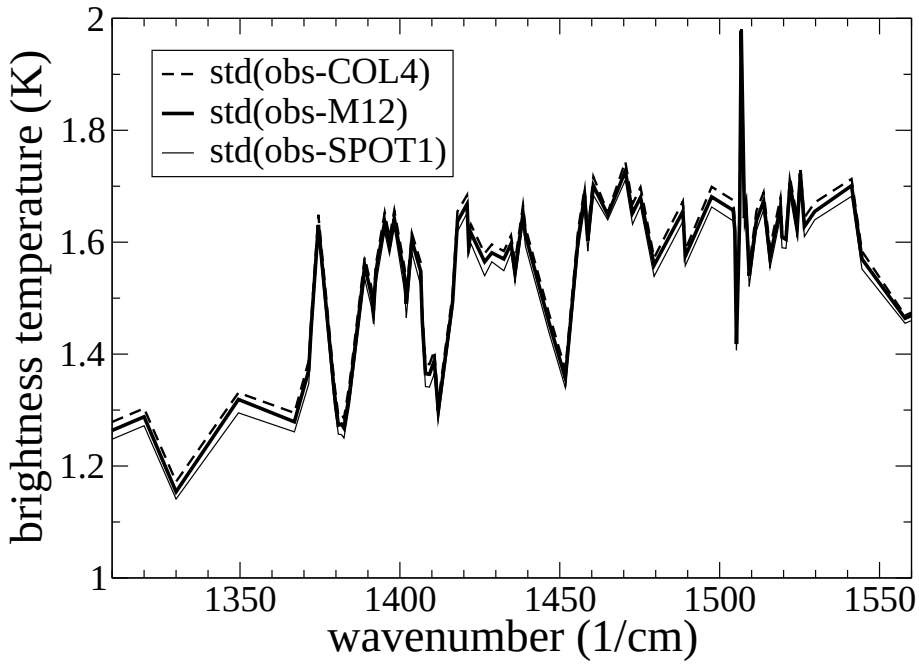


FIG. 2.14 – Ecart-types des différences de températures de brillance entre les observations et l'ébauche obtenues avec les 3 opérateurs d'observation pour les observations IASI claires sur la partie ouest de la Méditerranée en Septembre 2007 dans tous les canaux vapeur d'eau étudiés en fonction du nombre d'ondes.

2.7 Evaluation de l'impact des observations pour la prévision des tempêtes

2.7.1 Etude de l'impact des données d'une campagne de mesures

Je présente ici l'étude d'impact de l'article de Fourrié *et al* (2006) qui a été réalisée lors du stage de DEA d'un élève ingénieur de l'Ecole de la Météorologie, David Marchal que j'ai encadré. La campagne expérimentale NA-TReC (North Atlantic THORPEX Regional Campaign) s'est déroulée sur l'Océan Atlantique Nord en octobre-décembre 2003 dans le cadre du programme international de l'Organisation Mondiale Météorologique THORPEX. Ce programme de 10 ans a pour but d'améliorer la précision des prévisions de 1 à 14 jours des événements météorologiques à fort impact sur la société. THORPEX est divisé en plusieurs sous-programmes de recherche : influence de l'échelle globale vers l'échelle régionale sur l'évolution et la prévisibilité des systèmes météorologiques, étude et démonstration des systèmes d'observation, ciblage et assimilation de données. THORPEX a également pour objectif de faciliter les collaborations internationales entre les institutions de recherche académique, les services opérationnels de prévision du temps et les utilisateurs. La campagne NA-TReC avait pour but de montrer la faisabilité d'un ciblage quasi-opérationnel des observations sur l'Océan Atlantique Nord. Plusieurs centres, dont Météo-France, avaient participé à la détermination des zones sensibles. Les données supplémentaires étaient des dropsondes lâchées depuis des vols d'avions et des radiosondages effectués aux réseaux secondaires de 6h et 18h.

Dans un premier temps, l'impact des données de la campagne sur une période de trois semaines (27 Novembre-14 Décembre 2003) a été testé dans le 4D-Var ARPEGE. Les données supplémentaires de la campagne ayant été prises en compte dans le modèle opérationnel, une ré-analyse sans ces observations pendant les trois dernières semaines de l'expérience comportant plusieurs cas d'étude intéressants a été réalisée. Il s'avère que l'impact des données additionnelles de la campagne de mesure est très légèrement positif sur les grandes zones de vérification (partie ouest de l'Océan Atlantique et Europe). Les tests statistiques de Student ont montré que cet impact est toutefois peu significatif. Le seul bénéfice apporté par les données de la campagne concerne une meilleure description et évolution de la température de la basse stratosphère sur l'Europe. Cette amélioration est due aux radiosondages supplémentaires lachés aux réseaux de 6 et 18 UTC. Ces résultats sont en accord avec ceux obtenus dans les autres centres météorologiques (Rabier *et al.*, 2008) comme le CEPMMT (Cardinali et Buizza, 2004) et le UK Met-Office (Petersen et Thorpe, 2007). Si on s'intéresse à la zone de vérification pour laquelle la zone sensible a été calculée (Fig. 2.15), les résultats sont plus contrastés. En effet, l'impact des données peut être positif, neutre ou même dégrader la qualité de la prévision. De plus, dans le cas des améliorations, il n'y a pas de lien évident entre le nombre d'observations supplémentaires dans la zone de ciblage et la valeur de l'amélioration de la prévision. En effet la plus grande amélioration obtenue dans le cas 25 du 3 Décembre 2003, aucune observation supplémentaire n'avait été réalisée. Cette amélioration peut être expliquée par le cyclage des assimilations : les analyses servent à faire une ébauche qui va être utilisée pour assimiler les autres observations supplémentaires. L'amélioration de la qualité de l'ébauche permettra une meilleure analyse grâce à une meilleure simulation et assimilation des données. De plus l'assimilation de données sert à améliorer les ébauches dans leur ensemble mais il se peut qu'une analyse soit dégradée ponctuellement par l'assimilation de données.

La seconde partie de l'étude repose sur l'examen du contenu en information des différentes sources de données d'une analyse. Il a été mené à l'aide du DFS développé dans le cadre d'ARPEGE par Bernard Chapnik du GAME. Chapnik *et al* (2006) ont proposé une méthode qui

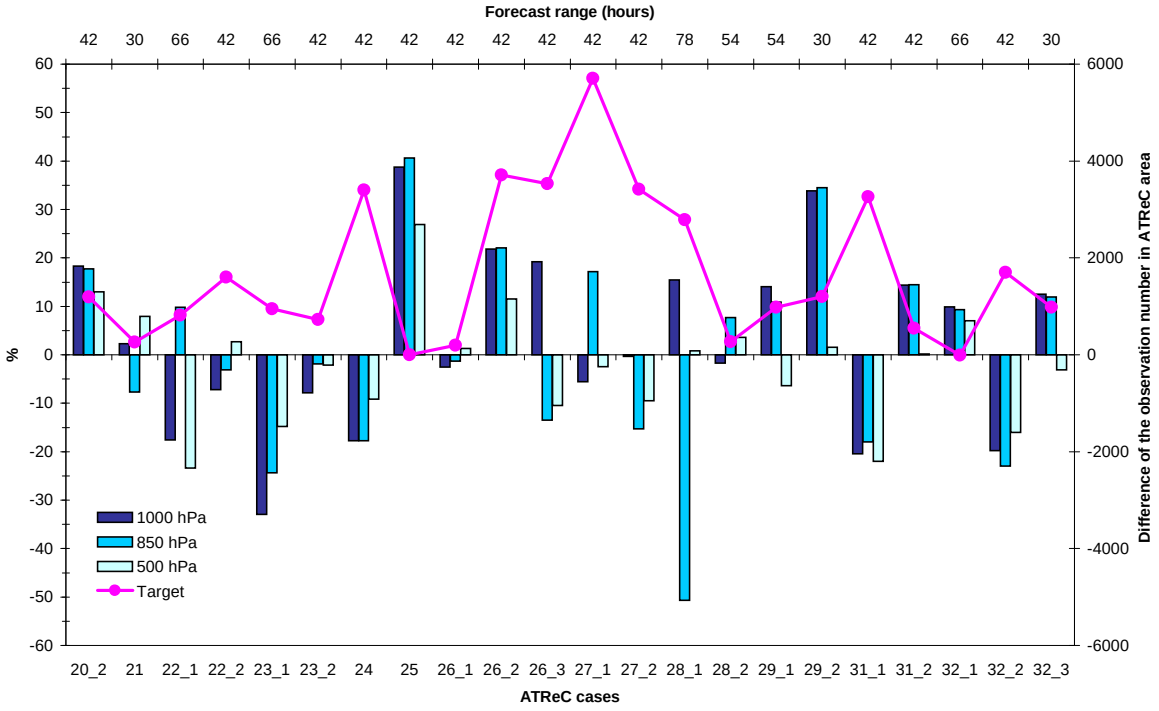


FIG. 2.15 – Pourcentage de différences d'erreur quadratique moyenne en géopotentiel à 1000, 850 et 500 hPa calculé sur la zone de verification au temps requis entre l'expérience avec les données de la campagne NA-TREC et celle sans ces données (un pourcentage positif indique un impact positif de l'expérience avec les données NA-TReC). Le nombre de données supplémentaires dans la région NA-TReC est donné par la ligne rose pour les 22 cas NA-TReC étudiés.

permet l'estimation d'une matrice connue seulement sous la forme d'un opérateur en réalisant deux analyses dont une est faite avec des observations perturbées. Ainsi on peut déterminer un DFS partiel correspondant à chaque type d'observations assimilées si celui si est non-corrélé avec le reste des observations. Pour plus de robustesse dans le calcul de ce diagnostic, nous avons utilisé 10 réalisations différentes des perturbations et appliqué la méthode de Chapnik *et al* (2006) sur un cas d'étude de la campagne NA-TReC pour connaître la répartition géographique de l'impact des observations sur l'analyse. Ce cas d'étude concerne la tempête de neige du 5 décembre 2003 qui a affecté la côte est des Etats-Unis et du Canada.

Avant de regarder les résultats et afin de mieux comprendre comment évolue ce paramètre, des expériences numériques ont été menées dans un modèle simple et uni-dimensionnel d'assimilation de données. Le DFS correspond à la trace de la matrice résultant du produit de la matrice de gain de l'assimilation par l'opérateur d'observation. Nous avons montré que le DFS a tendance à augmenter quand la longueur de corrélation d'erreur d'ébauche décroît (Figure 2.16), les observations pourront modifier l'analyse plus facilement indépendamment des autres observations. La variation du DFS en fonction du rapport entre l'erreur d'observation et l'erreur d'ébauche est illustrée par la figure 2.17. Le DFS diminue quand ce rapport augmente, c'est-à-dire quand l'erreur d'observation est plus grande que celle de l'ébauche. Une observation peut potentiellement plus modifier l'analyse si son erreur associée est petite. Enfin, quand la densité des observations augmente, le DFS augmente mais le DFS par observation diminue (Figure 2.18) car chaque contribution individuelle diminue même si la contribution globale

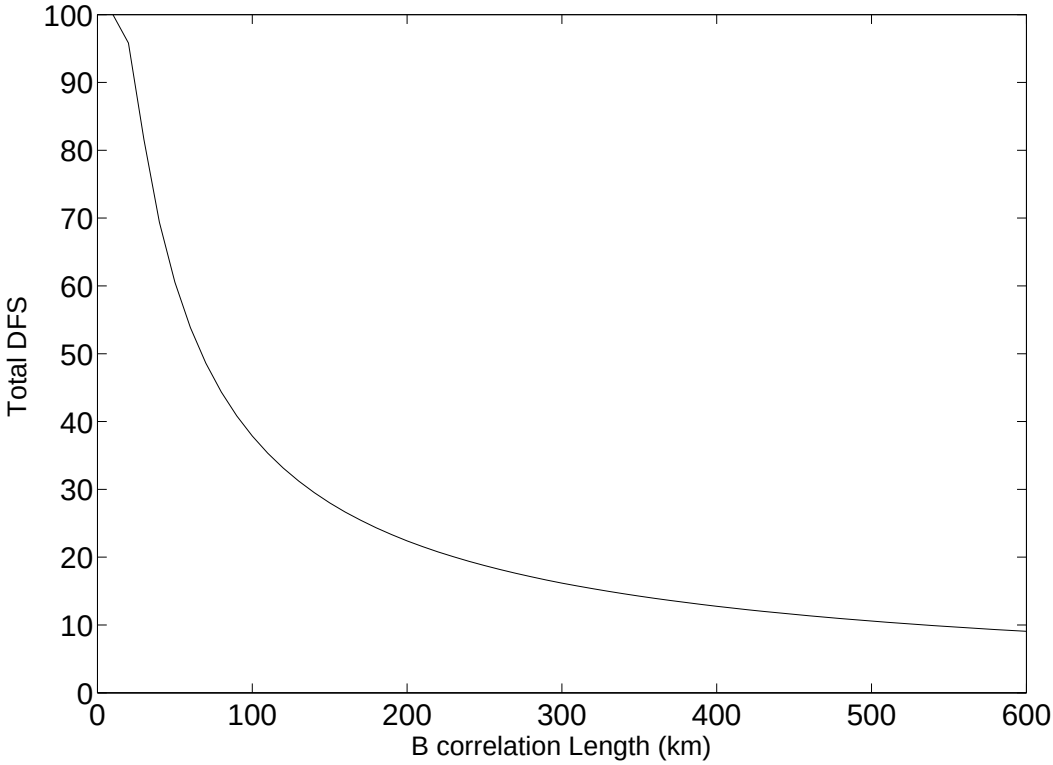


FIG. 2.16 – Influence de la longueur de corrélation sur le DFS total

augmente.

La méthode de calcul du DFS mise en oeuvre permet de le calculer pour chaque paramètre analysé, à savoir, la température, le vent, l'humidité et les températures de brillance. Nous avons montré que l'impact des données de vent est bien plus grand que celui de la température, notamment au niveau de la haute troposphère et de la basse stratosphère.

Dans le modèle ARPEGE, les observations satellitaires ont l'impact le plus important sur l'analyse sur le globe et plus particulièrement les données du sondeur infrarouge HIRS. On observe moins de disparité entre les différents canaux des sondeurs mais on constate que les canaux sensibles à l'humidité dans la haute troposphère de HIRS (n^{o}11 et 12) ont l'impact global et unitaire (c'est-à-dire par observation) le plus important (Fig. 2.19).

L'impact des observations a été ensuite étudié dans la zone sensible pour l'initialisation des tempêtes où les observations supplémentaires ont été déployées. Dans ce cas, les données fournies par les avions de ligne puis celles provenant des dropsondes ciblées ont le plus d'impact sur l'analyse. Si l'on s'intéresse à la répartition géographique du DFS dans la zone sensible, on constate que les dropsondes qui ont le poids le plus important localement, se situent dans le maximum principal de la zone sensible alors que les données avions qui contribuent le plus à l'analyse se situent au contraire dans le maximum secondaire. En étudiant l'évolution de l'impact de ces deux types d'observations avec le temps, on voit que celui des dropsondes croît le plus. Des expériences d'assimilation ont confirmé que le DFS donne une indication de comment les observations contribuent à l'analyse mais pas comment elles peuvent modifier la prévision subséquente.

Cette étude a eu lieu avant le lancement de IASI à bord de MetOp et où les données AIRS n'étaient pas encore assimilées dans le système de prévision numérique. Nous continuons à

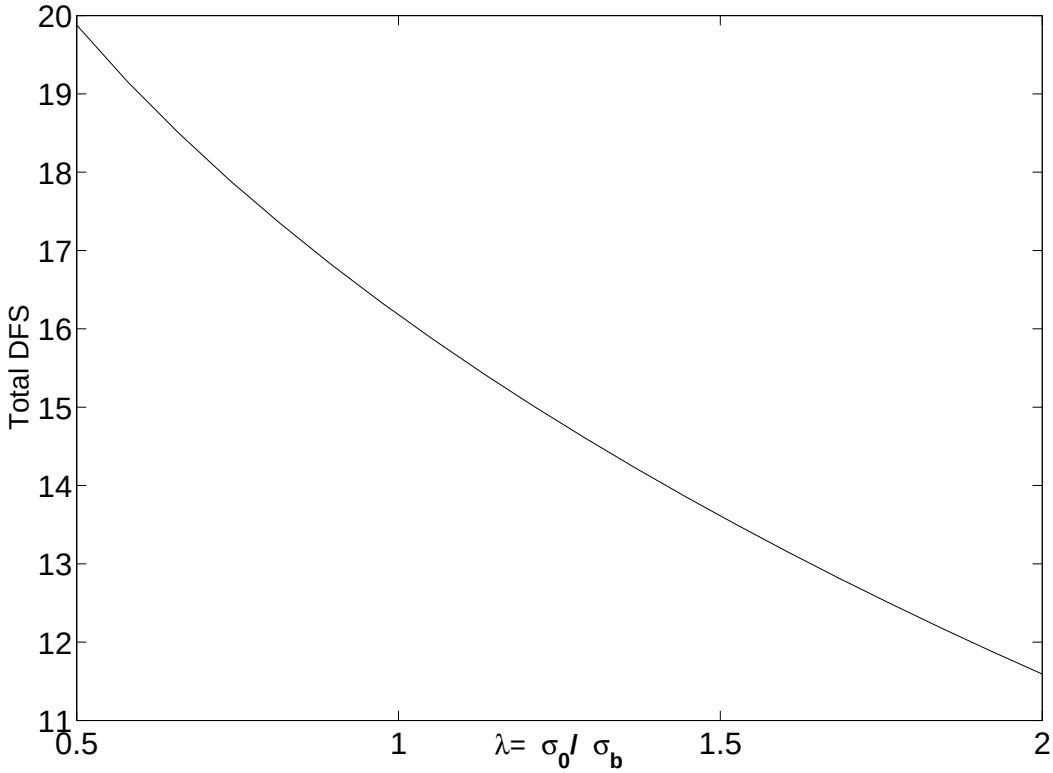


FIG. 2.17 – Influence du rapport σ_O/σ_B sur le DFS total

étudier les impacts des différents types d'observations avec ces nouvelles données maintenant disponibles.

2.7.2 Etudes des transitions extratropicales de cyclones avec les observations IASI

Cette étude (Anwender *et al.*, 2009) s'est déroulée dans le cadre d'une visite de 6 mois au CNRM-GAME de Doris Anwender de l'Institut pour la Recherche sur la Météorologie et sur le Climat de Karlsruhe en Allemagne. Si un cyclone connaît une transition extra-tropicale au large des Etats Unis, il peut influencer le temps sur l'Europe en étant repris par la circulation des moyennes latitudes. En se propageant, le système peut transporter une quantité importante d'humidité vers l'Europe. Les transitions extra-tropicales excitent ou modifient le train d'ondes de Rossby dans les latitudes tempérées. Ainsi une telle modification peut entraîner le creusement intense d'un thalweg et entraîner un temps sensible à fort impact. Ces traversées rapides de l'Océan Atlantique Nord et ces forts creusements potentiels des dépressions causent de sérieux problèmes aux prévisionnistes (Jones et coauthors, 2003). Pour améliorer la simulation des transitions extra-tropicales des cyclones, une bonne couverture en observations est importante en particulier sur les océans où se développent les cyclones. Le but de cette étude est de tester l'impact des données IASI sur les simulations de transitions extra-tropicales de cyclones atlantiques du mois de septembre 2008. Nous nous sommes donc intéressés à l'apport des canaux vapeur d'eau de l'instrument IASI sur trois cas de transitions extra-tropicales. Trois expériences d'assimilation ont été menées sur les trois premières semaines de septembre 2008 : une première sans observation IASI, une deuxième servant de référence n'assimilant que les canaux sensibles à la température et une troisième assimilant en plus 9 canaux sensibles à la

vapeur d'eau de la moyenne et haute troposphère.

Dans un premier temps, nous avons étudié l'apport de ces données sur le globe. Les scores globaux sur la période montrent un fort impact de IASI, particulièrement sur l'hémisphère sud où les résultats sont significatifs à plus longue échéance que pour l'hémisphère nord. L'impact des canaux vapeur d'eau est faible mais positif surtout pour l'humidité du modèle. Pour les domaines Atlantique Nord et Europe, les erreurs ont été calculées en utilisant l'énergie totale qui combine les variables importantes durant la transition extra-tropicale. Cette erreur est réduite pour plusieurs prévisions initialisées autour des périodes de transition extra-tropicale. Avec l'assimilation des canaux vapeur d'eau, on observe un nombre important de réduction d'erreurs de prévision. Nous avons également étudié l'erreur dans l'ébauche et dans l'analyse à partir des diagnostics calculés selon une méthode dérivée de Desroziers *et al*(2005). Cette erreur est réduite avec l'assimilation des canaux vapeur d'eau (Figure 2.20). Il en découle une plus faible erreur dans les analyses à partir desquelles les simulations sont initialisées. La réduction des erreurs d'analyses est particulièrement visible sur les structures dynamiques associées aux événements de transitions extra-tropicales (Figure 2.20).

Nous avons également montré dans l'article de Faccani *et al* (2009) dans le cadre de l'expérience AMMA (Analyse Multidisciplinaire de la Mousson Africaine) qu'il existait des rétroactions de l'assimilation d'observations dans les Tropiques sur la prévision aux moyennes latitudes. Ainsi la figure 2.21 montre l'évolution de la différence d'erreur de prévision issues d'une expérience assimilant les radiosondages de la campagne AMMA en août-septembre 2006 et une expérience assimilant les données du réseau en 2005 (avec le débiaisage opérationnel à l'époque). On voit que si l'impact des radiosondages AMMA avec un débiaisage approprié est encore présent sur l'Afrique de l'ouest à 24 heures d'échéance, cet impact se propage sur l'Océan Atlantique et vers l'Europe du Nord à 48 et 72 heures d'échéance. La propagation vers l'ouest au dessus de l'Afrique semble être liée aux ondes d'est. L'expérience utilisant les radiosondages supplémentaires étend le jet d'est vers le sud ce qui renforce les ondes d'Est africaines comme montré dans Leroux et Hall (2009). La propagation vers le Nord peut probablement être relié aux ondes de Rossby et est cohérente avec d'autres études dans lesquelles on a montré que la mousson africaine pouvait influencer l'Europe (Bielli *et al.*, 2009) et (Cassou *et al.*, 2005). Il apparaît donc que les radiosondages de la campagne AMMA ont un impact quelques jours plus tard sur la prévision sur l'Europe. Des scores de prévision calculés par rapport aux radiosondages de l'Europe montrent l'impact positif sur la prévision et sont statistiquement significatifs.

Cette étude illustre le lien entre les Tropiques et les moyennes latitudes. Même si je suis surtout intéressée par les dépressions des moyennes latitudes, il est ainsi utile d'utiliser un modèle global pour les étudier.

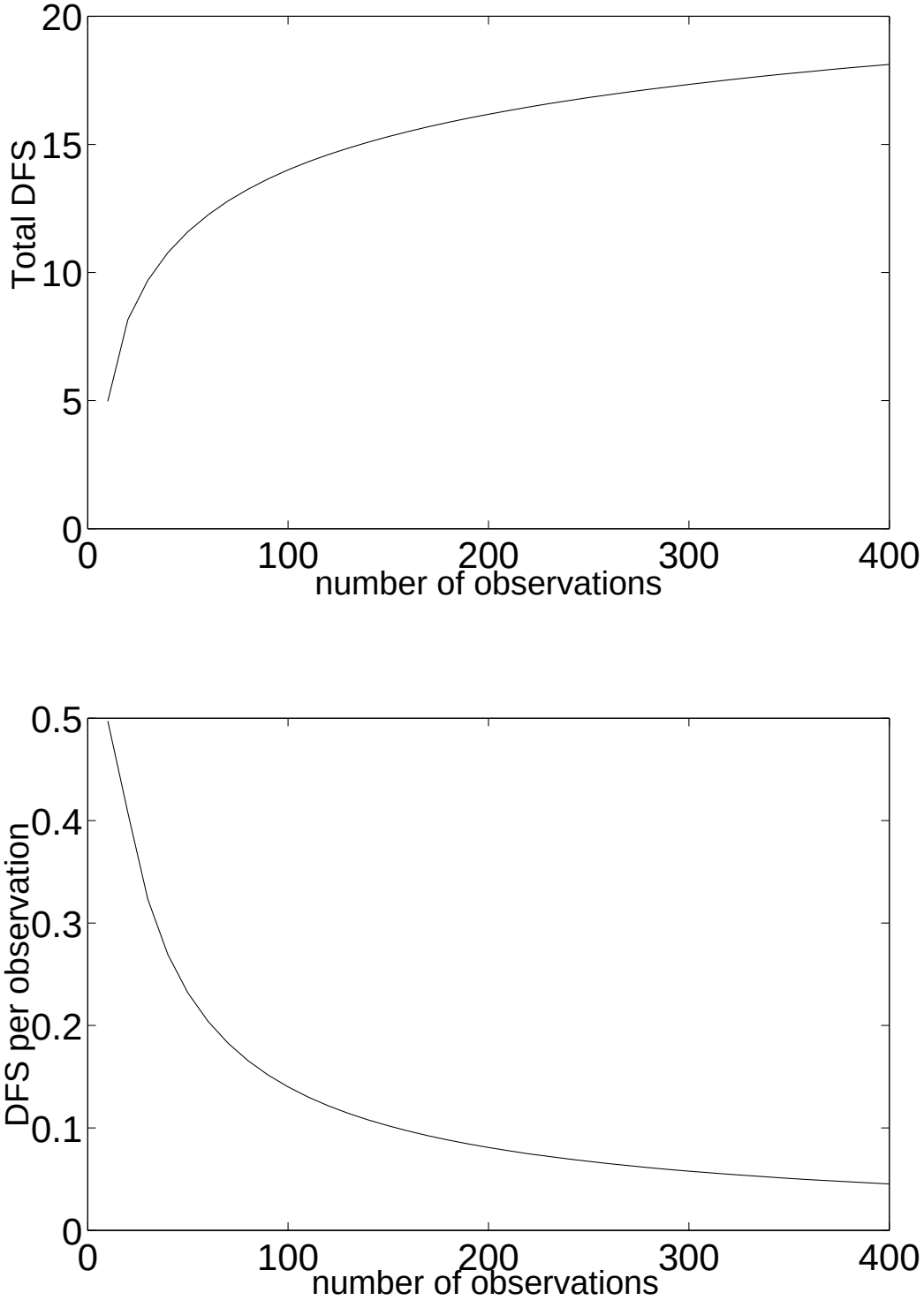


FIG. 2.18 – Influence du nombre d'observations sur le DFS total et sur le DFS unitaire.

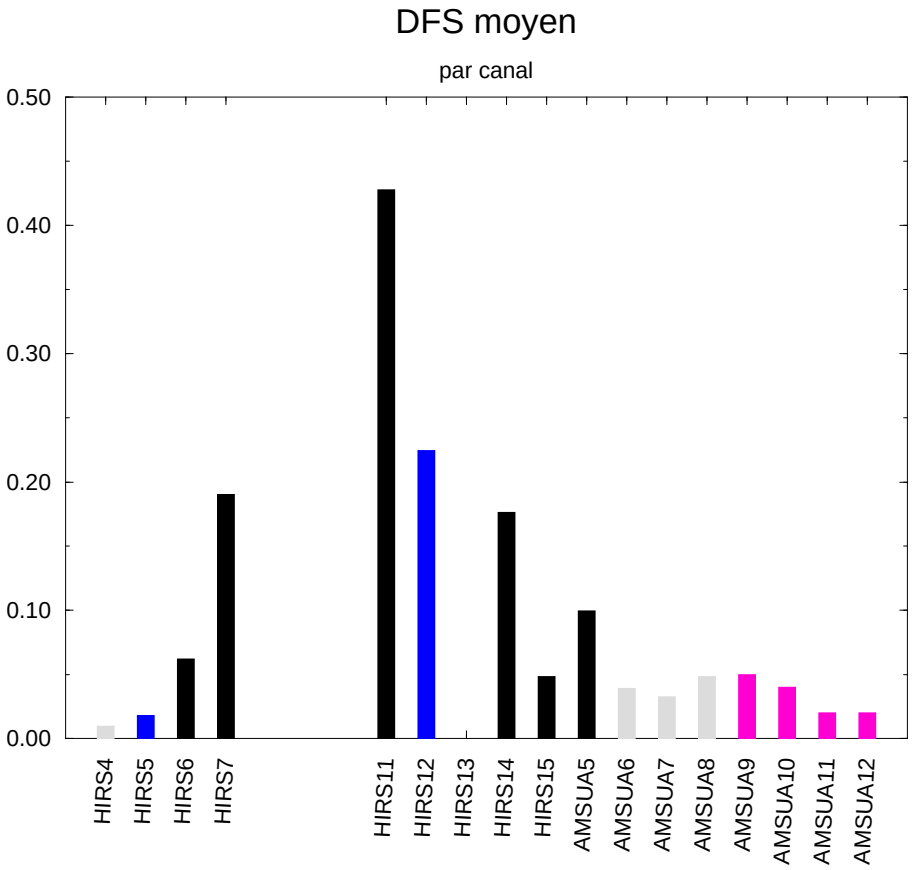
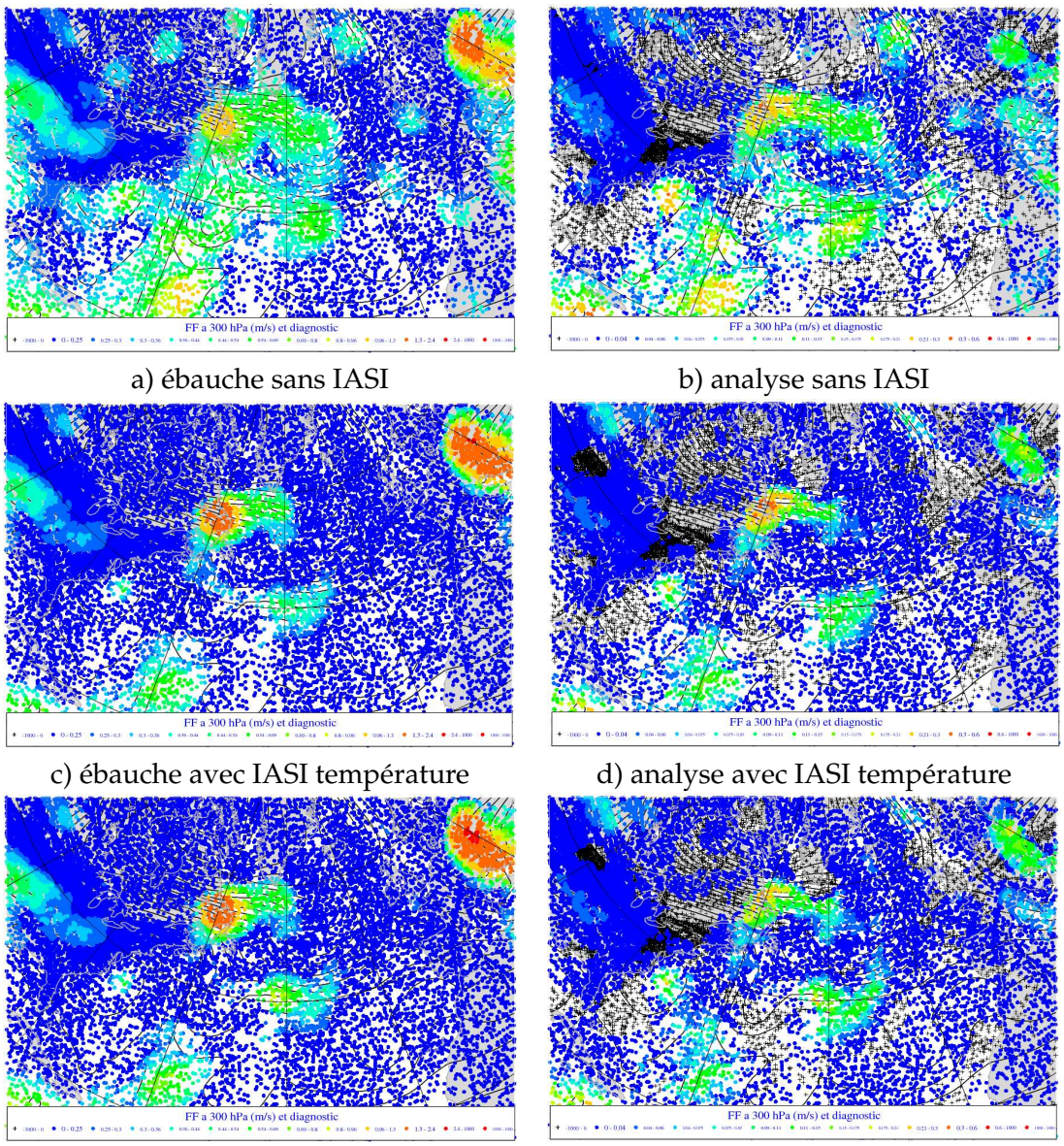


FIG. 2.19 – DFS unitaire des canaux des sondeurs TOVS assimilés



a) ébauche sans IASI

b) analyse sans IASI

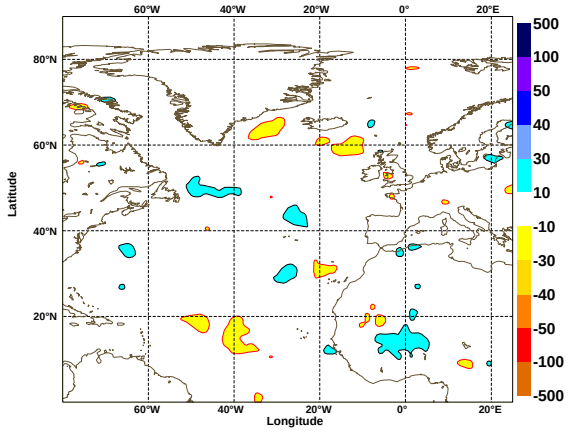
c) ébauche avec IASI température

d) analyse avec IASI température

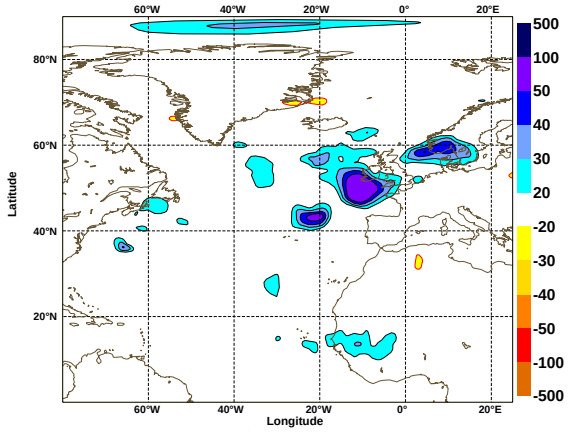
e) ébauche avec IASI T+WV

f) analyse avec IASI T+WV

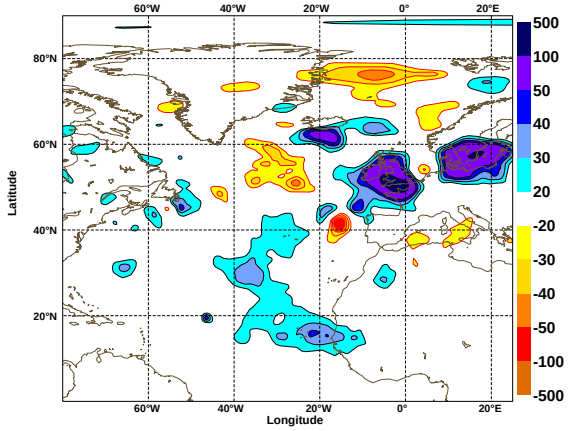
FIG. 2.20 – Diagnostics de variances d'erreurs d'ébauche (colonne de gauche) et d'erreurs d'analyse (colonne de droite) calculés dans l'espace des observations pour la couche 200-400 hPa pour les trois expériences : sans IASI (en haut), avec IASI canaux de température (au milieu) et avec IASI canaux de températude et vapeur d'eau (en bas) pour le 15 septembre 2008 à 00 UTC. Les couleurs bleues correspondent à de petites erreurs et les couleurs dans les rouges à de grosses valeurs en comparaison d'une erreur moyenne calculée sur 3 mois. Les + correspondent à des observations très proches des valeurs du modèle. Les lignes noires correspondent à l'analyse du géopotentiel à 200 hPa du CEPMMT



a) 24h



b) 48h



c) 72h

FIG. 2.21 – Différences d'erreur quadratique moyenne de prévision (24h (a), 48h (b) et 72h (c)) du géopotentiel à 500 hPa (en m) entre l'expérience assimilant les radiosondes de la zone AMMA corrigées du biais en humidité et l'expérience simulant un jeu d'observation avant la campagne AMMA calculée sur la période du 1^{er} août au 14 septembre 2006 par rapport aux analyses du Centre Européen.

2.8 Conclusions et perspectives

Une bonne partie de mes travaux de recherches a été consacrée à l'assimilation des observations des sondeurs infarouges hyperspectraux qui ont été lancés dans les années 2000. Ainsi j'ai participé à la préparation de leur assimilation en montrant qu'un nombre restreint de canaux pouvait suffire pour en tirer bénéfice, y compris dans les zones sensibles où se développent les tempêtes. Ensuite j'ai proposé une méthode pour assimiler les radiances en présence de nuages en inversant deux paramètres nuageux (pression de sommet des nuages et fraction nuageuse) à partir des observations. De plus avec la disponibilité du modèle de méso-échelle AROME, j'ai contribué à étudier l'apport de l'aggrégation de l'information du modèle dans le pixel de l'observation pour la simulation des radiances. Le but de l'analyse étant de fournir un état de l'atmosphère le plus précis possible, l'assimilation de données permet une inversion de l'information fournie par les différentes sources d'observations, notamment lors de campagnes de mesures. En dehors des campagnes de mesures, on peut tester l'impact des observations comme l'apport d'information sur la vapeur d'eau par le sondeur IASI. Mes travaux de recherches contribuent à l'effort constant d'amélioration des modèles de prévision numérique et j'ai participé à la mise en oeuvre dans la chaîne opérationnelle de l'assimilation des radiances claires et nuageuses AIRS et des canaux vapeur d'eau de IASI.

Les voies d'amélioration de l'assimilation des radiances sont nombreuses. Elles concernent les conditions d'observations sur les continents, en présence de nuages et l'augmentation de la densité du nombre de points d'observation assimilés.

D'une part avec le développement de modèles régionaux, l'**assimilation des radiances sensibles à la surface terrestre** est nécessaire car la part des continents devient prépondérante dans le domaine du modèle. La problématique repose sur la nécessité de bien connaître les paramètres de surface que sont la température de surface et l'émissivité de surface qui affectent le calcul du transfert radiatif.

Avec Abdelaziz Babqiqi ingénieur de la Météorologie Marocaine (Babqiqi *et al.*, 2006), nous avons commencé à étudier ce problème lors d'un stage de 3 mois en inversant l'émissivité de surface à partir de canaux fenêtre AIRS choisis à partir de la méthode de Karbou *et al.* (2006) développée pour les observations micro-ondes et en étudiant les simulations réalisés à partir de ce paramètre inversé. Ces travaux ont été réalisés alors que l'assimilation des radiances AIRS n'était pas encore opérationnelle et la détection nuageuse assez sommaire. En effet, une bonne détection nuageuse est nécessaire pour l'inversion des paramètres de surface afin d'être sûr que le signal que l'on observe est bien celui de la surface et non celui d'un nuage. Le principe de la méthode de Karbou *et al.* (2006) consiste à inverser l'équation du transfert radiatif pour quelques canaux fenêtres des sondeurs hyperspectraux et à utiliser cette valeur de l'émissivité pour la simulation des autres radiances. En effet l'émissivité de surface varie également en fonction de la longueur d'onde et il est nécessaire de représenter cette variation. Ainsi la figure 2.22 montre l'émissivité de surface inversée à partir de deux canaux très sensibles à la surface terrestre moyennée sur un mois (mi-août à mi-septembre) de 2005. On voit nettement que les déserts présentent une grande variabilité de l'émissivité de surface. Une bonne description de l'émissivité et de la température de surface devrait faciliter l'assimilation des radiances sur terre en améliorant les simulations à partir de l'ébauche. Ceci apparaît particulièrement important car ces déserts sont pauvres en observations conventionnelles et les systèmes affectant le pourtour méditerranéen peuvent se développer sur la région nord du Sahara comme nous l'avons vu dans la partie 2.6.1 avec l'étude du Medicane.

Anaïs Vincensini commence une thèse que je vais encadrer sur le sujet dans le contexte de la

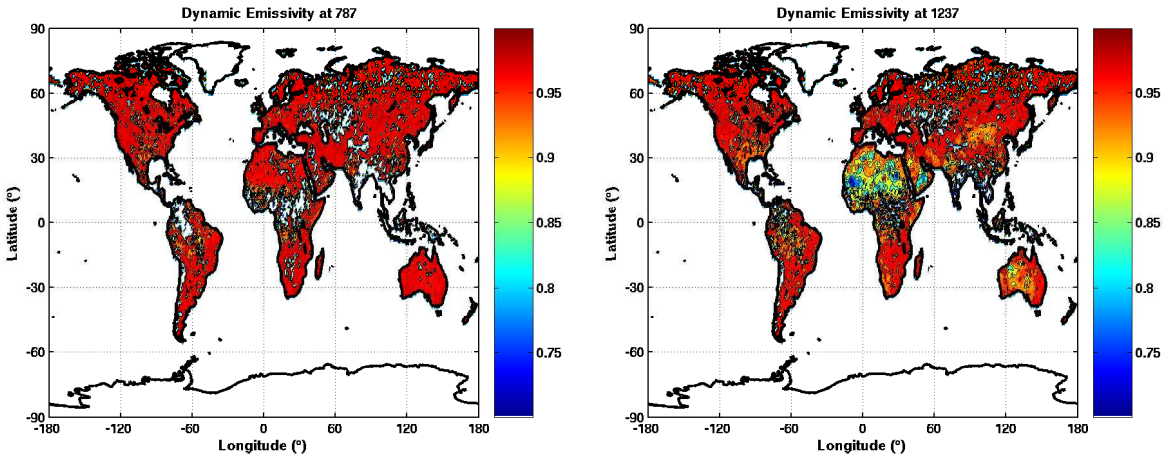


FIG. 2.22 – Émissivité de surface inversée moyenne à partir du canal 787 (a) et du canal 1237 (b) du sondeur AIRS calculée du 15/08/2005 au 14/09/2005.

préparation de la campagne HyMeX. Nous travaillerons avec le modèle développé pour cette campagne AROME WMED (West Mediterranean). Dans un premier temps, nous utiliserons des atlas d'émissivité infrarouge réalisées à partir de mesures satellitaires et de laboratoires du CIMSS (Cooperative Institute For Meteorological Satellite Studies). Nous les comparerons à des méthodes d'inversions de l'émissivité de surface et nous étudierons la température de surface inversée à partir des observations satellitaires suivant la méthode de Fatima Karbou. Stéphanie Guedj, dans le cadre de sa thèse, a choisi une approche similaire pour l'assimilation des radiances infrarouges SEVIRI à bord de MSG sur terre dans les modèles ALADIN et AROME. Nous pourrons donc comparer nos résultats, notamment les restitutions de température de surface et d'émissivité.

L'approche choisie pour l'**assimilation des radiances nuageuses infrarouges** est à ce jour assez simple (hypothèse d'un nuage mono-couche) et il est tout à fait envisageable d'essayer d'utiliser les profils d'hydrométéores disponibles dans les modèles météorologiques pour réaliser les simulations des observations nuageuses. Une approche similaire a été retenue au CEPMMT pour l'assimilation des observations des imageurs micro-ondes affectées par la pluie et les nuages (Geer *et al.*, 2010 et Bauer *et al.*, 2010). Ainsi, les radiances claires, nuageuses et pluvieuses sont assimilées avec une seule approche "tous ciels" qui utilise les paramétrisations de la physique "humide" et un modèle de transfert radiatif avec de la diffusion multiple pour l'opérateur d'observation. L'utilisation d'un module du modèle de transfert radiatif RTTOV prenant en compte les variables nuageuses du modèle d'analyse pour la simulation des radiances infrarouges nuageuses (Chevallier *et al.*, 2002) existe au CEPMMT et ce code devra être mis en place pour assimiler correctement les radiances nuageuses. Mohamed Dahoui (Dahoui, 2006) avait étudié l'assimilation des radiances nuageuses à partir de simulations effectuées avec des profils d'hydrométéores d'un schéma diagnostique de nuages du modèle ARPEGE avec un tel code de transfert radiatif. Depuis, ARPEGE a été doté d'un schéma pronostique de nuages ce qui améliore la représentation des hydro-météores, mais il faudrait disposer d'un schéma de nuages dans les modèles tangent-linéaire et adjoint. Plus récemment, des travaux exploratoires sont en cours dans les modèles à aire limitée. Ainsi un opérateur d'observation similaire à ce qu'avait proposé Mohamed pour les radiances nuageuses avec un schéma de physique simpli-

fiée a été développé dans le cadre du modèle régional HIRLAM (Stengel *et al.*, 2010) pour les observations SEVIRI. L'assimilation des radiances SEVIRI nuageuses dans le modèle HIRLAM est en cours de test (Stengel et Linskog, 2010).

L'assimilation des observations AIRS et IASI dans le modèle AROME depuis 2010 offre la possibilité de reprendre les travaux de Mohamed. Ceci sera le sujet de thèse de Pauline Martinet que je vais encadrer lors de sa formation complémentaire par la recherche. Une autre approche pourra consister en l'initialisation ou la correction de paramètres nuageux du modèle atmosphérique à partir des restitutions obtenues avec les observations satellitaires. Depuis mai 2010, nous disposons de la couverture nuageuse déduite de l'imageur AVHRR dans le pixel IASI. Cette information pourrait être utilisée pour discriminer les cas auxquels appliquer la méthode du CO2-slicing (c'est-à-dire pour des pixels complètement recouverts de nuages). Ceci permettrait d'étendre le nombre de cas nuageux dans l'assimilation des radiances nuageuses. À plus long terme, suivant l'évolution de la qualité des variables micro-physiques du modèle atmosphérique, on s'orientera soit vers une extension de la méthode du CO2-slicing avec la restitution de paramètres nuageux pour des nuages à une ou plusieurs couches à partir de la méthode variationnelle proposée par Auligné (2007), soit vers une simulation des radiances nuageuses à partir des variables micro-physiques atmosphériques dans le modèle de transfert radiatif.

Mes travaux sont essentiellement centrés sur le deuxième terme de l'équation de la fonction coût (Eq. 2.5) qui concerne les observations. Le premier terme représente l'information a priori, l'ébauche qui est sujet elle aussi à des erreurs que l'on représente par la matrice de covariances d'erreurs d'ébauche. Cette matrice a un impact important sur l'assimilation car elle répartit le poids entre l'ébauche et les observations, elle propage l'information des points d'observation vers les autres points du modèle en la lissant et elle répartit l'information entre les différentes variables du modèle. Une connaissance précise des statistiques des erreurs d'ébauche est donc requise pour l'assimilation de données. Différentes méthodes peuvent être utilisées pour les estimer comme les statistiques d'erreur entre des prévisions à différentes échéances (Berre *et al.*, 2006). Des études récentes (Montmerle et Berre, 2010) ont montré que les covariances d'erreurs d'ébauches variaient également entre les zone de ciel clair et les zones précipitantes. Il faudra suivre également ces développements autour de la matrice $\mathbf{B}$ pour améliorer l'assimilation des radiances nuageuses.

Actuellement le **nombre de canaux assimilés** dans le système est très faible (moins de 80) mais il devrait croître rapidement. Des travaux ont débuté pour prendre en compte les corrélations des erreurs d'observations (Borman et Bauer, 2010 ; Bormann *et al*, 2010). Les corrélations inter-canaux existent dans la bande de la vapeur d'eau et au niveau de la surface. Le traitement des radiances nuageuses dans l'opérateur d'observation introduit également une corrélation dans les erreurs. Depuis avril 2010, les versions opérationnelles d'ARPEGE, ALADIN et AROME bénéficient d'une densité d'assimilation des radiances satellitaires tous les 125 km. Des travaux au CEPMMT ont suggéré que cette densité pouvait encore être augmentée. Si l'on veut tirer profit d'une beaucoup plus haute densité horizontale des observations, notamment dans les modèles de méso-échelle, il faudra prendre également en compte les corrélations spatiales des erreurs d'observations.

Les campagnes de ciblage pour l'amélioration de la prévision des tempêtes à l'aide de moyens aéroportés sont onéreuses et difficiles à mettre en place en raison de la sécurité du traffic aérien, notamment au dessus de l'Océan Atlantique. Depuis quelques années, des études ont été menées pour accroître la densité des observations satellitaires dans les zones sensibles. Ainsi Dando *et al* (2007) ont étudié l'impact des données ATOVS ciblées dans les zones sen-

sibles sur quelques études de cas sur l'Europe du Nord et la nouvelle Zélande à partir d'une résolution d'assimilation dégradée des ATOVS afin de s'affranchir de potentielles corrélations d'erreurs d'observations. Ils ont trouvé un apport bénéfique de telles observations. Bauer *et al* (2010) ont également étudié l'augmentation de la densité des observations dans les zones sensibles. Ils ont trouvé un impact positif d'une telle augmentation par rapport à celle dans des zones climatologiquement actives. Ceci ouvre donc des perspectives pour renouveler le ciblage des observations pour la prévision des tempêtes.

Toutes ces recherches sur une meilleure exploitation des observations satellitaires devraient fortement contribuer à la mise en place du modèle AROME W-MED développé pour la campagne expérimentale d'**HyMeX**. Cette campagne a pour objectif de mieux comprendre et de quantifier le cycle hydrologique et les processus associés de la Méditerranée avec une attention particulière sur les événements météorologiques à forts impacts sociétaux comme les précipitations intenses, la variabilité intra-annuelle à décennale du système couplé méditerranéen et les tendances associés dans le contexte du changement climatique. Je coordonne avec Thomas Auligné (NCAR) la tâche de modélisation de l'assimilation de données atmosphériques du plan de mise en oeuvre d'HyMeX. Cette tâche associe l'assimilation des données de campagne de mesures, de radars et d'observations satellitaires ainsi que la modélisation de covariances d'erreurs d'ébauche. Ainsi, des expériences telles que le NA-TReC 2003 ont montré que la prévision d'un nombre significatif de phénomènes météorologiques ou hydrologiques à fort impact sur le pourtour septentrional du bassin Méditerranéen est sensible à ce qui survient en amont sur mer et sur la frange méridionnale du bassin méditerranéen. Cependant ces parties du bassin sont bien moins richement dotées que la partie nord en observations in-situ. Il est donc nécessaire de considérer une plus forte densité d'observations de ces zones, au moins pour les périodes d'observation intensive. Les données satellites infrarouges et micro-ondes sur terre et sur mer peuvent contribuer à cet effort et faire l'objet d'une exploitation plus intense dans les systèmes de prévision numérique. De plus, comme nos résultats sur la campagne AMMA l'ont montré, un meilleur traitement des données d'humidité qu'elles soient d'origine conventionnelle ou satellitaire permet d'améliorer l'analyse de ce paramètre dans le modèle atmosphérique.

Tous ces développements que j'ai menés bénéficieront à l'utilisation des sondeurs hyperspectraux IASI de la série MetOp dont les suivants devraient être lancés en 2012 et 2016. Par ailleurs, ces études permettront également de mieux utiliser et assimiler les observations du futur sondeur hyperspectral infrarouge IRS (InfraRed Sounding) qui sera à bord de Meteosat Troisième Génération. Il possédera une bonne définition spectrale ainsi qu'une forte répétitivité temporelle, que ce soit sur terre, sur mer en condition claire ou nuageuse, ce qui en fera un moyen d'observation privilégié pour améliorer la simulation des phénomènes à développement rapide.

Références bibliographiques

Anwender, D., Fourrié, N., Rabier, F., et Arbogast, P. (2009). Data impact experiments using IASI observations during the Ets of Hurricanes Gustav, Hannah et Ike (2008). Internal report, CNRM.

Auligné, T., McNally, T., et Dee, D. (2007). Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Met. Soc.*, 133 :631–642.

Auligné, T. (2007). *Assimilation variationnelle des observations de sondeurs infrarouges hyperspectraux : correction de biais et détection nuageuse*. PhD thesis, Thèse de l'université Toulouse III, soutenue le 8 juin.

Aumann, H., Chahine, M., Gautier, C., Goldberg, M., Kalnay, E., McMillin, L., Revercomb, H., Rosenkranz, P., Smith, W., Staelin, D., Strow, L., et Susskind, J. (2003). Airs/amsu/hsb on the aqua mission : design, science objectives, data products and processing systems. *IEEE Trans. Geosci. Remote Sens.*, 41(2) :253–264.

Babqiqi, A., Fourrié, N., Rabier, F., et Karbou, F. (2006). Détermination de l'émissivité de surface sur terre des données AIRS sur terre. Rapport de stage., CNRM.

Baker, N. L. et Daley, R. (2000). Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Q. J. R. Met. Soc.*, 126 :1431–1454.

Bauer, P., Buizza, R., Cardinali, C., et Thépaut, J.-N. (2010a). Impact of singular vector based satellite data thinning on NWP. *Q. J. R. Met. Soc.*, page submitted.

Bauer, P., Geer, A., Lopez, P., et Salmond, D. (2010b). Direct 4D-Var assimilation of all-sky radiances. Part I : Implementation. Tech Memo 618, ECMWF.

Bergot, T. et Doerenbecher, A. (2002). A study on the optimization of the deployment of targeted observations using adjoint-based methods. *Q. J. R. Met. Soc.*, 128 :1689–1712.

Bergot, T., Hello, G., Joly, A., et Malardel, S. (1999). Adaptive Observations : a feasability study. *Mon. Wea. Rev.*, 127 :743–765.

Berre, L., Stefanescu, S. E., et Pereira, M. B. (2006). The representation of the analysis effect in three error simulation techniques. *Tellus*, 58A :196–209.

Bielli, S., Douville, H., et Pohl, B. (2009). Understanding the West African monsoon variability and its remote effects : an illustration of the grid point nudging methodology. *Climate Dyn.*, pages DOI :10.1007/s00382–009–0667–8.

Bormann, N. et Bauer, P. (2010). Estimates of spatial and interchannel observation-error characteristics for current sounder radiances for numerical weather prediction. I : Methods and application to ATOVS data . *Q. J. R. Met. Soc.*, page DOI :10.1002/qj.616.

Bormann, N., Collard, A., et Bauer, P. (2010). Estimates of spatial and interchannel observation-error chartacristics for current sounder radiances for numerical weather prediction. II : Application to AIRS and IASI data. . *Q. J. R. Met. Soc.*, page DOI :10.1002/qj.615.

Cardinali, C. (2009). Monitoring the observation impact on the short-range forecast. *Q. J. R. Met. Soc.*, 135 :239–250.

Cardinali, C. et Buizza, R. (2004). Observations sensitivity to the analysis and the forecast : a case study during atrec targeting campaign. In *First THORPEX International Science Symposium, Montreal, Canada*, page WMO TD 1237 WWRP/THORPEX No 6.

Cassou, C., Terray, L., et Phillips, A. (2005). Tropical atlantic influence on european heat waves. *J. Climate*, 18 :2805–2811.

Cayla, F. (2001). L'interféromètre iasi ; un nouveau sondeur satellitaire à haute résolution. *La météorologie 8ème série*, 32 :23–39.

Chahine, M. (1974). Remote sounding of cloudy atmospheres. I : The single cloud layer. *J.Atmos. Sci.*, 31 :233–243.

Chapnik, B., Desroziers, G., Rabier, F., et Talagrand, O. (2006). Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. *Q. J. R. Met. Soc.*, 132.

Chevallier, F. (1999). Tigr-like sampled databases of atmospheric profiles from the ecmwf 50-level forecast model. Nwp saf research report 1, NWP SAF.

Chevallier, F., Chédin, A., Chéruy, F., et Morcrette, J.-J. (2000). Tigr-like atmospheric profile databases for accurate radiative flux computation. *Q. J. R. Meteor. Soc.*, 126 :777–785.

Chevallier, F., Lopez, P., Tompkins, A. M., Janiskova, M., et Moreau, E. (2004). The capability of 4D-Var systems to assimilate cloud-affected satellite infrared radiances. *Q. J. R. Met. Soc.*, 130 :917–932.

Collard, A. (2007). Selection of IASI channels for use in numerical weather prediction. *Q. J. R. Met. Soc.*, 133 :1977–1991.

Collard, A. et McNally, A. P. (2009). The assimilation of Infrared Atmospheric Sounding interferometer radiances at ECMWF. *Q. J. R. Met. Soc.*, 135 :1044–1058.

Collard, A. D. (1998). Notes on iasi performance. NWP Technical Report 253, Met Office, Bracknell,UK.

Courtier, P., Thépaut, J.-N., et Hollingsworth, A. (1994). A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Met. Soc.*, 120 :1367–1387.

Dahoui, M., Lavanant, L., Rabier, F., et Auligné, T. (2005). Use of MODIS imager to help dealing with AIRS cloudy radiances. *Q. J. R. Met. Soc.*, 131 :2559–2579.

Dahoui, M. L. (2006). *Vers une assimilation variationnelle des radiances satellitaires nuageuses*. Thèse de doctorat, Université Toulouse 3-Paul Sabatier, France, soutenue le 19 juin 2006.

Dando, M., Thorpe, A., et Eyre, J. (2007). The impact of targeted satellite observations on numerical weather prediction. *Q. J. R. Met. Soc.*, 133 :1945–1957.

Dee, D. (2004). Variational bias correction of radiance data in the ECMWF system. In *Proc. of ECMWF Workshop on Assimilation of high spectral resolution sounders in NWP, July2004*, pages 97–112.

Desroziers, G., Berre, L., Chapnik, B., et Poli, P. (2005). Diagnosis of observation, background and analysis error statistics in observation space. *Q. J. R. Met. Soc.*, 131.

Doerenbecher, A. et Bergot, T. (2001). Sensitivity to observations applied to FASTEX cases. *Nonlinear Processes in Geophysics*, 8(6) :467–481.

Duffourg, F., Ducrocq, V., Fourrié, N., Jaubert, G., et Guidard, V. (2010). Simulation of satellite infrared radiances for convective-scale data assimilation over the mediterranean. *J. Geophys. Res.*, 115 :D15107, DOI :10.1029/2009JD012936.

Faccani, C., Rabier, F., Fourrié, N., Agusti-Panareda, A., Karbou, F., Moll, P., Lafore, J.-P., Nuret, M., Hdidou, F., et Bock, O. (2009). The Impacts of AMMA Radiosonde Data on the French Global Assimilation and Forecast System. *Weather and Forecasting*, 24(5) :1268–1286.

Fourrié, N. et Rabier, F. (2004). Cloud charasteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Q. J. R. Met. Soc.*, 130 :1839–1856.

Fourrié, N. et Thépaut, J.-N. (2003). Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Q. J. R. Met. Soc.*, 129 :2425–2439.

Fourrié, N., Marchal, D., Rabier, F., Chapnik, B., et Desroziers, G. (2006). Impact study of the 2003 north atlantic thorpex regional campaign. *Q. J. R. Met. Soc.*, 132 :275–295.

Geer, A. J., Bauer, P., et Lopez, P. (2010). Direct 4D-Var assimilation of all-sky radiances. Part II : Assessment. Tech. memo 619, ECMWF.

Harris, B. et Kelly, G. (2001). A satellite radiance bias correction scheme for radiance assimilation. *Q. J. R. Met. Soc.*, 127 :1453–1468.

Heilliette, S. et Garand, L. (2007). A practical approach for the assimilation of cloudy infrared radiances and its evaluation using AIRS simulated observations. *Atmosphere-ocean*, 45(4) :211–225.

Hello, G., Lalaurette, F., et Thépaut, J.-N. (2000). Combined use of sensitivity information and observations to improve meteorological forecast : the " christmas storm" case. *Q. J. R. Met. Soc.*, 126 :621–647.

Joiner, J. et Poli, P. (2005). Note on the effect of horizontal gradients for nadir-viewing microwave eand infrared sounders. *Q. J. R. Met. Soc.*, 131 :1783–1792.

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., Gall, R., Hewson, T. D., Hildebrand, P. H., Jorgensen, D. Lalaurette, F., Langland, R. H., Lemaitre, Y., Mascart, P., Moore, J. A., Persson, P. O., Roux, F., Shapiro, M. A., Snyder, C., Toth, Z., et Wakimoto, R. M. (1999). Overview of the field phase of the Fronts and Atlantic Storm-Track EXperiment (FASTEX) project. *Q. J. R. Met. Soc.*, 125 :3131–3163.

Jones, S. et coauthors (2003). The extratropical transition of tropical cyclones : Forecast challenges, current understanding, and futures directions. *Weather and Forecasting*, 18 :16–56.

Kaplan, L., Chahine, M., Susskind, J., et Searl, J. (1977). Spectral band passes for a high precision satellite. *Applied Optics*, 16 :322–325.

Karbou, F., Gérard, E., et Rabier, F. (2006). Microwave Land Emissivity and Skin Temperature For AMSU-A and -B Assimilation Over Land. *Q. J. R. Met. Soc.*, 132 :2333–2355.

Klinker, E., Rabier, F., et Gelaro, R. (1998). Estimation of the key-analysis errors using the adjoint technique. *Q. J. R. Met. Soc.*, 124 :1909–1933.

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

Leroux, S. et Hall, N. (2009). On the relationship between african easterly waves and the african easterly jet. *J. Atmos. Sci.*, 66 :2303–2316.

Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. *Q. J. R. Met. Soc.*, 112 :1177–1194.

McNally, A. (2002). A note on the occurence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Met. Soc.*, 128 :2551–2556.

McNally, A. et Watts, P. (2003). A cloud detection algorithm for high spectral resolution infrared sounders. *Q. J. R. Met. Soc.*, 129 :3411–3423.

McNally, A. P. (2009). The direct assimilation of cloud-affected infrared satellites radiances for numerical weather prediction. *Q. J. R. Met. Soc.*, 135 :1214–1229.

Menzel, W., Smith, T., et Stewart, T. (1983). Improved cloud motion wind vector and altitude assignment using VAS. *J. appl. Meteorol*, 22 :377–384.

Montmerle, T. et Berre, L. (2010). Diagnosis and formulation of heterogeneous background error covariances at mesoscale. *Q. J. R. Met. Soc.*, 136 :in press.

Moscatello, A., Marcello Miglietta, M., et Rotunno, R. (2008). Numerical analysis of a Mediterranean "hurricane" over south-eastern Italy. *Mon. Wea. Rev.*

Palmer, T. N., Gelaro, R., Barkmeijer, J., et Buizza, R. (1998). Singular vectors, metrics and adaptive observations. *J. Atmos. Sci.*, 55 :633–653.

Pangaud, T., Fourrié, N., Guidard, V., Dahoui, M., et Fabier, F. (2009). Assimilation of AIRS radiances affected by mid to low level clouds. *Mon. Wea. Rev.*, pages 4276–4292 (17) DOI : 10.1175/2009MWR3020.1.

Pavelin, E., English, S., et Eyre, J. (2008). The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Q. J. R. Met. Soc.*, 134 :737–749.

Petersen, G. et Thorpe, A. (2007). The impact on weather forecast of targeted observations during A-TReC. *Q. J. R. Met. Soc.*, 133 :417–431.

Prunet, P., Thépaut, J.-N., et Cassé, V. (1998). The information content of clear-sky IASI radiances and their potential for numerical weather prediction. *Q. J. R. Met. Soc.*, 124 :211–241.

Rabier, F., Fourrié, N., Chafaï, D., et Prunet, P. (2002). Channel selection methods for Infrared Atmospheric Sounding Interferometer. *Q. J. R. Met. Soc.*, 128 :1011–1027.

Rabier, F., Gauthier, P., Cardinali, C., Langland, R., Tsyrulnikov, M., Lorenc, A., Steinle, P., Gelaro, R., et Koizumi, K. (2008). An update on THORPEX-related research in data assimilation and observing strategies. *Nonlinear Processes in Geophysics*, 15 :81–94.

Rabier, F., Klinker, E., Courtier, P., et Hollingsworth, A. (1996). Sensitivity of forecast errors to initial conditions. *Q. J. R. Met. Soc.*, 122 :121–150.

Rodgers, C. (1996). Information content and optimisation of high spectral resolution measurements. *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II*, 2830 :136–147.

Rodgers, C. (2000). *Inverse methods for atmospheres : Theories and practice*. World Scientific Publications, Singapore.

Saunders, R., Brunel, P., English, S., Bauer, P., O'Keeffe, U., Francis, P., et Rayer, P. (2006). RTTOV-8 – Science and Validation report. Technical report, SAF NWP.

Stengel, M., Lindskog, M., Undén, P., Gustafsson, N., et Bennartz, R. (2010). An extended observation operator in HIRLAM 4D-Var for the assimilation of cloud-affected satellite radiances. *Q. J. R. Met. Soc.*, 136 :DOI :10.1002/qj.621.

Stengel, M. et Linskog, M. (2010). Assimilation of Cloud-affected Infrared Radiances in HIRLAM 4D-Var. In *17th International ATOVS Study Conference, Monterey, California, United States*.

Susskind, J., Barnet, C. D., et Blaisdell, J. (2003). Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Remote Sensing*, 41(2) :390–409.

Publications essentielles

Fourrié N. and Thépaut J.-N. (2003) : Evaluation of the AIRS Near real Time channel selection for application to numerical weather prediction. *Quart. J. R. Met. Soc.*, 129, 2425-2439.

Fourrié N. and Rabier F. (2004) : Clouds Characteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Quart. J. R. Met. Soc.*, 130, 1839-1856.

Pangaud T., Fourrié N., Guidard V., Dahoui M. and Rabier F. (2009) : Assimilation of AIRS radiances affected by mid to low-level clouds. *Monthly Weather Review*, 4276-4292 (17) DOI : 10.1175/2009MWR3020.1.

Duffourg F., Ducrocq V., Fourrié N., Jaubert G. and Guidard V. (2010) : Simulation of satellite infrared radiances for convective-scale data assimilation over the Mediterranean. *Journal of Geophysical Research*.115, D15107, DOI :10.1029/2009JD012936

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

Q. J. R. Meteorol. Soc. (2003), **129**, pp. 2425–2439 doi: 10.1256/qj.02.210

Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction

By NADIA FOURRIÉ[1]* and JEAN-NOËL THÉPAUT[2]

[1]*CNRM/GMAP, Météo-France, Toulouse, France*
[2]*European Centre for Medium-Range Weather Forecasts, Reading, UK*

(Received 11 November 2002; revised 26 March 2003)

SUMMARY

The Atmospheric Infrared Sounder (AIRS) on board the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Aqua satellite provides 2378 channels for each field of view of the instrument. As it is neither feasible nor efficient to assimilate all the channels in a numerical weather-prediction system, a policy of channel selection has to be designed in this context. This paper attempts to assess the optimality of the selection of the AIRS radiance channels that are made available to the scientific community in near real time (hereafter called AIRS NRT) by the National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite Data and Information Service. This assessment is done by comparing this channel selection with a method preserving the information content of the instrument, the so-called 'global' method. It turns out that although the selected channels are different and the information content as measured by the entropy reduction (ER) and the degrees of freedom for signal (DFS) is slightly smaller for the AIRS NRT channel set than for the 'global' set, both channel selections give similar results in terms of analysis error for temperature, humidity and ozone. The robustness of the results is then evaluated by varying the range of input parameters to the channel-selection scheme, in particular the atmospheric training dataset on which the channel selection is based, and the background-error covariance matrix. It is found that the performance of the 'global' channel selection is sensitive to the training dataset, while the AIRS NRT channel selection remains robust, even, to some extent, for the retrieval of key analysis-error structures. Altogether, the 'manually selected' AIRS NRT channels provide a good compromise between robustness and quality.

KEYWORDS: Reduced AIRS near-real-time dataset Validation

1. INTRODUCTION

By measuring radiation in many thousands of different channels, advanced infrared sounders such as the Atmospheric Infrared Sounder (AIRS, 2378 channels) and the Infrared Atmospheric Sounding Interferometer (IASI, 8461 channels) have the potential to provide atmospheric temperature and composition information at a much higher vertical resolution and accuracy (Prunet *et al.* 1998) than can be achieved with the current generation of operational sounding instruments (e.g. High Infrared Radiation Sounder—HIRS, 20 channels). Successful exploitation of this new generation of satellite instruments is one of the major challenges for numerical weather prediction (NWP) centres for the next ten years.

It is neither feasible nor efficient to assimilate all of the channels, most centres are thus following effective channel-selection strategies for radiance assimilation in NWP. The challenge is to find a set of channels that is small enough to be assimilated efficiently in a global NWP system (within operational time constraints), and large enough to capture important atmospheric variability. Several information-content studies for advanced sounders aim to identify the 'best' channels for NWP in order to minimize the reduction of information from advanced infrared sounders. In particular, Rabier *et al.* (2002) have tested several methods in the context of the IASI instrument. They found that the channel-selection method following Rodgers (1996) selects the number of IASI channels (in clear-sky conditions) in an optimal way, which preserves the information content of the instrument.

* Corresponding author: CNRM/GMAP, Météo-France, 42 av. G. Coriolis, 31057 Toulouse Cedex, France. e-mail: Nadia.Fourrie@cnrm.meteo.fr

Following the launch of the NASA* Aqua satellite in May 2002, a reduced set of AIRS radiance channels have been selected by the AIRS science team (Susskind *et al.* 2003) and are available to the scientific community in near real time (NRT) by NOAA†/NESDIS‡. The main goal of this study is to apply the Rabier *et al.* (2002) methodology to the AIRS instrument in order to assess the quality of the AIRS NRT channel selection versus channel selections which optimize measurement information content (OMIC).

In section 2, the experimental framework of the study and the OMIC channel selection are briefly described. The efficiency of the AIRS NRT channel selection is then compared with the OMIC selection in terms of information content and linear one-dimensional variational assimilation (1D-Var) performance (section 3). The robustness of results to different inputs to the OMIC channel selection is then evaluated. Section 4 addresses the problem of representativeness of the training dataset used for the OMIC channel selection. The usefulness of the AIRS NRT channel set for the retrieval of key analysis-error structures is studied in section 5. Conclusions are discussed in section 6.

2. Experimental framework

(*a*) *Linear optimal estimation theory*

The general framework of this channel-selection study is linear optimal estimation theory in the context of NWP. We follow the framework presented at length by Rabier *et al.* (2002) from which we summarize the main elements. The atmospheric profiles of temperature, humidity and ozone, and surface temperature at a given location are represented by a vector $\mathbf{x}$ and the satellite observations by a vector $\mathbf{y}$. The observations are linked to the atmospheric state by an observation operator including the radiative-transfer model:

$$\mathbf{y} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}_{\mathbf{O}} + \boldsymbol{\varepsilon}_{\mathbf{F}} \tag{1}$$

where the measurement and the forward model errors $\boldsymbol{\varepsilon}_{\mathbf{O}}$ and $\boldsymbol{\varepsilon}_{\mathbf{F}}$ are each assumed to be gaussian noises with error covariance matrices $\mathbf{O}$ and $\mathbf{F}$. We denote $\mathbf{R} = \mathbf{O} + \mathbf{F}$ the resulting observation-error covariance matrix. The background state vector $\mathbf{x}^{\mathrm{b}}$ has an associated error covariance matrix denoted $\mathbf{B}$. The radiative-transfer equation is assumed to be weakly nonlinear in the vicinity of the background state, making the tangent linear assumption valid: $\mathcal{H}(\mathbf{x}) = \mathcal{H}(\mathbf{x}^{\mathrm{b}}) + \mathbf{H}(\mathbf{x} - \mathbf{x}^{\mathrm{b}})$ where $\mathbf{H}$ is the tangent linear model of the radiative-transfer model $\mathcal{H}$.

Rabier *et al.* (2002) use two additional concepts introduced by Rodgers (2000): the entropy reduction ($\mathrm{ER} = -(1/2)\log_2 \mathrm{de}(\mathbf{AB}^{-1})$) and the degrees of freedom for signal ($\mathrm{DFS} = \mathrm{Tr}(\mathbf{I} - \mathbf{AB}^{-1})$, where $\mathbf{A}$ represents the analysis-error covariance matrix). Let λ represent eigenvalues of $\mathbf{AB}^{-1}$ (a small λ corresponding to a direction where the measurements have significantly reduced the analysis-error variance). It can be shown that:

$$\mathrm{DFS} = n - \sum_{\lambda \in \sigma(\mathbf{AB}^{-1})} \lambda \tag{2}$$

$$\mathrm{ER} = -\frac{1}{2} \sum_{\lambda \in \sigma(\mathbf{AB}^{-1})} \log_2 \lambda \tag{3}$$

where n represents the total number of degrees of freedom of the analysis problem.

* National Aeronautics and Space Administration.

† National Oceanic and Atmospheric Administration.

‡ National Environmental Satellite Data and Information Service.

Both concepts are very useful in that they quantify the gain in information brought by the observations with respect to the background information (the larger the ER or the DFS the better).

The linear estimate of the analysis-error covariance matrix is used for the 1D-Var retrieval. The optimal analysed state $\mathbf{x}^a$ is given by $\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{y}^b)$ with $\mathbf{K} = \mathbf{A}\mathbf{H}^\top\mathbf{R}^{-1}$, $\mathbf{y}^b = \mathcal{H}(\mathbf{x}^b)$ and $\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^\top\mathbf{R}^{-1}\mathbf{H})^{-1}$. $\mathbf{K}$ is the Kalman gain matrix and $\mathbf{A}$ is the analysis-error covariance matrix introduced above. The parameter space is the temperature, humidity and ozone profile defined on the 43 pressure levels of the radiative-transfer model for AIRS (RTAIRS, Matricardi *et al.* 2001) and the surface temperature.

In this study, RTAIRS (now incorporated in RTTOV-7, which is the radiative-transfer model for the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder 7) described by Matricardi *et al.* (2001) was used for $\mathcal{H}$ and $\mathbf{H}$. This model uses a fixed vertical discretization with 43 pressure levels (see Matricardi and Saunders 1999). The $\mathbf{B}$ matrix was interpolated from the current (2002) operational 60-level European Centre for Medium-Range Weather Forecasts (ECMWF) background-error covariance matrix representing short-range forecast errors of the ECMWF model. The computation of the original ECMWF background-error statistics relies on an ensemble of data assimilation experiments in which the members differ because of random noise added to the observations, in agreement with the assumed observation errors (Fisher, personal communication). The interpolation of the covariance from 60 to 43 levels is linear and preserves the diagonal of $\mathbf{B}$. The surface-temperature background-error standard deviation was included (1.04 K corresponding to an inflation of 0.5^2 to the lowest-model-level temperature background-error variance) in the $\mathbf{B}$ matrix and a correlation with the lowest tropospheric level was added (it is assumed that the vertical correlation between the surface and the first model level is identical to the correlation between that level and the one above). The covariance matrix $\mathbf{O}$ was derived from the latest estimation of the AIRS instrument noise (Hannon, personal communication). An additional constant error of 0.2 K coming from the radiative-transfer-model noise (matrix $\mathbf{F}$) is added to this measurement error in order to obtain the $\mathbf{R}$ matrix.

Besides the quality criteria described above, the efficiency of each channel selection was evaluated in terms of linear 1D-Var.

(*b*) *Channel selections*

(i) *AIRS NRT channel selection.* The choice of these channels is described in Susskind *et al.* (2003). The 324 channels are selected in order to retrieve surface temperature and surface emissivity, and temperature, water-vapour and ozone profiles. One objective of the AIRS NRT selection is to have a good sampling of the window regions for determining the spectral variations in the surface emissivity, which is not retrieved in our study. In addition, another objective for this channel selection is that the selection is optimized to be applicable for all conditions, especially for sounding in cloudy conditions (Joiner, personal communication) where sensitive structures exist. The AIRS NRT set generally picks channels on wings of absorption lines with sharper weighting functions, as shown by Kaplan *et al.* (1977), allowing more channels peaking above clouds.

(ii) *Optimal selections.* Optimal information-content channel selections depend on atmospheric state through the dependence on the background-error covariance matrix and on the tangent linear of the radiative-transfer model. In an operational NWP

data assimilation context, it is unfeasible to perform channel selection atmosphere by atmosphere. Some representative channel sets are thus required. In this study, we design a channel selection based on an OMIC global set. This global set is derived from the iterative method for channel selection, proposed by Rodgers (1996) and used in Rabier *et al.* (2002). This iterative method entails performing successive analyses, each one using only one extra channel at a time. The resulting analysis-error covariance matrix is updated accordingly and used at the next iterative step. This ensures that all the information coming from previous channels is taken into account for the selection of the new channel. The channel selection in our case is based on maximizing the ER (and is therefore 'optimal' in that sense). It has been verified that the channel selection does not change if one considers the DFS as the choice criterion for the channel selection. Therefore, as in Rabier *et al.* (2002), a 'global' channel selection was computed as an average of the different optimal channel selections over a representative atmospheric-profile dataset.

This set is part of the ECMWF atmospheric database (Chevallier 1999; Chevallier *et al.* 2000) and forms a set of 108 profiles of temperature, humidity and ozone, and surface temperature and surface pressure covering most of the atmospheric variability. All atmospheric scenes are assumed to be cloud free, over sea and for nadir views. The 108 profiles are divided into 75 midlatitude (20°N–70°N, 20°S–70°S), 14 tropical (20°N–20°S) and 19 polar (70°N–90°N, 70°S–90°S) profiles. In this study, the background fields and the AIRS data were simulated from this atmospheric dataset.

As a starting point, we have assumed that 324 channels are available in the AIRS NRT selection. Since some channels of the original 2378 are known to be of poorer quality, they are excluded from the optimal selection. Applying this check, the total number of AIRS channels is reduced from 2378 to 2140.

The global channel selection was built as follows. For each of the 108 profiles, each channel picked by the channel selection is given a rank from 1 to 324 that corresponds to its rank in the iterative selection process. The 324 channels with the lower rank were chosen in order to build the 'global' channel set. This channel set is used here as a benchmark for the validation of the AIRS NRT selection. The 'global' channel selection is an average selection based on the set of 108 individual optimal channel selections. In fact, only 63 channels are always selected through the 108 channel selections (corresponding to very different atmospheric situations), whereas 1504 channels are never selected.

In addition, the 'OMIC' channel selections tailored to the three different air-mass classes were computed as the average of the representative profiles of each air-mass type (midlatitude, polar and tropical). Figure 1 displays the spectral location of the different 'OMIC' channel selections (global, and the ones corresponding to the three different air-mass classes). Eight spectral regions were designed and correspond to those described in Table 1. The different spectral bands are labelled by letters to facilitate the discussion.

The 'global' channel selection (averaged over the whole dataset) shares 319 channels with the 'midlatitude' selection, 288 channels with the 'tropical' selection and 254 channels with the 'polar' selection. The 'polar' and 'tropical' selections share only 226 channels. In particular, the 'polar' selection does not pick any channel in the short-wave side of band A (12.5 μm) in contrast to the 'tropical' selection, but it favours the pure window (10.5 μm, band B). In fact, when the selection is based on tropical profiles the 12.5 μm region acts as a water-vapour band and channels are selected in this area due to large specified humidity background errors, whereas the humidity signal is weak in this spectral region when dealing with very dry polar profiles which favour a selection in a cleaner window region and in band E (6 μm water-vapour band). As expected, the

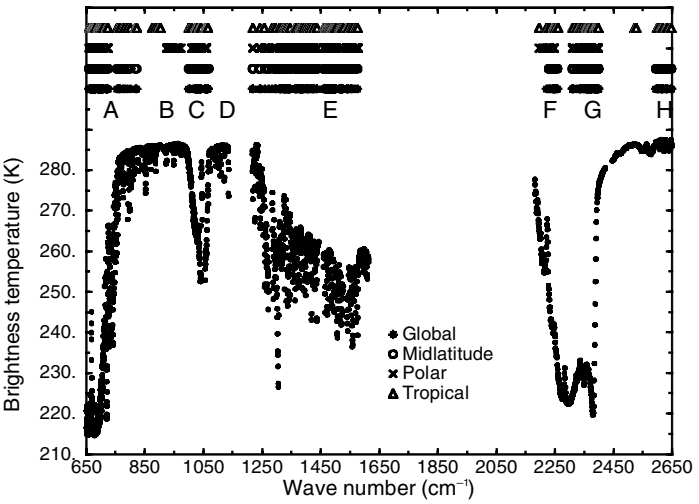


Figure 1. Typical spectrum of the Atmospheric Infrared Sounder (AIRS) and location of the channels selected to optimize measurement information content by the 'OMIC' method averaged over all profiles ('global') and over different air-mass profile classes ('midlatitude', 'polar' and 'tropical'). See Table 1 for the letters corresponding to spectral bands.

TABLE 1. DESCRIPTION OF THE SPECTRAL REGIONS AND DISTRIBUTION OF CHANNELS OF AIRS NRT AND GLOBAL SELECTIONS

Band	Description	Wave number (cm^{-1})	Number of AIRS NRT channels	Number of global channels
A	Temperature sounding	650–800	129	76
B	Surface window	800–1000	12	3
C	Ozone	1000–1080	25	38
D	Surface window	1080–1150	7	0
E	Water vapour	1080–1620	62	129
F	Temperature sounding	2180–2300	34	18
G	Temperature sounding	2300–2390	16	39
H	Surface window	2390–2670	39	26

AIRS NRT refers to the selection of Atmospheric Infrared Sounder radiance channels available in near real time.

'global' selection is dominated by a midlatitude signal due to a rather poor representation of polar and tropical air masses in the sampling dataset.

3. COMPARISON BETWEEN AIRS NRT AND 'GLOBAL' CHANNEL SELECTIONS

The resulting 'global' channel selection shares only 119 channels with the AIRS NRT selection. Figure 2 displays the AIRS spectrum corresponding to a midlatitude profile together with the location (at the top of the figure) of the AIRS NRT (stars) and 'global' selected channels (circles). At first sight, the AIRS NRT channel selection spans the infrared spectrum more evenly than the 'global' channel selection, which emphasizes specific small spectral domains. This choice for the AIRS NRT selection is obviously aiming to represent as uniformly as possible the whole AIRS spectrum. The fact that only 119 out of 324 channels are selected by both approaches could seem worrying. However, as can be seen from Fig. 2, large spectral areas are covered by both sets (long-wave carbon dioxide, ozone, water vapour, . . .) and obviously 'global' and AIRS NRT have selected neighbouring wave numbers.

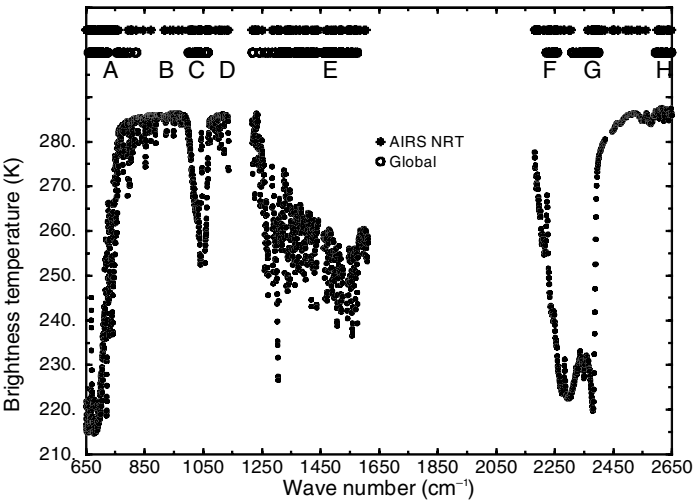


Figure 2. Typical spectrum of the Atmospheric Infrared Sounder (AIRS) superimposed at the top of the figure by the location of the channels chosen by the AIRS near real time (NRT) selection (stars) and picked by the 'global' channel selection (circles). See Table 1 for the description of the letters corresponding to different spectral bands.

A closer look at both selections indicates that 51% of AIRS NRT channels are located in bands A, B and C against 36% of the 'global' channels, while 40% of this latter set are chosen in band E. Band D is hardly covered by the 'global' channel selection, whereas both AIRS NRT and 'global' selected channels evenly represent band G. Band C is also well captured by the two selections. The study of the Jacobians of the associated selected channels in both selections indicates an over-representation of the upper stratosphere (bands A and G) in the 'global' selection. This over-representation of the upper stratosphere can be explained by the large temperature background-error variances at these levels. The lack of band-G channels in the AIRS NRT set was due to the expectation of non-local thermodynamic equilibrium in these channels (Susskind *et al.* 2003). These effects are seen in the real AIRS data and these channels are unusable in sunlight. Our 'global' selection is thus valid only for night-time clear conditions. There is also a dominating choice for band E (water vapour) in the 'global' channel selection, most likely driven by the vertical structure of the humidity background-error covariance matrix. It should also be noted that this band also contains temperature information. This is due to nonlinearity in the radiative-transfer coupling temperature and humidity in the water-vapour band ν_2. The induced uncertainty in Jacobians due to water vapour may not be well taken into account by the linear approach. Conversely, AIRS NRT selection has a large number of channels peaking in the low troposphere corresponding to bands A and H. The larger number of AIRS NRT channels selected in band H were selected on the basis of their information content during the day when reflected sunlight is considered and this gives a potentially large information content for lower-tropospheric humidity.

Figure 3 and Table 2 show that ER and DFS are slightly smaller for the AIRS NRT selection than for the 'global' selection. The number of independent pieces of information brought by the 324 channels (as defined in the section above) is also smaller (25.2 against 28.8) in the case of AIRS NRT selection. However, both selections greatly improve DFS, ER and the number of eigenvalues compared with the use of the 15 channels corresponding to the current HIRS instrument (see Table 2).

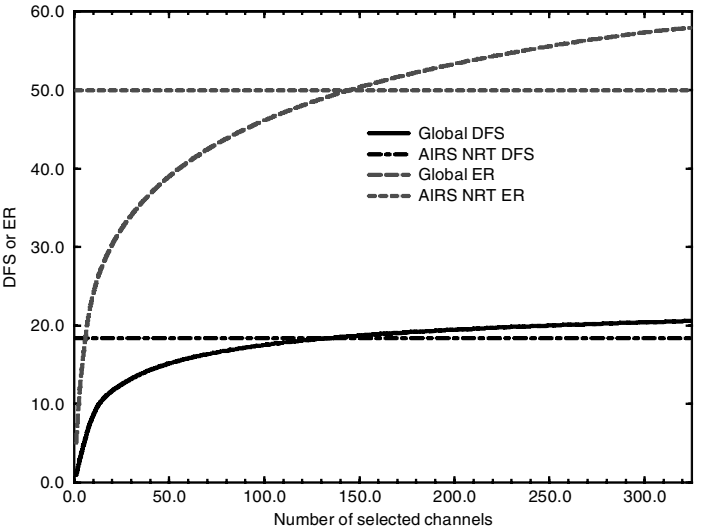


Figure 3. Evolution of the degrees of freedom for signal (DFS) and the entropy reduction (ER) for the 'global' channel selection with respect to the number of selected channels, averaged over the 108 profiles. The corresponding values of DFS and ER for the Atmospheric Infrared Sounder (AIRS) near real time (NRT) channel set is also appended.

TABLE 2. VALUES OF DEGREES OF FREEDOM FOR SIGNAL (DFS) AND ENTROPY REDUCTION (ER) FOR 'ATMOSPHERIC INFRARED SOUNDER (AIRS) NEAR REAL TIME (NRT)', 'GLOBAL', 'ALL CHANNELS' AND 'HIGH INFRARED RADIATION SOUNDER (HIRS) LIKE' SELECTIONS

Experiment	DFS	ER
AIRS NRT	18.36	49.93
Global	20.59	57.94
All channels	24.34	76.16
HIRS like	6.44	14.06

'All channels' corresponds to the experiment when 2140 channels are used. 'HIRS like' corresponds to the experiment when the 15 AIRS channels closest to the current channels of HIRS are used.

We examined the standard deviations of background error and of linear 1D-Var analysis error for temperature, humidity and ozone profiles as well as for surface temperature, averaged over the 108 profiles (Fig. 4). Both AIRS NRT and 'global' selections outperform the 'HIRS like' set in terms of analysis-error reduction for temperature, humidity and ozone. AIRS NRT and 'global' selections provide similar temperature retrieval errors in the troposphere and in the low and midstratosphere. The 'global' selection produces analysis errors smaller by 0.2–0.3 K compared with the AIRS NRT set in the upper stratosphere. This is certainly due to the massive selection of stratospheric channels by the 'global' selection. In addition, this reduction of analysis temperature error certainly explains the differences observed for the ER or the DFS between both selections. However, one might question whether these differences are

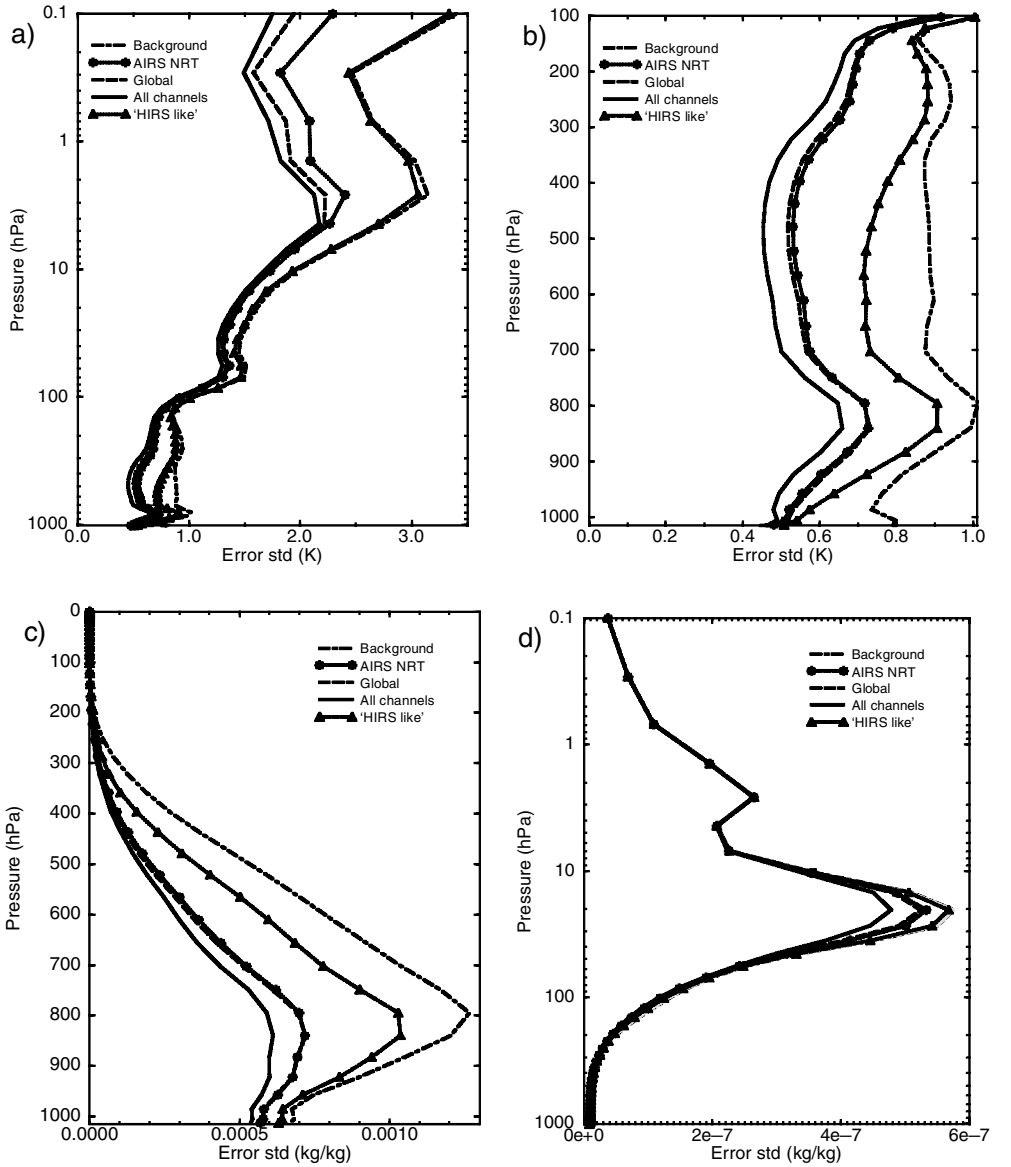


Figure 4. One-dimensional variational assimilation (1D-Var) averaged (over 108 profiles) error standard deviation, for temperature ((a) in log scale and (b) in linear scale), humidity (c) and ozone (d). Experiment names as explained in Table 2.

relevant in the context of short- or medium-term forecasts. Otherwise, both selections provide similar humidity and ozone analysis errors.

For surface temperature, the analysis-error standard deviation is decreased to 0.05 K for both selections to be compared with 1.04 K for the background and 0.3 K for the 'HIRS like' channel set. This result should be taken with great care since the very small analysis error is explained by the somewhat 'empirical' value specified for the surface-temperature background error. This quantity is poorly known in NWP models as is the

background-error vertical correlation between surface temperature and lowest-model-level temperature. In addition, the errors due to uncertainties on surface emissivity or cloud contamination are not taken into account in this study.

Comparisons with 'all channel' retrieval errors (Fig. 4) provide some guidance about the loss of information entailed in the two channel selections tested. The use of all channels leads to a further reduction of temperature analysis error of about 0.1 K in the troposphere and roughly of 0.2 K in the stratosphere. The improvement in the humidity analysis is very important in the lower troposphere when all the channels are used in the assimilation. Likewise, the ozone retrieval error is substantially decreased when 2140 channels are used. The further gains in ER and in DFS are 18.2 and 3.75, respectively, when all the channels are used in the analysis (Table 2).

We conclude that, even though a loss of information content is to be expected by using around one tenth of the total number of AIRS channels, the AIRS NRT selection seems reasonable for NWP applications. Even though the information content is slightly smaller than for the 'global' selection, the respective qualities of AIRS NRT and 'global' selections are very comparable in terms of linear 1D-Var retrieval error for all retrieved parameters. This shows that despite a different choice of channels (only 119 out of 324 are in common), the spectral bands (and therefore the atmospheric layers) are sufficiently similarly covered by the two selections to provide equivalent performance.

In the following sections, the robustness of the results to various inputs to the channel selection is assessed.

4. Impact of the air mass on the quality of the AIRS NRT selection

In this section, we investigate the robustness of the different channel selections to the air mass under consideration. Two questions are considered: (a) Is the AIRS NRT selection robust to the atmospheric air mass (e.g. polar, midlatitude, tropical)? (b) What happens to the OMIC selection if it is based on a given air mass and applied to a very different one?

(*a*) *Variation of the channel selection with respect to the air-mass type*

As already mentioned in section 2, the three channel selections depending on air mass were tailored according to the average of the representative profiles of each air-mass type. The outcome of the various 1D-VAR experiments using the different channel-selection combinations ('midlatitude', 'polar', 'tropical' and 'global', not shown) can be summarized as follows:

- For temperature, AIRS NRT and the different 'OMIC' selections provide similar results for all air-mass types.
- For humidity, AIRS NRT gives similar performance to 'midlatitude' and 'tropical'. 'Polar' channel selection provides slightly better results than AIRS NRT below 400 hPa for the polar air-mass class, in agreement with having more channels selected in band C (157 channels for 'polar' versus 61 for AIRS NRT).
- For ozone, the selections give similar results with the exception of the tropical air mass, where the 'tropical' selection provides slightly smaller analysis errors than AIRS NRT.

In terms of DFS and ER (Table 3), the AIRS NRT channel set provides slightly poorer results compared with the OMIC channel selection. An air-mass dedicated channel selection provides better DFS or ER for each air mass than the global channel selection. The gain when using all the channels is 3.0 to 3.5 units in term of DFS and

TABLE 3. VALUES OF DEGREES OF FREEDOM FOR SIGNAL (DFS) AND ENTROPY REDUCTION (ER) FOR THE DIFFERENT CHANNEL SELECTIONS WITH RESPECT TO THE AIR-MASS TYPE

Channel set	Air-mass class	DFS	ER
'Midlatitude'	Midlatitude	20.54	57.69
'Polar'	Polar	21.07	62.10
'Tropical'	Tropical	20.59	55.55
AIRS NRT	Midlatitude	18.34	49.88
AIRS NRT	Polar	18.10	50.77
AIRS NRT	Tropical	18.80	49.08
Global	Midlatitude	20.55	57.68
Global	Polar	20.73	60.87
Global	Tropical	20.67	55.39
2140 channels	Midlatitude	24.06	74.9
2140 channels	Polar	24.22	77.77
2140 channels	Tropical	24.21	73.24

AIRS NRT refers to the selection of Atmospheric Infrared Sounder radiance channels available in near real time.

more than 15 units for ER. It is worthwhile noting that the best result in terms of DFS or ER is obtained for the polar air mass. This may be due to an over-estimation of the climatological background-error covariance matrix for this air mass in comparison with the actual error. Altogether, the AIRS NRT selection seems to be fairly insensitive to the air-mass category, thus confirming the quality of careful 'manual' choice of these channels performed by the AIRS science team.

(*b*) *Importance of the training dataset for the quality of the channel selection*

It has been shown above that the channel selection can vary depending on the air-mass type the selection is trained with. To document further this issue, we illustrate here the impact of performing a 1D-Var analysis for a certain atmospheric situation using a channel selection based on a completely different atmospheric training dataset.

Figure 5 displays the 1D-Var performance of the AIRS NRT, the 'tropical', 'polar', 'global' and 'all channel' selections averaged over the 14 tropical profiles available in our dataset. One can clearly see the degradation in temperature and humidity retrievals below 700 hPa of the 'polar' selection, due to a poor sampling of the tropospheric water-vapour channels. In addition, this polar selection leads to a slight degradation of the stratospheric ozone retrieval. One also notes a loss of around 0.5 units in terms of DFS and 2.39 units in terms of ER. Moreover, the 'global' selection performs similarly to the 'polar' one for low-tropospheric humidity. This indicates the importance of an exhaustive training dataset if one plans to replace the day-1 AIRS NRT selection by a more optimal method in the near future. It is of particular importance to capture all the weather regimes for highly variable quantities.

5. RETRIEVAL OF KEY ANALYSIS-ERROR STRUCTURES

The current background-error covariance matrix used at ECMWF only provides a climatology of short-range forecast errors. On the other hand, previous information-content studies, such as those of Prunet *et al.* (1998) and Collard (1998), have suggested that advanced sounders such as AIRS and IASI could resolve in clear air some of the small-scale baroclinic structures that have been identified by sensitivity studies (Klinker

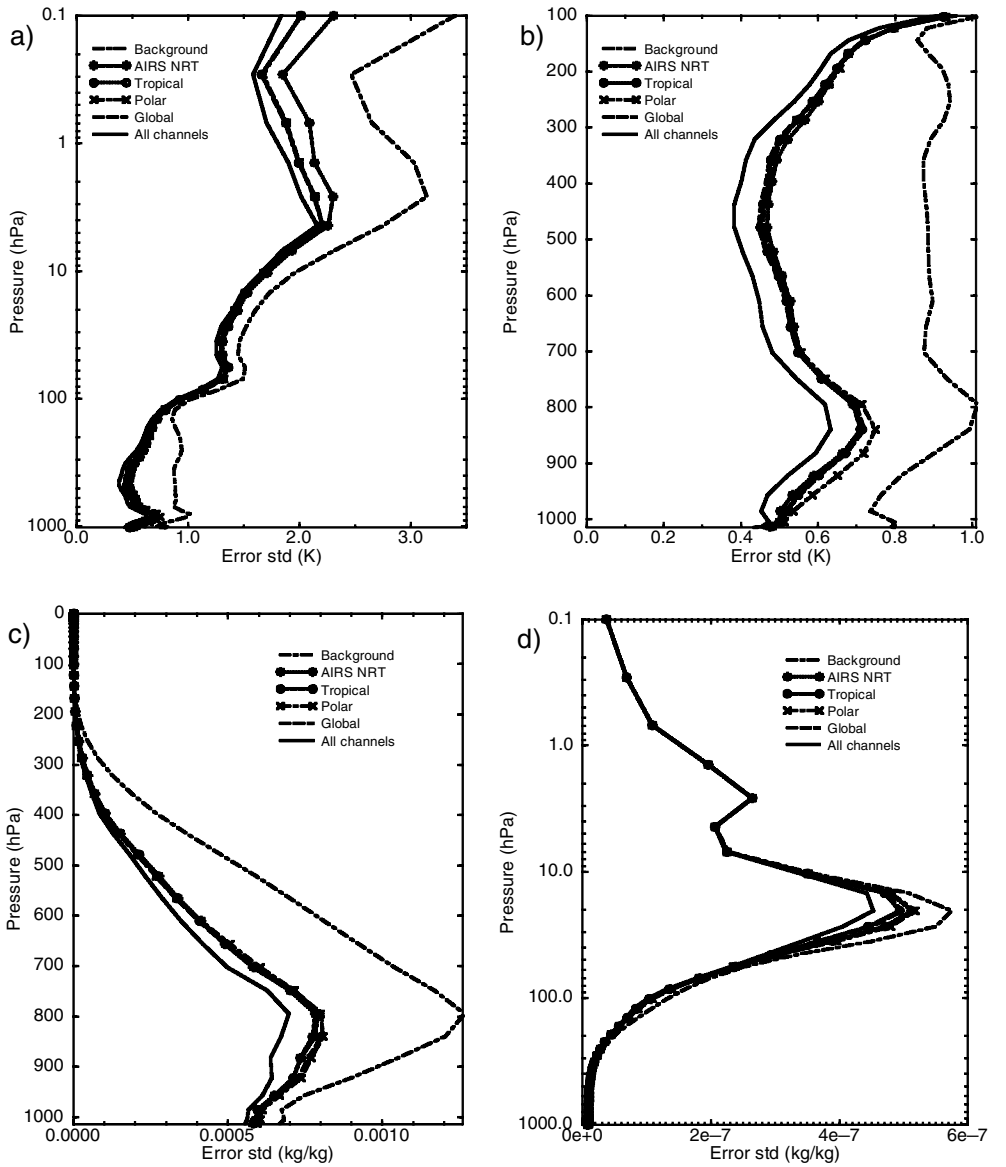


Figure 5. One-dimensional variational assimilation of temperature ((a) and (b)), humidity (c) and ozone (d) error standard deviation averaged over the tropical air-mass type (14 profiles). Errors are represented for 'background', 'AIRS NRT' (the selection of Atmospheric Infrared Sounder radiance channels available in near real time), 'tropical', 'polar' and 'global' channel selections. In addition, results for all the channels are given.

et al. 1998) as being crucial to forecast-error development. It is therefore important to assess the ability of the AIRS NRT selection to 'observe' such important atmospheric patterns, commonly called 'key analysis errors'.

One month of temperature 'key analysis errors' (as described in Klinker *et al.* 1998) was computed at a resolution of T159 (120 km): these 'errors' represent perturbations that, if added to the ECMWF operational analysis, reduce the 48-hour forecast error (defined as the global difference between the 48-hour forecast and the verifying analysis). Up to now, neither humidity nor ozone perturbations are considered in the

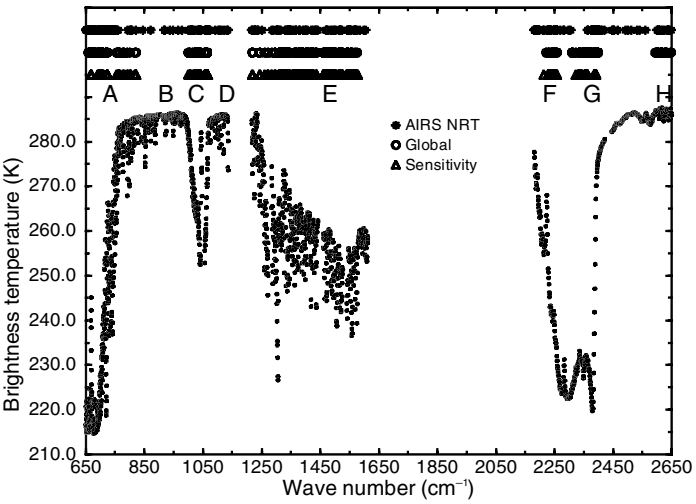


Figure 6. Typical spectrum of the Atmospheric Infrared Sounder (AIRS) superimposed on the figure by the location of the channels chosen by the AIRS near real time (NRT) selection (stars), picked by the 'global' channel selection (circles) and by the 'sensitivity' channel selection (triangles). See Table 1 for the letters corresponding to different spectral bands.

sensitivity computations run operationally at ECMWF, therefore standard climatological short-range forecast errors were used for these two fields. Also, surface temperature was excluded from the selection criterion. As a consequence, only results for atmospheric temperature are relevant in the sensitivity study described here. The sensitivity temperature structures are generally of small amplitude (meaning that a small atmospheric perturbation in this area can have a very large impact on the forecast quality) and can be fairly sharp both in the horizontal and in the vertical. From this one month worth of statistics, a new temperature covariance matrix was derived, hopefully capturing the variability of the key atmospheric elements that are essential to observe in order to dramatically improve the forecasts. This covariance matrix (averaged over one month) is sharper in the vertical and horizontal than the operational background covariance error (not shown). In addition, as can be seen from Fig. 7, the error standard deviations are proportionally large in the troposphere and in the high stratosphere.

The OMIC selections computed with the operational temperature background-error covariance matrix ('global') and 'key analysis error' covariance matrix ('sensitivity') share 234 channels. The main differences appear in bands A, G and H where channels are selected solely by the 'global' method using the operational background error (Fig. 6). The absence of channels in band H is explained by the removal of surface temperature from the 'sensitivity' experiment. Due to larger absolute values of standard deviations using the standard background-error covariances, 'global' selects more channels in the stratospheric parts of bands A and G. 'Sensitivity' selects more channels in bands C (ozone, 14 channels) and E (water vapour), channels obviously also having some sensitivity to temperature due to the nonlinearity in radiative transfer for this band and located in layers where the temperature background-error variance is proportionally large.

The 1D-Var temperature-error standard deviation is shown in Fig. 7 for AIRS NRT and 'sensitivity' selections, together with the 'key analysis error' standard deviation (used here as a background reference). The two selections share only 93 channels.

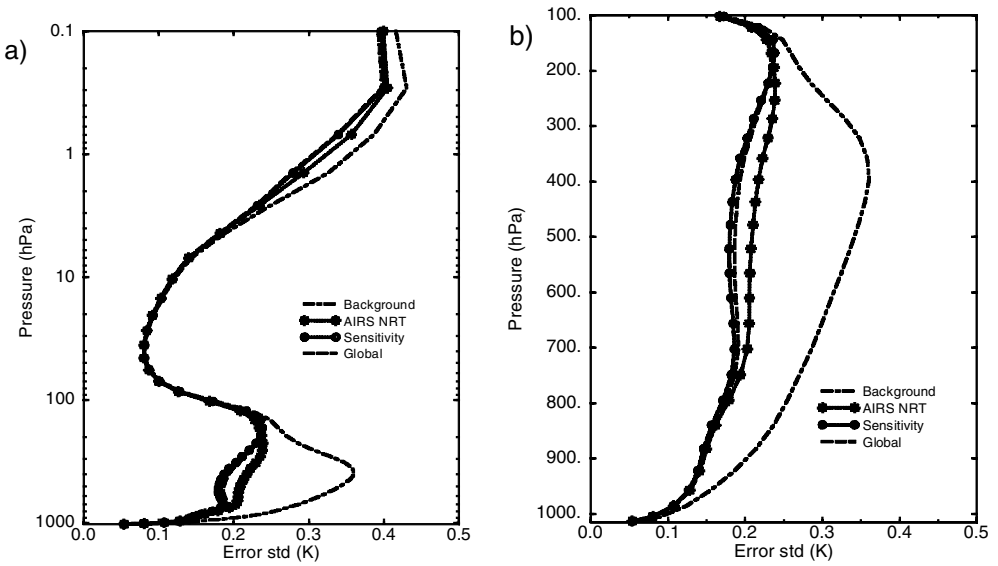


Figure 7. One-dimensional variational assimilation (1D-Var) of the temperature-error standard deviation averaged over 108 profiles ((a) in log scale and (b) in linear scale), for various channel selections. 'Key analysis error' corresponds to the standard deviation of the key analysis error. 'AIRS NRT' and 'sensitivity' correspond to 1D-Var analysis errors for AIRS NRT (the selection of Atmospheric Infrared Sounder radiance channels available in near real time) and 'global' selections using the 'key analysis' as the background error.

Even though the superiority of the 'sensitivity' channel selection is clear in the troposphere and high stratosphere, AIRS NRT selection still manages to substantially reduce the original 'key analysis error' variance in the troposphere. The 'global' channel selection provides comparable results to the 'sensitivity' channel selection with the exception of the 300–800 hPa atmospheric layer where this latter reduces the analysis error slightly more.

This only gives a flavour of the efficiency of the AIRS NRT channel selection to cope with sensitivity perturbations, since it is now well recognized that these sensitive areas are generally affected by clouds that will badly affect the performance in exploiting the AIRS data (McNally 2002; Fourrié and Rabier 2003). Our result implies that if these structures exist in reality and have some signature in clear-sky areas and if nonlinearity of the radiative transfer does not affect the retrieval of such structures too much by coupling humidity and temperature, the current AIRS NRT selection should be able to observe them.

6. Conclusions

The AIRS NRT channel selection to be provided to the operational weather centres was assessed via comparison with a 'global' selection deduced from an optimal iterative method following Rodgers (1996) for clear night-time conditions. AIRS NRT and 'global' channel selections are somewhat different, only 119 channels out of 324 are indeed identical in the two selections. However, both selected channels span sufficiently close spectral regions and lead to similar results in terms of DFS or ER, and also in terms of temperature, humidity, ozone and surface-temperature linear 1D-Var analysis errors.

The impact of the air mass on the channel selection was then studied. The outcome is that a poorly trained optimal channel selection can perform badly especially for low-tropospheric temperature and humidity. The performance of the OMIC channel selection is sensitive to the training data even though the impact on retrieval accuracy is not particularly detrimental. On the other hand, AIRS NRT selection seems to capture satisfactorily most of the variability of different atmospheric situations.

The robustness of the AIRS NRT channel selection for the retrieval of key analysis-error structures was assessed by the use of a covariance matrix representative of key analysis errors felt crucial for the quality of the short-range forecasts. It was shown that under those 'extreme' conditions, AIRS NRT again performs reasonably well compared with a dedicated channel selection.

All the experiments show that despite the overall slightly smaller information content of the AIRS NRT versus optimally derived channel selections, this selection seems very reasonable for NWP applications and appears to be robust. This gives some confidence in the day-1 strategy, which will use the AIRS NRT selection.

The limitations of this study lay in the fact that it is only valid for night-time clear conditions. In addition, the study does not allow for the nonlinearity of errors in the water-vapour band and it is dependent on the assumptions for background, forward-model and observational-error covariances. In particular, the forward model errors are assumed to be constant over the spectrum, which is not the case in reality. If they were taken into account, the optimal channel selection would be affected and the water-vapour band would be sampled less in the OMIC channel selection.

Acknowledgements

This work was partly supported by the NWP Satellite Applications Facilities via a scientific visit to ECMWF and by the Centre National d'Etudes Spatiales through the post-doctorate fellowship of Nadia Fourrié. The authors acknowledge Frédéric Chevallier for providing the atmospheric sampling dataset used to train the different channel selections, and Marco Matricardi, Tony McNally, Adrian Simmons, Tony Hollingsworth, Florence Rabier and Thomas Auligné for helpful discussions and suggestions.

References

Chevallier, F. 1999 'TIGR-like sampled databases of atmospheric profiles from the ECMWF 50-level forecast model'. NWP SAF Research Report No. 1

Chevallier, F., Chédin, A., Chéruy, F. and Morcrette, J.-J. 2000 TIGR-like atmospheric profile databases for accurate radiative flux computation. *Q. J. R. Meteorol. Soc.*, **126,** 777–785

Collard, A. D. 1998 'Notes on IASI performance'. NWP technical report No. 253, Met Office, Bracknell, UK

Fourrié, N. and Rabier, F. 2003 'Use of advanced sounders in cloudy conditions'. In proceedings of the twelfth international TOVS study conference, 27 February–5 March 2002, Lorne, Australia. In press

Kaplan, L. D., Chahine, M. T., Susskind, J. and Searl, J. 1977 Spectral band passes for a high precision satellite sounder. *Appl. Opt.*, **16,** 322–325

Klinker, E., Rabier, F. and Gelaro, R. 1998 Estimation of the key-analysis errors using the adjoint technique. *Q. J. R. Meteorol. Soc.*, **124,** 1909–1933

Matricardi, M. and Saunders, R. 1999 Fast radiative transfer model for simulation of infrared atmospheric interferometer radiances. *Appl. Opt.*, **38,** 5679–5691

Matricardi, M., Chevallier, F. and Tjemkes, S. 2001 'An improved general fast radiative transfer model for the assimilation of radiance observations'. ECMWF research department technical memorandum 345, European Centre for Medium-Range Weather Forecasts, Reading, UK

McNally, A. P. 2002 The occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Meteorol. Soc.*, **128,** 2551–2556

Prunet, P., Thépaut, J.-N. and Cassé, V. 1998 The information content of clear-sky IASI radiances and their potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **124**, 211–241

Rabier, F., Fourrié, N., Chafaš, D. and Prunet, P. 2002 Channel selection methods for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.*, **128,** 1011–1027

Rodgers, C. D. 1996 Information content and optimisation of high spectral resolution measurements. *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II*, **2830,** 136–147

2000 *Inverse methods for atmospheres: Theories and practice*. World Scientific Publ., Singapore

Susskind, J., Barnet, C. D. and Blaisdell, J. M. 2003 Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. *IEEE Remote Sensing*, in press

Q. J. R. Meteorol. Soc. (2004), **130**, pp. 1839–1856 doi: 10.1256/qj.03.27

Cloud characteristics and channel selection for IASI radiances in meteorologically sensitive areas

By NADIA FOURRIÉ* and FLORENCE RABIER
Météo-France, Toulouse, France

(Received 14 February 2003; revised 11 December 2003)

SUMMARY

The cloudiness in simulated Infrared Atmospheric Sounding Interferometer (IASI) pixels deduced from the Advanced Very-High-Resolution Radiometer (AVHRR) satellite imager is studied specifically in meteorologically sensitive areas during the Fronts and Atlantic Storm-Track Experiment. It is found that few clear AVHRR observations are located in the IASI pixels in these regions, which are covered by high-level and low-level clouds. The IASI channel selection is then studied in the context of the sensitive areas for the pixels with low-level clouds. The Entropy Reduction (ER) method, which was previously studied in a general context, is compared with two other channel selection methods using selection criteria based on adjoint sensitivity: the sensitivity to observations and the so-called Kalman-filter sensitivity. It is found that, even though the 'sensitive' methods give slightly better results than the ER one, the latter performs quite robustly and at a lower computational cost. The robustness to the specification of the background-error covariance matrix is then studied. It is shown that channel selection based on the ER method is particularly robust to the specification of the background-error covariance matrix, but the analysis step itself requires an accurate determination of the background-error covariance matrix. In addition it is shown that an independently computed constant channel set gives comparable results to the optimal channel set.

KEYWORDS: Advanced infrared sounders Cloud occurrence Satellite sounding Sensitive areas

1. INTRODUCTION

Advanced infrared sounders, such as the Atmospheric InfraRed Sounder (AIRS, launched on board the Aqua satellite in May 2002) and the Infrared Atmospheric Sounding Interferometer (IASI) scheduled to be on board the first operational European polar-orbiting satellite Metop, measure (or will measure) radiation in many thousands of channels. They have the potential to provide atmospheric temperature and composition information at a much higher vertical resolution and accuracy that can be achieved with the previous generation of satellite instruments such as the High-resolution Infrared Radiation Sounder (HIRS) (Smith *et al.* 1979). As it is neither feasible nor efficient to assimilate all the channels in a numerical weather-prediction (NWP) system, channel selection policies have been designed in the NWP context. Previous studies have proposed methods in order to reduce the number of channels to keep only the ones that provide the most useful information. In particular, Rabier *et al.* (2002) and Fourrié and Thépaut (2003) have used a method based on information content following Rodgers (1996), for clear-sky conditions.

If such advanced sounders are to lead to a very significant impact on the forecast quality, one should address in detail the question of retaining the maximum information possible in sensitive areas. Previous studies, such as those of Prunet *et al.* (1998) and Collard (1998), have suggested that IASI could resolve some of the small-scale baroclinic structures that have been identified by sensitivity studies as being crucial to forecast-error development (Rabier *et al.* 1996). Nevertheless, the main obstacle to using observations from advanced sounders is the presence of cloud, which can severely limit the information from infrared sounders. In this context, McNally (2002) investigated the

* Corresponding author: CNRM/GMAP, Météo-France, 42 av. G. Coriolis, 31057 Toulouse Cedex, France.
e-mail: Nadia.Fourrie@meteo.fr

occurrence of clouds in the sensitive areas using cloud fields from the ECMWF* model, and showed with an 'observable' sensitivity variable that there was a high correlation between meteorologically sensitive areas and cloud cover produced by the ECMWF model.

The first objective of this study is also to study the cloud cover and the cloud-top level, but from a satellite imager—the Advanced Very-High-Resolution Radiometer (AVHRR)—in simulated IASI pixels. This is done over the North Atlantic Ocean for cases of the Fronts and Atlantic Storm-Track Experiment (FASTEX) (Joly *et al.* 1999) and, more particularly, in the sensitive areas for the forecasts of storms of eight intensive observing periods (IOPs). Once the cloud characteristics of the sensitive areas are established, the robustness of the iterative channel selection method used by Rabier *et al.* (2002) (hereafter R02) and Fourrié and Thépaut (2003) is addressed for these regions. This method is compared with two channel selection methods using a criterion based on the adjoint sensitivity. These last two methods were developed in the context of the targeting problem, which consists in optimally selecting observations in the sensitive areas in order to reduce the error in the initial state of the forecast.

Section 2 presents the study of cloudiness in meteorologically sensitive areas for the eight FASTEX IOPs. The description of the channel selection framework is given in section 3. Results obtained in the simple context of perfect observations for profiles with cloud-top levels at 1000 hPa from the FASTEX IOP sensitive areas are discussed in section 4. The channel selection methods are then applied in a more general context for these particular synoptic situations, and the impact of the background-error covariance matrix and of a constant set are studied in sections 5 and 6, respectively. Finally, the results are summarized and the conclusions are given in section 7.

2. Study of the cloudiness in meteorologically sensitive areas

(*a*) *Experimental framework*

In this study, the cloud parameters (cloud cover and cloud-top pressure in simulated IASI pixels) have been retrieved from the AVHRR imager using the Mask AVHRR for Inversion ATOVS† (MAIA) method that was originally developed at the Centre de Météorologie Spatiale (CMS) France in the frame of the AVHRR and ATOVS processing package (Lavanant *et al.* 1999).

(i) *The cloud mask.* The MAIA method has been adapted to the IASI observations (Lavanant, personal communication). It processes the AVHRR observations mapped inside the simulated IASI pixel and determines the mean clear percentage cover in the IASI spot. With a succession of threshold tests applied to the AVHRR channels inside simulated IASI pixels, it is possible to determine some cloud parameters. The cloud cover is evaluated from the percentage of clear AVHRR pixels in each simulated IASI ellipse shape. A minimum threshold of 90% of clear AVHRR spots has been arbitrarily chosen for regarding IASI pixels as clear. In addition, for cloudy pixels, the cloud-top temperature (if the cloud can be considered as a black body), among other parameters, is available as output from the MAIA algorithm. The cloud-top pressure level is deduced from the cloud-top temperature with a series of differences between the cloud-top temperature from the MAIA method and a temperature profile interpolated onto the simulated observation location. The temperature fields are available every 50 hPa in the vertical and are provided by the FASTEX 4D-Var reanalysis (Desroziers *et al.* 2003),

* European Centre for Medium-Range Weather Forecasts.

† Advanced TIROS (Television InfraRed Observation Satellite) Operational Vertical Sounder.

which included additional observations from the field experiment. The cloud-top pressure is considered to be at the level where the difference between the cloud-top temperature and the atmospheric-profile temperature is a minimum. It is worth noting that only pixels over the sea for which a cloud-top temperature could be retrieved have been considered in this study. Indeed, since the MAIA method has been mainly tested over the ocean because of the CMS reception area, there is a lack of confidence in the method over land-ice in the Halifax reception area.

(ii) *Determination of the sensitive areas.* Short-term forecast errors are mainly due to errors in the initial state of the forecast. In order to determine the regions where small errors in the initial state may increase and lead to major forecast errors, the gradient of a diagnostic function of the final state with respect to the initial conditions can be used (e.g. Rabier *et al.* 1996). Sensitive areas can be defined with scaled gradients resulting from these adjoint calculations. In this study, the gradient is based on the diagnostic function used in the study by Hello *et al.* (2000). It is computed from the forecast of the mean-sea-level pressure over the area of the FASTEX weather systems studied. Eight dates in February 1997 have been examined. This period corresponds to a zonal weather regime, which is favourable to the development of cyclonic systems over western Europe (Joly *et al.* 1999; Baehr *et al.* 1999). The temperature fields of the gradient at 900, 800, 700, 600 and 500 hPa levels have been squared and summed at each geographical location in order to obtain a single horizontal sensitivity field. The sloping nature of the gradient pattern is not represented in the summation of fields at different levels. Although the calculation with squared quantities prevents the positive and the negative values of the sensitivity cancelling each other in the vertical direction, it provides an averaged horizontal sensitivity field that is wider than any of the individual sensitivity fields at each level. Once the sensitivity field has been computed, a threshold (arbitrarily fixed at 30% of the sensitivity-field maximum) is applied to it in order to obtain a horizontal mask for the observations located in the sensitive area for which cloud cover and cloud-top height are to be studied. It should be noted that the extent of the derived sensitive area depends slightly on this threshold.

Before presenting the overall results for the eight IOPs, a single case—the 'FASTEX cyclone'—is studied in the next subsection.

(*b*) *A case study: FASTEX intensive observing period 17*

The case study is FASTEX IOP 17, which is a well-sampled case (Cammas *et al.* 1999). It is an example of a low with a complex life cycle which ends as an explosive cyclogenesis (Baehr *et al.* 1999) at 00 UTC 20 February 1997, with a strong deepening rate of 40 hPa in 24 hours and a minimum central pressure of 943 hPa. We are mainly interested in the analysis made at 00 UTC 18 February and the subsequent 36-hour forecast. The verification area is defined by the bounding lines of latitude and longitude 63°N, 52°N, 20°W and 0°W at 12 UTC 19 February 1997.

Figure 1 displays the cloud-top pressure (deduced by the MAIA method) of the simulated IASI pixels located in the sensitive area. Most of the pixels have a cloud cover exceeding 90%. The observed cloud can be considered as opaque and no clear pixel is observed in this sensitive area. Even the high-level clouds are determined as opaque. This fact is due to an artefact of the MAIA method, which only processes the pixels for which is possible to determine a black-body temperature. A large area having cloud-top levels above 600 hPa is observed in a wide southern part of the sensitive area and low-level clouds (cloud-top pressures between 1000 and 700 hPa) are present in its northern part. This particular distribution of clouds can be explained

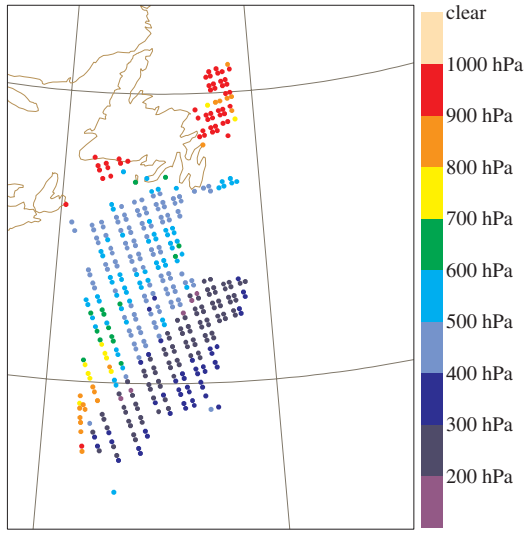


Figure 1. The cloud-top pressure of the IASI pixels for the sensitive area at 00 UTC 18 February 1997.

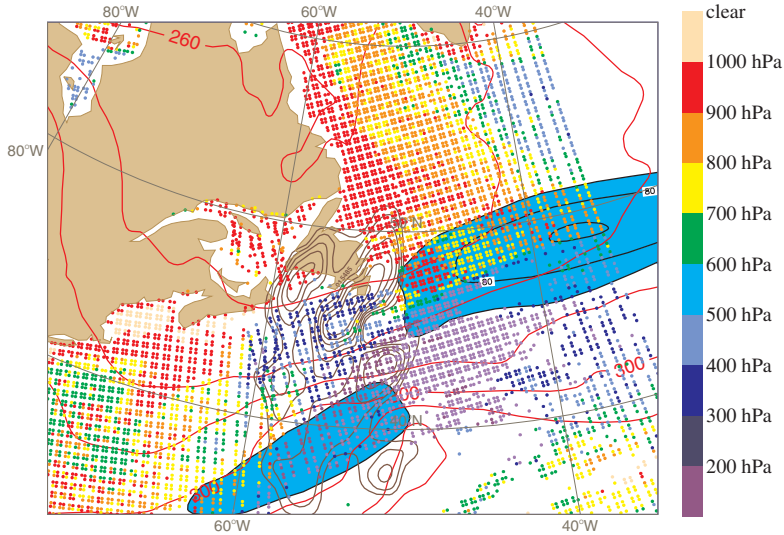


Figure 2. The cloud-top pressure of the IASI pixels over the whole area for IOP17 (00 UTC 18 February 1997). The blue shaded area corresponds to the 300 hPa wind velocity above 60 m s^{-1}. The red lines represent the 850 hPa equivalent potential temperature and are plotted every 10 K. The brown lines are associated with the sensitive area and are plotted every one tenth of the maximum value above the threshold; they have been computed from temperature fields of the gradient of the forecast mean-sea-level pressure over the weather system studied (see text for further explanation).

by the meteorological situation summarized in Fig. 2. In fact, the sensitive area is located south of Newfoundland in between two upper-level jets (in the left exit of the southern one and in the entrance of the northern one). These two jets come from the split of the upper-level jet, and a so-called cloud-head pattern with high-level cloud has developed (Cammas *et al.* 1999). This cloud shape is indicative of a rapid cyclogenesis; during the

TABLE 1. DESCRIPTION OF EIGHT FASTEX CASES

IOP	Initial time	Forecast range (hours)	Verification area	Level of maximum sensitivity (hPa)	Number of IASI pixels
12	12 UTC 8 February	36	(67.0°N, 55.0°N, 25.0°W, 8.0°W)	900	195
14	12 UTC 12 February	36	(56.5°N, 46.0°N, 28.5°W, 8.0°W)	800	801
15	12 UTC 14 February	36	(62.0°N, 43.0°N, 30.0°W, 5.0°W)	600	91
16	12 UTC 16 February	30	(62.5°N, 55.0°N, 12.0°W, 0.0°W)	700	183
17-a	00 UTC 18 February	36	(63.0°N, 52.0°N, 20.0°W, 0.0°W)	600	423
17-b	18 UTC 17 February	42	(63.0°N, 52.0°N, 20.0°W, 0.0°W)	800	355
18	12 UTC 22 February	24	(62.0°N, 46.0°N, 40.0°W, 10.0°W)	700	1387
19	00 UTC 26 February	36	(63.5°N, 57.0°N, 30.0°W, 13.0°W)	800	163

previous 12 hours the surface pressure deepened by 17 hPa (Cammas *et al.* 1999), an explosive rate according to Sanders and Gyakum (1980).

(*c*) *Averaged results of eight FASTEX IOPs*

A similar study has been conducted for eight FASTEX IOPs. A brief description of the synoptic cases corresponding to some cases studied by Hello and Bouttier (2001) and Hello (2002) is given in Table 1. The sensitive direction was computed from the forecast trajectory of the cyclones following the definition target and verification times, as described by Bergot (1999). The overall forecast range was mostly 36 hours and the verification area, situated offshore of western Europe, was centred on the location of the cyclones at the time of the verification. For six cases out of eight, the sensitivity maximum was located below, or at, the 700 hPa level.

The IASI pixels have been simulated in the same area as the AVHHR observations, which are available from the Halifax reception station. These observations are representative of a mid-latitude region. The averaged results of the distribution of the cloud-top pressure for the sensitive areas and the whole dataset are presented in Fig. 3. For the sensitive areas, two maxima in the distribution of the cloud-top level are identified for the atmospheric layers located between 1000 and 900 hPa (low-level clouds) and between 400 and 300 hPa (high-level clouds). Nonetheless, this synthetic table does not illustrate the case-to-case variability of the cloud-top pressure in the sensitive area. In contrast, for the whole IASI dataset, almost 60% of the cloud tops of the observations are situated below 600 hPa. Furthermore, the cloud-top pressure in observations of the mid-latitude area seems almost constant in time, mainly with low-level cloud. The sensitive areas are often located in the southern entrance area of an upper-level jet-streak. Their horizontal extension depends also on the weather system. For almost all cases, the northern part of the sensitive area corresponds to low-level cloud and the southern one to high-level cloud. The particular shape of the cloud cover strongly limits an adequate sampling of the sensitive areas, the maximum of which is located at or below 700 hPa. These results corroborate the study of McNally (2002), who found a high degree of correlation between high cloud and sensitive areas during the winter.

In conclusion, we have shown that parts of the simulated IASI pixels located in the sensitive areas were covered by either low-level clouds or by high-level clouds. We now tackle the question of the channel selection for these low-level-cloud conditions (noting that the high-level-cloud conditions are not likely to provide useful information from the infrared sounders).

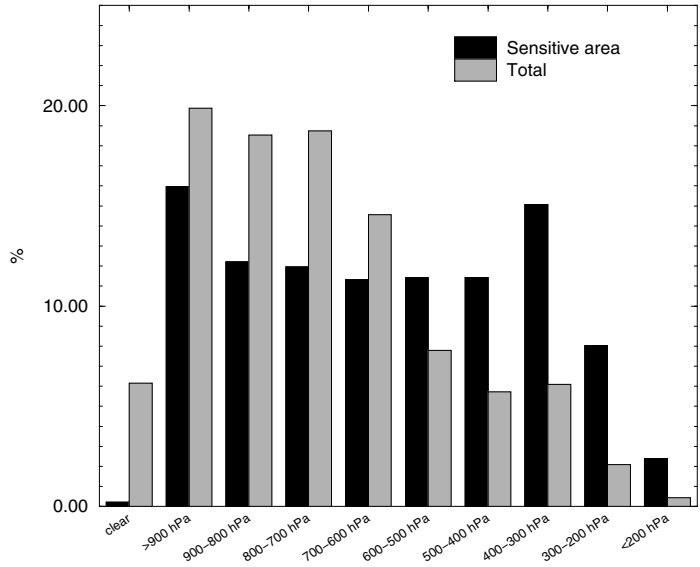


Figure 3. Averaged distribution of the cloud-top level of the IASI pixels over the eight sensitive areas of the FASTEX intensive observing periods and over the whole dataset corresponding to a mid-latitude region.

3. Description of the channel selections

(*a*) *General framework*

The general framework of this study is linear optimal estimation theory in the context of NWP. The atmospheric profile at a given location is represented by a vector $\mathbf{x}$ and the satellite observation by a vector $\mathbf{y}^{\mathrm{o}}$. The observations are linked to the atmospheric state by the radiative-transfer equation:

$$\mathbf{y}^{\mathrm{o}} = \mathcal{H}(\mathbf{x}) + \boldsymbol{\varepsilon}_{\mathbf{O}} + \boldsymbol{\varepsilon}_{\mathbf{F}}, \tag{1}$$

where $\mathcal{H}$ is the radiative-transfer model and the measurement and forward-model errors $\boldsymbol{\varepsilon}_{\mathbf{O}}$ and $\boldsymbol{\varepsilon}_{\mathbf{F}}$ are assumed to be Gaussian noise with covariance matrices $\mathbf{O}$ and $\mathbf{F}$. We will denote $\mathbf{R} = \mathbf{O} + \mathbf{F}$, the resulting observation-error covariance matrix. Our previous knowledge of the atmospheric state is represented by the background state vector $\mathbf{x}^{\mathrm{b}}$. Its error covariance matrix is denoted $\mathbf{B}$. All errors are assumed to be unbiased (zero mean). The radiative-transfer equation is assumed to be weakly nonlinear, making the tangent linear assumption valid:

$$\mathcal{H}(\mathbf{x}) = \mathcal{H}(\mathbf{x}^{\mathrm{b}}) + \mathbf{H}(\mathbf{x} - \mathbf{x}^{\mathrm{b}}), \tag{2}$$

where the so-called Jacobian matrix $\mathbf{H}$ is the tangent-linear approximation of $\mathcal{H}$. The optimal analysed state is given by

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}(\mathbf{y}^{\mathrm{o}} - \mathbf{y}^{\mathrm{b}}), \tag{3}$$

with $\mathbf{y}^{\mathrm{b}} = \mathcal{H}(\mathbf{x}^{\mathrm{b}})$, $\mathbf{K} = \mathbf{A}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$ and $\mathbf{A} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}$. The superscript T denotes the matrix transpose, $\mathbf{K}$ is the Kalman gain matrix, and $\mathbf{A}$ is the analysis-error covariance matrix. $\mathbf{A}$ is also linked to $\mathbf{B}$ following the classical relationship

$$\mathbf{A} = \mathbf{B} - \mathbf{K}\mathbf{H}\mathbf{B} = \mathbf{B} - \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{H}\mathbf{B}.$$

If one considers the assimilation of one single channel, **A** becomes

$$\mathbf{A} = \mathbf{B}\left\{\mathbf{I} - \frac{1}{1 + (\mathbf{Bh})^{\mathrm{T}}\mathbf{h}}\mathbf{h}(\mathbf{Bh})^{\mathrm{T}}\right\}, \tag{4}$$

with **h** being the vector corresponding to a line of the Jacobian matrix normalized by $\mathbf{R}^{-1/2}$. The parameter space is the temperature on the 43 pressure levels of the fast radiative-transfer model RTIASI (Matricardi and Saunders 1999). The IASI data have been simulated from atmospheric profiles (temperature, humidity and surface temperature) for nadir views, and noise has been added to the simulated data using the **O** covariance matrix provided by CNES* in 2001. As in the R02 study, a constant error of 0.2 K coming from the radiation-transfer-model noise is added to the measurement errors. The correlations existing between adjacent channels are ignored in the observation-error matrix.

Only low-level clouds were tackled here. We have studied atmospheric profiles with a cloud-top level at 1000 hPa because the MAIA method found a maximum of low-level cloud top for this level. We assume cloud tops at the 1000 hPa level with a surface emissivity of one. As our radiation-transfer model is not assumed to be very accurate for the cloudy brightness-temperature simulation, and because there is an uncertainty on the cloud information coming from the accuracy of the MAIA algorithm, the channels sensitive to the cloud effects have been removed before the channel selection. The choice criterion is the Jacobian; the test used for the determination of the cloud contamination is the ratio between the total Jacobian surface and the Jacobian surface below the cloud-top level. If this ratio shows that more than 2% of the Jacobian surface is below cloud level, which corresponds to the two lowest levels of RTIASI for a cloud at 1000 hPa, the channel is considered to be too sensitive to cloud to be kept in the channel selection. This allows the channels sensitive to the two last levels below 1000 hPa to be removed from the channel selection, and the number of channels is decreased from 8461 to 2300. This number would be reduced to about 1600 channels if one considered cloud-top levels at 950 hPa. This prior channel selection is intended to keep the simulations as realistic as possible, assuming that the channel contaminated by cloud effects will be more difficult to use for the day-1 assimilation of IASI.

(*b*) *Channel selection methods*

The problem of dynamical channel selection is tackled here. The iterative channel selection of R02 is compared with two methods developed in the framework of the targeting problem. The targeting issue consists of determining where to optimally select some observations in order to reduce the forecast-error variance.

(i) *Entropy reduction.* This iterative method for channel selection, as proposed by Rodgers (1996) and used by R02, consists in performing successive analyses, each using only one channel at a time. The channel selection in our case is based on maximizing the entropy reduction (ER) (Rodgers 2000):

$$\mathrm{ER} = -\tfrac{1}{2}\log_2 \det(\mathbf{AB}^{-1}) = \tfrac{1}{2}\log_2(1 + \mathbf{h}^{\mathrm{T}}\mathbf{Bh}). \tag{5}$$

The analysis-error covariance matrix **A** is updated accordingly and it is used at the next step as the background-error covariance matrix. This ensures that all the information brought by previous channels is taken into account for the selection of the new channel. It is referred to in the following as the 'ER method'.

* Centre National d'Etudes Spatiales.

(ii) *Sensitivity to observations.* The topic of sensitivity to observations (Baker and Daley 2000; Doerenbecher and Bergot 2001) has been developed in the context of adaptive observations, and these authors have suggested that the sensitivity with respect to observations could be an efficient tool for defining the location of targeted observations. An advantage of this kind of sensitivity is that takes into account the way in which the observations will be assimilated through the multiplication of the gradient by the **A** matrix. Another advantage is that the sensitivity to observations is significantly reduced in regions already sampled by existing observations. This method has been applied to our problem of channel selection. Following the philosophy of the method, the channel $\mathbf{y}_i$ selected is the one that maximizes the sensitivity to observation $\nabla_{\mathbf{y}_i^o}\mathcal{J}$:

$$\nabla_{\mathbf{y}_i^o}\mathcal{J} = \mathsf{r}_i \mathbf{h}_i'^{\mathrm{T}} \mathbf{A}_i \nabla_{\mathbf{x}^a}\mathcal{J}, \tag{6}$$

where $\mathcal{J}$ is the sensitivity cost function, r_i is the corresponding observation error value, $\mathbf{h}_i'$ is the vector corresponding to a line of the Jacobian matrix, $\nabla_{\mathbf{x}^a}\mathcal{J}$ is the sensitivity to the initial conditions and $\mathbf{A}_i$ is computed from Eq. (4) for the corresponding channel $\mathbf{y}_i$. As before, this matrix is used in the next step as the background matrix in order to take into account the information brought by the previously selected channels.

(iii) *Kalman-filter sensitivity.* The Kalman-filter sensitivity (KFS) method was proposed by Bergot and Doerenbecher (2002) in order to find the optimal deployment of targeted observations. This method is based on the maximum decrease of the variance of the error on a given scalar function of forecast. The reduction in the variance of the error on a given scalar function of the forecast, due to the inclusion of one channel $\mathbf{y}_i$ is given by:

$$(\quad_i)^2 = \nabla_{\mathbf{x}^a}\mathcal{J}\mathbf{B}_{i-1}\mathbf{h}_i'(\mathsf{r}_i + \mathbf{h}_i'^{\mathrm{T}}\mathbf{B}_{i-1}\mathbf{h}_i')^{-1}\mathbf{h}_i'^{\mathrm{T}}\mathbf{B}_{i-1}\nabla_{\mathbf{x}^a}\mathcal{J}, \tag{7}$$

where $\mathbf{B}_{i-1}$ corresponds to the analysis-error covariance matrix obtained at the previous step. As for the ER method, the **A** matrix is updated and is used as the **B** matrix for the next step. This method is the so-called 'KFS method'.

(*c*) *Key analysis-error matrix*

One month of 'key analysis errors' (KAEs), as described by Klinker *et al.* (1998), has been computed at a resolution of T159 (120 km). These 'errors' represent perturbations that, if added to the ECMWF operational analysis, reduce the 48-hour forecast error (defined as the global difference between the 48-hour forecast and the verifying analysis). Up to now, humidity perturbations have not been considered in the sensitivity computations, and so only temperature is included in the sensitivity study described here. These structures are generally of small amplitude (meaning that a small atmospheric perturbation in this area can have a very large impact on the forecast quality) and can be fairly sharp, both in the horizontal and in the vertical. The associated covariance matrix (averaged over one month) is sharper in the vertical and horizontal than the operational background covariance error. In addition, the error standard deviations are proportionally large in the troposphere (Fig. 4) in comparison with the ones in the stratosphere; the error maximum in the troposphere at about 300 hPa is of the same order as in the stratosphere, but the error maximum in the troposphere for the 'climatological' **B** matrix is about half the stratospheric maximum. The KAE covariance matrix has been chosen as the **B** matrix because our study is focused on channel selection in the sensitive areas.

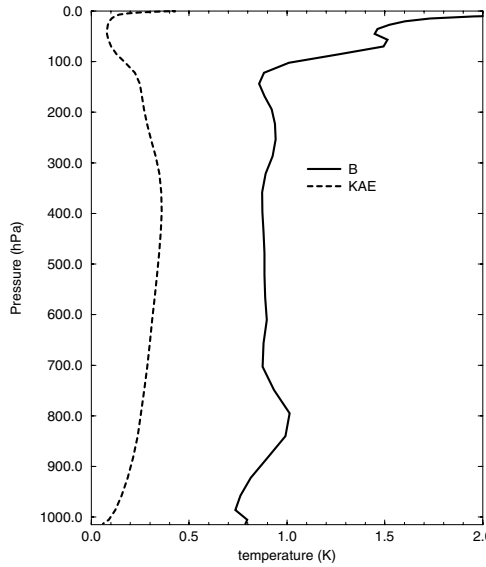


Figure 4. Standard deviation of temperature for a climatological background-error covariance matrix **B** (derived from the 60-level ECMWF covariance matrix representing short-range forecast errors projected onto the 43 radiative-transfer model levels) and for key analysis errors (KAE).

4. Results obtained with perfect observations

In this first study, the truth is a 'corrected state' which is a 3D-Var analysis modified with a scaled gradient perturbation. The coefficient applied to the perturbation was found by Hello (2002) in order to guarantee that, inside the sensitive area, the distance between the corrected state and the observations is minimal (Hello *et al.* 2000). Different observation sets can be used and the coefficient for which the forecast score, based on verifying observations from the corrected state, is the best is chosen. The IASI measurements are simulated from this 'corrected state'. In order to understand better the manner in which the three channel selection methods work, no noise has been added to the observations. The background profile comes from the 3D-Var analysis ('uncorrected' by the gradient), and one tries to retrieve the sensitive perturbation with perfect observations of the reality. The **B** matrix used in the computation comes from the KAE matrix. We assume that KAEs represent well the perturbations from the gradient perturbations. As in the R02 study, 300 channels have been selected.

(*a*) *Overall results*

The average result over several IOPs (142 profiles with cloud-top level at 1000 hPa) is displayed in Fig. 5(a) in terms of root-mean-square (rms) analysis error. Only four IOPs have been studied because profiles with cloud-top levels at 1000 hPa were observed for these IOPs. The mean rms background error has a dipole shape with maxima at 700 and 850 hPa, in good agreement with previous sensitivity studies that established the presence of a sensitivity maximum in the low- or mid-troposphere. The three methods provide similar rms analysis errors, with an advantage to the gradient-based selection methods at around 850 hPa. It should be noted that the analysis error is greater than the background error in the 100–450 hPa layer. This somewhat surprising

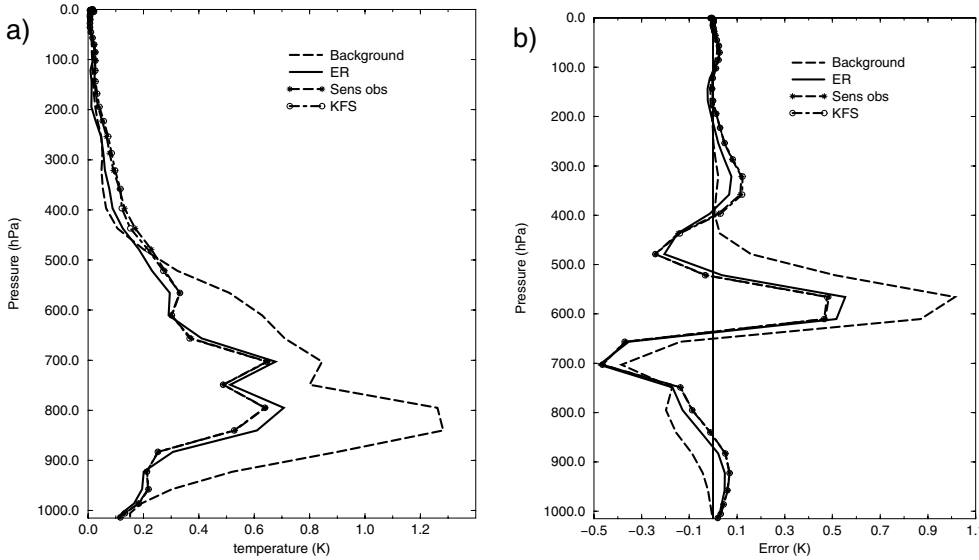


Figure 5. (a) The root-mean-square temperature analysis errors for three channel selection methods (ER, 'Sens obs', and KFS—see text) in the case of atmospheric profiles corresponding to a cloud-top level of 1000 hPa averaged over 142 profiles spanning several IOPs, together with the background errors. (b) Deviations from the true temperature for one particular representative profile for the background and the three different analyses. 'ER' refers to the analysis error obtained with the entropy-reduction method, 'Sens obs' to the analysis error given by the sensitivity-to-observation method and 'KFS' to the analysis error produced by the Kalman-filter sensitivity method.

TABLE 2. MEAN PERFORMANCES AVERAGED OVER 142 PROFILES OF THE THREE CHANNEL SELECTION METHODS (SEE TEXT)

Channel selection method	Degree of freedom for signal	Entropy reduction	Root-mean-square ratio
Entropy reduction	4.77	7.36	0.69
Sensitivity to observations	3.04	5.8	0.65
Kalman-filter sensitivity	3.09	5.9	0.66

result probably stems from the fact that the error variance in the background is not well described by the **B** matrix used for this small sample.

The degree of freedom for the signal (DFS) is given by

$$\mathrm{DFS} = \mathrm{Tr}(\mathbf{I} - \mathbf{A}\mathbf{B}^{-1}),$$

where Tr denotes the trace. This is useful in that it quantifies the gain in information brought by the observations with respect to the background information, and is larger for the ER method (Table 2). Nevertheless, the ratio between the rms analysis error and the rms of the background,

$$\left\{ \sum_{\text{levels}} \frac{(\mathbf{x}^{\mathrm{a}} - \mathbf{x}^{\mathrm{t}})^2}{(\mathbf{x}^{\mathrm{b}} - \mathbf{x}^{\mathrm{t}})^2} \right\}^{\frac{1}{2}},$$

is smaller for the gradient-based methods than for the ER method (0.65 and 0.66 instead of 0.69, see Table 2). This paradox is explained in the next subsection. 91 channels

are shared by the ER method and the KFS method, 100 channels are shared by the ER method and the sensitivity-to-observation method and 239 channels are commonly selected by the KFS and the sensitivity-to-observation methods.

(*b*) *Single profile study*

In order to study more precisely the channels selected by the three methods, attention has been paid to a single profile with a cloud-top level of 1000 hPa and close to the centre of the sensitive area. The results, in term of background and analysis errors, are shown in Fig. 5(b). The background error represents the difference between the original 3D-Var and the 'corrected' analysis. It is larger in the 400–900 hPa atmospheric layer and its maximum is located at about 550 hPa. Another relative maximum is present at 700 hPa. The three methods largely decrease the analysis error in comparison to the background one, although they slightly degrade the analysis in the 650–750 hPa atmospheric layer and around 350 hPa. The analysis error is slightly smaller for the two methods based on the adjoint gradient than for the ER method around the mean-error peak, while the ER method gives better results in the upper troposphere. This better performance of the ER method around 300 hPa can be explained by the fact that it uses more channels that peak high in the troposphere. The use of more (noise-free) observations helps to prevent the negative effects of extrapolation of information with a **B** matrix not absolutely adequate for this precise profile.

The DFS is larger for the ER method (4.71 instead of 2.86 and 2.95 for the KFS method and sensitivity-to-observation method, respectively) and, as found previously, the rms ratio is smaller for these last methods (0.64) than for the ER method (0.67), showing that the analysis error is decreased more in the case of adjoint-based channel selection methods. These results show that the theoretical reduction of uncertainty brought by the analysis, as measured by the DFS or the ER methods, is not the same information as the rms error of the analysed state. The DFS and the ER approaches indeed consider error correlations and smoothing in the retrieval system.

Figure 6 exhibits the normalized Jacobians of the 300 channels selected by each selection method. While the two methods based on the gradient preferably select channels peaking at 400 hPa, 500 hPa and 800 hPa, the channels selected by the ER method span the atmosphere more evenly. In contrast to both methods based on the gradient, the ER method selects a few channels above 300 hPa, owing to the large value of the corresponding standard deviation of the KAE in the upper stratosphere (Fig. 4). 83 channels are commonly selected between the ER and the KFS methods, 96 between the ER and the sensitivity-to-observations methods, and 240 between the two methods based on the gradient. These results are in good agreement with the study by Bergot and Doerenbecher (2002), who showed that the KFS and sensitivity-to-observations methods lead to the same optimal location of the targeted observation in most cases. The ER method, which does not take into account the sensitivity field, selects channels that are different from those selected by the adjoint selection methods, but gives comparable results, except for the major peak of sensitivity.

In conclusion, the ER method seems to be quite a robust channel selection scheme, even in sensitive areas. In addition, it is noteworthy that this method is less expensive in terms of computing time than the two other methods. As the very expensive computation of the **A** matrix for each channel and at each iteration is required by the selection method based on sensitivity to observations, and as the results provided by this method are not much better than the ones based on the KFS, the method based on sensitivity to observations is not discussed any further in this paper.

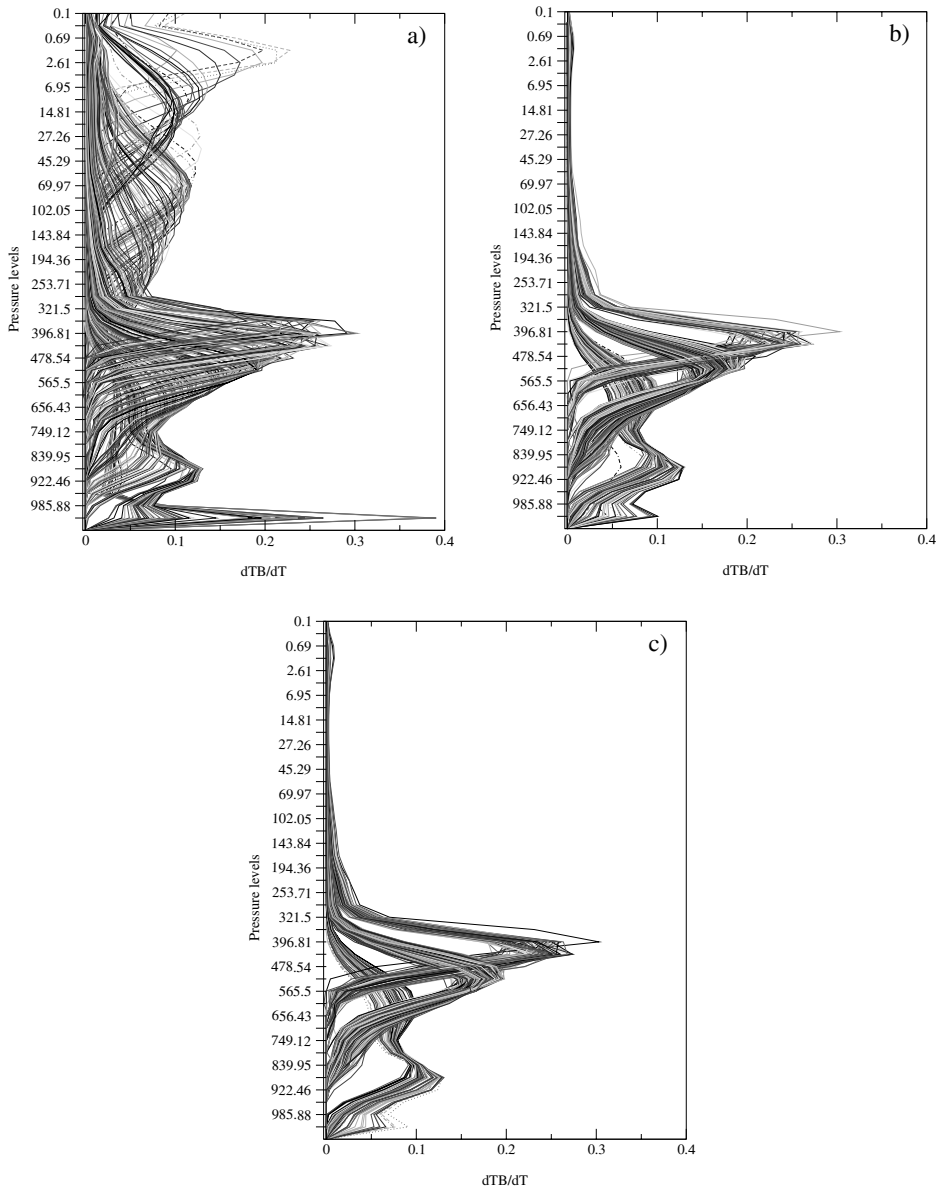


Figure 6. Jacobians (i.e. derivatives of the brightness temperature with respect to the temperature at the different levels of the radiative-transfer model) of the 300 channels selected by the channel selection methods based on different criteria: (a) entropy reduction, (b) sensitivity to observations and (c) Kalman-filter sensitivity. The y-axis varies from 1013.15 hPa to 0.1 hPa.

5. Influence of the background-error covariance matrix

Imagine that one wants to perform the best analysis in the sensitive areas without knowing accurately the associated background-error covariance. What is the impact of such a mis-specification on the analysis? In order to test the sensitivity of the selection methods to the background-error covariance matrix, two **B** matrices have been considered: the one computed from the KAE and the climatological-error covariance matrix (derived from the 60-level ECMWF covariance matrix representing short-range

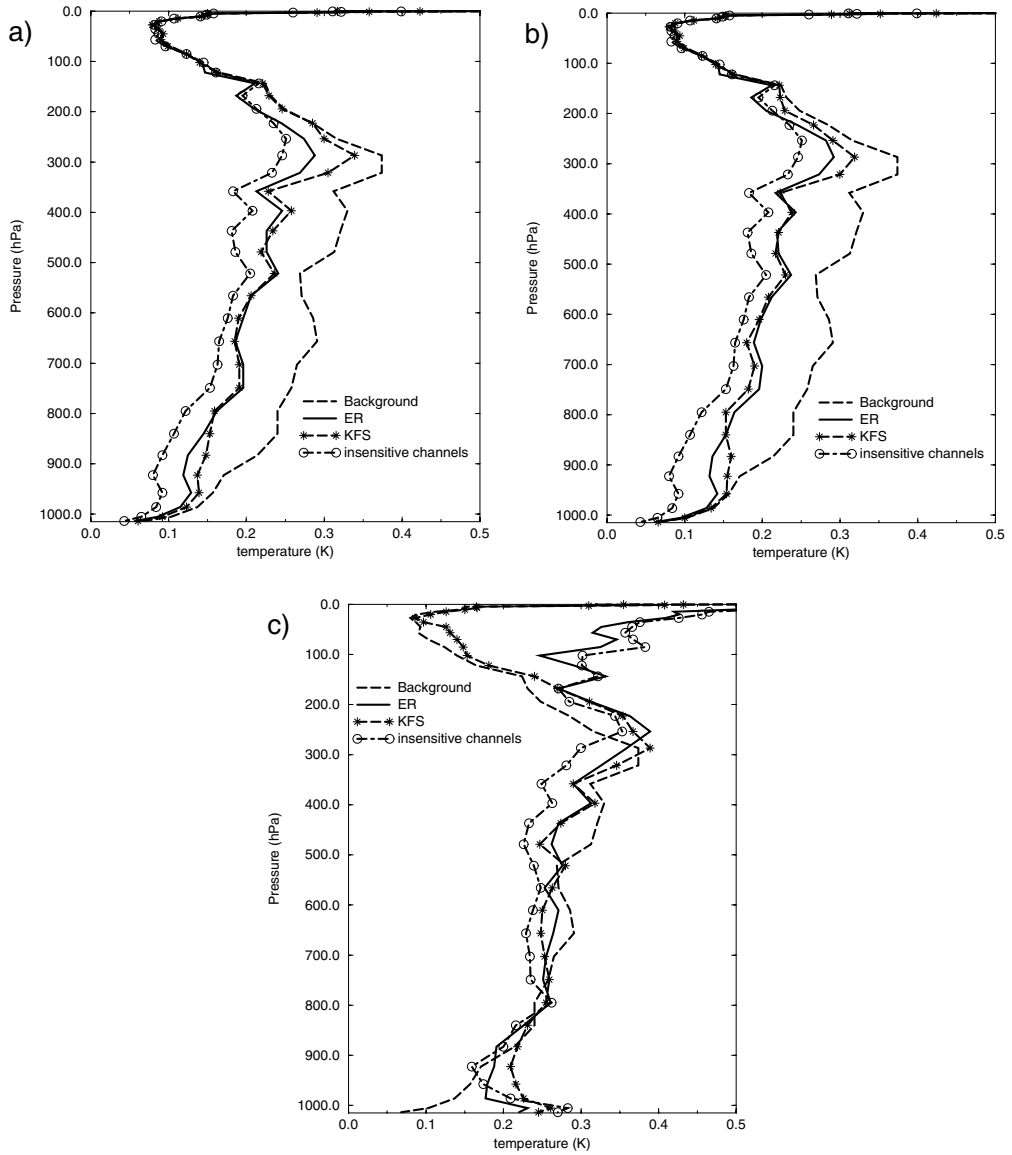


Figure 7. Root-mean-square temperature errors for two channel selection methods (ER and KFS) averaged over 142 profiles corresponding to a cloud-top level of 1000 hPa with respect to the specification of the background-error covariance matrix, together with the background errors and errors in assimilating all the channels not contaminated by cloud effects (insensitive channels): (a) good specification of the **B** matrix, (b) mis-specification of the **B** matrix during the channel selection process, and (c) mis-specification of the **B** matrix in the channel selection and the analysis. See Fig. 5 for explanation of the labels.

forecast errors projected on the 43 RT model levels). In this case the reference is the corrected 3D-Var analysis, and a noise derived from the KAE matrix is added in order to obtain the background state. Furthermore, a noise is added to the data simulated from the truth, in order to be able to discuss the results from a statistical point of view.

The resulting rms analysis errors are shown in Fig. 7. By construction, the rms of the background error corresponds quite well to the standard deviation of the **B** matrix shown in Fig. 4 because the bias is assumed to be close to zero. In the case of a good

TABLE 3. MEAN PERFORMANCES OF THE THREE CHANNEL SELECTION METHODS ACCORDING TO THE SPECIFICATION OF THE **B** MATRIX

B matrix method	Channel selection method	Degree of freedom for signal	Entropy reduction	Root-mean-square ratio
Key analysis errors	Entropy reduction	4.66	7.1	0.84
	Kalman-filter sensitivity	3.03	5.66	0.88
	Insensitive channels	6.23	12.35	0.78
Mis-specification of the **B** matrix during the channel selection	Entropy reduction	4.55	6.88	0.85
	Kalman-filter sensitivity	3.16	5.50	0.87
Mis-specification of the **B** matrix during the channel selection and the analysis	Entropy reduction	9.94	21.25	1.85
	Kalman-filter sensitivity	5.98	13.02	1.23
	Insensitive channels	11.94	29.46	2.08

specification of the background errors (Fig. 7(a)), the ER method is generally better than the KFS method in the troposphere. As a reference, the assimilation of all the channels not contaminated by the clouds (about 3000 channels, the number depending on each atmospheric profile) allows improvement of the rms analysis error by 0.03 K in the troposphere. The mis-specification of the **B** matrix in the channel selection process only (the KAE matrix is still used in the analysis) leads to similar results obtained by the use of the KAE matrix (Fig. 7(b) and Table 3). In fact, 240 channels are commonly shared by the two ER methods and 178 channels by the two KFS selections. The mis-specification of the background errors (Fig. 7(c)) in the channel selection and in the analysis leads to a rather poor analysis (which is not systematically better than the background). Although it is improved in the main sensitive layer by both methods (300–700 hPa), the addition of information in the high troposphere and near the surface degrades the analysis, as suggested by the rms error obtained by the assimilation of all the channels not contaminated by clouds. In an optimal case, the addition of information should improve the analysis. This is not the case here, which is a clear sign that our analysis is sub-optimal. Table 3 confirms this fact and, even though the ER or the DFS are larger than for the two previous analyses (a priori estimation of improvement upon the background), the rms ratio is larger in the third case. The large values of DFS and ER are erroneous because the equation used for the computation of **A** is wrong.

In conclusion, our results suggest the importance of a good specification of the **B** matrix for the analysis, but this is valid for all the methods; this result is in agreement with McNally (2000). However, for the channel selection process itself, a mis-specification of **B** does not lead to dramatic changes.

6. IMPACT OF A CONSTANT CHANNEL SET

Here one wants to study the impact of a non-optimal channel set. As in R02, a constant channel set has been built from the ER method. The 'constant' selection is computed as an average selection based on a set of representative mid-latitude situations. This set is part of the ECMWF atmospheric database (Chevallier *et al.* 2000) and forms a set of 170 vertical profiles of temperature, humidity, ozone together with values of surface pressure and surface temperature. The KAE matrix has been used as the **B** matrix during the channel selection process and a noise derived from the KAE matrix has been added in order to obtain the background state. Likewise, another constant channel set has been built from the climatological **B** matrix and a background state with noise derived

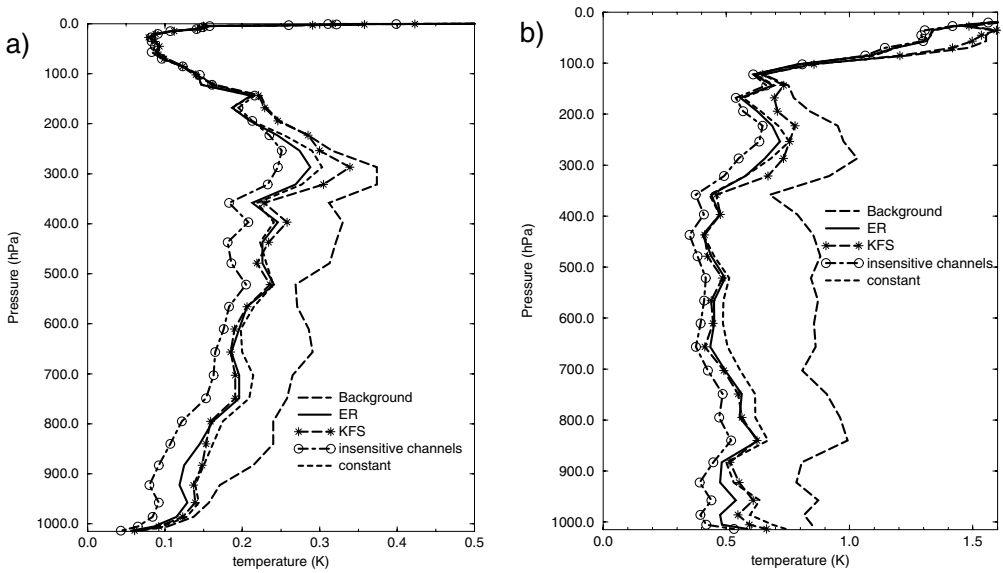


Figure 8. The root-mean-square temperature errors for four channel selection methods (ER, KFS, insensitive channels, and constant channel set—see text) averaged over 142 profiles corresponding to a cloud-top level of 1000 hPa with (a) a key analysis-error matrix and (b) a climatological background-error covariance matrix. See Fig. 7 for the an explanation of the labels.

TABLE 4. MEAN PERFORMANCES OF THE FOUR CHANNEL SETS DISCUSSED IN SECTION 6

	Channel selection method	Degree of freedom for signal	Entropy reduction	Root-mean-square ratio
Key analysis error	Entropy reduction	4.66	7.1	0.84
	Kalman-filter sensitivity	3.03	5.66	0.88
	Constant channel set	4.21	6.19	0.86
	Insensitive channels	6.23	12.35	0.78
Climatological **B**	Entropy reduction	9.96	21.26	0.82
	Kalman-filter sensitivity	5.89	12.99	0.92
	Constant channel set	9.24	19.22	0.83
	Insensitive channels	11.97	29.48	0.78

from this climatological **B** matrix. Both channel sets have been used for the analysis of the 142 profiles located in the sensitive areas.

Before using the channels in the assimilation, it was checked that they were not contaminated by clouds. The number of channels of the constant set actually used was close to 220 but varied with each profile. Figure 8 shows the results obtained with these two constant sets for the KAE matrix and the climatological **B** matrix. In both cases, a slight degradation is obtained in the troposphere in comparison with an optimal channel set. Moreover, the constant sets provide larger DFS or ER than the KFS set and a smaller rms ratio than the KFS set in both cases (Table 4). These results suggest that an optimal channel set is not necessarily required for the analysis, and that a constant set provides a good compromise between the computational cost and the quality of the analysis.

7. CONCLUSIONS

In this paper, the issue of reducing the number of IASI channels to be used in data assimilation has been specifically addressed in the context of sensitive areas and

for low-level clouds. In the first section, the cloud-top level in simulated IASI pixels for FASTEX sensitive areas was studied using the MAIA method with AVHRR imager observations. It was found that sensitive areas are usually covered by low-level clouds in their northern parts and by high-level clouds in their southern ones. These results corroborate the study of McNally (2002) who established a high correlation between the meteorologically sensitive areas and cloud cover produced by the ECMWF model.

As a result of this first study, the channel selection was then tackled in the context of low-level clouds (1000 hPa) for the eight FASTEX IOPs in order to retrieve information from the sensitive areas. The so-called iterative method of R02, based on the ER, has been compared with two methods derived from the targeting strategy consisting of adding optimal observations in sensitive areas in order to reduce the error in the analysis that can lead to large forecast errors. These latter techniques are called the sensitivity-to-observations method and the Kalman-filter-sensitivity method. In a first step, the three methods were compared for the retrieval of one sensitive scaled-gradient perturbation and it was shown that they provide similar results, even though the ER selection and the other methods share few channels. As the sensitivity-to-observation channel selection method is very expensive in terms of computer time, only the ER and KFS methods were then compared from a statistical point of view for profiles from the FASTEX IOPs with low-level clouds. The ER method seems robust, even when the profiles are located within sensitive areas. The impact of a good specification of the background-error covariance matrix and of a constant set has also been tested. Both channel selection methods are mainly independent of a good specification of the **B** matrix (although the analysis depends highly on it), and 80% to 60% of the channels are selected in common in both cases. The use of a constant channel set computed from an independent set of atmospheric profiles still leads to a significant improvement of the analysis compared to the background, even if this analysis error is slightly larger than the one obtained with an optimal channel set.

Our main conclusions are: the ER channel selection method is quite robust, even for the highly sensitive profiles studied here. We have shown that there was no large dependence on the background-error covariance matrix. In addition, the constant channel set deduced from this optimal channel selection can be computed from an independent atmospheric database, if it is representative of the profile air-mass. The ER channel selection method is very promising, as previous results had not been focused on such a characteristic set of profiles, a factor that is especially crucial for NWP. However, the first part of this study suggests that clouds are an important issue and that a radiative-transfer model taking into account the cloud effects is required in order to go further in channel-selection studies. One limitation of our work is that it took place out of the context of operational NWP assimilation, and that the results of cloud-level studies in sensitive areas are only valid for a limited number of synoptic situations represented by the eight FASTEX IOPs, and so may have no general applicability to other situations and seasons. Future work could involve the assimilation of IASI radiances in an NWP assimilation scheme. This would provide an opportunity to test the real impact of different channel selections on forecast scores for which gradient-based channel selections are optimized.

ACKNOWLEDGEMENTS

This work was partly supported by the Centre National d'Etudes Spatiales through the post-doctorate fellowship position of Nadia Fourrié. The authors wish to thank Gwenaëlle Hello (Centre National de Recherche Météorologique) who ran the computation of the cost-function gradients for the determination of the sensitive areas and

the calculation of the optimal perturbations of the gradient to be applied to the 3D-Var analyses. Tiphaine Labrot and Lydie Lavanant (Centre de Météorologie Spatiale) are also acknowledged for having provided the cloud determination using AVHRR observations interpolated onto IASI pixels for the Halifax area. The authors thank Jean Pailleux, Gérald Desroziers and Alexis Doerenbecher for helpful discussions and suggestions.

REFERENCES

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

Baker, N. L. and Daley, R. 2000 Observation and background adjoint sensitivity in the adaptive observation targeting problem. *Q. J. R. Meteorol. Soc.*, **126,** 1431–1454

Bergot, T. 1999 Adaptive observations during FASTEX: A systematic survey of upstream flights. *Q. J. R. Meteorol. Soc.*, **125,** 3271–3298

Bergot, T. and Doerenbecher, A. 2002 Study on the optimization of the deployment of targeted observations using adjoint-based methods. *Q. J. R. Meteorol. Soc.*, **128,** 1689–1712

Cammas, J.-P., Pouponneau, B., Desroziers, G., Santurette, P., Joly, A., Arbogast, P., Mallet I., Cagnaux, G. and Mascart, P. 1999 FASTEX IOP 17 cyclone: Introductory synoptic study with field data. *Q. J. R. Meteorol. Soc.*, **125,** 3393–3314

Chevallier, F. , Chédin, A., Chéruy, F. and Morcrette, J.-J. 2000 TIGR-like atmospheric profile databases for accurate radiative flux computation. *Q. J. R. Meteorol. Soc.*, **126,** 777–785

Collard, A. D. 1998 'Notes on IASI performance'. NWP Technical Report 253, Met Office, FitzRoy Road, Exeter, Devonshire EX1 3PB, UK

Desroziers, G., Hello, G. and Thépaut, J.-N. 2003 A 4D-Var reanalysis of the FASTEX experiment. *Q. J. R. Meteorol. Soc.*, **128,** 1301–1315

Doerenbecher, A. and Bergot, T. 2001 Sensitivity to observations applied to FASTEX cases. *Nonlinear Proc. Geophys.*, **8**(6)**,** 467–481

Fourrié, N. and Thépaut, J.-N. 2003 Evaluation of the AIRS near-real-time channel selection for application to numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **128,** 2425–2439

Hello, G. 2002 'Prise en compte de la dynamique associée aux dépressions des latitudes moyenne dans la détermination des conditions initiales des modèles météorologiques'. PhD Thesis, Université Paul Sabatier, Toulouse III, France

Hello, G. and Bouttier, F. 2001 Using adjoint sensitivity as a local structure function in variational assimilation. *Nonlinear Proc. Geophys.*, **8**(6)**,** 347–355

Hello, G., Lalaurette, F. and Thépaut, J.-N. 2000 Combined use of sensitivity information and observations to improve meteorological forecasts: A feasibility study applied to the 'Christmas storm' case. *Q. J. R. Meteorol. Soc.*, **126,** 621–647

Joly, A., Browning, K. A., Bessemoulin, P., Cammas, J.-P., Caniaux, G., Chalon, J.-P., Clough, S. A., Dirks, R., Emanual, K. A., Eymard, L., Gall, R., Hewson, T. D., Hildebrand, P. H., Jorgensen, D., Lalaurette, F., Langland, R. H., Lemaître, Y., Mascart, P., Moore, J. A., Ola, P., Persson, 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

Klinker, E., Rabier, F. and Gelaro, R. 1998 Estimation of the key-analysis errors using the adjoint technique. *Q. J. R. Meteorol. Soc.*, **124,** 1909–1933

Lavanant, L., Legléau, H., Derrien, M., Levasseur, S., Monnier, G., Ardouin, L., Brunel, P. and Bellec, B. 1999 'AVHRR cloud mask for sounding applications'. In Proceedings of the tenth international TOVS study conference, 27 January to 2 February 1999, Boulder, Colorado, USA

Matricardi, M. and Saunders, R. 1999 A fast radiative transfer model for Infrared Atmospheric Sounding Interferometer radiances. *J. Appl. Optics*, **38,** 5679–5691

McNally 2000 Estimates of short-range-forecast temperature error correlations and the implications for radiance-data assimilation. *Q. J. R. Meteorol. Soc.*, **126,** 361–373

2002 A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Q. J. R. Meteorol. Soc.*, **128,** 2551–2556

Prunet, P., Thépaut, J.-N. and Cassé, V. 1998 The information content of clear-sky IASI radiances and their potential for numerical weather prediction. *Q. J. R. Meteorol. Soc.*, **124,** 211–241

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. 1996 Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.*, **122**, 121–150

Rabier, F., Fourrié, N., Chafaï D. and Prunet, P. 2002 Channel selection methods for infrared atmospheric sounding interferometer radiances. *Q. J. R. Meteorol. Soc.*, **128,** 1011–1027

Rodgers, C. D. 1996 'Information content and optimization of high spectral resolution measurements'. Pp. 136–147 in *Optical spectroscopic techniques and instrumentation for atmospheric and space research II*, SPIE Vol. 2830. International Society for Optical Engineering, PO Box 10, Bellingham, Washington, 98227-0010, USA

2000 *Inverse methods for atmospheres: Theories and practice*. World Scientific Publications, Singapore

Sanders, F. and Gyakum, J. R. 1980 Synoptic–dynamic climatology of a 'bomb'. *Mon. Weather Rev.*, **108,** 1589–1606

Smith, W. L., Woolf, H. M., Hayden, C. M., Wark, D. Q. and McMillin, L. M. 1979 The TIROS-N Operational Vertical Sounder. *Bull. Am. Meteorol. Soc.*, **60**(10)**,** 1177–1187

Assimilation of AIRS Radiances Affected by Mid- to Low-Level Clouds

THOMAS PANGAUD, NADIA FOURRIE, AND VINCENT GUIDARD

CNRM-GAME, Météo-France, and CNRS, Toulouse, France

MOHAMED DAHOUI

European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom

FLORENCE RABIER

CNRM-GAME, Météo-France, and CNRS, Toulouse, France

(Manuscript received 30 March 2009, in final form 10 June 2009)

ABSTRACT

An approach to make use of Atmospheric Infrared Sounder (AIRS) cloud-affected infrared radiances has been developed at Météo-France in the context of the global numerical weather prediction model. The method is based on (i) the detection and the characterization of clouds by the CO_2-slicing algorithm and (ii) the identification of clear–cloudy channels using the ECMWF cloud-detection scheme. Once a hypothetical cloud-affected pixel is detected by the CO_2-slicing scheme, the cloud-top pressure and the effective cloud fraction are provided to the radiative transfer model simultaneously with other atmospheric variables to simulate cloud-affected radiances. Furthermore, the ECMWF scheme flags each channel of the pixel as clear or cloudy. In the current configuration of the assimilation scheme, channels affected by clouds whose cloud-top pressure ranges between 600 and 950 hPa are assimilated over sea in addition to clear channels. Results of assimilation experiments are presented. On average, 3.5% of additional pixels are assimilated over the globe but additional assimilated channels are much more numerous for mid- to high latitudes (10% of additional assimilated channels on average). Encouraging results are found in the quality of the analyses: background departures of AIRS observations are reduced, especially for surface channels, which are globally 4 times smaller, and the analysis better fits some conventional and satellite data. Global forecasts are slightly improved for the geopotential field. These improvements are significant up to the 72-h forecast range. Predictability improvements have been obtained for a case study: a low pressure system that affected the southeastern part of Italy in September 2006. The trajectory, intensity, and the whole development of the cyclogenesis are better predicted, whatever the forecast range, for this case study.

1. Introduction

The Atmospheric Infrared Sounder (AIRS) on board the *Aqua* satellite (Aumann et al. 2003) belongs to a new generation of advanced satellite sounding instruments that provide information about atmospheric temperature and humidity profiles with a spectral resolution far exceeding that of previous sounders such as the High Resolution Infrared Radiation Sounder (HIRS). The Infrared Atmospheric Sounder Interferometer (IASI) on board the *MetOp* satellite (Cayla 2001) with a better spectral resolution than AIRS, also belongs to this new generation of advanced satellite sounding instruments. These highly informative observations are used to improve numerical weather prediction (NWP) analysis and forecast accuracy.

A large amount of measurements from advanced infrared sounders are affected by clouds [i.e., 90% according to Fourrié and Rabier (2004)] and cloud-contaminated observations are currently rejected by the data assimilation system due to deficiencies in the representation of cloud processes. Since clouds affect the infrared satellite observations, clouds have to be detected before data are assimilated. Indeed, unfiltered cloud-affected observations can have a negative impact on the quality of NWP analyses if they are assimilated as clear observations.

Corresponding author address: Thomas Pangaud, Météo-France, 42 avenue Gaspard, Coriolis, Toulouse CEDEX 31057, France.
E-mail: thomas.pangaud@cnrm.meteo.fr

DOI: 10.1175/2009MWR3020.1

Several cloud-detection schemes have been used: the most conservative method consists in rejecting all fields of view classified as cloudy (English et al. 1999) and then only to focus on completely clear locations. This technique leads to a yield in terms of exploited soundings of about 5% (Pavelin et al. 2008). Another technique developed at the European Centre for Medium-Range Weather Forecasts (ECMWF) by McNally and Watts (2003), hereinafter called the ECMWF scheme, is based on the detection of cloud-affected channels and consists in using only channels that are thought to be unaffected by clouds, inside clear or cloud-affected soundings. The yield in terms of exploited sounders is higher than by the previous method and the benefit in terms of additional assimilated channels is far from being negligible (McNally and Watts 2003). However, only a small proportion of the available soundings is being exploited, especially for tropospheric channels: roughly 15% for a stratospheric channel and 8% for an upper-tropospheric channel. A third alternative approach, the so-called cloud-clearing method, consists in simulating a clear observation by using multiple fields of view and assuming that only cloud fraction changes between these adjacent fields of view (Eyre and Watts 1987; Li et al. 2005), but the assumption of homogeneous cloud characteristics is not verified in the majority of cases (Pavelin et al. 2008) and the successful exploitation of cloud-cleared radiances by NWP is thus not obvious.

Nevertheless, the assimilation of only clear observations can lead to a poor description of some particular atmospheric situations such as regions with high humidity (e.g., low-latitude regions). Finally, it is well known that the sensitive regions, where cyclogenesis start to develop, are often cloudy (McNally 2002). The underexploitation of advanced sounders and the wish to better describe the atmosphere and to limit forecast error growth in sensitive regions motivate our research efforts to assimilate cloud-affected radiances. The observed radiance spectra that contain temperature and humidity information can also provide information on clouds. Two main methods are actually investigated in NWP centers to deal with cloud-affected infrared radiances: the first approach is based on the insertion of a physical cloud scheme into the observation operator. An approach based on a diagnostic cloud scheme has been tested by ECMWF, for instance, with an encouraging impact on temperature and humidity fields (Chevallier et al. 2004). This diagnostic cloud scheme estimates cloud parameters from the model variables: a cloud is assumed here to be a multilayer gray body. A radiative transfer model (RTM) is used as an observation operator, then it simulates cloud-affected radiances. The second approach is based on the use of the cloud-top pressure (CTP) and the effective cloud fraction N_e calculated by a cloud-characterization algorithm. These parameters are then provided to the RTM to simulate cloud-affected radiances. Several techniques are commonly used to retrieve cloud parameters from radiance measurements, such as the minimum residual method (Eyre and Menzel 1989). Pavelin et al. (2008) have demonstrated that a one-dimensional variational data assimilation (1D-Var) processing of CTP and N_e before being used by the RTM results in a more accurate cloud parameter retrieval. This approach constitutes the operational configuration at the Met Office to assimilate AIRS cloud-affected radiances.

For the purpose of this study, cloud parameters are estimated with the second above-described approach. This method has been chosen because preliminary cloud-affected radiance assimilation experiments have shown that this method leads to promising impacts on analysis and forecast (Dahoui 2006). It also presents the vantage to provide a better computational efficiency. The cloud parameter retrieval scheme used in this study is the CO_2-slicing method based on radiative transfer principles (Chahine 1974; Menzel et al. 1983). The RTM used is a fast RTM, the 8.5 version of the Radiative Transfer for Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (RTTOV), which is the operational version of RTTOV at Météo-France (Saunders et al. 2002). In this framework, once a hypothetical cloud-affected pixel is detected by the CO_2-slicing scheme and its parameters are retrieved, the cloud-affected radiances are simulated by RTTOV. Furthermore, channels of this pixel are flagged to be clear or cloudy by the ECMWF scheme. Finally, cloud-affected channels whose CTP is included between 600 and 950 hPa, and clear channels are assimilated over the sea in the French NWP global model, Action de Recherche Petite Echelle Grande Echelle (ARPEGE) through a four-dimensional variational data assimilation (4D-Var) scheme. Our assimilation period, which includes among others a strong mesoscale storm in the Mediterranean Sea, begins at 0000 UTC 1 September 2006 and ends at 1800 UTC 30 September 2006. At that time, the *MetOp* satellite was not launched yet and IASI data were thus not available. For this reason, the processing of IASI has not been mentioned in this paper and we have just focused on AIRS data processing.

The plan of this paper is as follows. Section 2 focuses on a brief description and evaluation of two cloud-detection schemes applied to AIRS: the ECMWF and the CO_2-slicing scheme. The method used to directly assimilate cloud-affected radiances in the 4D-Var assimilation scheme of ARPEGE is presented in section 3. Experiments assimilating clear and cloudy AIRS radiances have

been performed. Section 4 focuses on these experiments in terms of AIRS observation usage and characterization. Results in terms of impact on the quality of the analysis and on the skill of the forecasts are discussed in section 5. Finally the impact on the forecast is further evaluated in a case study of a strong mesoscale storm in section 6. Concluding remarks follow in section 7.

2. Evaluation of AIRS cloud-detection schemes

a. AIRS cloud-detection schemes

1) ECMWF SCHEME

The ECMWF scheme (McNally and Watts 2003) aims at detecting clear channels within a measured spectrum rather than the detection of totally clear pixels. If the background spectrum (defined as the spectrum computed from the best available background estimate of the atmospheric state) is close enough to the true state of the atmosphere, the cloud signature is identified by the first-guess departure of the observed spectrum from clear-sky background values. The term ''background'' represents here a short-range NWP forecast (6 h) from the previous analysis. In this study, a set of 319 channels, chosen among the 324 available ones in real time to NWP users, is used. Channels are first reordered into a vertically ranked space according to the lower tail of their respective weighting function. The ranking consists in assigning to each channel, a pressure level (in RTTOV coordinates) at which the radiation effect of a one-layer blackbody is less than 1%. A low-pass filter is then applied to the ranked departures to reduce the instrument noise and the cloud emissivity effect. Finally, a search for the channel at which a monotonically growing departure can first be detected, determines the first significant cloud contamination. Once this channel is found, all other less sensitive channels are flagged cloud free and the more sensitive ones are flagged cloudy. A pixel is declared clear if all of the 319 channels are flagged cloud free.

2) CO_2-SLICING SCHEME

The CO_2-slicing method (Chahine 1974; Menzel et al. 1983), based on radiative transfer principles, is widely used to retrieve CTP and N_e. Here N_e conceptually represents the product of the geometrical cloud fraction and the gray body emissivity of the cloud. This method uses a simplistic cloud model: the cloud is considered as a single layer of opaque or semitransparent thin cloud with a homogeneous emissivity. The algorithm uses observed radiances of a set of 124 AIRS channels selected in the CO_2 absorption band (between 649.612 and 843.913 cm^{-1}), which is very sensitive to clouds. For each AIRS pixel, and each channel in the set, the following function is calculated:

$$F_{k,p} = \frac{(R^{k}_{\text{clear}} - R^{k}_{\text{meas}})}{(R^{k_{\text{ref}}}_{\text{clear}} - R^{k_{\text{ref}}}_{\text{meas}})} - \frac{(R^{k}_{\text{clear}} - R^{k,p}_{\text{cld}})}{(R^{k_{\text{ref}}}_{\text{clear}} - R^{k_{\text{ref}},p}_{\text{cld}})}, \qquad (1)$$

where p is the pressure level number, k is the channel in the CO_2 band, k_{ref} represents the reference window channel (979.1279 cm^{-1}), R^{k}_{meas} is the measured radiance for channel k, R^{k}_{clear} is the simulated clear radiance for channel k, and $R^{k,p}_{\text{cld}}$ represents the simulated blackbody radiance for channel k at the cloud level p. The cloud-top pressure level assigned to each channel k is the pressure level $p_{c,k}$, which minimizes the function $F_{k,p}$. Before ascertaining the CTP of a hypothetic cloud, a filter that distinguishes channels with δTB*s* (difference between observed brightness temperature and simulated brightness temperature) lower than the radiometric noise is applied to the algorithm. If all channels are filtered out, the pixel is flagged clear. If the pixel is cloud contaminated, the CTP is then calculated by the following expression:

$$\text{CTP} = \frac{\sum p_{c,k} w_k^2}{\sum w_k^2}, \qquad (2)$$

where $w_k = \delta F_{k,p}/\delta \ln p$ is the derivative of the cloud pressure function.

The effective cloud fraction is obtained for the reference window channel by the following expression:

$$N_e = \frac{(R^{k_{\text{ref}}}_{\text{clear}} - R^{k_{\text{ref}}}_{\text{meas}})}{(R^{k_{\text{ref}}}_{\text{clear}} - R^{k_{\text{ref}},p_c}_{\text{cld}})}. \qquad (3)$$

If the algorithm produces a retrieved N_e smaller than 0.1, the pixel is flagged clear. Both above-calculated parameters are used by RTTOV to simulate cloud-affected radiances.

b. Results and discussion

The evaluation of both cloud-detection schemes is fundamental as unfiltered cloud-affected observations can have a detrimental impact on the quality of the analysis if they are assimilated as clear. For this study, a cloud-characterization product based on the Moderate Resolution Imaging Spectroradiometer (MODIS) data has been used as a reference. The MODIS imager is a key instrument on board the Earth Observation System satellites. We have used data product collected from the *Aqua* platform from the Interactions Clouds Aerosols Radiations Etc. (ICARE) center (more information available online at http://www.icare.univ-lille1.fr/archive/archive.php?dir=MODIS/MYD06_L2/), which

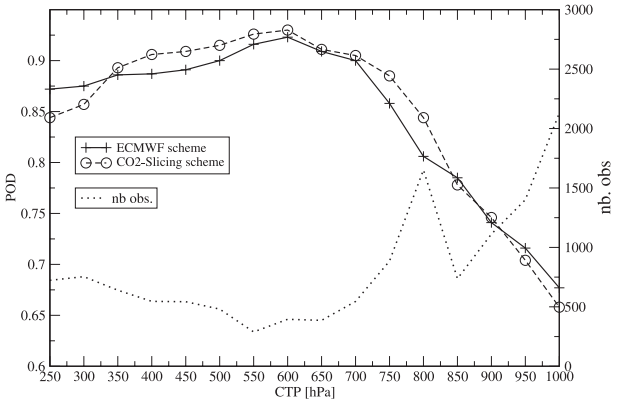


FIG. 1. Efficiency of the cloud detection according to the retrieved cloud-top pressure from MODIS. Evaluation period lasted from 1 to 10 Sep 2006. The CO_2 slicing is represented by a dashed line with circles and the ECMWF scheme is represented by a solid line with crosses. The number of observations is represented by the dotted line.

are produced at an horizontal resolution of 5 km at nadir and cover a 5-min time interval. Because of downloading resource limitations, the comparison is limited to the Atlantic region. The validation is performed within a 10-day period: from 1 to 10 September 2006. A total of 15 706 AIRS pixels have been processed. As the evaluation of both cloud-detection schemes has already been performed (Dahoui et al. 2005) and results are briefly discussed in this paper.

Performances for both schemes are similar in terms of detection of cloudy and clear pixels. These results confirm previous studies (Dahoui et al. 2005). Figure 1 highlights the accuracy of the cloud detection for the CO_2-slicing and the ECMWF schemes for several ranges of CTP inferred from MODIS. The detection of medium clouds (between 400 and 800 hPa) delivers the best results. In contrast, the detection of low-level and high clouds (CTP $>$ 800 hPa and CTP $<$ 400 hPa, respectively) is not as efficient as for medium clouds for both schemes. For low-level clouds, this must be due to the similarity of the measured signal from clear-sky and low-cloud scenes in some particular situations which make the detection of these low clouds more difficult. Concerning high clouds, the problem is mostly due to the insensitivity of the sounding-based method to thin clouds (identical thresholds are applied for the detection of each type of clouds). Furthermore, clouds whose CTP varies between 250 and 300 hPa and those whose CTP varies between 900 and 1000 hPa are slightly better detected by the ECMWF scheme. Clouds whose CTP varies between 300 and 850 hPa are slightly better detected by the CO_2-slicing scheme.

The retrieved CTP from CO_2-slicing exhibits a correlation of about 0.79 with the CTP inferred from MODIS. Correlation between the N_e retrieved from the CO_2-slicing scheme and the ones inferred from MODIS is quite low for each range of N_e (about 0.51). This low correlation may first be explained by the presence of clouds distributed in the vertical (which are assumed to be single-layer clouds by the RTTOV) whose radiative effect is comparable to a cloud with a higher cloud fraction located at a single level. Moreover, the relative bad detection of low clouds that exhibit a low contrast with the surface in some particular situations may also account for this low correlation. These CTP and N_e correlations are of the same order of magnitude as those found by Pavelin et al. (2008) with the minimum residual method: between 0.67 and 0.81 for the CTP and between 0.46 and 0.78 for the N_e.

3. Framework of the study

a. Methods to determine background cloud parameters

The observation operator used here for the assimilation of cloud-contaminated radiances is the 8.5 version of RTTOV, complemented by horizontal and vertical interpolations. It simulates infrared and microwave radiances observed by satellite sounders (Saunders et al. 2002). For each channel, the optical depth is obtained from linear regressions based on RTTOV input variables: temperature and humidity profiles and surface data. Although RTTOV has initially been designed to be used in clear-sky conditions, it is able to simulate cloud-contaminated radiances. The radiative transfer in cloudy conditions can be performed by two different techniques:

- The cloud is assumed to be a single-layer blackbody of negligible depth that CTP has been provided to the RTM. The cloud-affected radiance is calculated using the following expression:

$$R_\nu^{\mathrm{cld}}(\theta) = B_\nu[T(z_{\mathrm{cld}})]\tau_\nu(z_{\mathrm{cld}},\theta) + \int_{z_{\mathrm{cld}}}^{\infty} B_\nu[T(z)][d\tau_\nu(z,\theta)\, dz]/dz, \qquad (4)$$

where $R_\nu^{\mathrm{cld}}(\theta)$ is the cloudy monochromatic radiance of frequency ν going away from the top of the cloud toward space with an incidence angle θ, $\tau_\nu(z_{\mathrm{cld}}, \theta)$ represents the transmittance between the cloud top and space, $T(z_{\mathrm{cld}})$ represents the cloud-top temperature, $T(z)$ is the temperature of an atmospheric layer at altitude z, and B_ν represents the Planck function. Equation (4) defines the overcast radiance. In reality, most of the pixels are partially cloud contaminated.

These kinds of pixels are defined by a linear blending of overcast and clear radiances using the following equation:

$$R_\nu(\theta) = (1 - N)R_\nu^{\mathrm{clr}}(\theta) + NR_\nu^{\mathrm{cld}}(\theta), \qquad (5)$$

where $R_\nu^{\mathrm{clr}}(\theta)$ represents the clear-sky component of the radiance and N represents the cloud cover.

- A more realistic multilayer parameterization of clouds is in the RTTOVCLD module: clouds are defined by several vertical levels and for different cloud water phases (liquid or ice) from vertical profiles provided to the RTM (temperature profiles, humidity, cloud cover, liquid cloud water, and cloud ice) by a physical cloud scheme. Clouds are assumed to be multilayer gray bodies.

We made the choice to use the first method although the use of a physical cloud scheme (second method) is potentially valuable (Chevallier et al. 2004). Unfortunately, the physics of the model in its current state is far from the required accuracy. According to Dahoui (2006), the first method provides a better computational efficiency and leads to promising impacts on analysis and forecast. The cloud emissivity is assumed to be equal to 1, which is acceptable for thick opaque clouds but not valid for thin clouds like cirrus. The latter are high clouds whose impact is not taken into account in our assimilation scheme (as only radiances affected by mid- to low-level clouds are treated here). Thus, this assumption has no impact in the simulation of cloud-affected radiances in this study context.

b. Experiment setup

The NWP model used in this work is ARPEGE (Courtier et al. 1991), which is the Météo-France operational global model. It uses a stretched grid with a horizontal resolution of 23 km over France and 133 km over the antipodes. The assimilation scheme is a 4D-Var assimilation (Courtier et al. 1994; Rabier et al. 2000). It uses the following data sources: land surface stations, sea surface station (drifting buoys, ship reports, etc.), aircraft data, in situ sounding data (radiosonde and pilot balloon reports, etc.), wind profiler radar data, Global Positioning System (GPS) ground-based data, geostationary satellite winds (atmospheric motion vectors), and polar-orbiting satellite radiances and winds. Radiances from polar-orbiting satellites provide ARPEGE with roughly 50% of the total volume of assimilated data. AIRS data represented roughly 65% of the volume of polar-orbiting satellite data and more than 30% of the total volume of assimilated data. All these data sources are subject to various quality control checks and suspect ones are rejected. The quality control applied to AIRS is based on differences between observations and background values, with a rejection of the observation if the difference is larger than a given threshold. Moreover, a geographical thinning of 250 km is applied to all radiance data. Satellite data are bias corrected with an adaptive variational method (viz., VarBC), using geometric and flow-dependent predictors to remove systematic errors in satellite radiance data (Dee 2004; Auligné et al. 2007). In this study, three types of predictors were used: a global offset applied to each channel, a scan angle correction based on the nadir-viewing angle (to the power of 1, 2, and 3) and an air mass correction (predictors are the four atmospheric thicknesses 1000–300, 200–50, 10–1, and 50–5 hPa). We have used the operational subset of 54 channels located in the temperature longwave band chosen among the 324 channels available on the Global Transmission System. This subset contains 19 stratospheric channels with weighting functions peaking from 56 to 85 hPa, 12 channels with weighting functions peaking from 102 to 222 hPa sampling the tropopause, and 23 upper-tropospheric channels with weighting-functions peaking from 253 to 478 hPa (Fig. 2). These channels are only assimilated over seas. The assimilation period of this study covers a full month, from 1 to 30 September 2006. In this approach, once a cloud-affected pixel is detected by the CO_2-slicing scheme, cloud parameters (CTP and N_e) are retrieved from infrared radiance measurements out of the observation operator and are then directly used as inputs of RTTOV, which will then simulate the cloud-affected spectrum. Thus, N_e and CTP are determined at the beginning of the assimilation cycle with no adjustments during the minimization.

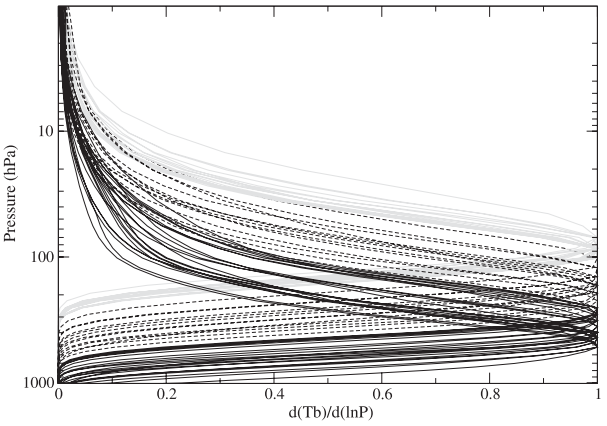


FIG. 2. Weighting functions of the 54 assimilated AIRS channels. Solid black lines represent tropospheric peaking channels, dashed lines represent tropopause peaking channels, and solid gray lines represent stratospheric peaking channels.

Furthermore, the ECMWF scheme identifies each channel of the pixel as clear or cloudy. Finally, channels affected by clouds whose CTP ranges between 600 and 950 hPa are assimilated together with clear ones and all other cloud-contaminated channels are rejected from the assimilation scheme. Two assimilation experiments have been run to test the impact of the direct assimilation of cloud-contaminated radiances: the first one is a reference experiment that only assimilates clear AIRS radiances with the others above-described data. This reference experiment is called REF in the following study. The second one assimilates cloud-affected radiances on top of clear radiances and other above-described data sources. In the following study this experiment is called EXP. The adaptive variational bias correction from REF is applied to both clear- and cloud-affected radiances in EXP. An experiment similar to EXP (figures not shown) has been run with its own adaptive bias correction but differences with EXP remain small (e.g., no drift in number of active assimilated observations). To disentangle effects of assimilating cloud-affected radiances from those of bias correction variation, only EXP is discussed in this paper.

c. Application for mid- to low-level clouds

As previously seen, cloud-affected radiances are assimilated here for clouds whose CTP is included between 600 and 950 hPa. On one hand, the rejection of clouds situated higher than 600 hPa is motivated by the assimilation of tropospheric channels (up to 478 hPa) into our assimilation scheme and only clouds situated lower than peaks of the weighting function are selected. Indeed, the assimilation of clouds situated higher than the weighting function peak of a given channel can lead to large biases and root-mean-square (RMS) errors for both background and analysis as shown by Pavelin et al. (2008). On the other hand, it has been shown in section 2b that the detection of low-level clouds ($800 < \text{CTP} < 950$) was not as efficient as the detection of higher clouds and their assimilation could thus appear delicate. A sensitivity study has been performed to check whether or not the AIRS spectrum was affected by the nondetection of these low clouds. It consists in comparing EXP (which simulates cloud-contaminated radiances) and REF (which only simulates clear radiances) averaged innovations (observations minus simulated radiances using the background state of the atmosphere) with respect to the channel number at a given CTP range (800–950, 600–800, 400–600, and 200–400 hPa) for all the observations (Fig. 3). The AIRS spectrum is almost not affected by the simulation inside RTTOV of clouds whose CTP is included between 800 and 950 hPa (Fig. 3a). Difficulties of cloud-detection schemes to detect some low clouds have thus a negligible impact on the background simulation, contrary to other level clouds where the simulated AIRS spectrum is more affected by the simulation of clouds inside RTTOV (Figs. 3b–d). The assimilation of such low-level clouds is also motivated by the large number of retrieved CTP between 900 and 950 hPa (Fig. 4). However, very low clouds ($\text{CTP} > 950$ hPa) were rejected because the huge number of these cloud types (Fig. 4) does not seem physically plausible. The CO_2 slicing is obviously not able to accurately treat these very low clouds, most probably because of their weak impact on the signal measured by the sounder.

Our choice to assimilate cloud-affected channels whose CTP is included between 600 and 950 hPa appears thus to be a good compromise between, on one hand, the rejection of doubtful observations and those potentially spoiling the analysis, and, on the other hand, the assimilation of a nonnegligible volume of additional cloud-affected observations.

4. Impact on AIRS observations

a. Number and localization of additional assimilated observations

Figure 5 shows the number of active observations for each assimilated channel for the first assimilation window (0000 UTC 1 September 2006). This assimilation window is particularly interesting because the background used is the same for EXP and REF. The number of active observations is roughly the same for EXP and REF until channel 138, which is one of the first channels whose weighting function is pointing into the tropopause. Previous channels (from 15 to 117) are stratospheric channels, thus seldom cloud contaminated, and exhibit the same behavior for both experiments. Channels 138 to 251 are upper-tropospheric channels and are thus potentially contaminated by clouds. This hypothetical contamination leads to a rejection of these contaminated channels in REF whereas some of these contaminated channels are assimilated in EXP. The lower the channel points in the atmosphere, the larger the probability of a cloud contamination and the larger is the difference between the assimilated observation numbers for the two experiments. Furthermore, the number of active clear observations is lower for EXP than for REF for each tropospheric channel. This may be due to the geographical thinning of the assimilated pixels. This thinning is based on the maximum number of nonrejected channels (monitored and assimilated channels) per profile. REF keeps the profile with the maximum number of clear channels as cloud-contaminated channels are rejected from the assimilation. It is the opposite for EXP that has a maximum number of nonrejected channels in the profile in the

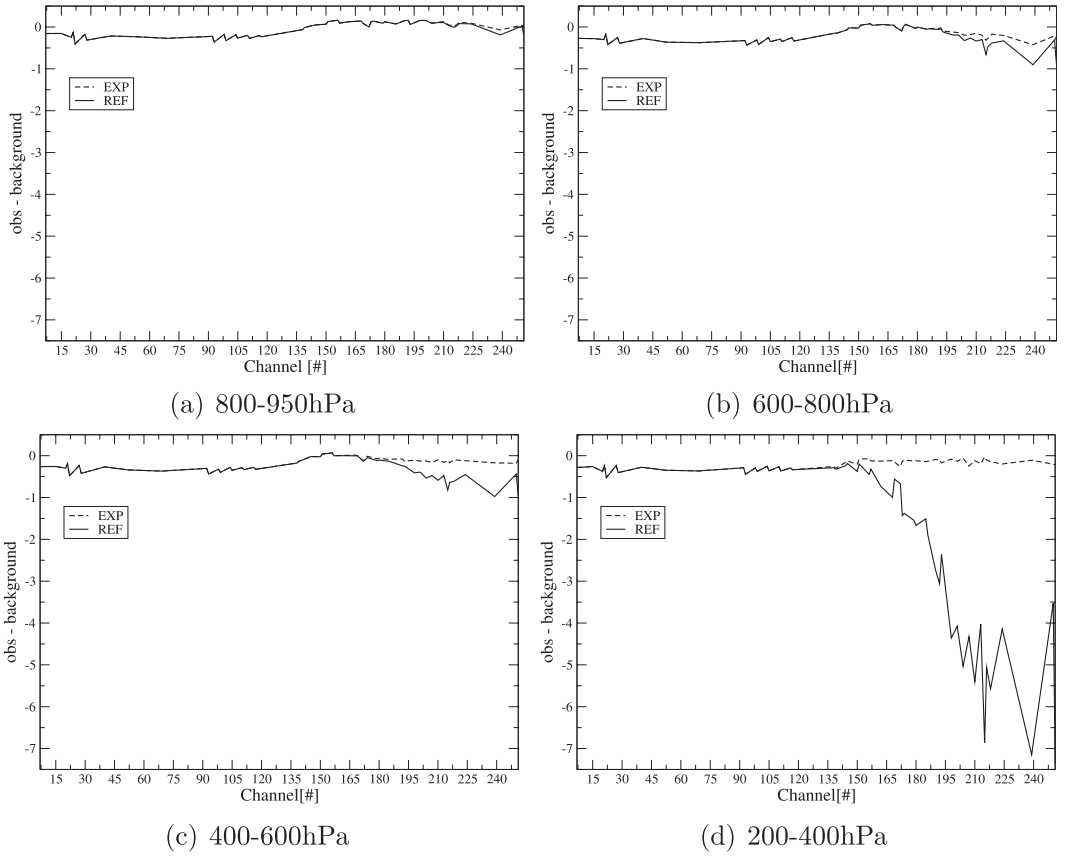


FIG. 3. Averaged innovations (observations−background) after bias correction of all observations for the 54 assimilated channels 2006 according to cloud-top pressure for 1 Sep. Solid line is for REF and dashed line is for EXP.

presence of a cloud whose CTP is between 600 and 950 hPa (in that case, all channels are nonrejected, except for the gross error).

Figures 6a,b, respectively, exhibit the number of assimilated pixels and assimilated channels (many channels are potentially assimilated in a given pixel) for both experiments with respect to their latitudinal position and for the whole study period. Most of the active pixels and channels are situated in the Southern Hemisphere, which is consistent with the assimilation of overseas observations (the oceanic surface is larger in the Southern Hemisphere than in the Northern Hemisphere). EXP assimilates between 1% and 4% more pixels than REF (Fig. 6a) with a slightly larger rate at mid- to high latitudes. The 250-km geographical thinning restricts these rates that would be higher if the geographical density of assimilated radiances were higher. The difference of assimilated channels between EXP and REF in mid- to high latitudes (Fig. 6b) is much larger than the difference of assimilated pixels (Fig. 6a): on average, 10% of additional channels against 3.5% of additional pixels. These differences are not visible for low latitudes (1.5% of additional pixels and channels on average). Thus, even if the assimilation of cloud-affected radiances yields additional active pixels to ARPEGE with quite a uniform repartition on the globe, the additional information is

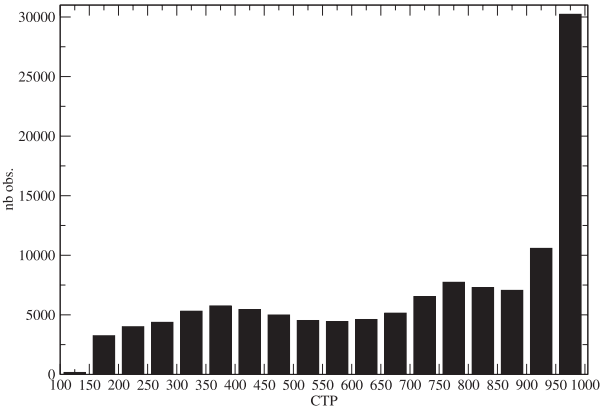


FIG. 4. Distribution of cloud-top pressure retrieved from CO_2 slicing on AIRS cloud-affected pixels ($N_e > 0.1$) used in the validation period (1–10 Sep 2006).

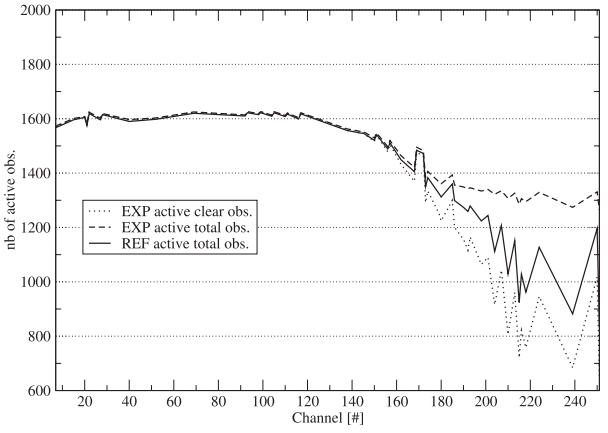


FIG. 5. Number of active channels at 0000 UTC 1 Sep 2006 with respect to channel number. The total number of active channels of REF is represented by the solid line, the total number of active channels of EXP is represented by the dashed line, and the number of active clear channels of EXP is represented by the dotted line.

much larger for mid- to high latitudes than for low latitudes. This could be due to the nonassimilation of high-level clouds (CTP $<$ 600 hPa) that are most frequent in the tropics [see the International Satellite Cloud Climatology Project (ISCCP) Web site online at http://isccp.giss.nasa.gov].

The daily evolution of the number of assimilated observations has also been studied at 1200 UTC (not shown) for assimilated channels peaking at the tropopause, upper troposphere, and midtroposphere in order to check whether the quality control was stable all along the assimilation period. For these channels, the number of active observations is constant for both experiments all along the study period, which proves that the quality control (and the cloud detection) is reliable.

b. Residual bias of observations

Figure 7a compares innovations of REF and EXP for each channel and for all observations (used and rejected observations) for the first assimilation window. Only the impact of the brightness temperature simulation by CO_2 slicing is thus seen in this figure. Innovation amplitudes fluctuate from -8 to 3 K for REF, with an absolute amplitude maximum for longwave surface peaking channels (from 759.6 to 1030.5 cm^{-1} and from 1061.8 to 1135.6 cm^{-1}) and from 1.5 to 4 K for EXP (with an isolated peak at -5.5 K for channels 72 and 73). Innovations obtained by EXP are thus globally 4 times smaller than the ones obtained by REF for surface channels. It globally represents a cooling of the background estimation of about 6 K. Another significant cooling (from 6 to 3 K) is observed for lower-tropospheric peaking channels sensitive to water vapor (from 1218.5 to 1392.2 cm^{-1}). A smaller

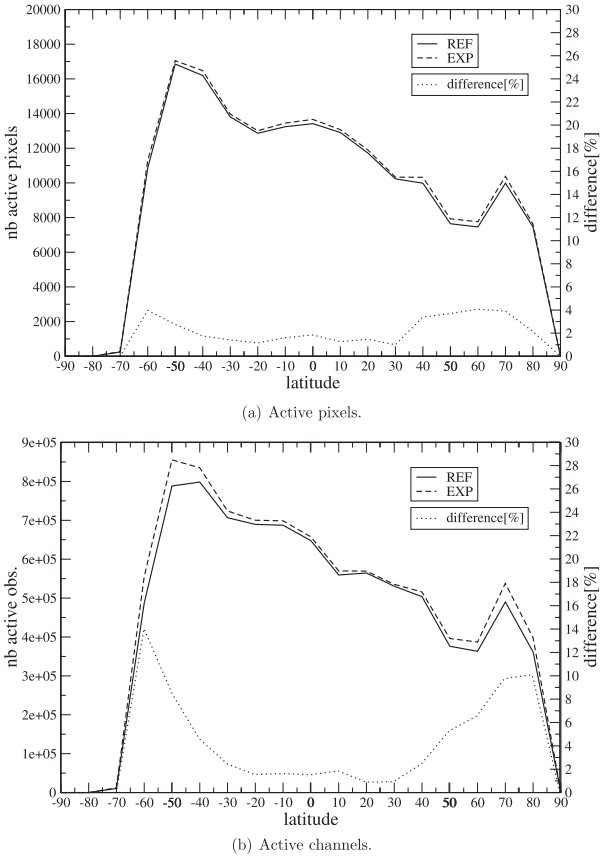


FIG. 6. Geographical distribution of (a) active pixels and (b) active channels with respect to latitude over the entire assimilation period (1–30 Sep 2006). The solid line represents EXP and the dashed line represents REF. The difference (%) between both experiments is the dotted line.

cooling (1–2 K) is observed for upper-tropospheric peaking channels sensitive to water vapor (from 1397.1 to 1605.1 cm^{-1}). Each shortwave channel (except shortwave stratospheric channels) presents also a significant cooling of about 4 K (from 2181.9 to 2298.7 cm^{-1} and from 2380.4 to 2664.1 cm^{-1}). These differences show that to take into account cloud parameters into the radiative transfer calculations leads to model equivalents more consistent with observations. Figure 7b is a zoom on the active channel part of the previous figure. Innovations for EXP and REF are similar from 651.1 to 701.1 cm^{-1} (stratospheric channels). Innovation amplitude differences are larger from 702.5 to 721.5 cm^{-1}, which are tropospheric peaking channels and are potentially more cloud contaminated. We have verified for the last day of our assimilation period (30 September 2006) at 0000 UTC that the variational bias correction was efficient throughout the assimilation period: innovation amplitudes are of the same order than those observed in the first assimilation window for both experiments.

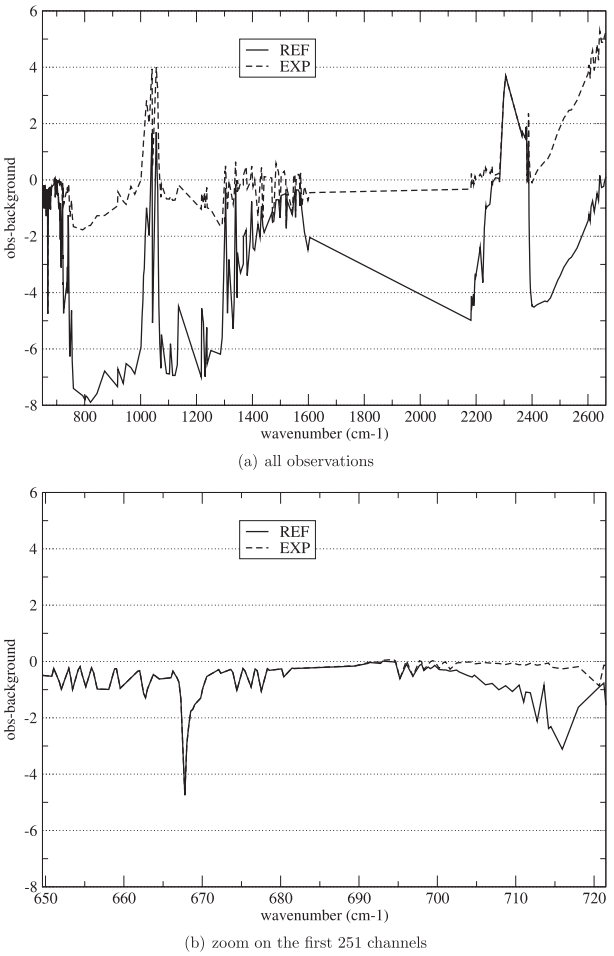


FIG. 7. Averaged innovations (observations−background) of all observations for the 324 available AIRS channels at 0000 UTC 1 Sep 2006. REF is represented by the solid line and EXP by the dashed line.

5. Impact on analysis and forecasts

a. Impact on analysis

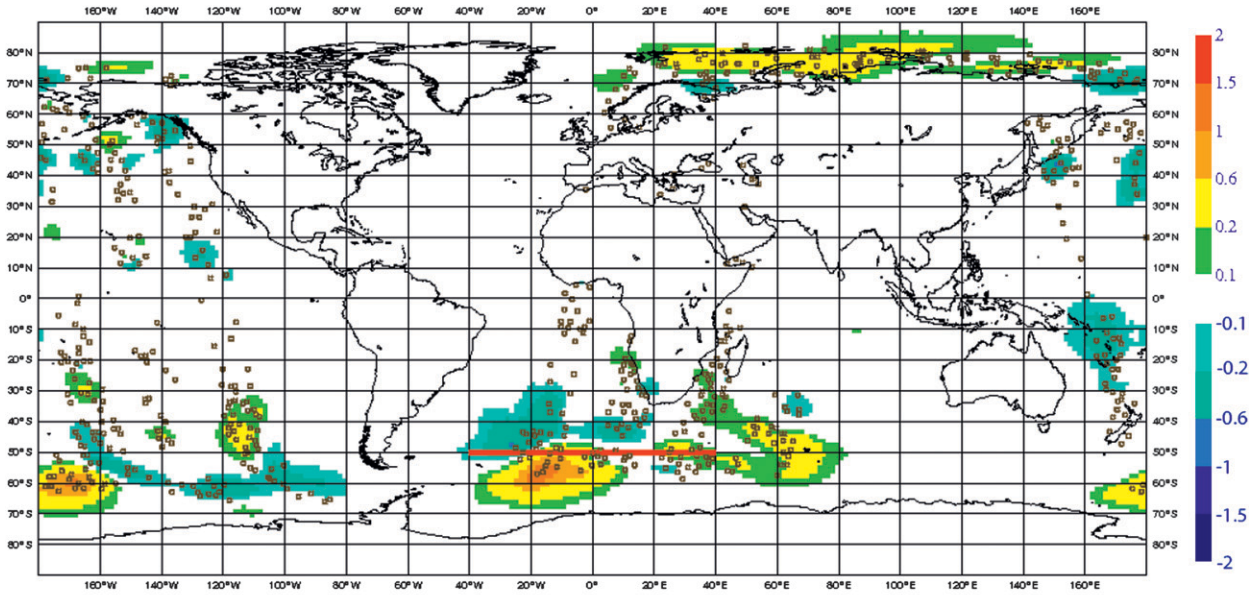
EXP and REF statistics (differences in terms of biases and RMS) from various types of assimilated observations (conventional and satellite data) have been studied over the whole assimilation period. Both background and analysis statistics are in the majority of cases equivalent for both EXP and REF (not shown), but slight improvements have been noticed. With regards to conventional data, a slight reduction of, respectively, the bias and the RMS is observed for winds from profilers in the stratosphere over the Southern Hemisphere and in the upper-troposphere over the Northern Hemisphere. Regarding satellite data, a better fit to microwave sounders sensitive to atmospheric humidity has been noticed: analysis and background bias reduction for the Special Sensor Microwave Imager (SSM/I) in the tropics and over the Southern Hemisphere in the wider atmosphere and a background bias reduction of mid-tropospheric channels of the Advanced Microwave Sounding Unit-B (AMSU-B) in the tropics.

Figure 8a highlights differences between EXP and REF analyses for the temperature at 500 hPa for the first analysis. Differences in model fields are, as expected, located in places where cloud-affected observations are assimilated in EXP showing that these differences are mainly due to the assimilation of cloud-affected observations. These differences are globally located in the Northern Hemisphere between 40° and 90°N and in the Southern Hemisphere between 30° and 70°S. By contrast, the intertropical zone is almost not affected by the assimilation of cloud-affected observations. Furthermore, our assimilation period spreads out the austral winter, where oceans below 70°S are fully covered by sea ice, which is not the case for oceans above 70°N, mostly free from ice. Our assimilation system includes a test that rejects observations situated over sea ice, therefore explaining the negligible impact from 80° to 65°S.

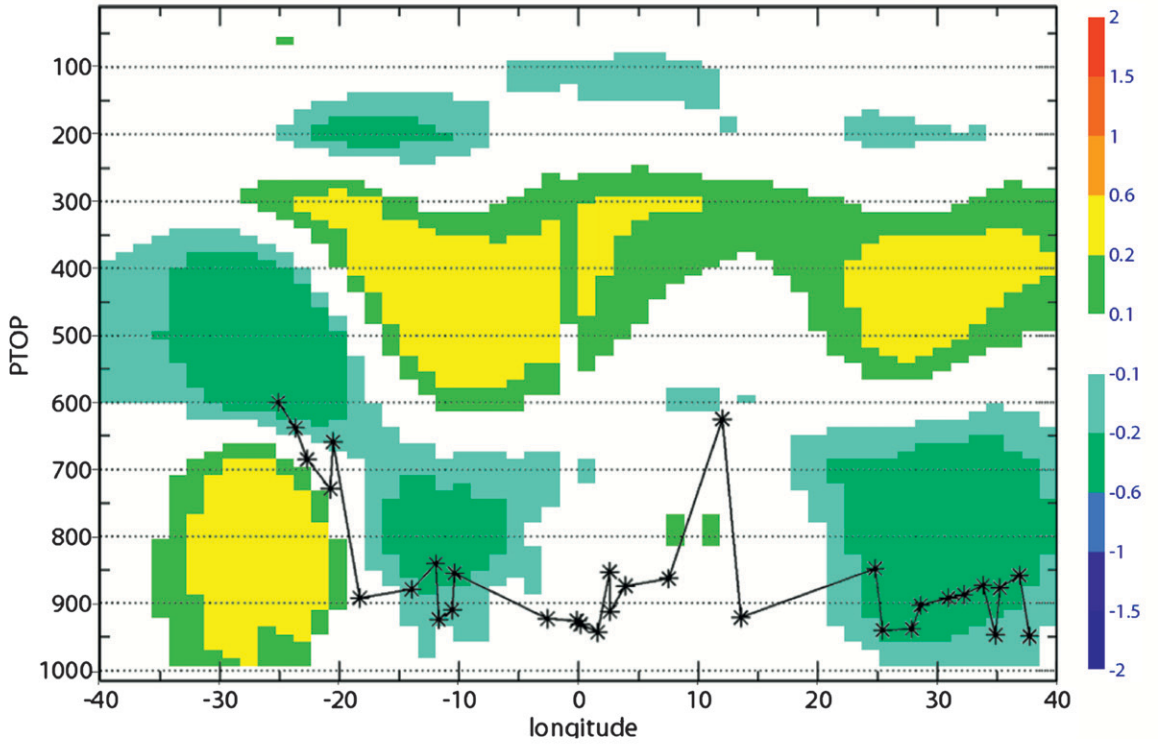
A longitudinal cross section of analysis temperature differences between EXP and REF is shown (Fig. 8b) for the first analysis to characterize the impact of the assimilation of some cloud-affected observations on the analysis with respect to their diagnosed top pressure (from the CO_2 slicing). The cross section at 50°S from 40°W to 40°E is shown because this area displays many active cloud-contaminated observations with different retrieved CTP. From 30° to 20°W, medium clouds are assimilated ($600 < \text{CTP} < 700$ hPa): a warming spreading from the CTP down to the surface is visible. A small cooling is also visible above the CTP up to the upper troposphere (350–400 hPa). From 20°W to 0° and from 20° to 40°E, very low clouds are assimilated ($900 < \text{CTP} < 950$ hPa): a small cooling is visible just above the CTP spreading within the whole lower troposphere (until 650–700 hPa). The upper troposphere is also affected from 600 to 300 hPa: a warming is indeed observed. For very low clouds, these changes can also be accompanied by a slight cooling of the tropopause. Even if these changes highlight the capacity of the cloud-affected radiances to modify the analysis, the lack of verifying data prevents us from assessing which analysis better fits the real state of the atmosphere.

b. Impact on forecasts

The impact of the assimilation of cloud-affected radiances on global forecasts was determined by comparing the objective forecast scores in terms of RMS errors with respect to radiosonde observations. Table 1 shows the RMS forecast error differences for the geopotential



(a) Analysis differences at 500hPa



(b) Cross-section of analysis differences.

FIG. 8. Analysis differences between REF and EXP for temperature at 0000 UTC 1 Sep 2006. (a) The 500-hPa level. Brown circles indicate places where cloud-contaminated pixels are assimilated and the red line at 50°S indicates the location of the cross section shown in (b). The longitudinal cross section is at 50°S and from 40°W to 40°E (shaded colors) and shows the corresponding cloud-top pressure for the assimilated cloud-affected radiances.

TABLE 1. RMS errors of the differences (%) between the geopotential forecast by EXP and REF for a 4-day forecast range with respect to radiosonde observations. Results are averaged from 0000 UTC 1 Sep to 1800 UTC 4 Oct 2006. Boldface indicates a positive impact and regular font indicates a negative impact. Italics indicate values that are statistically significant.

Verifications	Level	Forecast range			
Domain	(hPa)	24	48	72	96
	100	**0.2**	***0.5***	**0.6**	**0.4**
20°N	250	**0.31**	***1.5***	**1.9**	**0.9**
(>20°N)	500	**0.0**	**0.35**	**2.3**	**0.8**
	850	−0.2	−0.2	**2.2**	**0.3**
	100	**0.5**	***1.1***	***0.8***	−0.2
20°S	250	***0.2***	***1.2***	−1.8	−0.5
(<20°S)	500	**0.7**	**1.1**	−1.0	**1.4**
	850	***1.3***	***2.4***	−0.1	**0.0**
	100	***0.7***	***0.7***	**0.4**	**0.0**
Tropics	250	**0.0**	−0.1	**0.0**	−0.5
(20°N, 20°S)	500	**0.5**	**0.4**	**0.6**	−0.1
	850	***0.2***	***0.7***	**0.8**	**0.6**
	100	**1.6**	***2.0***	**2.0**	**0.1**
Europe	250	**0.2**	**4.1**	**4.7**	−0.4
	500	0.0	**3.5**	**6.9**	−0.9
	850	−0.3	**2.6**	**6.7**	−1.2

height. These errors have been computed with respect to radiosonde observations for the validation period ranging from 0000 UTC 1 September 2006 to 1800 UTC 4 October 2006. The impact is rather neutral to slightly positive. The most significant improvement was obtained for the geopotential height according to a statistical Student's *t* test with respect to radiosonde observations with a level of confidence of 95%. Over the Northern Hemisphere, positive impacts are found for a 24–96-h forecast range in the stratosphere and in the mid- to upper troposphere. Improvements at 100 and 250 hPa for a 48-h forecast range are statistically significant. Over the Southern Hemisphere, a positive impact is found for a 24–48-h forecast range in the whole atmosphere. These improvements are significant at the 24–48-h forecast range at 850 and 250 hPa and for the 48–72-h forecast range at 100 hPa. A negative impact has been observed at longer forecast ranges but it is not statistically significant. Over the tropics, positive impacts are found for the 24–72-h forecast range in the stratosphere and in the mid- to low troposphere. These improvements are significant at 850 and 100 hPa for the 24–48-h forecast range. Over Europe, the impact is mainly positive until the 72-h forecast range in the whole atmosphere and negative for the 96-h forecast range. These positive impacts are larger than for the other domains but only significant at 100 hPa at the 48-h forecast range. Scores with respect to ECMWF analyses are rather neutral to slightly positive.

Regarding the other parameters (temperature, humidity, and wind), impacts are roughly neutral to slightly positive but not significant with respect to radiosonde observations and ECMWF analyses. Similar results were found over another 3-week period (from 0000 UTC 29 May to 1800 UTC 20 June 2008) in a different context: more satellite data were assimilated (ATOVS from *MetOp*, IASI, and GPS radio occultation observations) and the model resolution was higher.

Figures 9a–d represent RMS errors and biases of the 48-h forecast over Europe for the geopotential, temperature, humidity, and wind, respectively. The geopotential is improved in terms of RMS in the whole atmosphere and in terms of bias in the stratosphere. These improvements are significant from 10 to 200 hPa. The temperature is slightly improved in terms of bias and RMS errors from the surface to the tropopause with significant impacts at 200 hPa. The humidity is slightly improved in terms of RMS in the upper troposphere and in the tropopause but not significantly. The wind is improved in terms of RMS errors from the surface to the tropopause and from the upper troposphere to the tropopause in terms of bias. These improvements are significant from 100 to 200 hPa.

The impact of the assimilation of cloud-affected observations on forecasts will now be further discussed by an evaluation of the predictability on a meteorological situation: the so-called Mediterranean hurricane (MEDICANE) case.

6. Case study: The MEDICANE case

a. Synoptic description of the event by ECMWF analyses

On 26 September 2006, a strong mesoscale storm with some resemblance to a polar low hit the southeast part of Italy and caused intense precipitation. This event has been investigated from an observational and a numerical (analysis and forecasts from WRF model) point of view by Moscatello et al. (2008). In this paper, this description has been completed by analyses from the ECMWF model, which will be used as a verification to display the relative performance in terms of predictability of both experiments. The trajectory and the evolution of the mean sea level pressure (MSLP) of the low with respect to each time step are summarized in Fig. 10. ECMWF analyses show that a small-scale cyclone was generated in the southeast of the atlas Tunisian Ridge near the Algerian–Tunisian border at around 0000 UTC 25 September 2006 with a MSLP of about 1007 hPa (Fig. 11a). During the following 18 h, the low moved through an east–northeastward direction without intensifying to reach eastern Malta Island at some 800 km away from its initial position (Fig. 11b) at 1800 UTC 25 September. At 0000 UTC 26 September, the low entered the Ionian Sea near the Ionian coast of Calabria. At this stage, it began

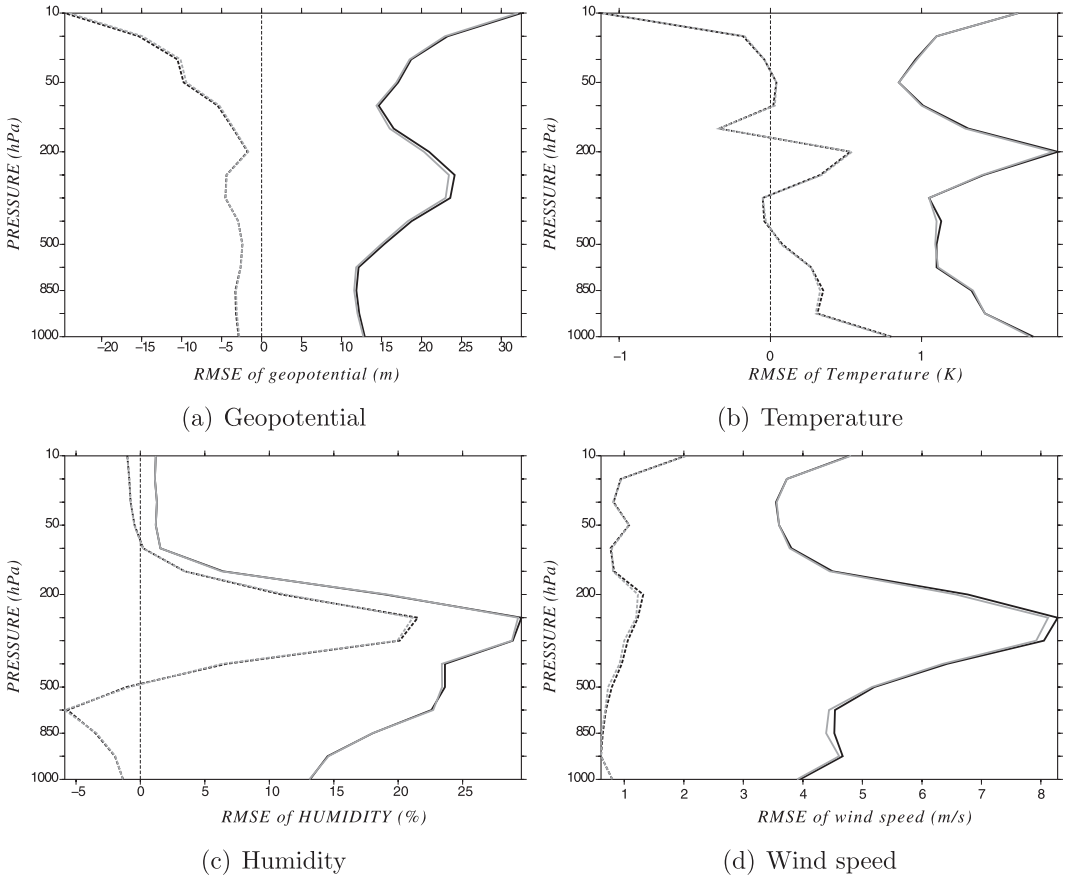


FIG. 9. RMS and mean (bias) of the (a) geopotential, (b) temperature, (c) humidity, and (d) wind differences between 48-h forecasts and radiosonde observations for EXP (gray lines) and REF (black lines) as a function of the pressure. RMS and mean are represented by solid and dashed lines, respectively. Statistics are calculated over the European region and averaged over the period ranging from 0000 UTC 1 Sep to 1800 UTC 4 Oct 2006.

to intensify reaching a MSLP of 1005 hPa. Between 0000 and 0600 UTC, the low moved northeasterly certainly in response to a low pressure system (1001 hPa) visible in the Tyrrhenian Sea (Moscatello et al. 2008) to reach the southern Torente Gulf at 0600 UTC (Fig. 11c). All along this time, the low strongly intensified, reaching a MSLP of about 996 hPa at 0600 UTC. Between 0600 and 1200 UTC, the trajectory of the low curved northward and the depression then crossed the Apulia regions to reach the Adriatic Sea with a MSLP of about 994 hPa, the absolute minimum of the depression at 1200 UTC 26 September. For the next 6 h, the low moved along a northwestward direction and reached the Gargano promontory at 1800 UTC (Fig. 11d) with decreasing intensity (997 hPa). The low then moved inland where it finally died.

b. Predictability of the event

Four sets of numerical simulations have been performed for both experiments in order to compare the predictability, position, and intensity of the cyclone during its whole development, from 0000 UTC 25 September 2006 to 1800 UTC 26 September 2006.

1) The first set, hereinafter called FCST2300, begins at 0000 UTC 23 September 2006 and ends at 1800 UTC 26 September 2006 after a 90-h forecast.
2) The second set, FCST2312, begins at 1200 UTC 23 September and ends after a 78-h forecast.
3) The third set, FCST2400 begins at 0000 UTC 24 September and ends after a 66-h forecast.
4) The fourth set, FCST2412 begins at 1200 UTC 24 September and ends after a 54-h forecast.

Figures 12a–d exhibit the trajectory of the MEDICANE depression forecast by FCST2300, FCST2312, FCST2400, and FCST2412, respectively, ranging from 0000 UTC 25 September 2006 to 1800 UTC 26 September 2006. Forecasts provided by EXP display a better positioning of the low with respect to the ECMWF analyses than REF forecasts for each forecast set and globally for each

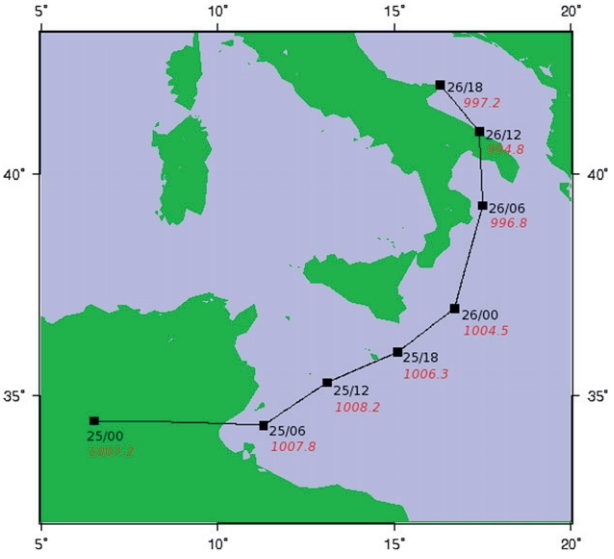


FIG. 10. Trajectory of MEDICANE from ECMWF analysis from 0000 UTC 25 Sep to 1800 UTC 26 Sep 2006. Black numbers indicate the time step and red numbers indicate the corresponding MSLP pressure (hPa).

forecast range. This assessment is true for each time step except for forecasts whose validity date is 1800 UTC 26 September. The position of the low at 1800 UTC 25 September is not forecasted by REF contrary to EXP whichever the considered set. Moreover, the position of the low is neither forecasted by REF at 0000 and 0600 UTC 26 September for the FCST2300 contrary to EXP.

Figures 13a,b highlight the evolution of forecast MSLP minimum by FCST2300 and FCST2312 and FCST2400 and FCST2412, respectively, with respect to the ECMWF analysis. The evolution of the intensity of the low is better forecast by EXP than REF throughout its whole evolution period for each forecast set. FCST2400 displays the largest averaged difference between EXP and REF (Fig. 13b): differences of MSLP between EXP forecasts and ECMWF analysis are 1.5 hPa on the mean against 3.5 hPa for REF. Furthermore, these differences between EXP and REF tend to grow with the forecast range (at least for FCST2400).

c. *Accumulated precipitation*

According to Moscatello et al. (2008), the storm caused intense precipitation over the southeastern part of Italy. The first MEDICANE landfall occurs between 0600 and 1200 UTC 26 September 2006 over the Salentine Peninsula. Equitable threat scores (ETS), which are defined as the intersection of the forecast and observations divided by the union of the forecast and observations at a given threshold, have thus been calculated (Fig. 14). These scores highlight forecast performance of both experiments in terms of cumulated precipitation (within a 6-h window). The closer the score is to 1, the better the forecast. These scores are calculated with respect to cumulated precipitation observations (SYNOP situated between 28° and 55°N and between 0° and 30°E) available from 0600 to 1200 UTC 26 September 2006 for four differents threshold (0.1, 1, 5, and 10 mm) and according to seven different forecast ranges (6, 12, 36, 48, 60, 72, and 84 h).

For very low precipitation (threshold = 0.1 mm), performances of both experiments are quite bad at each forecast range (not shown) but they are slightly improved in EXP. The same assessment can be made regarding a threshold of precipitations of 1 mm (not shown). However, for longtime forecast ranges (from 72 to 84 h), performances of REF are slightly better than EXP for this threshold. As can be seen in Figs. 14a,b, regarding thresholds of 5 mm (steady precipitation) and 10 mm (strong precipitation), respectively, EXP globally exhibits better scores with respect to observations than does REF. For steady precipitation (threshold = 5 mm), performances of both experiments are better than for lower precipitations (Fig. 14a) with better performances of EXP at each forecast range except for the 84-h forecast range. Regarding stronger precipitation (threshold = 10 mm), both performances are reasonable (Fig. 14b) but EXP also exhibits better performances at each forecast range. Cumulated precipitation forecasts are thus improved whatever the rainfall rate but especially for steady and strong precipitation.

7. Conclusions and future developments

The ARPEGE 4D-Var assimilation system has been modified to directly assimilate some AIRS cloud-affected radiances in addition to clear ones. Once a cloud-affected pixel is detected by the CO_2-slicing scheme, cloud parameters (CTP and N_e) are retrieved and directly provided to RTTOV, which will then simulate the cloud-affected spectrum. Furthermore, the ECMWF scheme identifies each clear and cloudy channel of the pixel. Finally, channels affected by clouds whose CTP ranges between 600 and 950 hPa are assimilated in addition to clear ones. All other cloud-contaminated channels are rejected from the assimilation scheme.

ECMWF and CO_2-slicing schemes have been evaluated using independent MODIS data. It has been found that both schemes were performing well in terms of cloud detection. Both cloud-detection performances fluctuate with respect to the diagnosed height of the considered cloud: medium clouds ($400 < \text{CTP} < 800$ hPa) are best detected whereas the detection of low and very low clouds is not very efficient. However, it has also been

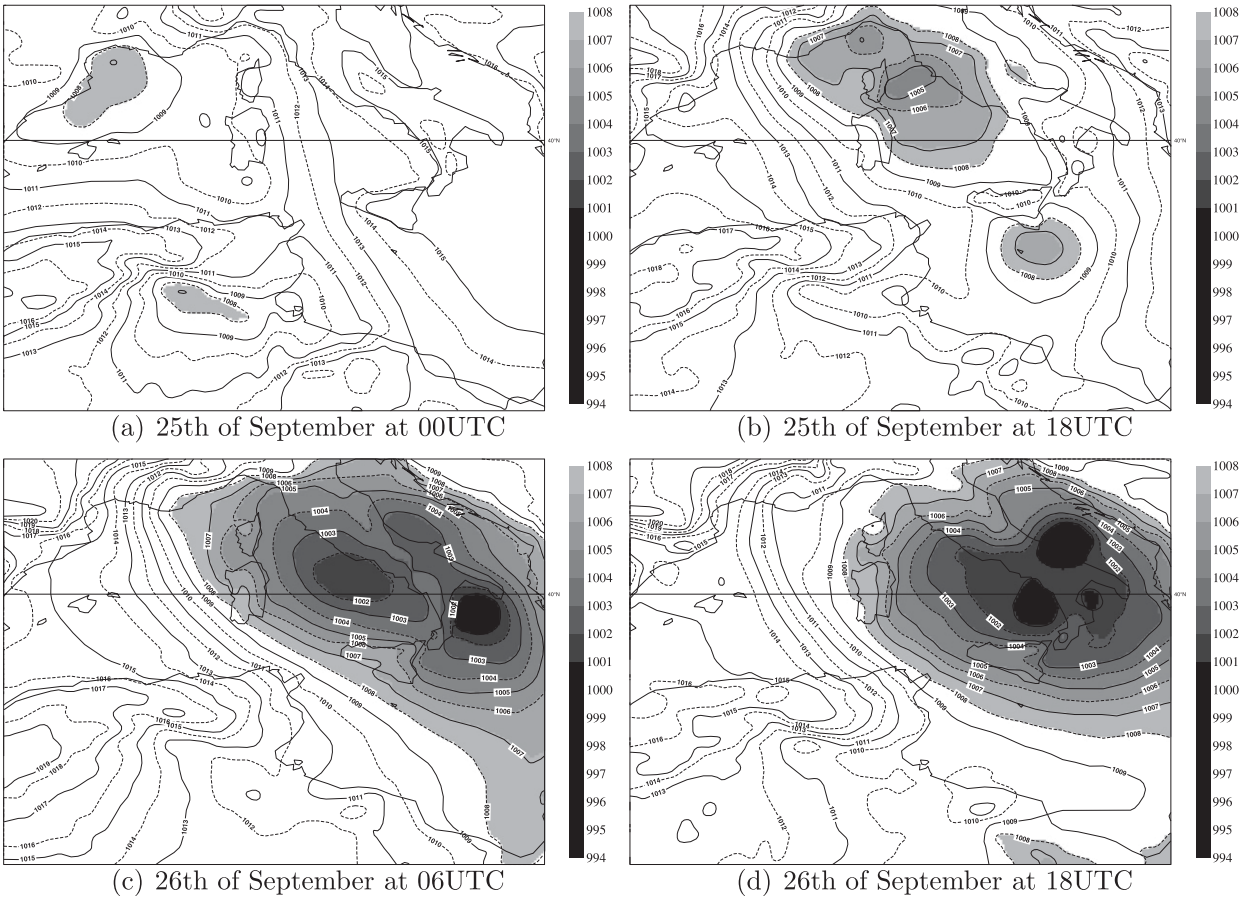


FIG. 11. Analyses of MEDICANE by the operational ECMWF model for 0000 UTC 25 Sep to 1800 UTC 26 Sep 2006. Contours are MSLP (interval of 1 hPa) and shading indicates values <1006 hPa.

demonstrated that difficulties of cloud-detection schemes to detect some low-level clouds have a negligible impact on the simulation of observations contrary to the other levels of clouds. The retrieved cloud-top pressure from CO_2 slicing exhibits a good correlation with the cloud-top pressure inferred from MODIS. The retrieved effective cloud fraction exhibits lower correlations with the effective cloud fraction inferred from MODIS.

Assimilation experiments have been run with the new assimilation scheme: on average, 3.5% of additional pixels are assimilated over the globe with quite a uniform geographic repartition but additional assimilated channels are much more numerous for mid- to high latitudes (10% of additional assimilated channels on average). These extra assimilated data first affect the simulation of model equivalents, which are more consistent with observations as shown by the significant reduction of innovations, especially for longwave surface peaking channels and shortwave channels. Second, the assimilation of cloud-affected observations in addition to clear ones permits the analysis and the background to better fit some conventional observations such as wind profilers and some satellite data sensitive to the atmospheric humidity, such as SSM/I and AMSU-B in terms of bias and RMS errors. Finally, analysis model fields are affected, especially where the extra data volume is important (e.g., mid- to high latitudes). Changes within the atmosphere column differ in accordance with the diagnosed height of the assimilated cloud.

Global forecasts using cloud-contaminated observations have been performed for two validation periods: impacts are rather neutral to slightly positive for each parameter over all domains and for both validation periods. However, they are on the whole not significant except for the geopotential where significant improvements have been highlighted for the 24–48-h forecast range in the stratosphere and in the high troposphere over each domain and in the low troposphere over the tropics and the Southern Hemisphere.

The new assimilation scheme has also been tested on a practical study case: MEDICANE, a strong mesoscale storm, which hit the southeast part of Italy on the

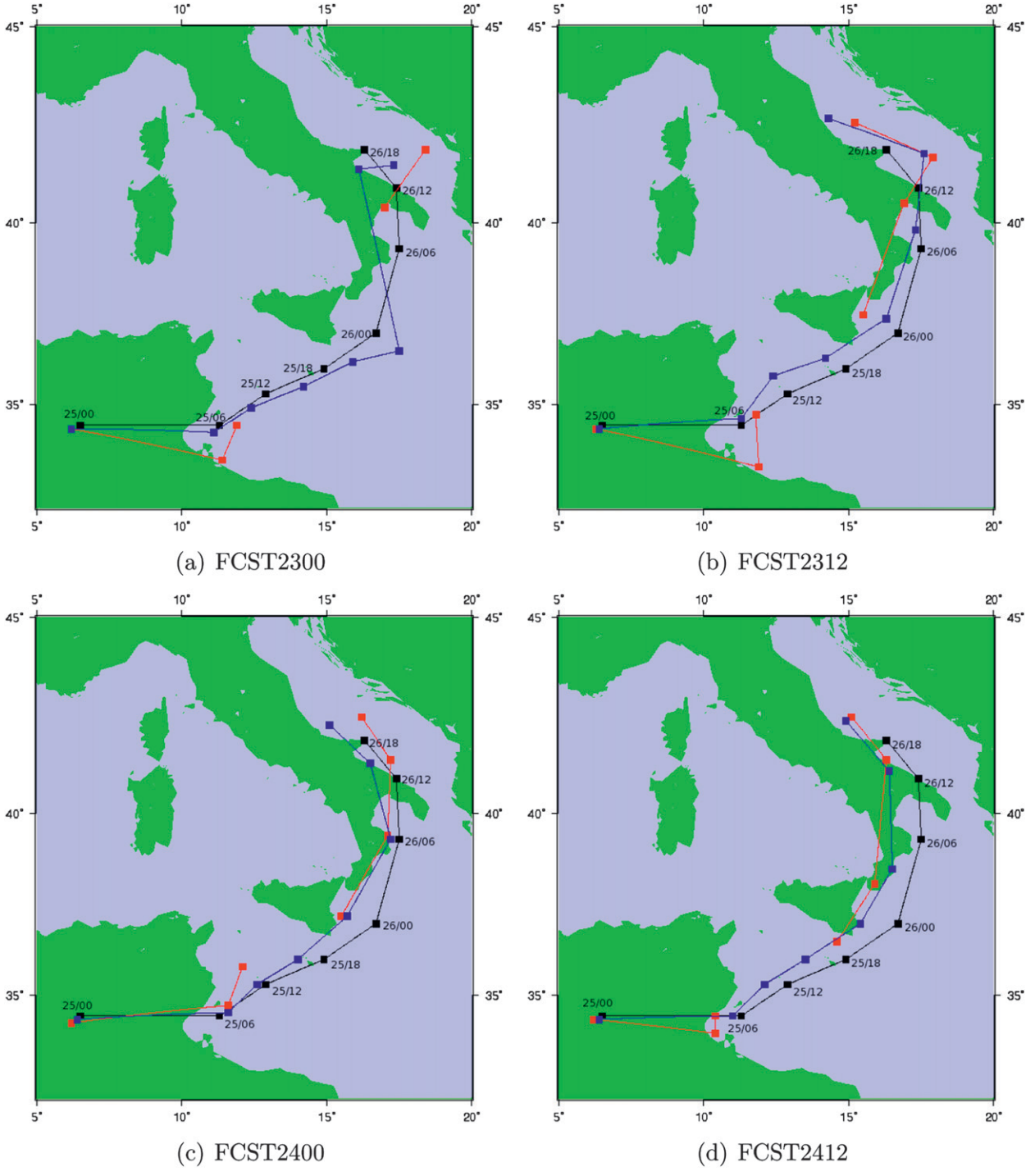
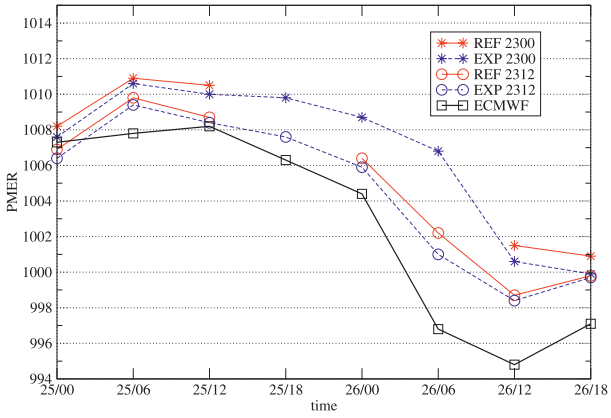
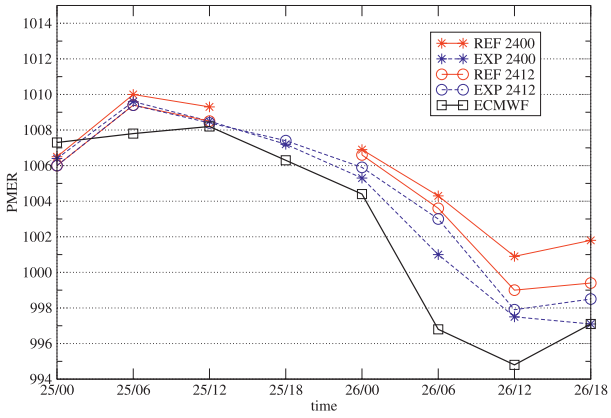


FIG. 12. Trajectory of MEDICANE from (a) FCST2300, (b) FCST2312, (c) FCST2400, and (d) FCST2412. The lines represent the trajectories forecasted by REF (red) and EXP (blue), and given by the ECMWF analysis (black). Black numbers indicate the corresponding time step of MEDICANE.

26 September 2006. Forecasts run from the analysis using cloud-affected radiances are improved with respect to those assimilating only clear ones: the trajectory, intensity, and whole development of the low are better predicted for all the forecast range. Cumulated precipitation forecasts are also improved whatever the rate of rainfall but especially for steady and strong precipitation.

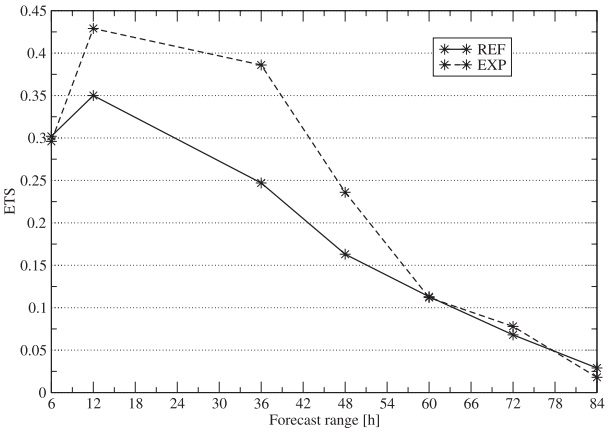


(a) FCST2300 and FCST2312

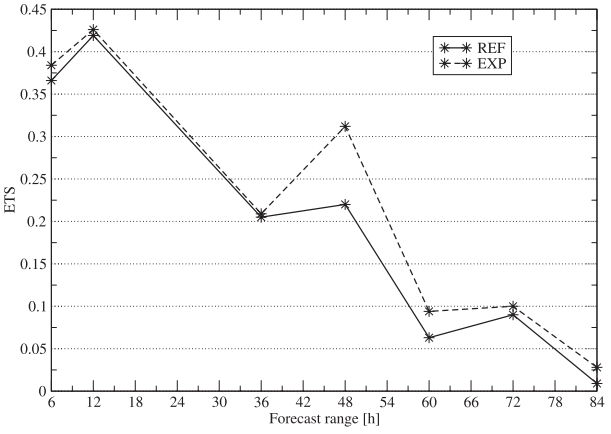


(b) FCST2400 and FCST2412

FIG. 13. Evolution of the minimum of MSLP from 0000 UTC 25 Sep to 1800 UTC 26 Sep 2006. Black line with squares represents the trajectory from ECMWF, dashed blue lines represent the evolution from EXP, and solid red lines represent the evolution from REF. (a) Stars (circles) represent FCST2300 (FCST2312). (b) Stars (circles) represent FCST2400 (FCST2412).



(a) Threshold=5mm



(b) Threshold=10mm

FIG. 14. ETS for EXP and REF with respect to the forecast range for thresholds of 5 and 10 mm between 0600 and 1200 UTC 26 Sep 2006. Solid line represents the REF and the dashed line represents the EXP.

Although results presented in this paper are encouraging, several issues are still to be adressed. First, the bias correction used for clear radiances is applied to cloud-affected radiances and the corrections required should differ from cloud-affected to clear observations. Second, cloud parameters are directly provided to RTTOV in the current assimilation scheme. A prior adjustment by a 1D-Var scheme could admittedly slightly improve the assimilation scheme (Dahoui 2006; Pavelin et al. 2008) but this method, besides being expensive in calculation time, may create discrepancies between cloud parameters and other control variables. The adjustment of cloud parameters inside the iterative process of the 4D-Var could lead to more consistent cloud parameters with other variables. Finally, the rejection of cloud-affected channels whose CTP is situated above 600 hPa leads to a rather low yield in terms of additional active data. Experiments assimilating clouds up to 400 hPa (with a rejection of channels whose weighting-function peak is below the CTP) have been run with a rather neutral (to negative) impact on forecasts. The assimilation of these kinds of clouds thus requires additional developments.

This new assimilation method is obviously valid for other advanced sounders and regarding the short-term development of the approach described in this paper, the next step will be to extend this cloud-affected radiance assimilation scheme to IASI data. Its higher spectral resolution could lead to promising results.

On February 2009, the new assimilation scheme described here was implemented in the operational data assimilation configuration at Météo-France.

Acknowledgments. The authors express their thanks to the ICARE Centre for providing MODIS cloud data products used to evaluate the cloud-detection schemes. In addition, Jean Maziejewski and Fatima Karbou are warmly thanked for their helpful comments on

a previous version of this manuscript. The authors are also grateful to Alexis Doerenbecher and Frank Guillaume for providing many improvements on the quality of figures presented in this work. Finally, we thank the two anonymous reviewers for their constructive comments.

REFERENCES

Auligné, T., T. McNally, and D. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.,* **133,** 631–642.

Aumann, H., and Coauthors, 2003: AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products and processing systems. *IEEE Trans. Geosci. Remote Sens.,* **41** (2), 253–264.

Cayla, F., 2001: L'interféromètre iasi; un nouveau sondeur satellitaire haute résolution (The IASI interferometer: A new high-resolution satellite sounder). *La Météorologie,* **32,** 23–39.

Chahine, M., 1974: Remote sounding of cloudy atmospheres. I. The single cloud layer. *J. Atmos. Sci.,* **31,** 233–243.

Chevallier, F., P. Lopez, A. M. Tompkins, M. Janiskova, and E. Moreau, 2004: The capability of 4d-Var systems to assimilate cloud-affected satellite infrared radiances. *Quart. J. Roy. Meteor. Soc.,* **130,** 917–932.

Courtier, P., C. Freydier, F. Rabier, and M. Rochas, 1991: The ARPEGE Project at Météo-France. *Proc. ECMWF Seminar on Numerical Methods in Atmospheric Models,* Vol. 7, Reading, United Kingdom, ECMWF, 193–231.

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

Dahoui, M. L., 2006: Vers une assimilation variationnelle des radiances satellitaires nuageuses (Toward a variational assimilation of cloudy satellite radiances). Thèse de doctorat, Université Toulouse 3-Paul Sabatier, France, 160 pp.

——, L. Lavanant, F. Rabier, and T. Auligné, 2005: Use of MODIS imager to help dealing with AIRS cloudy radiances. *Quart. J. Roy. Meteor. Soc.,* **131,** 2559–2579.

Dee, D., 2004: Variational bias correction of radiance data in the ECMWF system. *Proc. ECMWF Workshop on Assimilation of High Spectral Resolution Sounders in NWP,* Reading, United Kingdom, ECMWF, 97–112.

English, S., J. Eyre, and J. Smith, 1999: A cloud-detection scheme for use with satellite sounding radiances in the context of data assimilation for numerical weather predictions. *Quart. J. Roy. Meteor. Soc.,* **125,** 2359–2378.

Eyre, J. R., and P. Watts, 1987: A sequential estimation approach to cloud-clearing for satellite temperature sounding. *Quart. J. Roy. Meteor. Soc.,* **113,** 1349–1376.

——, and W. P. Menzel, 1989: Retrieval of cloud parameters from satellite sounder data: A simulation study. *J. Appl. Meteor.,* **28,** 267–275.

Fourrié, N., and F. Rabier, 2004: Cloud characteristics and channel selection for IASI radiances in meteorologically sensitive areas. *Quart. J. Roy. Meteor. Soc.,* **130,** 1839–1856.

Li, J., C. Liu, H. L. Huang, T. J. Schmit, X. Wu, and W. P. Menzel, 2005: Optimal cloud-clearing for AIRS radiances using MODIS. *IEEE Trans. Geosci. Remote Sens.,* **43,** 1266–1278.

McNally, A., 2002: A note on the occurrence of cloud in meteorologically sensitive areas and the implications for advanced infrared sounders. *Quart. J. Roy. Meteor. Soc.,* **128,** 2551–2556.

——, and P. Watts, 2003: A cloud detection algorithm for high spectral resolution infrared sounders. *Quart. J. Roy. Meteor. Soc.,* **129,** 3411–3423.

Menzel, W., T. Smith, and T. R. Stewart, 1983: Improved cloud motion wind vector and altitude assignment using VAS. *J. Climate Appl. Meteor.,* **22,** 377–384.

Moscatello, A., M. Marcello Miglietta, and R. Rotunno, 2008: Numerical analysis of a Mediterranean ''hurricane'' over southeastern Italy. *Mon. Wea. Rev.,* **136,** 4373–4397.

Pavelin, E., S. English, and J. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **134,** 737–749.

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

Saunders, R. W., P. Brunel, G. Chevallier, F. Deblonde, M. English, S. Matricardi, and P. J. Rayer, 2002: RTTOV-7 science and validation report. NWP Forecasting Research Tech. Rep. 387, 51 pp.

Simulation of satellite infrared radiances for convective-scale data assimilation over the Mediterranean

Fanny Duffourg,[1] Véronique Ducrocq,[1] Nadia Fourrié,[2] Geneviève Jaubert,[1] and Vincent Guidard[2]

Received 31 July 2009; revised 2 February 2010; accepted 18 February 2010; published 7 August 2010.

[1] This paper focuses on the simulation of satellite infrared passive observations for their assimilation in high horizontal resolution (2.5 km) numerical weather prediction systems. In order to better represent the sensitivity of the satellite measurement to the whole atmosphere within its footprint, new observation operators are designed. They aggregate the model information contained within the satellite field of view. The different observation operators are evaluated for the simulation of Infrared Atmospheric Sounding Interferometer (IASI) and Atmospheric Infrared Sounder (AIRS) observations over a whole month. The new observation operators are found to improve the simulation of water vapor channels and to have neutral to slightly negative impact for temperature channels. For most channels, the standard deviation of the observation minus guess departures is reduced. The modifications of the simulations are substantial for water vapor channels for which the weighting functions peak at pressures greater than 340 hPa for IASI and between 340 hPa and 800 hPa for AIRS. The most important ones appear where fine-scale humidity gradients occur in dry sounded layers. The new observation operators improve the simulation of IASI and AIRS observations by filtering out these fine-scale patterns that are not detected by the instruments. With improvements in observations minus guess reaching 2 K, the new observation operators may avoid the rejection of some observations by quality control procedures during the assimilation process. Single observation assimilation experiments are then carried out using the different observation operators. They show that even large modifications in the observation simulation have almost no impact on the final analysis.

Citation: Duffourg, F., V. Ducrocq, N. Fourrié, G. Jaubert, and V. Guidard (2010), Simulation of satellite infrared radiances for convective-scale data assimilation over the Mediterranean, *J. Geophys. Res.*, *115*, D15107, doi:10.1029/2009JD012936.

1. Introduction

[2] A number of meteorological centers have recently developed convective-scale numerical weather prediction (NWP) systems with the specific aim of improving the forecasting of high-impact weather events. Their kilometer-sized grid mesh, nonhydrostatic equations and improved microphysics parameterizations enable atmospheric deep convection to be resolved explicitly. The representation of the precipitating systems is thus significantly improved. Case studies focusing on Mediterranean heavy rainfall events have shown significant forecast improvements using nonhydrostatic convective-scale research models [*Ducrocq et al.*, 2002]. However, these studies have also pointed out the necessity to improve, in particular, the representation in the initial conditions of the moisture field as Mediterranean heavy rainfall simulations are very sensitive to the mesoscale structure of this field that is highly variable in space and time. A better ability to simulate the dynamical and physical processes at fine scale is not always sufficient to prevent bad forecasts. If some mesoscale key mechanisms are missing in the initial conditions, the model fails to reproduce the precipitating systems.

[3] Convective-scale assimilation of observations is a way to improve the initial conditions of kilometer-scale models. Such improvement over the sea is of particular importance because some Mediterranean heavy rainfall is triggered off shore [*Nuissier et al.*, 2008]. Over the sea, satellite data are practically the only routinely available observations. The new Infrared Atmospheric Sounding Interferometer (IASI) [*Cayla*, 2001; *Chalon et al.*, 2001] and Atmospheric Infra-Red Sounder (AIRS) [*Pagano et al.*, 2002; *Aumann et al.*, 2003] now offer high-resolution and accurate information

[1]Groupe de Météorologie de Moyenne Echelle, Centre National de Recherches Météorologiques, Groupe d'étude de l'Atmosphère Météorologique, Météo-France, Centre Nationale de la Recherche Scientifique, Toulouse, France.

[2]Groupe de Modélisation pour l'Assimilation et la Prévision, Centre National de Recherches Météorologiques, Groupe d'étude de l'Atmosphère Météorologique, Météo-France, Centre Nationale de la Recherche Scientifique, Toulouse, France.

0148-0227/10/2009JD012936

on temperature and humidity. IASI sounds the atmosphere with a horizontal resolution of 12 km at nadir. The accuracy of its measurements is expected to be better than 1 K for temperature retrievals and 10% below 500 hPa for relative humidity retrievals with a vertical resolution finer than 1 km [*Diebel et al.*, 1996]. The initial model analysis may therefore benefit from the assimilation of such new information over the sea.

[4] However, the assimilation of satellite data in a convective-scale data assimilation system is not straightforward. New problems arise, in particular because of the scale difference between model and satellite measurements: the model mesh is about 1 order smaller than any satellite observation spot. This problem is frequently encountered in hydrological applications when the satellite soil moisture data need to be retrieved at the finer scale of the hydrological model. In the past, several studies have addressed this issue, proposing that the satellite measurements should be disaggregated at the model scale by characterizing subpixel variability [*Reichle et al.*, 2001; *Kim and Barros*, 2002; *Merlin et al.*, 2005]. This approach cannot, however, be applied when the satellite measurements are directly assimilated without any previous retrieval. The variational data assimilation methods of the NWP systems which assimilate the satellite radiances directly, require indeed equivalent observations simulated from the values provided by the model to be compared with the real observations. For such simulations, model information from the different grid columns contained within the satellite field of view has to be aggregated. This issue has already been investigated by *Kleespies* [2009] for the assimilation of microwave sounder measurements: surface model parameters are aggregated to improve the simulation in case of nonuniform scenes. The aim of our study is to address this question by gathering atmospheric model information and focusing on the accurate, high-resolution IASI and AIRS infrared sounders.

[5] More precisely, the main purpose of this paper is to evaluate different methods for aggregating atmospheric model information contained in the satellite field of view to simulate IASI and AIRS observations. This issue is examined more specifically for the newly developed convective-scale 3D-Var data assimilation system of Météo-France called AROME (Application of Research to Operations at MEsoscale) and for the Mediterranean. We follow the ideas of the studies of *Brenot et al.* [2006], which evaluated the sensitivity to various formulations of Global Positioning System (GPS) zenith delay simulations, and *Caumont et al.* [2006] and *Caumont and Ducrocq* [2008] for the effect of how radar reflectivity and Doppler wind simulations, respectively, were formulated. Different ways of aggregating the model information within a IASI or AIRS spot are developed. They are described in section 2 together with the AROME system and the IASI and AIRS measurements. Section 3 evaluates the maximum differences between the various methods of aggregation for the simulation of IASI and AIRS brightness temperatures. The different simulations are then compared with real observations in section 4. Lastly, section 5 discusses the impact on the AROME analysis of the different methods for computing the model equivalent radiances by performing single observation assimilation experiments. Conclusions and outlooks follow in section 6.

2. Description of the Assimilation of Satellite Infrared Radiances in AROME

2.1. AROME 3D-Var

[6] AROME [*Ducrocq et al.*, 2005; *Yan et al.*, 2009] is a Limited Area Model (LAM) with a 2.5 km grid covering the territory of France including the northwestern part of the Mediterranean Sea (5.2°W to 11.2°E, 40.5°N to 51.7°N). It is coupled hourly with the ALADIN (Aire Limitée, Adaptation dynamique, Développement InterNational) forecast [*Radnóti et al.*, 1995] on its lateral boundaries. The 41 unequally spaced vertical levels cover the troposphere (with 30 levels) and, more loosely, the stratosphere up to about 1 hPa. AROME is a three-dimensional nonhydrostatic model, the dynamics of which is based on the ALADIN nonhydrostatic equations [*Bubnová et al.*, 1995]. The prognostic equations of the six water species (water vapor, cloud water, rain water, cloud ice crystals, snow and graupel) as well as the physical parameterizations are shared with the nonhydrostatic Meso-NH model [*Lafore et al.*, 1998].

[7] AROME has its own 3D-Var data assimilation system [*Brousseau et al.*, 2008], based on that of ALADIN-FRANCE [*Fischer et al.*, 2005; *Guidard et al.*, 2006]. As such, it adopts the incremental formulation originally introduced in the ARPEGE/IFS (Action Recherche Petite Echelle Grande Echelle/Integrated Forecast System) global data assimilation system [*Courtier et al.*, 1994]. The AROME background error covariances are based on the same multivariate formulation as in ALADIN-FRANCE [*Berre*, 2000]. The design of the observation-error covariance matrix assumes that there are no observation-error correlations.

[8] The operational AROME 3D-Var data assimilation system uses a 3 h forward intermittent cycle: observations within a ±1 h 30 min assimilation window are used to perform the analysis from which a 3 h forecast is obtained to serve as a first guess for the next cycle. The two wind components, temperature, specific humidity and surface pressure are analyzed. Other model fields are cycled from the previous AROME guess. All conventional observations are assimilated (surface observations from land stations and ships, vertical soundings from radiosondes and pilot balloons, buoy and aircraft measurements), together with wind profilers, winds from Atmospheric Motion Vectors (AMV) and scatterometers, Doppler winds from radars, satellite radiances (such as IASI and AIRS measurements) as well as ground-based GPS measurements. Satellite radiances (or the corresponding brightness temperatures) are directly assimilated, using an observation operator that simulates the model-equivalent radiances. This observation operator is described in section 2.3. The differences between real and model-equivalent observations are bias-corrected with an adaptive variational method [*Dee*, 2004; *Auligné et al.*, 2007] before entering in the minimization of the variational cost function for the computation of the analysis.

2.2. IASI and AIRS Measurements

[9] IASI and AIRS are hyperspectral infrared passive nadir scanning radiometers, onboard the European MetOp and the

Table 1. IASI and AIRS Geometrical Characteristics[a]

Instrument	Angular FOV	Nadir Spot Size	Maximum Scan Angle	Maximum Spot Size
IASI	1.1°	12 km	48.3°	38 km × 20 km
AIRS	0.825°	13.5 km	49.5°	40 km × 22 km

[a]Nadir spot size corresponds to the diameter of the circular spot, and maximum spot size corresponds to the major and minor axes of the ellipsoidal spot. FOV stands for field of view.

American Aqua polar orbiting satellites, respectively. They measure the radiation coming out of the atmosphere on thousands of channels in the infrared spectrum (8461 for IASI and 2378 for AIRS). The measured outgoing radiation results from all the radiative emissions of each layer of the atmosphere and the surface, weakened by the atmospheric absorption (if the atmospheric diffusion is neglected which is realistic in clear-sky conditions). The measured radiances can thus be expressed by the following radiative transfer equation:

$$L_\lambda(\mathrm{sat}) = \varepsilon_\lambda(\mathrm{surf})\mathrm{B}_\lambda(\mathrm{T_s})\tau_\lambda(\mathrm{surf}) + \int_{z_{\mathrm{surf}}}^{z_{\mathrm{sat}}} \varepsilon_\lambda \mathrm{B}_\lambda(\mathrm{T(z)})\tau_\lambda(\mathrm{z})\mathrm{dz}, \quad (1)$$

where λ is the wavelength of the considered radiation, L_λ is the radiance in this wavelength, ε_λ(surf) the surface emissivity, ε_λ the atmospheric emissivity, $B_\lambda(T)$ the black body radiance, T_s the surface temperature, z_{surf} the altitude of the surface, z_{sat} the altitude of the satellite and $\tau_\lambda(z)$ the transmission (the proportion of the radiation which has not been absorbed when passing through the atmospheric layer between the altitude z and the satellite). Radiation absorption and emission in the different layers of the atmosphere are sensitive to the meteorological conditions (humidity and temperature). The measured radiance is therefore strongly linked with the meteorological parameters in some specific highly sensitive zones of the atmosphere depending on the wavelength. For each channel, the vertical layers contributing to the radiances are represented by weighting functions.

[10] IASI and AIRS spectra can be divided into three major bands: (1) from 645 cm^{-1} to 1210 cm^{-1} (wavelength from about 8.25 μm to about 15.5 μm): CO_2 band mainly sensitive to temperature, named Temperature Long Wave band (TLW); (2) from 1210 cm^{-1} to 2040 cm^{-1} (wavelength from about 4.9 μm to about 8.25 μm): band mainly sensitive to humidity, named Water Vapour band (WV); and (3) from 2040 cm^{-1} to 2700 cm^{-1} (wavelength from about 3.7 μm to about 4.9 μm): band mainly sensitive to temperature, named Temperature Short Wave band (TSW).

[11] A subset of IASI and AIRS channels was selected for this study. Within the classical sets of 314 IASI and 324 AIRS channels defined for operational NWP applications ([*Collard*, 2007] for IASI, [*Susskind et al.*, 2003] for AIRS), only the channels assimilated operationally at ECMWF in 2007 were kept. Although very few data from IASI WV and non-Sun-sensitive TSW channels are currently assimilated by operational meteorological centers, these channels were all kept in this study to prepare for their future use. Within these subsets, we selected only the channels sounding below the tropopause as the vertical resolution within the stratosphere in AROME is not sufficient to simulate the stratospheric sounding channels correctly. This selection led to a subset of 120 AIRS and 163 IASI channels.

[12] IASI and AIRS are nadir-viewing sounders: they scan the atmosphere for different look positions along a plane which is perpendicular to the satellite orbit track. When looking at an off-nadir position, the atmosphere is scanned along a slanted line of sight. Table 1 gives the maximum angle formed between the line of sight and the nadir direction. As IASI and AIRS fields of view are 0.825° and 1.1°, respectively, the horizontal resolution of their measurements varies with the scan angle. The minimum (at nadir) and maximum (at swath edge) sizes of the satellite observation spot are given in Table 1 and range from 12 km to 40 km along the major axis of the ellipsoidal spot.

[13] The data used in this study were level 1c products provided for operational purposes by EUMETSAT (EUropean organisation for the exploitation of METeorological SATellites) for IASI and NESDIS (National Environmental Satellite Data and Information Service) for AIRS. The NESDIS AIRS product is available only for every ninth scan position.

2.3. IASI and AIRS Observation Operators

[14] The simulation of satellite brightness temperature from the model atmospheric profile is performed in two main steps by the observation operator. The first step forms a model column that represents the sounded atmosphere. The second step uses this model column to compute the simulated brightness temperature using a radiative transfer model. The brightness temperature calculation is performed with the Radiative Transfer for TOVS (RTTOV) model [*Saunders and Brunel*, 2005] which models the complex relationship linking the atmospheric profile with the radiances measured by satellite nadir scanning radiometers (as expressed by equation (1)). We used the version 8.5 of RTTOV with the default 43 levels. In our case, RTTOV was based on the line-by-line transmittance model GENLN2 [*Edwards*, 1992].

[15] In the operational AROME data assimilation system, the model column representing the sounded atmosphere is estimated at the center of the satellite observation spot by interpolating the four closest model columns surrounding this point. This procedure, called COL4 hereafter, comes from the previous larger-scale assimilation systems for which the model grid mesh is larger than the observation spot. However, with a 2.5 km horizontal resolution like the one used for AROME, a single IASI or AIRS observation spot covers more than 12 model grid points at nadir and about a hundred at swath edge. The instrument's point spread function is quasi-uniform over the spot (*Blumstein* [2005] for IASI and *Elliott et al.* [2006] for AIRS), so it is legitimate to consider that every model grid point in the spot contributes similarly to the measurement. The same weight for each grid point was thus used in the new aggregation methods developed in this study. Three new observation operators aggregating the model information contained within the satellite field of view were defined (Table 2 and Figure 1). In the observation operator called SPOT1 hereafter, the sounded atmosphere is represented by the mean of all the model columns located in the observation spot. This mean model column is then used to estimate the brightness temperature with RTTOV. For M12, the sounded atmosphere is estimated as the mean of the 12 model columns surrounding the center

Table 2. Characteristics of the Different Observation Operators

Operator	Method
COL4	four-point interpolation
M12	mean over 12 points before radiative transfer
SPOT1	mean over the spot before radiative transfer
SPOT2	mean over the spot after radiative transfer

of the observation spot. Finally, a more realistic observation operator averages the estimated brightness temperatures from each model column in the spot rather than averaging the model columns before the radiative transfer. This was found to be almost the same as averaging the radiances from each model column within the spot. This third observation operator, called SPOT2, requires much more computing time however, which is an important drawback for operational use.

3. Intercomparison of the Observation Operators

3.1. Methodology

[16] In this section, we examine what the maximum impact of a change in the observation operator may be on the simulation of IASI and AIRS brightness temperatures. For the whole month of September 2007, the four observation operators (Table 2) were applied to simulated observation spots centered on each grid point of every 3 h AROME analysis provided by a data assimilation cycle as described in section 2.1. Only clear simulated spots situated over the western Mediterranean Sea were considered. A simulated observation spot was clear if the mixing ratio of the total hydrometeors was less than 10^{-6} kg/kg over all the model columns included in the spot. The operational cloud detection system (the "cloud detect" software by *McNally and Watts* [2003]) could not be used here as no real observations were considered. Also, a simulated spot was defined as over the sea if, according to the model land-sea mask, the whole spot was entirely over sea. The observation operators were applied considering the maximum size of the observation spot (i.e., simulating swath edge measurements), in order to evaluate the maximum differences between the various observation operators.

[17] Mean biases and standard deviations for IASI and AIRS brightness temperature differences between the calculations with the various operators over the whole of September 2007 were computed using COL4 as the reference. The significance of the results was evaluated by comparison with the instrument noise, estimated in terms of Noise Equivalent Differential Temperature (NEDT). *Blumstein* [2007] gives the NEDT values for each IASI channel at the reference temperature of 280 K, and *Pagano et al.* [2002, 2003] give the values for AIRS at 250 K. With these reference values, NEDT was computed at the mean brightness temperature over the month for each channel, following *Chalon et al.* [2001]. It was verified that the variations of NEDT over the month remained within ±10% of the value computed with the monthly mean temperature (not shown).

3.2. Results Over the Month of September 2007

[18] For all the channels studied, Figure 2 shows the standard deviation over the whole month of September 2007 of the brightness temperature differences between calculations using the three new observation operators (M12, SPOT1 and SPOT2; see Table 2) and the current one (COL4), together with the NEDT. Figure 3 zooms in the WV channels.

[19] The standard deviations for SPOT1 and SPOT2 are remarkably close: the corresponding curves are almost superimposed. As expected, brightness temperature differences between SPOT1 or SPOT2 and COL4 are greater than differences between M12 and COL4 for most channels.

[20] For the three new observation operators, the standard deviation of the brightness temperature differences with COL4 is lower than the instrument noise for all TLW and TSW IASI and AIRS channels and also for the short-wave part of the IASI WV band (wave number between 1920 cm^{-1} and 2040 cm^{-1}). Substantial differences between SPOT1 (or SPOT2) and COL4 (i.e., larger than the NEDT) appear mainly for IASI and AIRS WV channels for which the weighting function peaks at pressures greater than 340 hPa for IASI, and between 340 hPa and 800 hPa for AIRS (wave number from 1210 to about 1455 cm^{-1} plus channel 3263, 1460.50 cm^{-1}, for IASI and from about 1315 to 1600 cm^{-1}, except three channels around 1550 cm^{-1} and channel 1449, 1330.98 cm^{-1}, for AIRS). These differences are smaller for AIRS channels with broad weighting functions (e.g., channel 1471, 1342.24 cm^{-1}). The differences between M12 and COL4 calculations are smaller; they remain substantial only for the low to middle troposphere (under 400 hPa) sounding channels with low NEDT (e.g., IASI channel 2889, 1367.00 cm^{-1}, and AIRS channel 1627, 1427.23 cm^{-1}).

[21] The impact of the new observation operators is larger on the simulation of WV channels because these channels are much more sensitive to atmospheric temperature and humidity than temperature channels. Figure 4 shows this difference in sensitivity for a IASI TLW channel and a IASI WV channel. It exhibits the Jacobians in temperature and specific humidity of both channels.

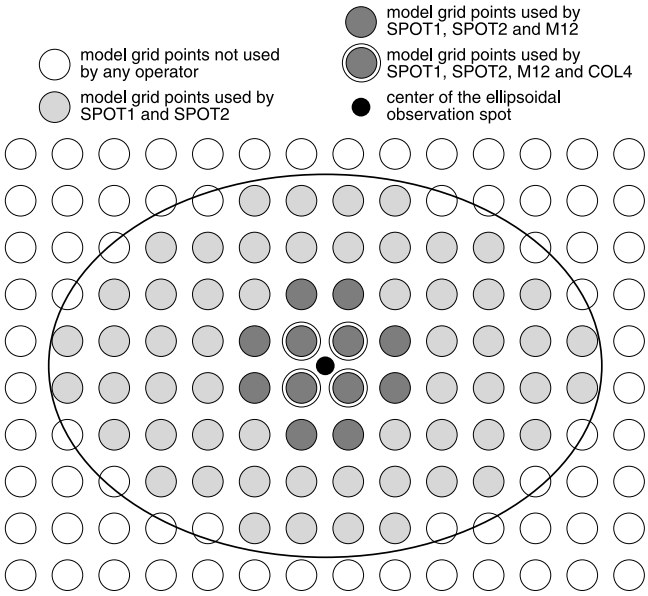


Figure 1. The model grid points aggregated by the different observation operators.

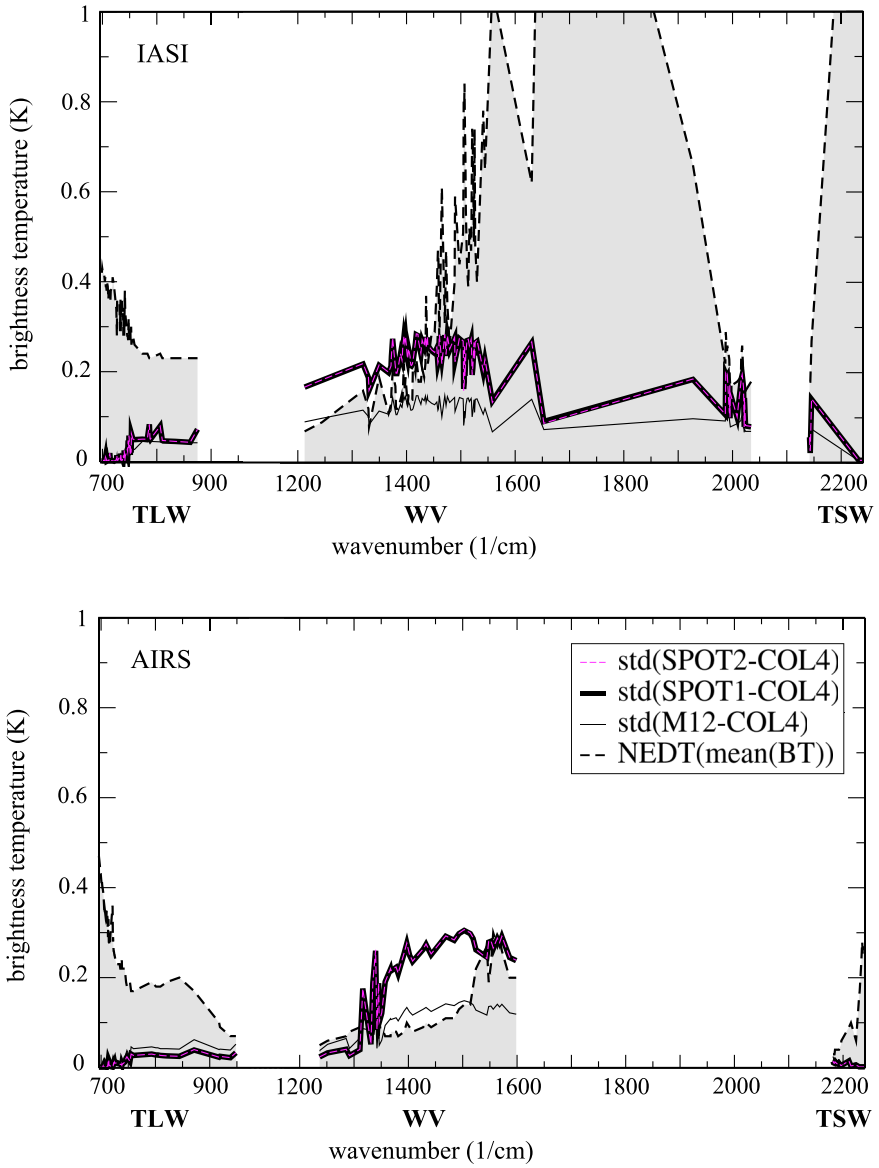


Figure 2. Standard deviation (in K) of the brightness temperature (BT) differences between simulations with a new observation operator (M12 in thin solid black line and SPOT2 in thin dashed violet line plotted on top of SPOT1 in thick solid black line) and COL4 for all the IASI and AIRS channels studied, displayed according to increasing wave number along the abscissa. The instrument NEDT estimated for each channel at the monthly mean brightness temperature is represented by the dashed black line, with the area under this line colored in grey.

[22] For the three new observation operators, the average of the brightness temperature differences with COL4 is very close to 0 K for each channel: no new biases have been introduced.

[23] Figure 5 shows the standard deviation of the brightness temperature differences between SPOT2 and SPOT1. Its comparison with the instrument noise confirms that using a spot-averaging method closer to the measurement (SPOT2) does not bring results substantially different from averaging first the model information over the spot before computing the brightness temperature (SPOT1). The standard deviation of their differences is indeed lower than the NEDT for all the IASI and AIRS WV channels. Even smaller differences between the two operators were found for IASI and AIRS TLW and TSW channels and for IASI channels of the short-wave part of the WV band (not shown).

[24] The histograms of the brightness temperature differences between SPOT1 or SPOT2 and COL4 for IASI WV channel 2919 and AIRS WV channel 1627 are displayed in Figure 6. Histograms for other WV channels (not shown) have a similar shape. They all show that the largest differences with COL4 are slightly weaker using SPOT2 rather than SPOT1. This is the main difference between calculations with SPOT1 and SPOT2 as both observation operators were globally in good agreement with each other for situa-

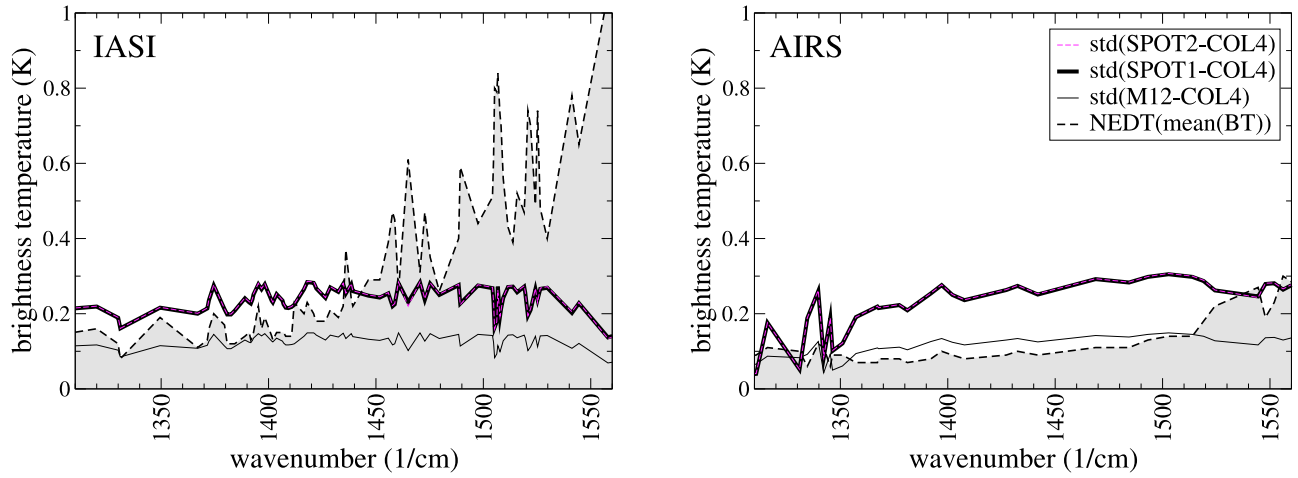


Figure 3. Zoom in on the WV channels of Figure 2.

tions where their simulations were closer to the COL4 calculations. This is not important enough to justify the use of a much more computing-time-consuming observation operator in an operational and real time context. In consequence, SPOT1 will be preferred.

[25] Finally, to assess the temporal variability of the brightness temperature differences between the various observation operators, the standard deviations of these differences were computed on each 3 h AROME analysis. Time series are shown on Figure 7 for some of the WV channels identified above as having substantial differences between SPOT1 and COL4 simulations: four IASI WV channels (2271 (1212.50 cm^{-1}), 2889 (1367.00 cm^{-1}), 2919 (1374.50 cm^{-1}) and 3244 (1455.75 cm^{-1})) and three AIRS WV channels (1455 (1334.60 cm^{-1}), 1627 (1427.23 cm^{-1}) and 1794 (1563.71 cm^{-1})). The corresponding weighting functions (Figures 7b and 7d) show that the part of the atmosphere sounded is quite different among the selected channels. The temporal variability of the brightness temperature differences between SPOT1 and COL4 is much greater than the instrument noise time variation. The amplitude of the instrument noise time variations ranges indeed from 0.015 K for the least variable channels, such as IASI channel 2271 and AIRS channel 1455, to 0.14 K or 0.19 K for the most variable ones, such as AIRS channel 1794 or IASI channel 3244 (not shown), whereas the var-

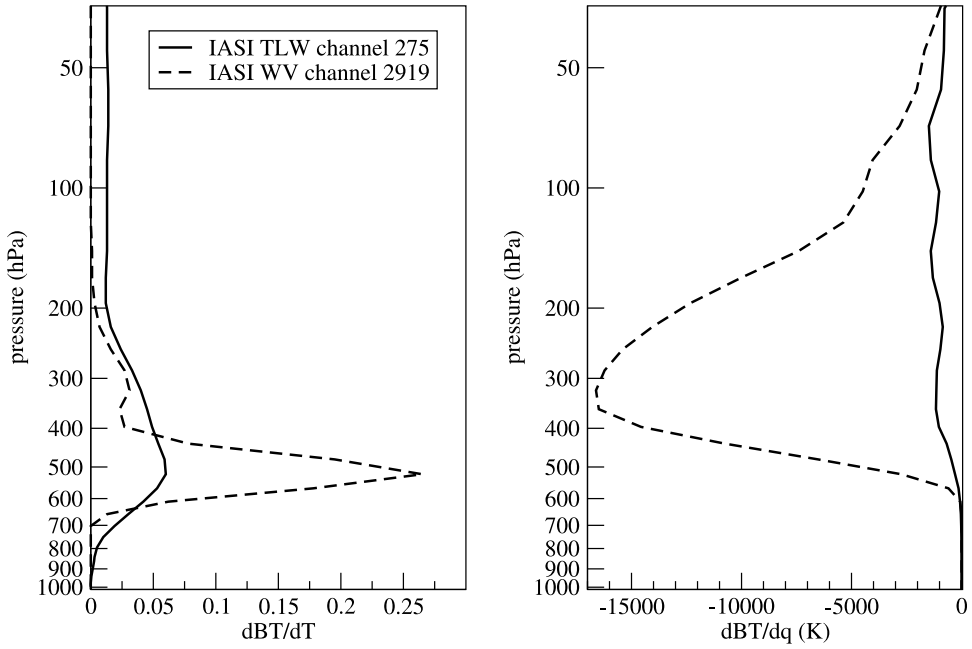


Figure 4. Jacobians in (left) temperature and (right) specific humidity for IASI TLW channel 275 (713.50 cm^{-1}) and IASI WV channel 2919 (1374.50 cm^{-1}), calculated for the observation of 2100 UTC, 6 September 2007, considered in section 5 (see Figure 12a for its location).

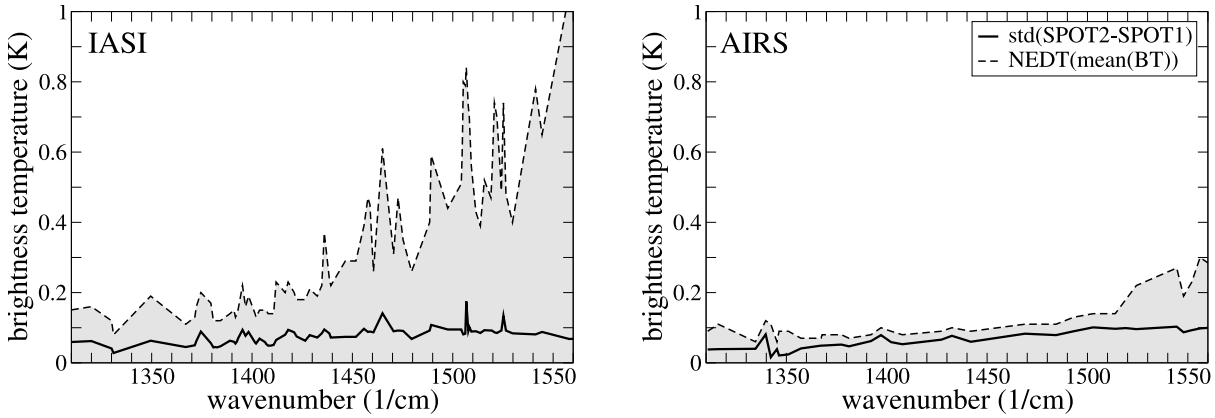


Figure 5. Same as Figure 3 but for brightness temperature differences between simulations with SPOT2 and SPOT1.

iations of the brightness temperature differences between SPOT1 and COL4 reach more than 0.4 K for the least variable channels and 1 K for the most variable ones (see Figure 7). The time variations are consistent among the different channels. The channels exhibit the largest differences for the same days (for instance 6, 16 or 24 September: the results for the 26 to 28 September period are not reliable as very few points were clear over the Mediterranean). Standard deviations as large as 0.5 to 1 K were obtained between SPOT1 and COL4 for these days for upper-level sounding channels.

3.3. Case Study: 24 September 2007

[26] As stated above, using SPOT1 instead of the current observation operator COL4 may substantially modify the simulated brightness temperatures of IASI and AIRS WV channels, especially for some specific days. Some particular fine-scale meteorological patterns present on these days can explain why averaging the model information over the whole spot leads to an estimation of the sounded atmosphere substantially different from that obtained using only four-point interpolation. We focus here on one of these situations, on 24 September 2007 at 2100 UTC, in order to determine the reasons for such large differences in the simulation of brightness temperatures.

[27] Figure 8 shows the simulation, at this date, of IASI 2919 water vapor channel brightness temperatures with COL4, M12 and SPOT1 together with the model relative humidity at 7000 m where the IASI channel 2919 peaks. The largest differences between the observation operator calculations are found in an area of fine-scale elongated filaments of moisture in the atmospheric layers observed. The same fine-scale filament pattern is evidenced in COL4 simulations, whereas it is partially smoothed with M12 by the 12-point average. With SPOT1, the spot-average filters out the fine-scale gradients of a size smaller than the observation resolution. So we can conclude that, in areas where strong fine-scale humidity gradients occur, COL4 and, to a lesser extent, M12 simulate brightness temperature structures that cannot be measured by the satellite. This can result in large differences between observed and first-guess equivalent radiances, which will lead to rejection of the observations or to production of too large analysis incre-

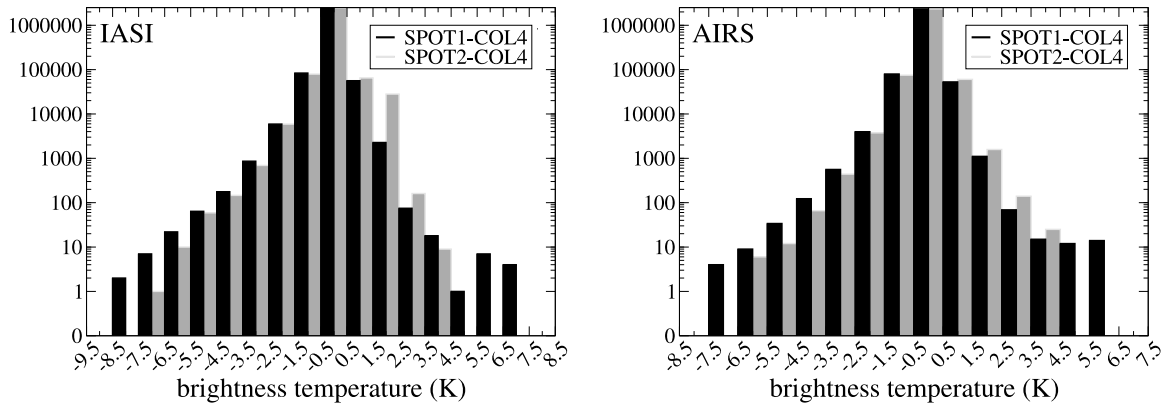


Figure 6. Histograms of the brightness temperature differences between simulations with SPOT1 and COL4 (black) and between simulations with SPOT2 and COL4 (grey) for IASI WV channel 2919 (1374.50 cm^{-1}) and AIRS WV channel 1627 (1427.23 cm^{-1}). The y axis is logarithmic.

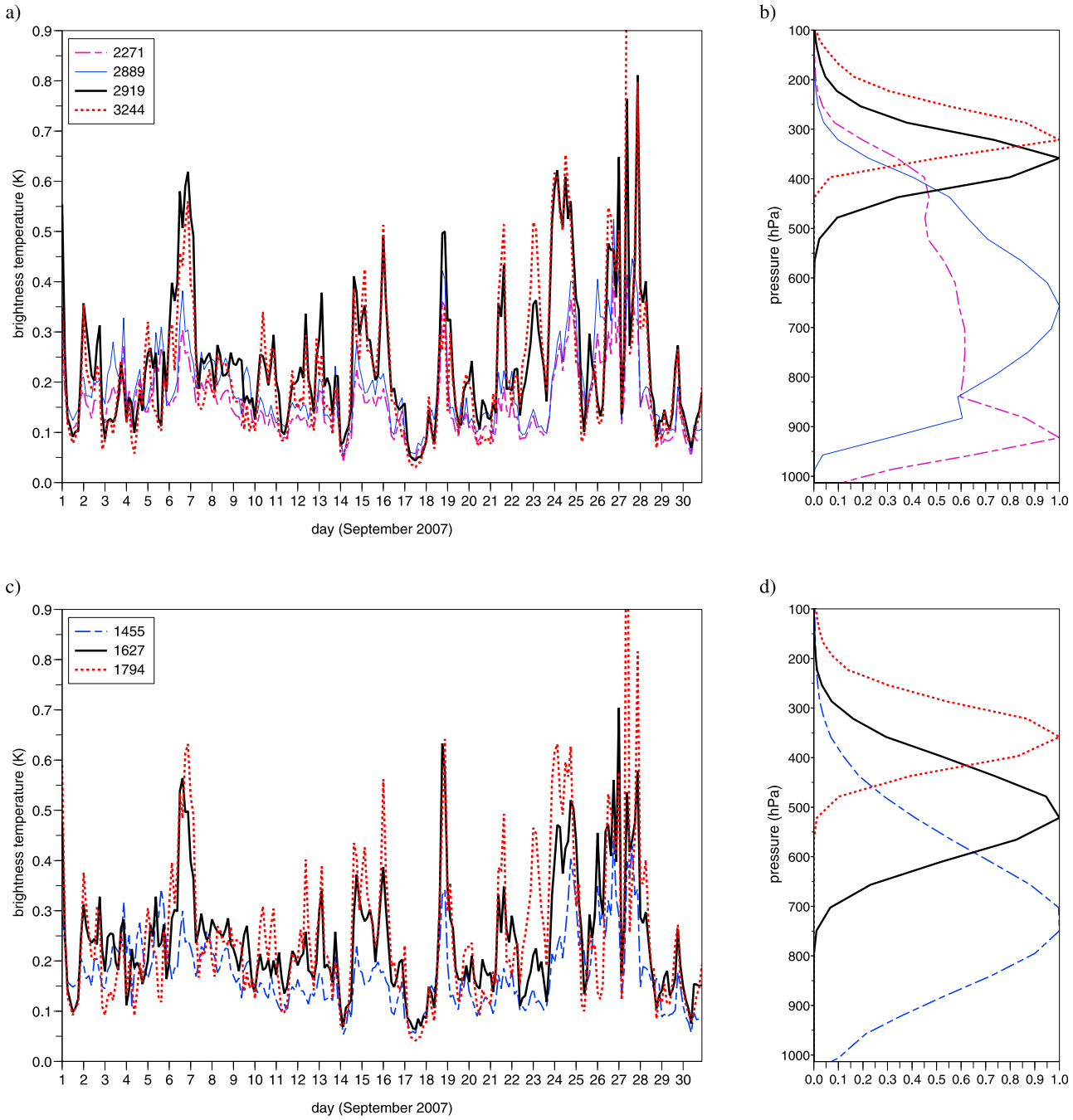


Figure 7. Time series over September 2007 of the standard deviation of the brightness temperature differences between simulations with SPOT1 and COL4 of (a) four IASI WV channels, 2271 (1212.50 cm^{-1}), 2889 (1367.00 cm^{-1}), 2919 (1374.50 cm^{-1}), and 3244 (1455.75 cm^{-1}), and (c) three AIRS WV channels, 1455 (1334.60 cm^{-1}), 1627 (1427.23 cm^{-1}), and 1794 (1563.71 cm^{-1}), in clear sky over western Mediterranean. The weighting functions of these channels (b) for IASI and (d) for AIRS were estimated with the U.S. Standard Atmosphere.

ments in the data assimilation process. Using SPOT1 might thus avoid these drawbacks.

[28] It is worth mentioning that the main brightness temperature differences between the three observation operators occur when very dry air conditions prevail. The other important differences encountered during September 2007 (situations identified on Figure 7) also arose in areas exhibiting spatial humidity variations in a very dry atmosphere. In such dry environment, moisture variations that are not particularly large in absolute terms become important relatively to the moisture content. Past studies such as those by *Berg et al.* [1999] for the High-resolution Infrared Radiation Sounder (HIRS) channel 12 and Special Sensor Microwave/Temperature 2 (SSM/T2) channel 2 or *Soden and Bretherton* [1996] for Geostationary Operational Environmental Satellites (GOES) channels also state that the

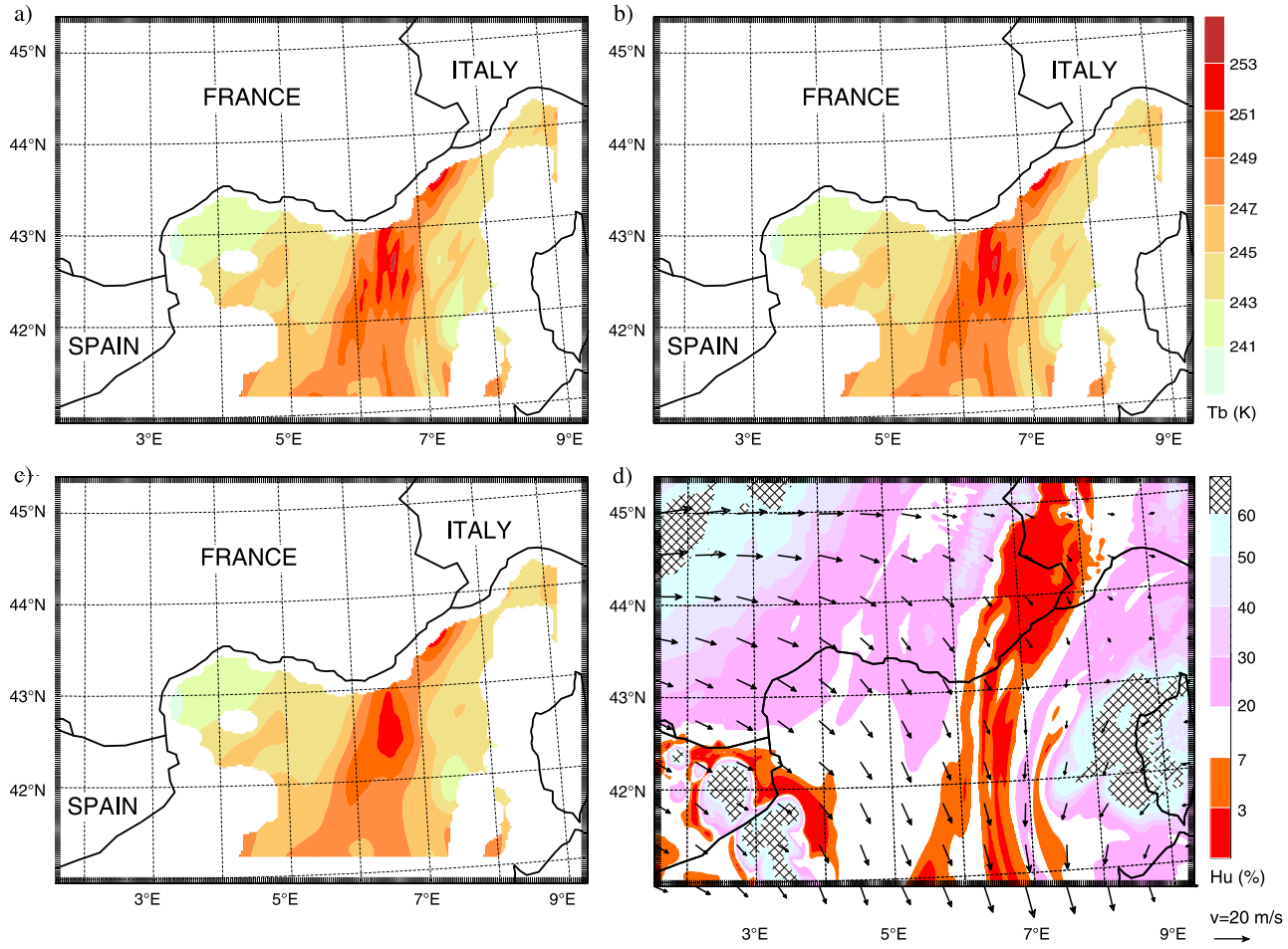


Figure 8. Simulation of IASI WV channel 2919 brightness temperature (in K, scale at the top right) at 2100 UTC, 24 September 2007 using (a) COL4, (b) M12, and (c) SPOT1; (d) 7 km ASL wind vector and relative humidity (in percent, scale at the bottom right). Observation operators are applied only over the sea and for cloud free areas. Wind arrows are drawn every 20 grid points with a scale in m/s given by the length of the arrow at the bottom right of Figure 8d.

sensitivity of the brightness temperature to atmospheric relative humidity increases with the dryness of the sounded air mass.

4. Comparison to the Observed Radiances

[29] So far, the impact of a change in the observation operator on the simulation of IASI and AIRS brightness temperature has proved substantial in the special case of a maximum scan angle and with specific meteorological conditions. This impact will now be assessed more generally for real observations in order to determine whether one observation operator provides model equivalent observations that are globally closer to the observations than another operator does.

[30] For this purpose, the screening stage of the 3 h AROME data assimilation cycle was performed throughout the month of September 2007. All clear IASI and AIRS observations situated over the Mediterranean Sea were thus simulated with COL4, M12 and SPOT1 from the 3 h AROME forecast which served as the first guess in the data assimilation cycle. An observation was considered to be over sea if, according to the model land-sea mask, all the model grid points covered by the observation spot were over the sea. For each observation, the "cloud detect" software of *McNally and Watts* [2003], which is used for the operational data assimilation of satellite observations, was applied here to distinguish between clear and cloudy channels.

[31] The departures between the observations (without bias correction) and their simulations with the three observation operators were calculated in order to see if they decreased when the new observation operators were used. As only one AIRS observation out of nine was provided, there were too few clear AIRS observations over the Mediterranean Sea to compute significant statistics. The following therefore focuses only on IASI observations. The number of clear IASI observations used is of about a hundred for the near-surface channels, and ranges from about 500 for the channels sounding around 950 hPa to more than 5000 for the highest peaking channels (near-tropopause channels).

[32] Figure 9 shows the histograms of the differences between $|obs - guess|$ obtained with SPOT1 or M12 and $|obs - guess|$ obtained with COL4 for the IASI WV channel

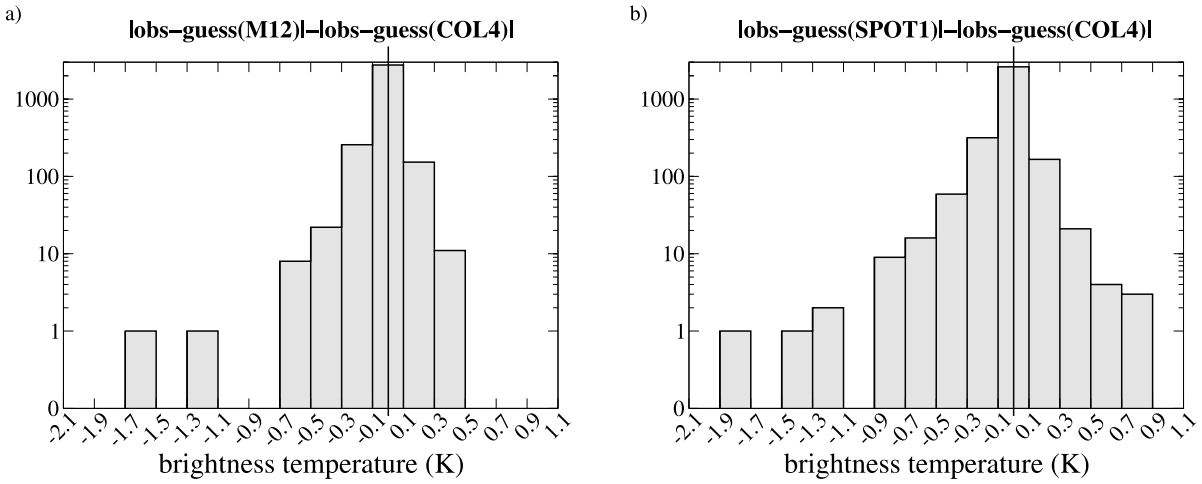


Figure 9. Histograms of the brightness temperature differences between |obs − guess| obtained with a new observation operator, (a) M12 or (b) SPOT1, and |obs − guess| obtained with COL4 for all IASI WV channel 2919 clear observations over the western Mediterranean in September 2007.

2919. There are more negative differences than positive differences, which indicates that SPOT1 and, to a lesser extent, M12 provide model equivalent brightness temperatures more often closer to the observations. This is also true for the other IASI WV channels while, in contrast, for TLW channels, the simulations with the new observation operators are more often slightly further from the observations and, for TSW channels, no clear signal can be found (not shown). Figure 9 also shows that the negative differences are stronger than the positive differences. This is corroborated for all IASI WV channels by Figure 10, which shows some statistics of these differences. This means that the decreases in the observation-guess departures of WV channels with the new observation operators are larger than the increases. The greatest decreases are as large as 1 K to 2.5 K with SPOT1 and 0.5 K to 1.9 K with M12 according to the WV channel considered. Still, in most cases, these decreases do not reach more than 0.5 K with SPOT1 and 0.25 K with M12. It is interesting to note that the size of the changes in the observation-guess departures does not depend on the magnitude of the initial departure obtained using COL4 (not shown). The improvements in the fit to WV channels are much larger than the degradations of the fit to TLW channels: with both SPOT1 and M12, the increases in the observation-guess departures of TLW channels remain in most cases below 0.04 K and do not reach more than 0.05 K to 0.15 K according to the TLW channel considered. The overall observation-guess biases of TLW channels are consequently only very slightly increased: only by 1% to 4% according to the channel considered (increase of less than 0.005 K for all TLW channels except the near-surface ones for which the increase can reach up to 0.016 K). This is consistent with

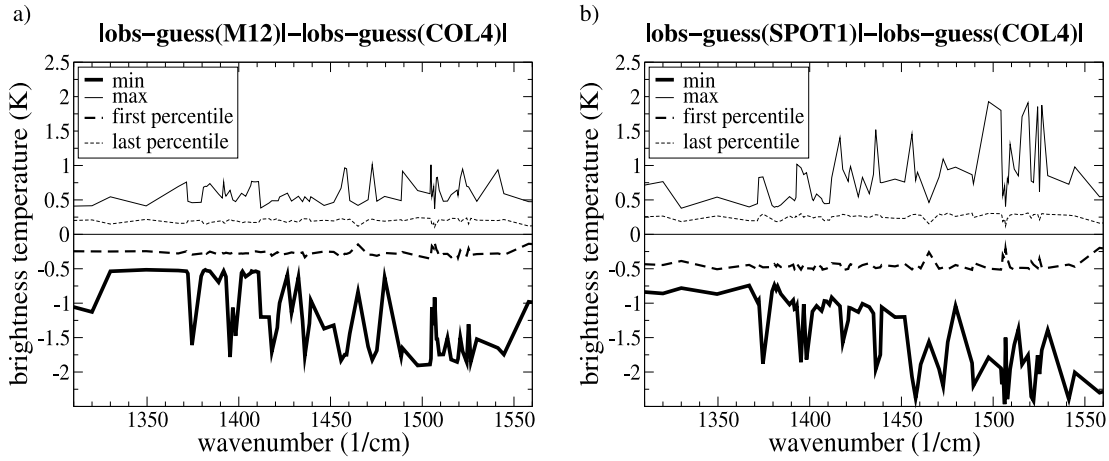


Figure 10. Minimum, maximum, first, and last percentiles of the brightness temperature differences between |obs − guess| obtained with a new observation operator, (a) M12 or (b) SPOT1, and |obs − guess| obtained with COL4 for clear IASI observations over western Mediterranean of September 2007 in all the WV channels studied displayed according to increasing wave number along the abscissa.

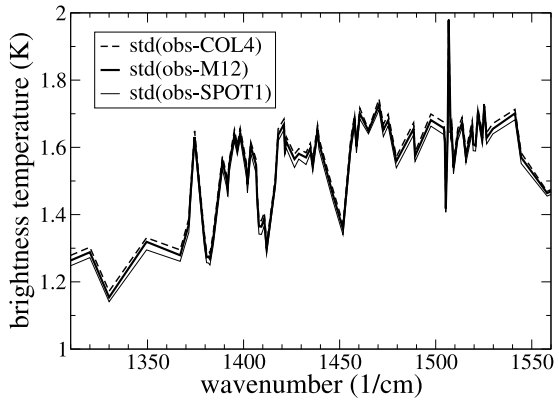


Figure 11. Standard deviation of the brightness temperature differences between observation and guess obtained with COL4, M12, and SPOT1 for clear IASI observations over the western Mediterranean of September 2007 in all the WV channels studied, displayed according to increasing wave number along the abscissa.

the results of section 3.2 showing, for off-nadir observation simulation, that the impact of the new observation operators in TLW channels is most of the time lower than the instrument noise and does not introduce new bias.

[33] Whereas the bias of the observation minus guess departures is almost unchanged in the WV channels (not shown) and very slightly increased in the TLW channels (as stated above), the standard deviation is slightly reduced for both WV and TLW channels (Figure 11 for WV channels) when the simulation is performed with SPOT1 and, to a lesser extent, M12. For TSW channels, these statistics remain mostly unchanged (not shown). All these results show that using the new observation operators should help to reduce the number of WV channel observations rejected during the assimilation process.

5. Single Observation Assimilation Experiment

5.1. Experiment Setup

[34] The impact of a change in the observation operator formulation is now examined in terms of analysis increment. We determine what impact the differences in the simulation of IASI or AIRS observations identified above may have on the final analysis. To do this, we performed assimilation experiments of a single IASI observation. In these experiments, we assimilated one specific channel of an IASI observation (see Figure 12 a for its location) at 2100 UTC, 6 September 2007. This observation was chosen because it was clear for the cloud-detect software, over the sea and in a region with mesoscale humidity gradients in a relatively dry environment (Figure 12). Quite large differences were thus found between its simulations with the various observation operators. Either a WV or a TLW IASI channel of this observation spot was assimilated. The WV and the TLW channels chosen were the ones having the strongest differences between the new observation operators and COL4: WV channel 3244 or TLW channel 275.

[35] Twin single observation assimilation experiments were performed using COL4 and M12. The differences between these twin experiments are not only their direct observation operator but also their jacobian and adjoint used to compute the analysis increments (model corrections applied to the first guess). With COL4, the elements of the jacobian are the partial derivatives of the brightness temperature with respect to the interpolated model variables whereas with M12, they are the derivatives with respect to the 12-point mean model variables. To assimilate an observation, the COL4 and M12 adjoints use the innovation (observation minus model equivalent simulation) and these

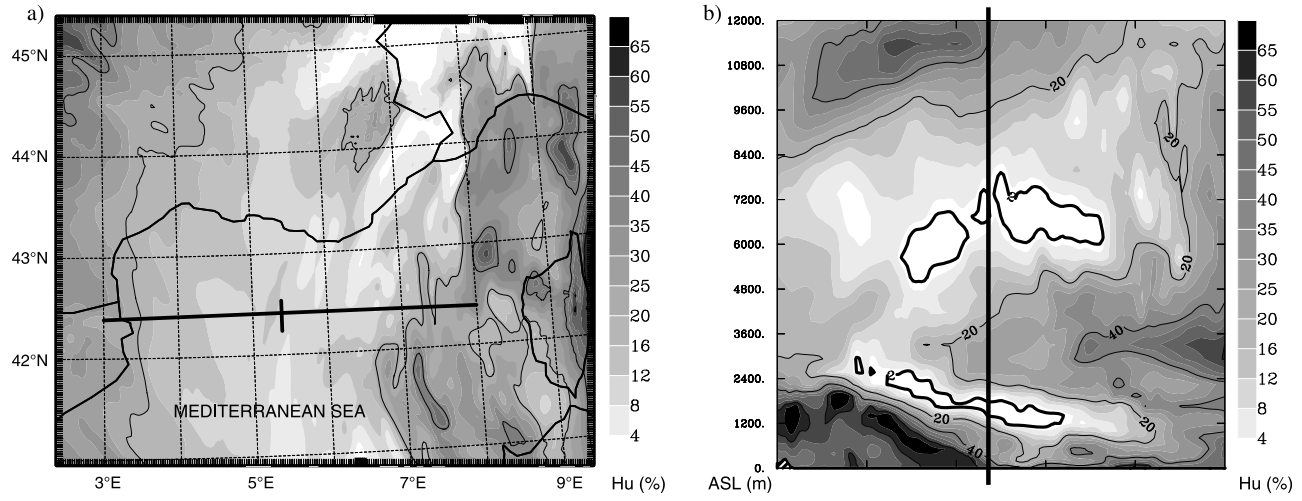


Figure 12. Location of the assimilated IASI observation (a) in the working area with the background relative humidity at 8700 m (grey scale on the right of Figure 12a and contours at 20% and 40%) and (b) in the vertical cross section along the line drawn in Figure 12a of the background relative humidity at 2100 UTC, 6 September 2007 (grey scale on the right of Figure 12b and contours at 2%, 20%, 40%, and 60%). The altitude is given in meters on the left of Figure 12b. The assimilated IASI observation is marked by the vertical line.

Table 3. Departures Between Observation and First Guess or Analysis Obtained With Both COL4 and M12 for the Two Single Observation Assimilation Experiments of IASI Channel 3244 and IASI Channel 275 Observations

		Observation-Guess		Observation-Analysis	
	Observation	COL4	M12	COL4	M12
Channel 3244	241.90 K	2.85 K	2.64 K	0.03 K	0.02 K
Channel 275	241.81 K	1.30 K	1.29 K	0.12 K	0.12 K

single Jacobians to determine the model correction to be applied. They distribute this correction over the four model grid points with weights proportional to their distance to the center of the observation spot or over the 12 model grid points with the same weight. The experiments assimilate a near-nadir observation having an observation spot that covers 18 grid points, so that SPOT1 and M12 are very similar and no additional assimilation experiment was carried out with SPOT1.

5.2. Results

[36] Table 3 gives the departures between the observation and the first guess and between the observation and the analysis for both twin assimilation experiments. As was to be expected according to the results of section 4, it shows that the first guess and the analyses are closer to the observations when the observation operator M12 is used. This is more obvious for channel 3244. According to the statistics over the whole of September 2007 discussed in section 4, using M12 improves the model equivalent estimation in TLW channel 275 by more than 0.012 K and in WV channel 3244 by more than 0.11 K for only about 10% of the observations. The reductions of the observation minus guess departure obtained here are therefore quite substantial with 0.21 K for channel 3244 and 0.012 K for channel 275.

[37] The impact of the observation operator formulation on the humidity and temperature analyses was evaluated. The differences between the analyses using M12 or COL4 were negligible for temperature and for both channels as the maximum difference was less than 0.006 K. Figure 13 displays the differences in terms of relative humidity between the analyses obtained using COL4 and M12 for the two channels. The results are perturbed in the area with less than 2% relative humidity as the drying induced by the assimilation of the observation was limited. Still, the differences are clearly weak, less than 0.4%. These differences result from slightly weaker increments due to smaller observation minus first-guess departures obtained using M12 (see Table 3). Modifying the first-guess error covariance matrix or the observation errors did not modify the conclusion of this study (not shown). All the single observation assimilation experiments performed here show that even the largest differences in the model equivalent estimation obtained using different observation operators have almost no impact on the final analysis of temperature and humidity.

6. Conclusions and Outlook

[38] This study has aimed to evaluate the impact of the observation operator formulation on the model-equivalent simulation for the assimilation of satellite infrared radiances in AROME. In particular, the question of the aggregation of the model information when the model grid mesh is 1 order smaller than the satellite spot has been addressed. Three new observation operators, M12, SPOT1 and SPOT2, have been defined in order to come closer to the way the satellite measurement is achieved. To characterize the sounded atmosphere, they aggregate part or all of the model information contained within the satellite field of view with the same weight instead of using only the single model column

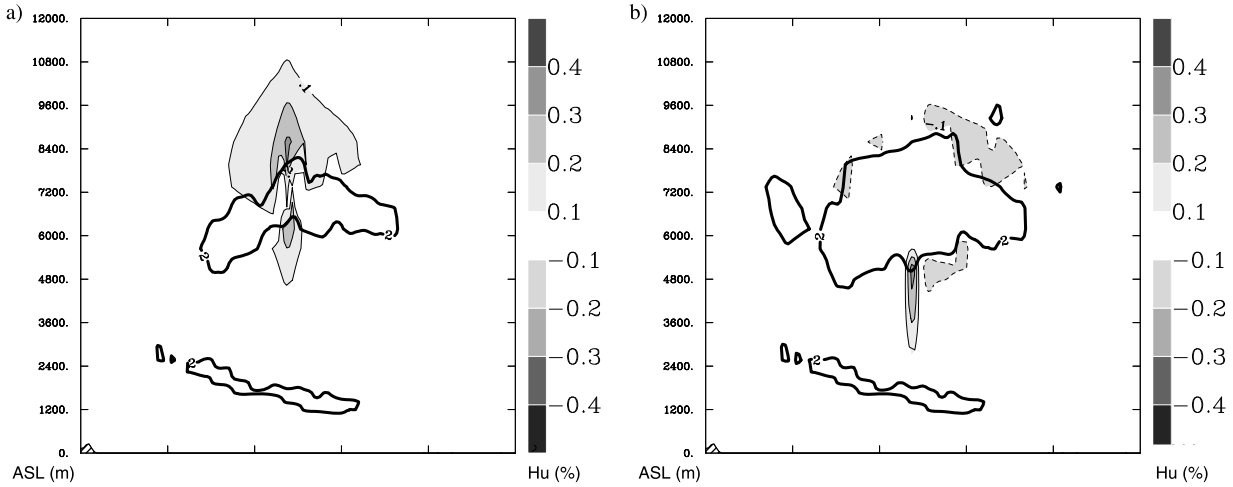


Figure 13. Vertical cross section along the axis drawn in Figure 12a of the differences in relative humidity between the analyses formed by the assimilation with M12 and COL4 (a) of the IASI WV channel 3244 observation and (b) of the IASI TLW channel 275 observation, at 2100 UTC, 6 September 2007. The grey scale in percent is to the right of each plot. Contours delimit the areas with less than 2% relative humidity in the analysis.

situated at the center of the observation spot as is done in the current observation operator (COL4).

[39] A comparison among the four observation operators was made over a 1 month period using the 3-hourly AROME analyses. Substantial differences were found between the various operators for water vapor channels only (peaking under 340 hPa and also above 800 hPa for AIRS) and where fine-scale horizontal humidity gradients occurred in the sounded layer. With the new observation operators, these fine-scale model humidity variations that cannot be measured by the satellite because they have a smaller size than the observation resolution, are filtered out. The greatest differences occur for very dry observed layers. Aggregating model information contained within the satellite spot before (SPOT1) or after (SPOT2) applying the radiative transfer led to no significant differences in most cases. For an operational use, the observation operator that consumes less computing time, SPOT1, can be given preference without substantial loss of accuracy. The M12 operator (aggregating model information over 12 grid points) provided results that are intermediate between SPOT1 and COL4.

[40] The model equivalent observations simulated by the various operators have then been compared to real observations by performing the screening step of the AROME data assimilation cycle. The observation operators were thus applied to the 3 h forecast that served as first guess in the AROME rapid update cycle. The comparison was carried out over a full 1 month period to reduce the weight of the model errors in this comparison. Using the new observation operators proved to reduce (up to 2 K) most of the time the departures between observations and simulations of WV channels while their impact was neutral for TSW and slightly negative for TLW channels. The standard deviation of these departures was slightly reduced for WV channels but also for most TLW ones. These improvements in WV channels were independent of the departure between observation and simulation with COL4. So, aggregating the model information over the whole spot rather than taking only the model information at the center of the satellite spot improves the estimation of the model equivalent observation by filtering out the model fine-scale patterns that are not detected by the satellite because of its lower spatial resolution. A better estimation of the model equivalent observation should lead to fewer observation rejections and reduce incorrect analysis increments due to representativeness errors.

[41] To evaluate the impact of the observation operator formulation in terms of analysis increment, single observation assimilation experiments were performed. Situations and channels having shown the largest differences between the observation operators were selected. They showed, however, that even large differences in the simulation of IASI observations have a negligible impact on the analysis of temperature and relative humidity.

[42] Even though the impact on the analysis increment of the differences between the model equivalent estimations from the various observation operators seems neutral, the modifications of the observation operator may have a larger influence on the whole data assimilation suite (bias correction, cloud detection, selection of assimilated observations, …). Future work will evaluate the overall impact of such modifications of the observation operator on a complete data assimilation cycle. The impact may be enhanced because of cycling and because the modifications of the observation operator affect many elements of the assimilation process, which in turn modify the analysis.

[43] **Acknowledgments.** The authors thank Lydie Lavanant for providing the IASI and AIRS weighting functions and for her help on the IASI products. Pascal Brunel is also recognized for his assistance on the Jacobian calculations.

References

Auligné, T., A. P. McNally, and D. Dee (2007), Adaptive bias correction for satellite data in a numerical weather prediction system, *Q. J. R. Meteorol. Soc.*, *133*(624), 631–642.

Aumann, H. H., et al. (2003), AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 253–264.

Berg, W., J. J. Bates, and D. L. Jackson (1999), Analysis of upper-tropospheric water vapor brightness temperatures from SSM/T2, HIRS and GMS-5 VISSR, *J. Appl. Meteorol.*, *38*, 580–595.

Berre, L. (2000), Estimation of synoptic and mesoscale forecast error covariances in a limited area model, *Mon. Weather Rev.*, *128*(3), 644–667.

Blumstein, D. (2005), IASI on METOP: On ground calibration of the FM2 instrument, paper presented at 14th International TOVS Study Conference, VCS, Beijing.

Blumstein, D. (2007), In-flight performance of the infrared atmospheric sounding interferometer (IASI) on METOP-A, *Proc. SPIE Int. Soc. Opt. Eng.*, *6684*, 12 pp., doi:10.1117/12.734162.

Brenot, H., V. Ducrocq, A. Walpersdorf, C. Champollion, and O. Caumont (2006), GPS zenith delay sensitivity evaluated from high-resolution numerical weather prediction simulations of the 8–9 September 2002 flash flood over southeastern France, *J. Geophys. Res.*, *111*, D15105, doi:10.1029/2004JD005726.

Brousseau, P., F. Bouttier, G. Hello, Y. Seity, C. Fischer, L. Berre, T. Montmerle, L. Auger, and S. Malardel (2008), A prototype convective-scale data assimilation system for operation: The AROME-RUC, *Tech. Rep. 68*, EUMETNET, Norrköping, Sweden.

Bubnová, R., G. Hello, P. Bénard and J.-F. Geleyn (1995), Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system, *Mon. Weather Rev.*, *123*(2), 515–535.

Caumont, O., and V. Ducrocq (2008), What should be considered when simulating Doppler velocities measured by ground-based weather radars?, *J. Appl. Meteorol. Climatol.*, *47*(8), 2256–2262.

Caumont, O., V. Ducrocq, G. Delrieu, M. Gosset, J.-P. Pinty, J.-P. du Châtelet, H. Andrieu, Y. Lemâitre, and G. Scialom (2006), A radar simulator for high-resolution nonhydrostatic models, *J. Atmos. Oceanic Technol.*, *23*(8), 1049–1067.

Cayla, F. (2001), L'interféromètre IASI, un nouveau sondeur satellitaire à haute résolution, *Meteorol. Ser. 8*, *32*, 23–39.

Chalon, G., F. Cayla, and D. Diebel (2001), IASI: An advanced sounder for operational meteorology, paper presented at 52nd Congress, IAF, Toulouse, France.

Collard, A. (2007), Selection of IASI channels for use in numerical weather prediction, *Q. J. R. Meteorol. Soc.*, *133*(629), 1977–1991.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth (1994), A strategy for operational implementation of 4D-Var using an incremental approach, *Q. J. R. Meteorol. Soc.*, *120*(519), 1367–1387.

Dee, D. (2004), Variational bias correction of radiance data in the ECMWF system, paper presented at Workshop on Assimilation of High Spectral Resolution Sounders in NWP, ECMWF, Reading, U. K.

Diebel, D., F. Cayla, and T. Phulpin (1996), IASI mission rationale and requirements, *CNES/EUMETSAT Tech. Rep. IA-SM-0000-10-CNE/EUM–4b*, 35 pp., EUMETSAT, Darmstadt, Germany.

Ducrocq, V., D. Ricard, J.-P. Lafore, and F. Orain (2002), Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity fields, *Weather Forecasting*, *17*, 1236–1256.

Ducrocq, V., F. Bouttier, S. Malardel, T. Montmerle, and Y. Seity (2005), Le projet AROME, *Houille Blanche*, *2005-2*, 39–43.

Edwards, D. P. (1992), GENLN2: A general line-by-line atmospheric transmittance and radiance model, *NCAR Tech. Note NCAR/TN-367 +STR*, 147 pp., NCAR, Boulder, Colo.

Elliott, D. A., T. S. Pagano, and H. H. Aumann (2006), The impact of AIRS spatial response on channel-to-channel and multi-instrument data analyses, *Proc. SPIE Int. Soc. Opt. Eng.*, *6296*, 30 pp.

Fischer, C., T. Montmerle, L. Berre, L. Auger, and S. E. Stefanescu (2005), An overview of the variational assimilation in the ALADIN/France numerical weather prediction system, *Q. J. R. Meteorol. Soc.*, *131*(613), 3477–3492.

Guidard, V., C. Fischer, M. Nuret, and A. Dziedzic (2006), Evaluation of the ALADIN 3D-Var with observations of the MAP campaign, *Meteorol. Atmos. Phys.*, *92*, 161–173.

Kim, G., and A. P. Barros (2002), Space-time characterization of soil moisture from passive microwave remotely sensed imagery and ancillary data, *Remote Sens. Environ.*, *81*, 393–403.

Kleespies, T. J. (2009), Microwave radiative transfer at the sub-field-of-view resolution, paper presented at 16th International TOVS Study Conference, Petrobras, Angra dos Reis, Brazil.

Lafore, J.-P., et al. (1998), The Meso-NH Atmospheric Simulation System. Part I: Adiabatic formulation and control simulations, *Ann. Geophys.*, *16*, 90–109.

McNally, A. P., and P. D. Watts (2003), A cloud detection algorithm for high-spectral resolution infrared sounders, *Q. J. R. Meteorol. Soc.*, *129*(595), 3411–3423.

Merlin, O., A. G. Chehbouni, Y. H. Kerr, E. G. Njoku, and D. Entekhabi (2005), A combined modeling and multi-spectral/multi-resolution remote sensing approach for disaggregation of surface soil moisture: Application to SMOS configuration, *IEEE Trans. Geosci. Remote Sens.*, *43*(9), 2036–2050.

Nuissier, O., V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin (2008), A numerical study of three catastrophic precipitating events over western Mediterranean region (southern France). Part I: Numerical framework and synoptic ingredients, *Q. J. R. Meteorol. Soc.*, *134*(630), 111–130.

Pagano, T. S., H. H. Aumann, S. L. Gaiser, and D. T. Gregorich (2002), Early calibration results from the Atmospheric InfraRed Sounder (AIRS) on Aqua, *Proc. SPIE Int. Soc. Opt. Eng.*, *4891*, 76–83.

Pagano, T. S., H. H. Aumann, D. E. Hagan, and K. Overoye (2003), Prelaunch and in-flight radiometric calibration of the Atmospheric InfraRed Sounder (AIRS), *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 265–273.

Radnóti, G., et al. (1995), The spectral limited area model ARPEGE-ALADIN, *WMO TD 699*, pp. 111–118, WMO, Geneva.

Reichle, R., D. Enthekabi, and D.B. McLauglin (2001), Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach, *Water Resour. Res.*, *37*(9), 2353-2364, doi:10.1029/2001WR000475.

Saunders, R., and P. Brunel (2005), RTTOV_8_7 user's guide, *NWPSAF-MO-UD-008*, 45 pp., EUMETSAT, Darmstadt, Germany.

Soden, B. J., and F. P. Bretherton (1996), Interpretation of TOVS water vapor radiances in terms of layer-average relative humidities: Method and climatology for the upper, middle and lower troposphere, *J. Geophys. Res.*, *101*(D5), 9333–9343.

Susskind, J., C. D. Barnet, and J. M. Blaisdell (2003), Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, *IEEE Trans. Geosci. Remote Sens.*, *41*(2), 390–409.

Yan, X., V. Ducrocq, G. Jaubert, P. Brousseau, P. Poli, C. Champollion, C. Flamant, and K. Boniface (2009), The benefit of GPS zenith delay assimilation to high-resolution quantitative precipitation forecasts: A case-study from COPS IOP 9, *Q. J. R. Meteorol. Soc.*, *135*(644), 1788–1800, doi:10.1002/qj.508.

V. Ducrocq, F. Duffourg, and G. Jaubert, Groupe de Météorologie de Moyenne Echelle, Centre National de Recherches Météorologiques, Météo-France, 42 av. Gaspard Coriolis, F-31057 Toulouse CEDEX 1, France. (fanny.duffourg@meteo.fr)

N. Fourrié and V. Guidard, Groupe de Modélisation pour l'Assimilation et la Prévision, Centre National de Recherches Météorologiques, Météo-France, 42 av. Gaspard Coriolis, F-31057 Toulouse CEDEX 1, France.

Q. J. R. Meteorol. Soc. (2006), **132**, pp. 275–295 doi: 10.1256/qj.05.31

Impact study of the 2003 North Atlantic THORPEX Regional Campaign

By NADIA FOURRIÉ*, DAVID MARCHAL, FLORENCE RABIER, BERNARD CHAPNIK and GERALD DESROZIERS
CNRM/GAME, CNRS and Météo-France

(Received 28 February 2005; revised 13 July 2005)

SUMMARY

An experiment took place during autumn 2003 with the aim of testing the feasibility of an operational targeting of observations over the North Atlantic Ocean in the context of the international programme THORPEX. The purpose of this paper is to evaluate the impact of these additional observations in the French operational model ARPEGE during the last three weeks of the campaign. Results are shown for large regions over and around the North Atlantic Ocean and for specific verification areas. Over Europe, the addition of observations is slightly beneficial for the forecast, mostly in the low troposphere over wide areas and above 100 hPa. However, the impact of extra data is more significant but also more mixed for the dedicated verification areas: they are case, forecast-range and level dependent. In addition, the information content is studied with the Degrees of Freedom for Signal (DFS) for the evaluation of the observation impact on the analysis of one case of December 2003. Firstly, the variations of the DFS have been illustrated in a simplified data assimilation system. It has been found for that case that satellite data have the most important global contribution to the overall analysis, especially the humidity sensitive infrared radiances. For the conventional data, the wind measurements of the aircraft and from the geostationary satellites are the most informative. For the targeted area, the data from aircraft and the dropsondes have the largest DFS. It has been noted that the DFS of the dropsondes located in the sensitivity maximum is larger than the other one even if there is no link between the DFS and the forecast. However, the impact of the dropsondes grows with respect to the forecast range and leads to an improvement of the forecast for this case.

KEYWORDS: Dropsondes Forecast score Observing-system experiment Sensitive areas

1. INTRODUCTION

The forecast of some midlatitude events, such as rapid cyclogenesis, remains difficult and as a result may have dramatic societal impact. Forecast errors may come from errors in the initial conditions. The meteorologically sensitive areas are known to be atmospheric regions where small errors in the initial conditions can develop and lead to major forecast failures (Rabier *et al.* 1996). A way to reduce the uncertainty of the forecast over a specific region is to add observations in these sensitive areas. Many methods can be used for the determination of the sensitive area where observations should be added. This is the aim of the targeting. Different targeting techniques have been developed: some involve either the adjoint of the linearized version of the forecast model or of the assimilation scheme, while others manipulate ensembles of forecasts. Majumdar *et al.* (2002) provide a detailed comparison of various targeting techniques. The Fronts and Atlantic Storm-Track Experiment was the first opportunity to test the concept of adaptive observation (Joly *et al.* 1999) followed by The North Pacific Experiment (Langland *et al.* 1999) and the Winter Storm Reconnaissance Program (WSRP) (Szunyogh *et al.* 2000) that is currently operational in the United States over the North Pacific Ocean during each winter.

The international World Meteorological Organisation programme THORPEX is a response to the challenge of improving the skill of high impact weather forecasts. This programme plans to contribute to the design and demonstration of interactive forecast systems, which include the concept of targeted observations. It also performs THORPEX observing-system tests and THORPEX Regional field Campaigns (TReCs)

* Corresponding author: CNRM/GAME, CNRS and Météo-France, 42 av. G. Coriolis, 31057 Toulouse Cedex 1, France. e-mail: Nadia.Fourrie@meteo.fr

to test and evaluate experimental remote sensing and *in situ* observing systems, and when feasible, demonstrate their impact on weather forecasts.

The Atlantic TReC (ATReC) that took place during autumn 2003 (13 October to 14 December 2003) aimed to test the feasibility of a quasi-operational targeting of observations over the North Atlantic Ocean and was sponsored by EUCOS*. A description of the operation plan by Stringer and Truscott (2004) can be found on the THORPEX web site†.

In this paper, we present the evaluation of the additional observation impact on the French operational forecast system for the last three weeks of the ATReC. This assessment is made through a partial re-analysis of the campaign and by the use of a diagnosis of the observation information content for one case-study. The latter aims to test the impact of the observations on the analysis whereas the former provides the possibility of testing the forecast quality. The paper is arranged as follows. In the second section, the general framework of the study, including the assimilation and forecast scheme and the available observations, are presented. Overall results on the forecast quality and forecast scores on the verification areas are discussed in section 3. Section 4 presents an evaluation of the observation impact on the analysis and on the forecast for one case. The conclusions of this study are discussed in section 5.

2. Experimental framework

(*a*) *Assimilation and forecast framework*

The forecast model is the French ARPEGE‡ operational model based on a stretched geometry on the globe (Courtier *et al.* 1991). The assimilation framework used for this study was the operational 4D-Var assimilation scheme of Météo France during the 2003 ATReC. A description of the 4D-Var developed at Météo France can be found in Janiskovà *et al.* (1999) and Geleyn *et al.* (2001). It is based on the minimization of a cost function that includes the distance between the evolution of the state variable during the six-hour assimilation window and the observations. The operational analysis was performed with a T358 spectral truncation on the stretched sphere with a stretching factor of 2.4. This leads to a resolution varying between 20 and 50 km over the North Atlantic Ocean. The conventional observations consist of surface observations (from surface stations (SYNOP) and buoys (BUOY)) and upper-air data. The latter come either from AIRcraft REPort (AIREP), wind observations (PILOT) or radiosondes. The radiosondes profiles are launched from terrestrial stations (TEMP) and ships (SHIP) at 00 and 12 UTC. In addition, satellite observations assimilated in ARPEGE are atmospheric motion vectors from geostationary satellites (SATOB) and polar-orbiting satellite radiances. At the date of the ATReC, only infrared radiances from the High-resolution InfraRed Sounder (HIRS) and from the Advanced Microwave Sounding Unit A (AMSU-A) were assimilated in the ARPEGE analysis.

(*b*) *Additional observations*

The last three weeks of the experimental campaign (from 00 UTC 27 November until 18 UTC 14 December 2003) have been studied. This period has been chosen because of interesting meteorological situations and the amount of targeting flights. Default verification areas were determined over the North Atlantic region and are described

* EUMETNET (European meteorological network) Composite Observing System.

† http:www.wmo.int/thorpex/pdf/atlantic_ob_system.pdf

‡ Action de Recherche Petite Echelle Grande Echelle.

TABLE 1. ADDITIONAL ATREC DATA NUMBER WITH RESPECT TO THE ANALYSIS TIME AND FOR THE WHOLE PERIOD RANGING FROM 00 UTC 27 NOVEMBER TO 18 UTC 14 DECEMBER 2003. NUMBERS IN BRACKETS REPRESENT THE NUMBER OF PROFILES.

Time (UTC)	Total	TEMP	DROP	SHIP	AIREP	BUOY
00	11 605	0 (0)	4 716 (11)	0 (0)	4 298	2 591
06	87 162	74 593 (421)	336 (3)	5 931 (45)	3 818	2 484
12	8 444	0 (0)	3 522 (34)	0 (0)	2 249	2 673
18	95 160	62 190 (336)	15 631 (137)	9 236 (65)	5 832	2 471
Total	202 301	136 783 (759)	24 205 (215)	15 167 (110)	15 947	10 269

See text for details of data source names.

in the appendix (Table A.1). Our study included 22 targeting dates out of the 45 ones of the whole period (Table A.2). The mean targeting forecast range is 42-hour but it can vary from 30-hour to 78-hour. The default verification areas are mostly west Atlantic (WATLAN) or northern Europe (NEUROPE).

The additional data collected during the campaign were *in situ* observations and remote-sensed data. *In situ* data consisted of dropsondes launched from aircraft (DROP), radiosonde ascents from terrestrial stations (TEMP) and ships (SHIP) at 06 and 18 UTC, drifting buoys data (BUOY), data from aircraft flights (AIREP) activated over specific regions. The remote-sensed data were rapid scan winds from Meteosat and GOES geostationary satellites (SATOB). These data were distributed on the Global Transmission Service and were assimilated operationally in the ARPEGE system except for the rapid scan wind. Table 1 shows the distribution of the additional data of ATReC with respect to the analysis time and the global number of extra observations. The extra observations represent 6% of the observations assimilated in ARPEGE at 18 and 06 UTC. The dropsondes were mostly launched for the 18 UTC analysis time and the radiosondes were added at 06 and 18 UTC in the same proportion. The additional AIREP and BUOY observations are less numerous, whose quantity of data are similar to those of the SHIP.

The operational run took into account the additional observations from the ATReC. A re-analysis of the three weeks from 00 UTC 27 November until 18 UTC 14 December 2003 period without any added data (called hereafter WITHOUT) was performed in order to test the impact of the data from the campaign.

In the WITHOUT experiment, the AIREP were removed over specific regions depending on the date, that were provided by S. Stringer from the ATReC operation team in March 2004. Identified ATReC BUOY have been removed together with all the dropsondes over the area 25°N–75°N, 95°W–30°E. The TEMP and the SHIP were omitted for 06 and 18 UTC over the same region. Therefore, we will compare the forecasts provided by the operational run (called WITH experiment) and by this re-analysis.

3. OVERALL RESULTS

(*a*) *Large verification regions*

The quality of both forecasts has been quantified with respect to the analyses and radiosondes for many parameters and levels. Only the comparisons to the analyses are shown in this section. These analyses were operational and have taken into account a maximum of data, including the supplementary data of the ATReC.

Figure 1 shows the root-mean square (r.m.s.) of the geopotential errors at 850 and 500 hPa for all the forecasts run from the four analysis times (00, 06, 12, 18 UTC)

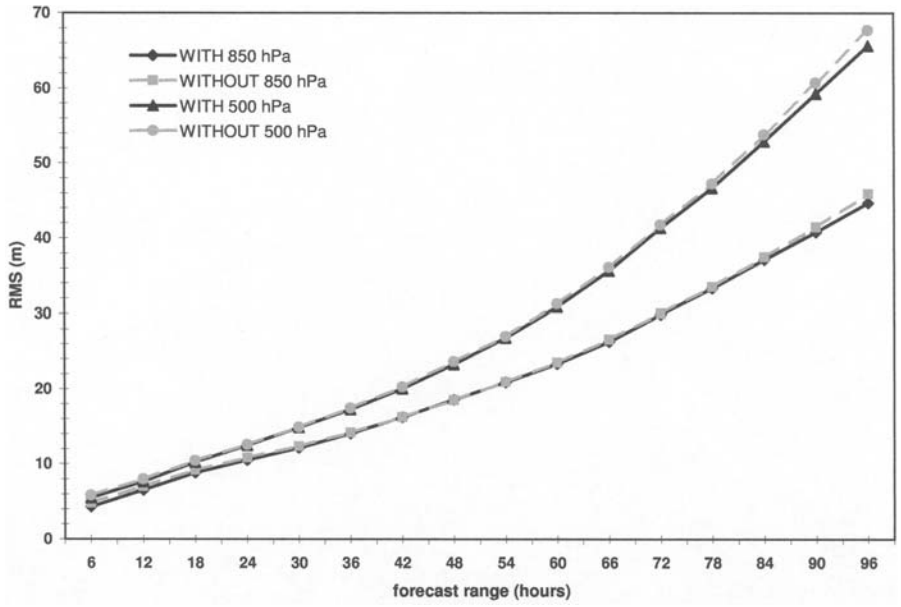


Figure 1. Root-mean-square error for geopotential at 850 and 500 hPa for the forecasts from the WITH and WITHOUT experiments (see text) over the ATReC area (25°N–75°N; 90°W–30°E) for the period ranging from 00 UTC 27 November to 18 UTC 14 December 2003.

and for the period ranging from 00 UTC 27 November to 18 UTC 14 December 2003, over the ATReC area. The inclusion of the campaign data improves the short-range (6-hour, 18-hour) forecast very slightly (it is also true when the comparison is made with respect to the radiosondes, not shown). It is neutral for the 24- to 72-hour forecast range and it improves the long-range forecast (beyond 72-hour). For the 18 UTC analysis time, the small improvement starts from 42 hours (not shown) whereas it starts from 60 hours at 00 and 06 UTC and from only 84 hours for 12 UTC. The slight benefit on the forecasts is more important at 00 and 18 UTC. This improvement is due to the inclusion of observations at 18 UTC and to the better quality of the resulting background at 00 UTC.

The results have been examined for all the default verification regions and we present the most significant results here. Over the WATLAN area (Table A.1), the impact of the assimilation of the supplementary data is positive in terms of r.m.s. from the 24-hour forecast range at 850 hPa (Fig. 2(a)) and everywhere in the troposphere (not shown).

On the contrary, a significant improvement is observed in the first steps of the forecast at 100 hPa (Fig. 2(b)) and in the atmospheric layer included between 200 and 50 hPa over the CEUROPE area. An explanation of this improvement can be given by Fig. 3, which displays the comparison of the temperature of the radiosondes with the one of the background and of the analysis for the WITH and WITHOUT experiments. The r.m.s. and the bias of the background departure are smaller for the WITH experiment in the atmospheric layer included between 300 and 70 hPa than for the WITHOUT experiment. This improvement is mainly due to the addition of the radiosonde observations in the analyses. To check this finding, an experiment similar to the WITHOUT experiment, but including the AIREP observations in the assimilation, has been performed and the resulting forecasts show that the AIREP observations do not

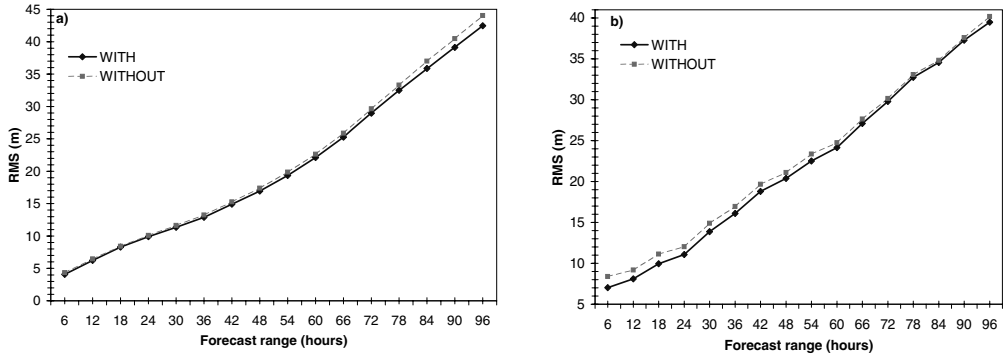


Figure 2. Root-mean-square error of the WITH and the WITHOUT experiments for (a) geopotential at 850 hPa over the WATLAN area (30°N–65°N; 15°W–35°E) and (b) geopotential at 100 hPa over the CEUROPE area (30°N–65°N; 15°W–35°E) for the period ranging from 00 UTC 27 November to 18 UTC 14 December 2003. See text for further explanation.

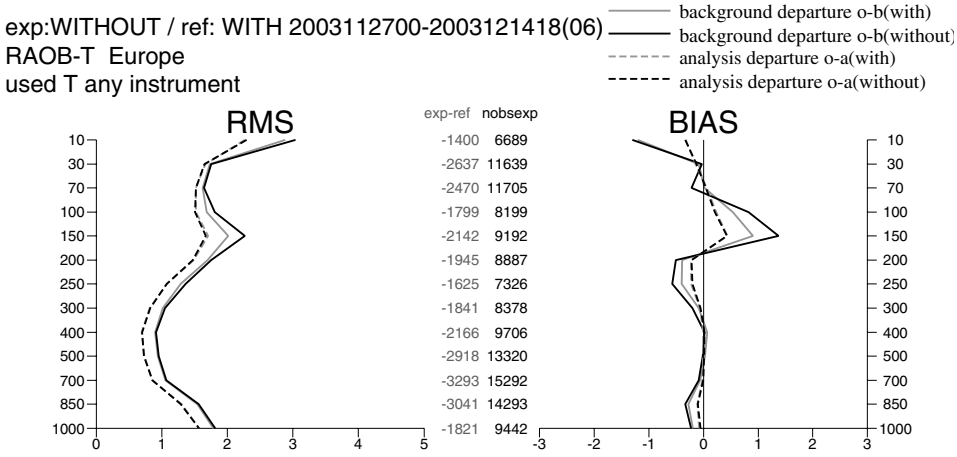


Figure 3. Comparison of the radiosonde temperature to the background and the analysis with respect to the pressure level over EUROPE for the period ranging from 00 UTC 27 November to 18 UTC 14 December. In the centre, the total observation number of the WITH experiment (in black) and the difference between the WITHOUT and the WITH experiments (in grey) are given. The 'with' and 'without' analysis lines coincide on the graphs. See text for details.

contribute to the improvement of the short-range forecast in the low stratosphere over Europe.

The small impact found in the Fig. 1 comes from averaging the r.m.s. over the large verification area of the ATReC that acts like a filter. The statistical relevance of the results has been studied with the statistical *t*-test of Student. It allows an assessment of whether the means of two groups are statistically different from each other with a certain degree of confidence. We estimated that both experiments were different when the degree of confidence is equal or superior to 90%. The statistical relevance of the benefit over the ATReC area is very weak. The study of the time series highlights that the positive impact comes from the period from 6–7 December (not shown). Before this period (from 18 UTC 5 December to 0 UTC 7 December) a lot of dropsondes were deployed for each analysis time and the forecasts were globally improved for each forecast range and each level. The relevance of the benefit of the targeting for

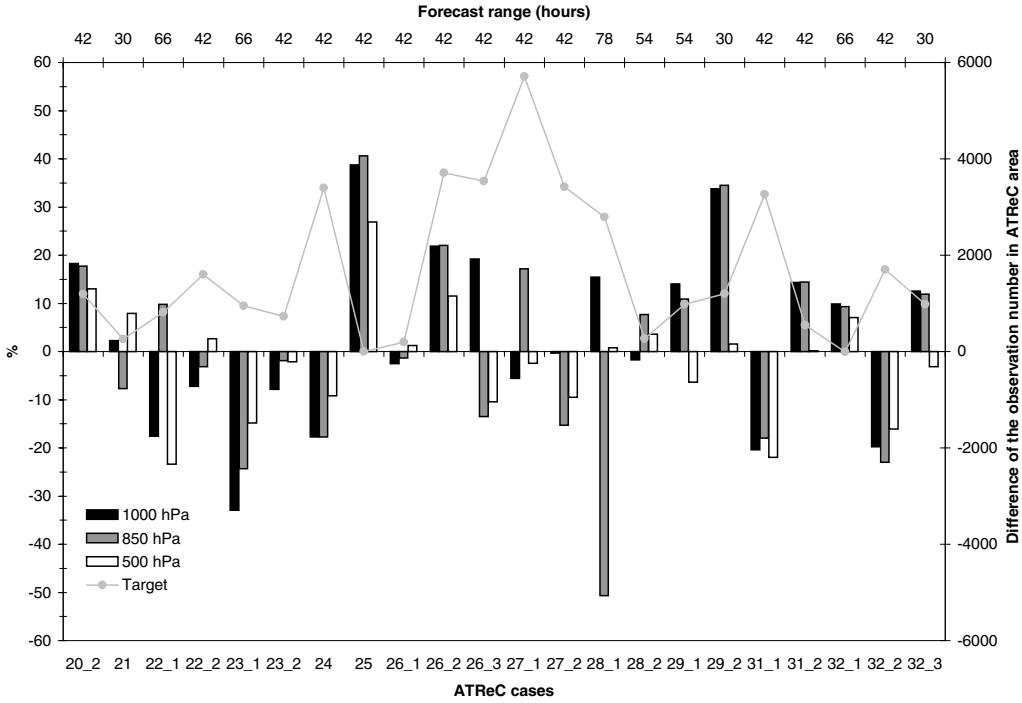


Figure 4. Percentage of the difference of the geopotential r.m.s. error at 1000, 850 and 500 hPa between the WITHOUT and the WITH experiments on the verification area at the required time (grey bars) and number of additional observations in the ATReC region (grey line) for the 22 studied ATreC cases. See Table A.2 for the description of the cases.

the WATLAN domain at 850 hPa for the 36-hour forecast range is relative (only 85%) because it is related to medium range. On the contrary, the improvement in the CEUROPE is significant with a degree of confidence of 90% until the 54-hour forecast range.

(*b*) *Verification area*

The principle of targeting is to improve the quality of the forecast on a verification area at a specific time. This aspect is more precisely studied in this subsection.

Figure 4 displays the percentage of improvement or degradation of the forecast (difference of r.m.s. error of geopotential between WITHOUT and WITH experiments divided by the error of the WITHOUT experiment) at 1000, 850 and 500 hPa, respectively, for the dedicated verification time and area. The forecast range is also given as it depends on the studied case. Table A.2 gives the list of the verification areas for the 22 studied cases. One observation time may correspond to several verification areas and times. A positive score points out that the additional data are beneficial (the WITH forecast r.m.s. error is lower than the WITHOUT one). For the 1000 hPa level, 9 cases out of 22 have been improved by more than 10% by the addition of data and 5 have been degraded by more than 10%, 8 cases are quite neutral. At 850 hPa, the cases are more evenly distributed: 8 are improved, 7 are neutral and 7 are degraded. At 500 hPa, the number of neutral cases increases to 14, 3 improved cases and 5 degradations are observed. The strongest degradation at 850 hPa (50%) corresponds to a long-term forecast, i.e. 78-hour (case 28_1). As is well known, the linear assumption, on which some

techniques of targeting are based, is only valid up to the 48-hour forecast range. The determination of the sensitive areas with adjoints of the linear models might be false, due to the long-term range of the targeting. Moreover, it is difficult to establish a relation between the score of the forecast and the number of supplementary data in the targeted area or in the ATReC region: the largest positive impact (case 25) is obtained with no additional observation in the sensitive area. For the other cases, there is no correlation between the observation number and the forecast score. Let us remember that we used a background for the analysis in which additional data have been included. The results show the cumulative effects of assimilating the additional observations in successive assimilation cycles. The impact of the additional data is contained in the analysis but also in the resulting background. No attempt was made to differentiate between the effect of additional observations in the sensitive area at the initial forecast time and the effect of improvements in the background forecast in this location.

(*c*) *Conclusions*

The larger the verification area, the smoother the results, because they are averaged over a larger domain. The results of this study agree with the results found by Desroziers *et al.* (2003) for the FASTEX re-analysis. The only significant positive impact of the additional ATReC observations is observed over Europe for the geopotential and the temperature forecast in the high troposphere and in the lower stratosphere up to the 54-hour forecast range. For the verification area, the results are more contrasted: at 1000 hPa, 40% of the cases are positive, the same proportion are quite neutral and 23% are negative, but at 500 hPa 63% are neutral, 23% are degraded and only 14% are improved. However, the improvement observed at 42-hour forecast range in the lower atmosphere close to 850 hPa over the verification areas due to the additional observation is 90% statistically significant.

4. CASE-STUDY: IMPACT OF THE OBSERVATIONS ON THE ANALYSIS AND ON THE FORECAST

(*a*) *Framework*

The previous results were valid for the forecast. In this section, the information content brought by the observations in the analysis is now studied. The analysis $\mathbf{x}^{\mathrm{a}}$ obtained from the background $\mathbf{x}^{\mathrm{b}}$ and the observations $\mathbf{y}^{\mathrm{o}}$ in the linear case is:

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

or

$$\mathbf{x}^{\mathrm{a}} = \mathbf{x}^{\mathrm{b}} + \mathbf{K}\mathbf{d}, \tag{2}$$

where $\mathbf{B}$ represents the background-error covariance matrix, $\mathbf{R}$ the observation-error covariance, $\mathbf{H}$ the linearized observation operator of $\mathcal{H}$ that allows the computation of the model equivalents in the space of observations, $\mathbf{K}$ is the so-called Kalman gain matrix ($\mathbf{K} = (\mathbf{B}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{R}^{-1}$) and $\mathbf{d}$ the innovation vector ($\mathbf{d} = \mathbf{y}^{\mathrm{o}} - \mathcal{H}(\mathbf{x}^{\mathrm{b}})$). To measure the impact of observations on analysis, the Degrees of Freedom for Signal (DFS) has been used (Rodgers 2000; Fisher 2003; Cardinali *et al.* 2004). It is algebraically defined by:

$$\mathrm{DFS} = \mathrm{Tr}\,(\mathbf{KH}). \tag{3}$$

The trace of the $\mathbf{KH}$ matrix quantifies the gain in information brought by the observations. The DFS characterizes how the assimilation system uses the observations

to pull the signal from the background. Chapnik *et al.* (2005) used a method introduced by Desroziers and Ivanov (2001) for the evaluation of the trace of the **KH** matrix, known only as an operator. It consists in performing two analyses (one normal analysis and one perturbed analysis $\mathbf{x}^{a*}$ with perturbed observations $\mathbf{y}^{o*}$). The quantity $(\mathbf{y}^{o*} - \mathbf{y}^{o})^{T}\mathbf{R}^{-1}\mathbf{H}(\mathbf{x}^{a*} - \mathbf{x}^{a})$ provides an approximation to Tr (**HK**). A partial DFS, associated with a particular subset of observations can also be determined, if these observations are not correlated with the rest of observations. However, if one wants to get the DFS at a single observation location, many analyses with randomized perturbations have to be performed in order to reduce the uncertainty of the estimated analysis variance reduction. It has been found that ten perturbations give a good estimation of the error reduction even at a single location. The DFS has thus been averaged as proposed in the paper by Chapnik *et al.* (2005). In the following, the averaged DFS over ten perturbations is presented as the DFS.

(*b*) *Example with a simplified assimilation scheme*

In order to illustrate the theoretical properties of the DFS with respect to the various parameters, the simplified assimilation scheme of Desroziers and Ivanov (2001) and Chapnik *et al.* (2005) has been used. It consists of a circular domain with perimeter of 40 000 km; 200 observations are equally spaced over the domain and the background is taken to be null. The true signal is generated in spectral space with the horizontal correlation and variance in the $\mathbf{B}_t$ matrix ($\mathbf{x}^{t} = \boldsymbol{\epsilon}^{b} = -\mathbf{B}_t^{1/2}\boldsymbol{\zeta}$, with $\boldsymbol{\zeta} \in \mathcal{N}(0, \mathbf{I}_n)$, where n is the dimension of $\mathbf{x}$, and a Gaussian correlation ($c(r) = \exp(-r^2/2L_B^2)$) for two points separated by the distance r. Assuming a diagonal correlation matrix for observation error $\mathbf{R}_t = \sigma_o^2\mathbf{I}_p$, where p is the number of observations, and a linear observation operator **H** (the operator is perfect and gives the true value of the field at observation location), the observation vector $\mathbf{y}^{o}$ is designed as $\mathbf{y}^{o} = \mathbf{H}\mathbf{x}^{t} + \mathbf{R}_t^{1/2}\boldsymbol{\zeta}$, with $\boldsymbol{\zeta} \in \mathcal{N}(0, \mathbf{I}_p)$. In our experimental framework, $\sigma_o = \sigma_b = 1$, the correlation length $L_B =$ 300 km and the truncation is 200 that leads to a distance of around 100 km between two grid points.

As indicated by its definition, the DFS depends on the structure function associated with the background error through **B**, on the ratio between the observation error and the background error and on the density of the observations through **H**. Figure 5 illustrates the influence of the background correlation length L_B on the DFS. When the correlation length decreases, the DFS tends to increase. It will be easier for the observations to modify the analysis independently of the other observations.

The second aspect is highlighted by Fig. 6, that displays the variation of the DFS with respect to the σ_o/σ_b ratio. The DFS decreases with the increase of this ratio, that is when the observation error is larger than the background one. An observation has more potential to modify the background when its associated error is small. However, σ_b is not explicitly computed in ARPEGE and one does not have this information straightforwardly in the ARPEGE data assimilation scheme.

When the number of observations increases (Fig. 7), the DFS increases but the DFS per observation decreases. Each individual contribution will diminish even if the global contribution increases.

As shown by

$$\mathbf{H}\mathbf{x}^{a} = \mathbf{H}\mathbf{x}^{b} + \mathbf{H}\mathbf{K}(\mathbf{y}^{o} - \mathcal{H}(\mathbf{x}^{b})), \tag{4}$$

the **HK** matrix also quantifies the sensitivity of the analysis to the observations:

$$\partial_{y^{o}}\mathbf{H}\mathbf{x}^{a} = (\mathbf{H}\mathbf{K})^{T}. \tag{5}$$

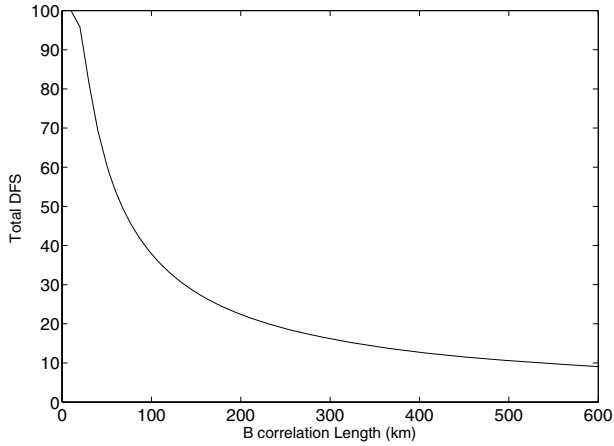


Figure 5. Influence of the background correlation length on the DFS.

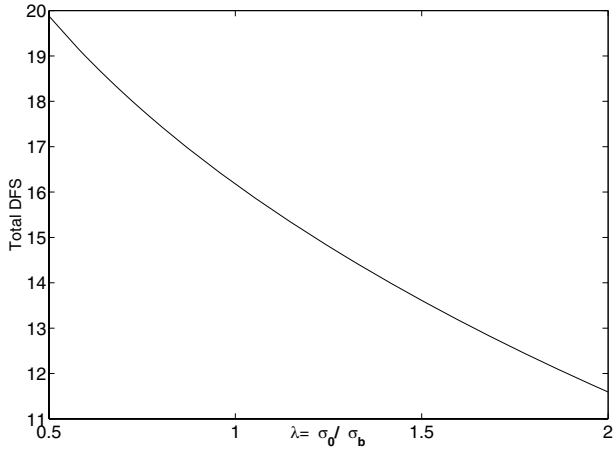


Figure 6. Influence of the σ_o/σ_b ratio on the DFS (see text).

As mentioned earlier, the DFS gives a measure of the self-sensitivity:

$$\mathrm{DFS} = \mathrm{Tr}\,(\mathbf{HK}) = \mathrm{Tr}\,(\partial_{y^o}\mathbf{H}\mathbf{x}^a). \tag{6}$$

The self-sensitivity is a measure of the impact of the observations on their analysed equivalents. This diagnostic was implemented and studied by Cardinali *et al.* (2004) in the ECMWF operational data assimilation scheme. It does not give any indication of the actual average analysis variance reduction.

This small discussion will allow results presented hereafter to be interpreted. It has been shown that the DFS per observation decreases when the considered observation number increases and when the observations are close to each other.

(*c*) *Over the globe*

As the method is quite central processor unit cost expensive, only one case-study is shown. Another case has also been studied and the results found are similar to the ones

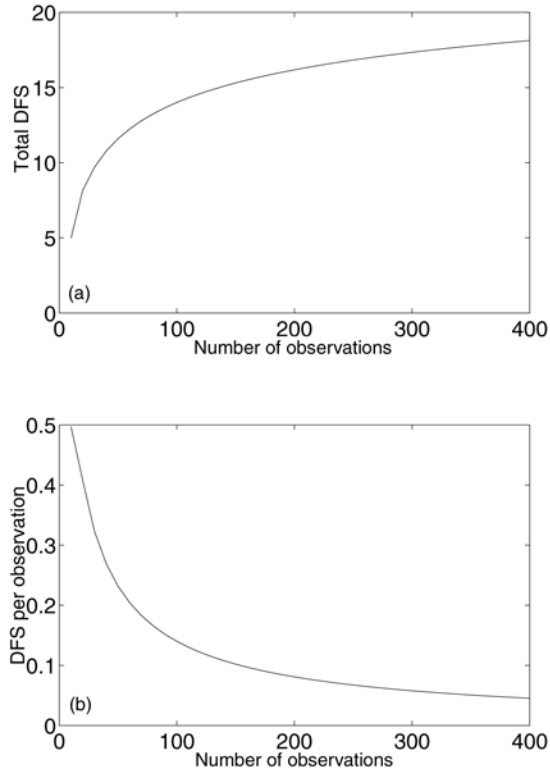


Figure 7. Influence of the number of observations on (a) the DFS and (b) the DFS per observation.

presented in this paper. It is the targeting case TREC_26 of the 18 UTC 5 December. This deployment aimed to improve the forecast of a severe snowstorm that affected the east coast of the North America two days later.

(i) *DFS with respect to the observation types.* Figure 8 shows the total DFS of each observation type assimilated over the globe in the ARPEGE analysis for 18 UTC 5 December for the WITH experiment. For this graphic, TEMP means the dropsondes in addition to the radiosondes. Satellite data, i.e. HIRS, SATOB and AMSU-A, but also AIREP data, are the most informative observations for the analysis. In contrast to Cardinali *et al.* (2004), AMSU-A data have a relatively weak information content for this case, and the HIRS radiances have a large information content. This different result may be explained by the ratio between HIRS and AMSUA observation numbers. In our case, AMSU-A observations represent only 1.6 times the HIRS number, whereas in the European Centre for Medium-Range Weather Forecasts example they are six times more numerous.

(ii) *DFS with respect to the assimilated parameters.* Figure 9 shows the DFS computed with the summation of the temperature and the wind measurements for the various observation types with respect to four atmospheric layers. These layers represent the low, mid- and high troposphere and the top of the atmosphere.

For the temperature observations, the AIREP bring the most important part of the information, especially in the 400–100 hPa layer (due to the flight levels). This dominance comes from the analysis time where fewer radiosondes have been launched

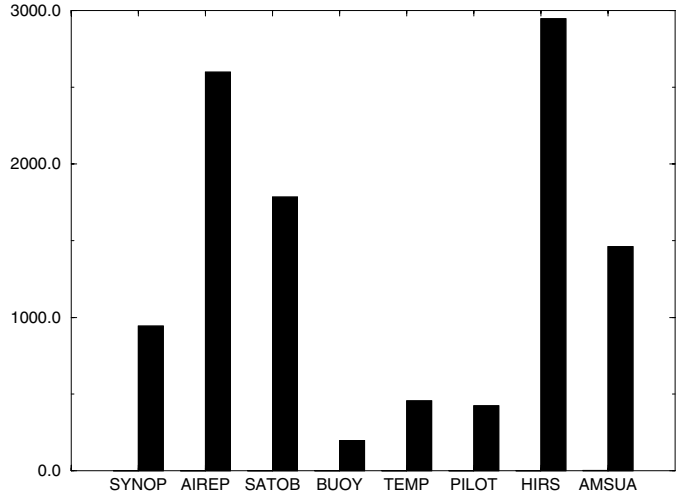


Figure 8. DFS of observation types for the analysis of 18 UTC 5 December 2003.

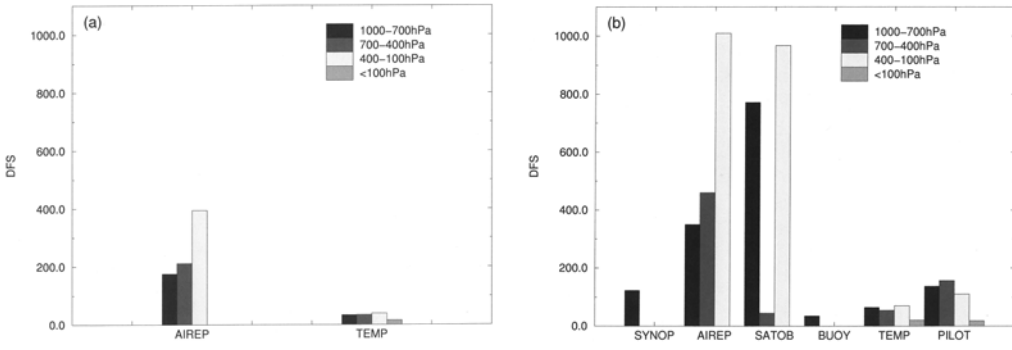


Figure 9. DFS of the observation types for the analysis of 18 UTC 5 December. (a) DFS with respect to the temperature measurements, and (b) DFS with respect to the wind measurements.

than at the 00 and 12 UTC analysis times. For the wind data, the AIREP of the 400–100 hPa layer and the SATOB give the most important contribution to the DFS (around 900). The SATOB DFS for low-troposphere wind observations is also large (about 800). This proportion between the DFS computed with the temperature observations and the one obtained with the wind measurements can be explained by the fact that the structure functions are shorter for the wind than for the temperature and tend to increase the partial DFS.

The two other measurement types that are assimilated in ARPEGE, but in a smaller proportion, have been studied. The assimilated specific-humidity measurements in ARPEGE come only from the TEMP for the conventional observations. Their DFS represents a small part of the global DFS (120 for all the levels, not shown). The major part of the DFS for geopotential is given by SYNOP observations (around 800, not shown) whereas the BUOY bring only 150 to the DFS.

Table 2 presents the partial averaged DFS (equal to the total DFS of an observation type divided by the number of observations in this observation type) and the corresponding number of observations (in brackets). The partial averaged DFS gives an indication

TABLE 2. PARTIAL AVERAGED DFS FOR EACH ASSIMILATED PARAMETER WITH RESPECT TO FOUR ATMOSPHERIC LAYERS (hPa). NUMBERS IN BRACKETS REPRESENT THE NUMBER OF OBSERVATIONS.

Parameter	Observations	1000–700	700–400	400–100	100–0
Temperature	AIREP	0.027 (6 437)	0.035 (6 111)	0.040 (9 851)	
	TEMP	0.052 (640)	0.055 (631)	0.064 (618)	0.056 (270)
Wind	SYNOP	0.059 (2 084)			
	AIREP	0.027 (13 100)	0.037 (12 504)	0.051 (19 842)	
	SATOB	0.131 (5 862)	0.178 (254)	0.111 (8 730)	0.072 (2)
	BUOY	0.098 (360)			
	TEMP	0.073 (876)	0.082 (660)	0.070 (1 000)	0.027 (772)
	PILOT	0.066 (2 070)	0.052 (3 018)	0.099 (1 112)	0.035 (542)
Humidity	TEMP	0.072 (543)	0.094 (575)	0.134 (120)	
Geopotential	SYNOP	0.064 (13 257)	0.127 (80)		
	BUOY	0.124 (1316)			

of the potential impact of an observation type of its capacity to modify the analysis. For the averaged DFS associated with the temperature, the TEMP observations bring the largest part of the information for all levels. For wind observations, SATOB of the 700–400 hPa layer have the greatest individual potential to modify the analysis. It is also noteworthy that the averaged DFS of the TEMP observations is larger for the wind than for the temperature. Again, this may come from the structure functions that are slightly shorter for the wind than for the temperature. This tends to increase the capability of a wind observation to modify the analysis and therefore its DFS. As shown in section 4(b), larger structure functions for the temperature errors imply a decrease of the DFS of the temperature observations. Another explanation may be given by the ratio σ_o/σ_b, but σ_b is not explicitly known in ARPEGE. The averaged DFS is larger for specific humidity than for temperature and of the same order as for wind. Again, this may be related to shorter structure functions for humidity than for temperature. For geopotential observations, the averaged DFS is of the same order of magnitude for BUOYS and SYNOP observations in the 700–400 hPa layer. One might think that they are remote observations and are thus less constrained by other observations. Note that from Table 2 the partial averaged DFS generally increases when the corresponding observation numbers decrease. That is logical since sparse observations have more individual impact on the analysis than dense observations.

(iii) *DFS with respect to channels.* Figure 10 shows the DFS and the averaged DFS of the assimilated channels from the National Oceanic Atmospheric Administration satellites. In contrast to the other conventional observations, the shape of the averaged DFS diagram is quite similar to the total DFS one. This is due to the number of satellite observations that are evenly localized over the globe. However, the mean DFS of the AMSU-A channels is smaller and only the AMSU-A channel 5 is relatively large. The maximum of its weighting function is close to 850 hPa. HIRS 11 and HIRS 12 have the most important DFS, either total or averaged (over 0.4). These channels are sensitive to the humidity in the mid- and low troposphere. Their weighting function maxima are at 700 and 500 hPa, respectively. It is interesting to note that the observation error is twice as large for the HIRS channel 12 as the one for HIRS 11. The surface channels HIRS 14 and HIRS 7 are in second position for information content. AMSU-A channel 9 sounding in the high atmosphere has also a relatively large DFS that can be explained by a large observation number (about 5400). Other HIRS and AMSU-A channels have a

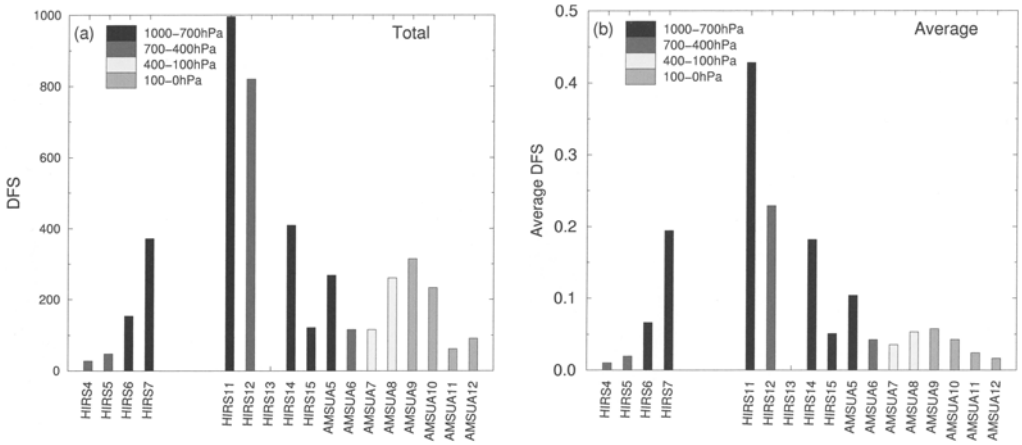


Figure 10. (a) Total DFS and (b) averaged DFS for the NOAA radiances for the analysis of the 18 UTC 5 December.

weak averaged information content. These results are in good agreement with the study of Cardinali *et al.* (2004), who found that HIRS 11 and 12 have a large self-sensitivity and AMSU-A channels a weak average influence.

(*d*) *Over the targeted area*

To complement the results obtained in the previous section, the DFS limited to the observations placed inside the sensitive area (30°N–50°N; 85°W–60°W) have also been computed and are shown in Fig. 11. The AIREP and the DROP observations have the most important contribution to the analysis. The DFS of the dropsondes of the low troposphere represents almost 50% of the DFS of the atmosphere while the DFS of the AIREP is twice as large in the high troposphere as in the other layers (not shown). The DFS has also been computed over the area (30°N–60°N; 90°W–55°W) proposed by the operation team to activate the specific AIREP observations dedicated to the ATReC. The ratio of the DFS between the different observation types does not change significantly. The DFS of the AIREP is now 399 whereas the dropsonde one is 143.

Figure 12 shows the sensitive area computed with total energy singular vectors of the operational ARPEGE model (in blue shading) and the one proposed by the operations of the ATReC (orange box). This latter is centred over the sensitive area of ARPEGE. Let us remember that the verification time is 42 hours for the studied case. To study the geographical distribution of the partial DFS of each observation, Fig. 12 also displays the value of the DFS for the dropsondes located off the eastern coast of North America and the radiosondes over the continent and over sea at 32°N and 64°W, and for the AIREP observations included in the area (30°N–60°N; 90°W–55°W). The AIREP observations are mostly located over the secondary maximum of the sensitive area and the dropsondes have been launched over the main maximum. The DFS associated with each AIREP measurement is very weak (less than 1) and the one for the dropsondes is over 2. Note that this difference can be explained by the computation of the DFS: only three observations (temperature, meridional and zonal winds) are taken into account for the AIREP, whereas for the dropsondes and the radiosondes these three measurements and the specific humidity are taken into account for each vertical level and summed at each observation location. Therefore, one cannot compare Figs. 12(a) and (b) in terms of DFS. When focusing on the averaged DFS, the contribution of the AIREP is much

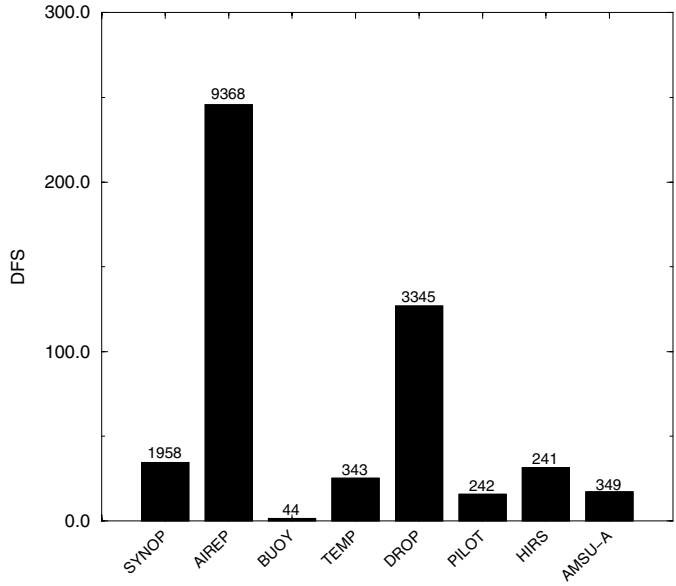


Figure 11. Total DFS over the targeted area (30°N–50°N; 85°W–60°W) for 18 UTC 5 December, and associated observation number.

smaller (0.024) than the one of the dropsondes (0.041). It is interesting to note that the radiosondes have an averaged contribution of 0.075 and a large DFS (over 10, for three radiosondes out of four). This result may be explained by the fact that they are remote from the other observations, whereas for the other observation types the observations are closer.

Even though the DFS is not related to any forecast field or sensitivity field, it is worth noticing that the DFS of the north-eastern dropsondes are larger and that they are located in the main maximum of the sensitivity. It has been checked that the background-error variance is also larger for the observations located in the north-eastern part of the flight. That may explain this increase in the DFS due to smaller σ_o/σ_b ratio as shown in section 4(b). In addition, the DFS of the transversal flight dropsondes, that are very close together, is lower than the others.

(*e*) *Impact of the observations on the forecast quality*

Now the impact of the observations on the forecast is addressed for this particular case. Following the results of the previous paragraph, two denial experiments have been performed. In the first one, the dropsondes have been removed from the analysis, in the second one, all the AIREP observations have been removed over the targeted area (30°N–50°N; 85°W–60°W).

Figure 13 shows the difference of analyses and 78-hour forecasts between the WITH and the without dropsondes experiments and between the WITH and the without AIREP experiments for the geopotential at 500 hPa. The impact of the AIREP is twice as large as the DROP data one at 500 hPa and for the other levels (not shown) for the analysis. Note that their impacts are of opposite signs. However, these differences do not propagate in the forecast in the same manner. The extent and the intensity of the WITH minus without dropsonde difference grow up with time while the ones of the WITH minus without AIREP are more localized. This result may be explained by

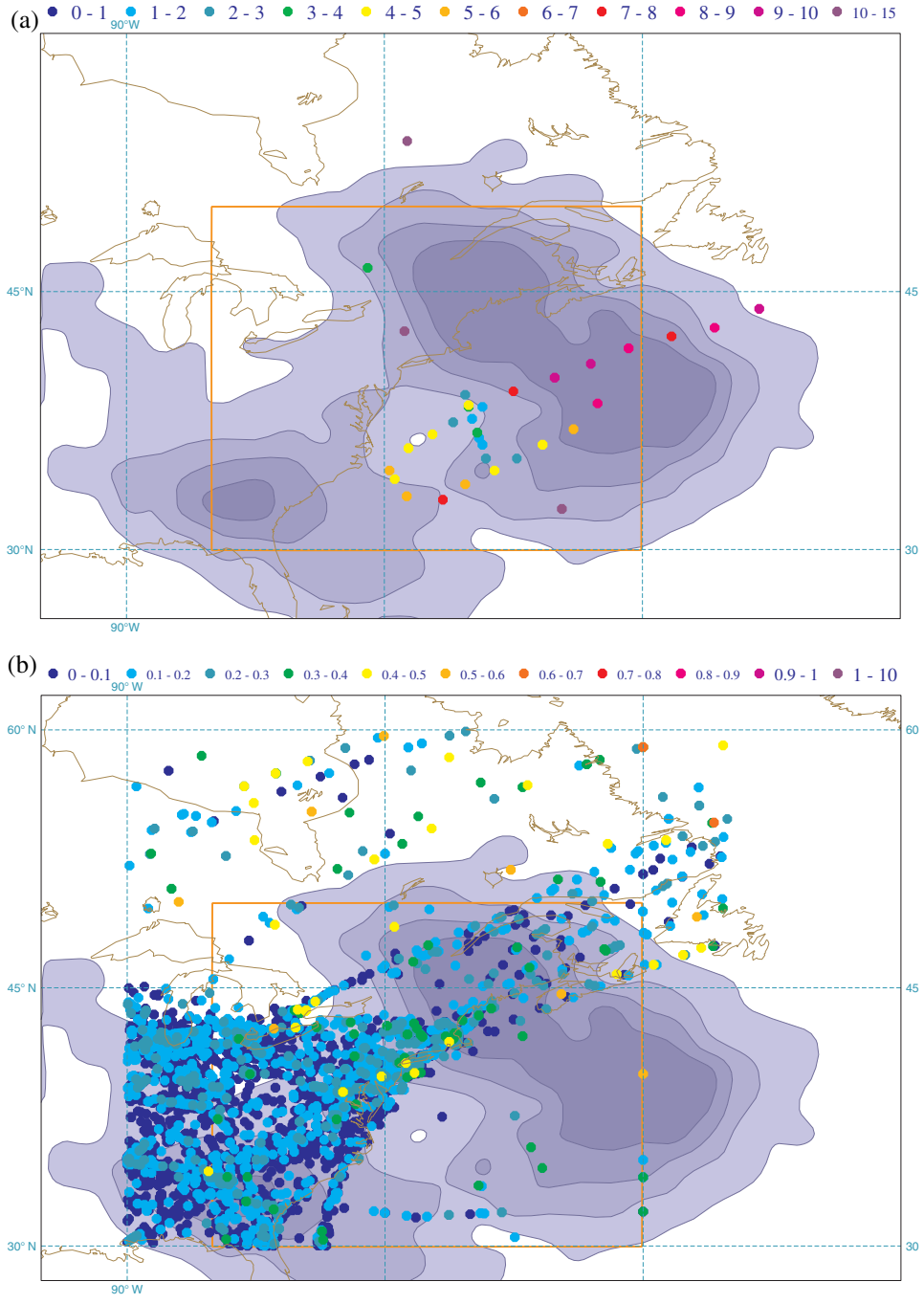


Figure 12. Sensitive area and DFS over the targeted area (30°N–50°N; 85°W–60°W, orange box) for the 18 UTC 5 December. The blue colour scale represents from the lightest to the darkest the sensitive area with sizes of 8×10^6, 4×10^6, 2×10^6, 1×10^6 km^2. (a) DFS values of the radiosondes and the dropsondes, and (b) DFS of the AIREP observations.

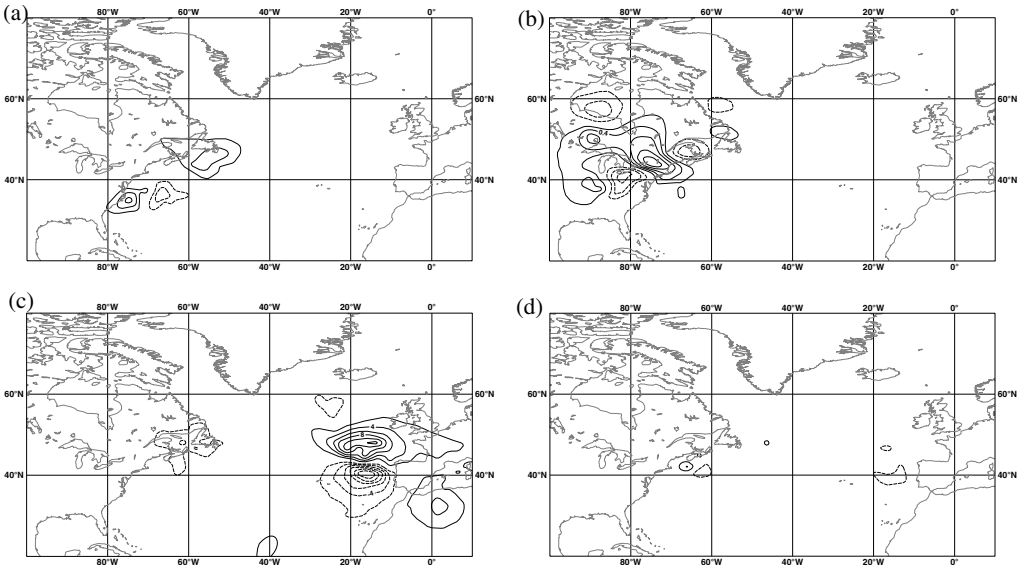


Figure 13. Difference of analyses for (a)–(b) 18 UTC 5 December 2003 and (c)–(d) 78-hour forecasts, for the geopotential at 500 hPa between (a) and (c) the WITH and the without dropsondes experiment, and (b) and (d) the WITH and the without AIREP experiment. Isolines are plotted every 0.2 m for the analyses and every 2 m at 78-hour forecast range. Positive values are shown with solid lines and negative values with dashed.

the location of the dropsondes that were deployed over the sensitive area and in the mid- and low troposphere, where sensitivity maximum is observed. On the contrary, the low-tropospheric AIREP observations are situated on the secondary maximum and out of the sensitive area.

The result in terms of forecast score is shown in Fig. 14, which displays the r.m.s. of the geopotential 78-hour forecast error over Europe computed with respect to the radiosondes for both experiments. The forecast error of the WITH experiment is also drawn. The forecast without the dropsondes in the analysis is worse in the whole troposphere than the one of the WITH experiment. The forecast without the AIREP is slightly improved (or similar) when compared with the one from the WITH experiment. This is just one case and does not indicate a malfunction of the variational assimilation for this kind of observations. Note that, for the verification time (00 UTC 7 December) over North America (not shown), the geopotential r.m.s. error of the experiment without the AIREP is smaller in the low troposphere than the one for the WITH experiment and it is larger in the 500–100 hPa layer. In contrast, the r.m.s. for the experiment without the dropsondes is larger below 400 hPa than the one of the WITH experiment. We thus observe the opposite effect of the observations on the 42-hour forecast.

(*f*) *Conclusions*

Over the globe, the infrared radiances sensitive to the humidity field and the AIREP observations have the largest DFS. The DFS has also been computed with respect to the assimilated parameters and it has been found that the wind measurements were the most informative. The DFS per observation has also been studied and it has been found that it decreases when the number of observations increases. For the observations located in the sensitive area, the AIREP and the dropsondes are the most informative. The study of the DFS of the case of the 2 December (case TREC_24) was also performed.

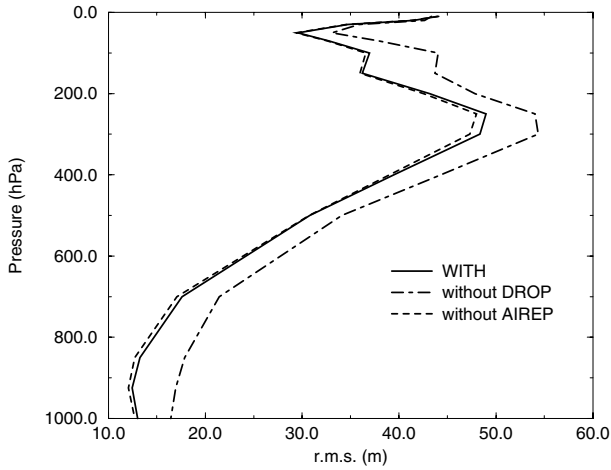


Figure 14. Root-mean square of the geopotential 78-hour forecast error from the analysis of 18 UTC 5 December 2003 over Europe computed with respect to the radiosondes of 00 UTC 9 December 2003.

It is a case of degradation (see Fig. 4). It comes from the background and the observations decrease the forecast error but do not succeed in correcting it. Over the globe the same results for the DFS are found: the most informative data come from the satellite observations and the AIREP. Over the targeted area, the dropsondes and the AIREP contribute the most to the DFS.

The DFS gives an indication as to how the observations contribute to the analysis but it does not indicate how these observations will modify the subsequent forecast as shown by the two denial experiments.

5. Conclusion and discussion

A partial re-analysis without the additional observations of the ATReC that was implemented during autumn 2003 has been performed: 22 cases out of 45 were studied and the impact of their additional data was evaluated. The forecast skill of this re-analysis was compared to the operational suite, which included the observations of the campaign. The forecast skill of the operational suite is identical or slightly better than the re-analysis one over the North Atlantic Ocean. The observation impact is more beneficial at 18 and 00 UTC but not very significant. However, a noticeable improvement is observed in the 18–54 hour forecast range over Europe in the upper troposphere/lower stratosphere due to the inclusion of radiosondes. When focussing on the verification area the improvement of the forecast is case, forecast-term and level dependent. The results consist of a mix of improvements and worsenings. However, the improvement in the lower troposphere at around the 42-hour forecast range is statistically significant.

We can note that the proportion of improved/degraded cases is different from the one of the 1999 and 2000 WSRPs (Szunyogh *et al.* 2000, 2002). The number of neutral cases is more important: it is similar to the improved case number (9 improved, 8 neutral, 5 degraded at 1000 hPa). This may be due to the fact that the WSRP globally sample remote regions over the ocean whereas the sampled regions of the ATReC were often localized just off the eastern North American coast and they were generally well sampled by the conventional observations at the previous analysis time.

The observation impact on the analysis has then been studied for one synoptic case corresponding to a severe snowstorm over North America at the beginning of December 2003. This has been addressed with the information content. The DFS has been studied over the globe and more precisely in the sensitive area. Over the globe, the infrared radiances sensitive to the humidity or sounding close to the surface have the most important information content. The wind measurements are also very informative. Over the targeted region, the AIREP and the dropsondes have the most important information content. DFS for each platform has been also studied. It happens that, for our case, dropsondes with the largest DFS are located in the main maximum of the sensitive area whereas AIREP are mostly over the secondary one. The properties of the DFS have been checked: the DFS per observation decreases when the considered observation number increases and when the observations are close to each other. Another case has also been studied and the same results were obtained. However, the DFS gives an indication of what the system does and not what it should do. Regarding the forecast, the increment due to the observations grows in time only for the dropsondes. Two examples of denial experiments have illustrated that the DFS cannot be straightforwardly linked to the forecast impact. To have the actual modification of the background, the DFS which is a tool for the diagnostic of the observation impact on the analysis, must be complemented with the study of other diagnostics.

Let us remember that the goal of the targeting is to decrease the forecast error on average and not each forecast individually. One cannot hope to improve the forecast for each case but on average the forecast will be improved. Morss and Emanuel (2002) have shown in idealized systems, even with a perfect observations and perfect model, that to add observations can degrade analyses and forecasts. These degradations may be due to deficiencies in the data assimilation system, but their risk is inherent in statistical data assimilation and in nonlinear prediction. These authors suggested that, due to the risk of degradation in adding observations, assimilating supplementary data and using the resulting initial conditions to produce forecasts should be considered as an alternative forecast and should rather be interpreted as probabilistic rather than deterministic processes.

The aim of the ATreC was to prove the validity of the concept of a near-operational targeting campaign with a component of the routine observing system switched to an adaptive mode. This study highlights the context in which observation targeting is operated that is not simple and straightforward. No effort was made to maximize the impact of the targeted observations in the sense that some techniques were still in test and employed sometimes at a sub-optimal level of their predicting capacities. Many points in the targeting methods and deployments not investigated in this study could therefore be improved: the optimality of the targeting methods, the way of targeting deployment and a finer usage of satellite observations. Moreover, the assimilation of the extra data should be also assessed in order to increase the impact of the data and to diagnose it.

The coherence of the sensitive areas computed with the various methods used during the ATReC (Ensemble Kalman Filter, Hessian vectors, singular vectors) has to be verified. This has been done by Doerenbecher *et al.* (2005) who showed that at small scale, when planning flight tracks for dedicated aircrafts, it was difficult to get a consensus between the different sensitive areas. The vertical structure of the sensitive areas should also be taken into account as it is well known that the sensitivity maximum is often located in the low and mid-troposphere (Rabier *et al.* 1996). For the Ensemble Transform Kalman Filter and the Hessian singular vectors, summary maps used during the ATReC were not the most advanced targeting guidance.

In particular, more accurate information could be obtained by testing virtual deployments within the targeting regions in order to compute which combination of observations decreases the uncertainty most. All techniques have a different degree of optimality and the efficiency of the techniques have not been checked in the various prediction systems involved in the ATreC. However, many studies such as Leutbecher *et al.* (2002) or Gelaro *et al.* (2000) have shown that adding observations in the sensitive areas has more impact than adding observations elsewhere because the increment projects on the structure of the dynamical sensitivities. Moreover, the additional observations should be associated with the corresponding sensitive areas and for each data assimilation system the impact of only these observations should be assessed.

The sensitive areas may have been correctly identified but the observations may not have been deployed in the heart of the sensitive areas. Indeed, the limited flexibility of most of the adapted observation platforms (when they are not dedicated to targeting such as jet aircrafts with dropsondes) implies that most of the targeted data have been deployed only close to the sensitive areas (e.g. radiosonde and the air traffic routes on AIREP data). In addition, the air traffic control limits the capability of targeting with aircrafts.

The small impact of the ATReC observations may also come from the small number of added observations deployed and assimilated in the WITH suite (6% for the 06 and 18 UTC analysis times, 2.5% globally averaged over all analysis times). These extra observations mainly consist in dropsondes that are launched very close to each other. This can induce correlation that is not taken into account in the assimilation process. The adaptive assimilation of the additional data should also be considered. In addition the targeting techniques do not take into account the assimilated observations between the time of the sensitive areas and the targeting time which could modify the forecast errors. A lot of satellite data are currently assimilated and Bouttier and Kelly (2001) have shown that they have a strong impact on the forecast skill. The targeted areas could not be considered as data-sparse areas. Finally, there have been few cases of high-impact weather during the ATreC and the one concerning the snowstorm of the 5 December has been well sampled, whereas the one concerning severe floods in south of France (beginning of December) has not be sampled with dropsondes. To increase the number of studied high-impact weather events, the adaptive use and assimilation of the satellite data in the sensitive areas could be a response.

Finally, let us point out that the goal of the targeting is to improve the forecast over a specific region and at a particular time. That is why some diagnosis on the potential and the ability of the observations to modify the forecast are interesting to study. This case-study shows that the diagnosis of observation impact on the forecast is required for targeting. Langland and Baker (2004) proposed the sensitivity of the forecast to the observations that is an adjoint-based procedure for estimating the impact of the observations on short-range forecast errors. Desroziers and Chapnik (2005) proposed, in a simple assimilation context, a randomized method to investigate the error variance reduction of the analysis and of the subsequent forecast. This method should soon be implemented in the ARPEGE model.

ACKNOWLEDGEMENTS

We acknowledge Dominique Puech, Véronique Mathiot and Francis Pouponneau for their technical support. Alexis Doerenbecher is acknowledged for providing the map of the sensitive area. He, Philippe Arbogast and Jean Antoine Maziejewski are also warmly thanked for their helpful comments on a previous version.

APPENDIX

TABLE A.1. DEFINITION OF THE VERIFICATION REGIONS

Region	Latitude	Longitude	Definition
ATReC	25°N–75°N	90°W–30°E	General area of the ATReC
WATLAN	30°N–65°N	85°W–40°W	Default, over the west Atlantic
NEUROPE	45°N–65°N	15°W–35°E	Default, over the north of Europe
SEUROPE	30°N–50°N	15°W–35°E	Default, over the south of Europe
CEUROPE	30°N–65°N	15°W–35°E	Default, over Europe
RVERIF	Required verification area (case-dependent, see Table A.2)		

TABLE A.2. DESCRIPTION OF THE 22 STUDIED ATREC CASES

	Observation			Verification				
		Region					Required area	
TReC	Date	Latitude	Longitude	Date	Time	Default region	Latitude	Longitude
20_2	27 Nov	45°N–65°N	75°W–30°W	29 Nov	12	NEUROPE	40°N–60°N	25°W–5°W
21	28 Nov	45°N–65°N	60°W–10°W	30 Nov	00	WATLAN	40°N–60°N	75°W–45°W
22_1	28 Nov	45°N–55°N	75°W–35°W	1 Dec	12	CEUROPE	45°N–60°N	15°W–15°E
22_2	29 Nov	30°N–60°N	55°W–5°W	1 Dec	12	CEUROPE	35°N–55°N	5°W–20°E
23_1	30 Nov	25°N–45°N	35°W–5°W	3 Dec	12	CEUROPE	35°N–55°N	5°W–20°E
23_2	1 Dec	25°N–45°N	35°W–5°W	3 Dec	12	CEUROPE	45°N–60°N	15°W–15°E
24	2 Dec	35°N–65°N	70°W–35°W	4 Dec	12	NEUROPE	55°N–75°N	10°W–30°E
25	3 Dec	none		5 Dec	12	CEUROPE	52°N–70°N	50°W–10°W
26_1	4 Dec	20°N–45°N	90°W–55°W	6 Dec	12	WATLAN	30°N–50°N	85°W–65°W
26_2	5 Dec	30°N–50°N	85°W–60°W	7 Dec	12	WATLAN	30°N–50°N	85°W–65°W
26_3	6 Dec	30°N–50°N	85°W–55°W	8 Dec	12	WATLAN	35°N–55°N	90°W–45°W
27_1	6 Dec	30°N–70°N	30°W–30°E	8 Dec	12	SEUROPE	30°N–50°N	10°E–30°E
27_2	7 Dec	30°N–70°N	30°W–30°E	9 Dec	12	SEUROPE	30°N–50°N	10°E–30°E
28_1	6 Dec	35°N–60°N	70°W–40°W	10 Dec	00	NEUROPE	45°N–65°N	20°W–10°E
28_2	7 Dec	35°N–60°N	70°W–35°W	10 Dec	00	NEUROPE	45°N–65°N	15°W–10°E
29_1	8 Dec	30°N–60°N	70°W–15°W	11 Dec	00	NEUROPE	45°N–65°N	20°W–10°E
29_2	9 Dec	40°N–65°N	50°W–20°W	11 Dec	00	NEUROPE	45°N–65°N	20°W–10°E
31_1	10 Dec	40°N–70°N	60°W–20°W	12 Dec	12	NEUROPE	45°N–65°N	20°W–15°E
31_2	11 Dec	40°N–70°N	60°W–20°W	13 Dec	12	NEUROPE	50°N–70°N	20°W–15°E
32_1	12 Dec	28°N–45°N	90°W–60°W	15 Dec	00	WATLAN	30°N–50°N	90°W–65°W
32_2	13 Dec	28°N–45°N	90°W–60°W	15 Dec	12	WATLAN	30°N–50°N	85°W–65°W
32_3	14 Dec	25°N–45°N	90°W–60°W	16 Dec	00	WATLAN	35°N–55°N	80°W–60°W

All times UTC. Observations made at 18 UTC.

REFERENCES

Bouttier, F. and Kelly, G. 2001 Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Q. J. R. Meteorol. Soc.*, **127,** 1469–1488

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. 2006 Diagnosis and tuning of observational error statistics in a quasi-operational data assimilation setting. *Q. J. R. Meteorol. Soc.* (in press)

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. European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK

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

Desroziers, G. and Ivanov, S. 2003 Diagnosis and adaptive tuning of information error parameters in a variational assimilation. *Q. J. R. Meteorol. Soc.*, **129,** 1433–1452

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

Doerenbecher, A., Leutbecher, M. and Richardson, D. S. 2005 'Comparison of observation targeting predictions during the (North) Atlantic TReC 2003'. In Proceedings of the first international THORPEX science symposium, 6–10 December 2003, Montreal, Canada

Fisher, M. 2003 Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. ECMWF Tech. Memo. No. 397, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, Berkshire RG2 9AX, UK

Gelaro, R., Reynolds, C. A., Langland, R. H. and Rohaly, G. D. 2000 A predictability study using geostationary satellite wind observations during NORPEX *Mon. Weather Rev.*, **128,** 3789–3807

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. American Meteorological Society, 45 Beacon Street, Boston MA02108-3693, USA

Janiskovà, 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

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., Gall, R., Hewson, T. D., Hildebrand, P. H., Jorgensen, D., Lalaurette, F., Langland, R. H., Lemaitre, Y., Mascart, P., Moore, J. A., Persson, P. 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

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

Langland, R. H., Toth, Z., Gelaro, R., Szunyogh, I., Shapiro, M. A., Majumdar, S. J., Morss, R. E., Rohaly, G. D., Velden, C., Bond, N. and Bishop, C. H. 1999 The North Pacific Experiment (NORPEX-98): Targeted observations for improved North American weather forecast. *Bull. Am. Meteorol. Soc.*, **80,** 1363–1384

Leutbecher, M., Barkmeijer, J., Palmer, T. N. and Thorpe, A. J. 2002 Potential improvement to forecasts of two severe storms using targeted observations. *Q. J. R. Meteorol. Soc.*, **128,** 1641–1670

Majumdar, S. J., Bishop, C. H., Buizza, R. and Gelaro, R. 2002 A comparison of ensemble-transform Kalman-filter targeting guidance with ECMWF and NRL total energy singular vector guidance. *Q. J. R. Meteorol. Soc.*, **128,** 2527–2549

Morss, R. E. and Emanuel, K. A. 2002 Influence of added observations on analysis and forecast errors: Results from idealized systems. *Q. J. R. Meteorol. Soc.*, **128,** 285–321

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A. 1996 Sensitivity of forecast errors to initial conditions. *Q. J. R. Meteorol. Soc.*, **122,** 121–150

Rodgers, C. D. 2000 *Inverse methods for atmospheres: Theories and practice.* World Scientific Publ., Singapore

Stringer, S. and Truscott, B. 2004 Atlantic-THORPEX Regional Campaign. Operations plan. Available at http:www.wmo.int/thorpex/pdf/atlantic_ob_system.pdf

Szunyogh, I., Toth, Z., Morss, R., Majumdar, S., Etherton, B. J. and Bishop, C. H. 2000 The effect of targeted dropsonde observation during the 1999 Winter Storm Reconnaissance Program. *Mon. Weather Rev.*, **128,** 3520–3537

Szunyogh, I., Toth, Z., Zimin, A. V., Majumdar, S. J. and Persson, A. 2002 Propagation of the effect of targeted observations: The 2000 Winter Storm Reconnaissance Program. *Mon. Weather Rev.*, **130,** 1144–1165