THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Université Toulouse III - Paul Sabatier*
Discipline ou spécialité : *Sciences de l'Univers, de l'Environnement et de l'Espace*

Présentée et soutenue par *Xin YAN*
Le *10 Mars 2009*

Titre : *Assimilation de données GPS pour la prévision de la convection profonde*

JURY

M. Sylvain COQUILLAT Président
M. Olivier BOCK Rapporteur
M. Frédéric MASSON Rapporteur
M. Cyrille FLAMANT Examinateur
M. Paul POLI Examinateur
Mme Véronique DUCROCQ Directrice de thèse
Mme Andrea WALPERSDORF co-directrice de thèse

Ecole doctorale : *Ecole doctorale des Sciences de l'Univers et de l'environnement*
Unité de recherche : *Centre National de Recherches Météorologiques*
Directeur(s) de Thèse : *Véronique DUCROCQ, Andrea WALPERSDORF*
Rapporteurs : *Olivier BOCK, Frédéric MASSON*

Remerciements

First of all, I would like to give my greatest thanks to my supervisor Dr. Veronique Ducrocq for her support and advice during the three years of my PhD in Meteo France. Thank you for your patience and guidance for me and the time you spend on in revising the articles and thesis. Deepest thanks !

I would like to thank my parents for their love and non-stop support to me since I was born. Without your encouragement and trust, I would never be able to travel so far in the journey of life.

Thank you Genevieve Jaubert, for patiently explaining to me all the necessary knowledge and tools for my work. Also many thanks to my dear colleagues in the team of Micado, Olivier Caumont, Laurent Labatut, Olivier Nuissier, Didier Ricard, Anne-lise Beaulant, Fanny Duffourg,Emily Bresson, Beatrice Vincendon and Karen Boniface and all the gentle ladies in MesoNH team ! Whenever I had a tough problem, I could always have a reasonable explanation from you. Even though I never drink a single cup of coffee, but I did enjoy the atmosphere of the coffee breaks in the team with all of you !

Thanks to Andrea Walpersdorf, Paul Poli, Marouan Hakam, Cedric Champollion for all the great jobs and the exchanging of ideas in writing articles. Thank you so much for putting ideas together.

Special thanks to my dear friends who has always cheered me up and helped me out : Claudia Faccani, Bora Ucar, Ernani Nascimento, Pere Quintana, Thomas Lauvaux, Julien Pergaud, Julia Hildago, Xueping Guo, Tomas Kral.

Résumé

Ce travail de thèse visait à exploiter le potentiel des observations GPS sol pour l'assimilation de données à mesoéchelle et la prévision numérique du temps à haute résolution. Nous avons tout particulièrement examiné l'impact de l'assimilation des données GPS sur la prévision à l'échelle convective des systèmes orageux. Les systèmes d'assimilation utilisés sont les systèmes d'assimilation à mesoéchelle de Météo-France : le 3DVAR/ALADIN et le 3DVAR/AROME. Deux cas d'étude ont été traités, avec pour chacun des cas un nombre important de données GPS assimilées au cours de cycles d'assimilation continus sur de longues périodes. Les situations météorologiques des deux cas d'étude sélectionnés sont caractérisées par plusieurs évènements convectifs précipitants. Pour le premier cas d'étude (5-9 septembre 2005) avec plusieurs épisodes de pluie intense affectant les régions Méditerranéennes françaises, le système d'assimilation utilisé est le 3DVAR/Aladin à la résolution de 9,5 km. Un cycle d'assimilation long d'un mois est d'abord réalisé puis les analyses ALADIN issues de ce cycle d'assimilation sont utilisées comme conditions initiales et aux limites de simulations à 2,4 km de résolution horizontale avec le modèle de recherche MesoNH. Pour le second cas d'étude (18-20 juillet 2007), les observations GPS du réseau dense déployé au cours de la campagne COPS en complément des observations du réseau opérationnel européen E-GVPAP sont assimilées directement avec le 3DVAR/AROME à la résolution de 2,5 km. Pour la réalisation de ces expériences, une attention particulière a été portée à la sélection et au pré-traitement des observations GPS qui entrent dans les systèmes d'assimilation ; une telle procédure peut être vue comme un premier contrôle de qualité des observations GPS. Pour les deux cas, les résultats des expériences d'assimilation et de prévision avec ou sans assimilation de données GPS montrent un impact positif sur la prévision des précipitations intenses ; l'impact est plus significatif sur le second cas d'étude.

Abstract

The aim of this thesis work is to exploit the potential of ground based GPS observations for mesoscale data assimilation and high-resolution numerical weather prediction. Our main interest lies in the investigation of the impact of GPS observation assimilation in improving the forecast of convective scale weather phenomena such as convective storms. The data assimilation systems we have used are the Météo-France 3DVAR/ALADIN and 3DVAR/AROME systems. Two case studies were conducted with large numbers of GPS observations being assimilated. The cases selected are both convective rainfall events. For the first case (5-9 September 2005) with heavy precipitation over the French Mediterranean regions, the system we used is the 9.5km 3DVAR/ALADIN assimilation system. Analyses produced by one month of assimilation cycle experiments are used later as the initial and boundary conditions for starting the simulation of the 2.4km high resolution research model MesoNH. For the second case study (18-20 July 2007), high density GPS observations obtained from the COPS observations field campaign together with GPS observations coming from the European operational E-GVAP network are assimilated directly into the 3DVAR/AROME assimilation system with 2.5km horizontal resolution. Special attentions have been paid on the selection and pre-treatment of the GPS observations before they enter in the assimilation system. Such procedure can be viewed as a quality control for the observations. For both cases, results of twin experiments with and without assimilating GPS observations have suggested a positive impact on the prediction of heavy precipitation, the impact being more significant on the second case.

Table des matières

Introduction

Les systèmes convectifs, accompagnés de précipitations intenses et de foudre, et provoquant dans certains cas des inondations, sont souvent responsables de dégâts et de pertes humaines. Certaines régions sont plus particulièrement soumises à ce risque orageux. C'est en particulier le cas des zones montagneuses où les reliefs permettent le soulèvement des masses d'air nécessaire au déclenchement de la convection profonde. Le pourtour méditerranéen nord-occidental est aussi soumis à ce type de systèmes précipitants qui peuvent prendre un caractère stationnaire sur la région entraînant des cumuls de précipitation importants et des crues rapides. Des cumuls de précipitation de l'ordre de 200 mm en 24 heures ne sont pas rares et engendrent des crues éclairs lorsqu'elles s'abattent sur les bassins versants pentus de la région. La région Cévennes-Vivarais est particulièrement soumise à ce risque hydrométéorologique.

Afin de réduire les dégâts causés par ces évènements de précipitations convectives, il est important que la capacité des modèles de prévision numérique du temps à prévoir ce type d'évènements progresse pour permettre une alerte plus précoce et plus précise. La qualité d'une prévision numérique dépend de plusieurs facteurs, et en particulier de la représentation des processus physiques dans le modèle, des conditions initiales et, pour un modèle à domaine limité, des conditions aux limites latérales. Parallèlement à l'augmentation de résolution des modèles de prévision du temps et aux progrès réalisés dans leurs représentations des processus physiques, l'assimilation de données à mesoéchelle s'est développée au cours des dernières années pour améliorer la description des conditions initiales. A Météo-France, une assimilation de données à mesoéchelle a ainsi été développée pour le modèle à aire limitée ALADIN couvrant l'Europe de l'ouest à la résolution de 9.5 km. Depuis juillet 2005, le modèle opérationnel ALADIN-France est désormais initialisé à partir des analyses issues de son propre cycle d'assimilation. AROME, le nouveau modèle de Météo-France, constitue une nette avancée en terme de résolution et de représentation du cycle de l'eau et de la convection profonde par rapport aux autres modèles opérationnels ARPEGE et ALADIN. Sa haute résolution (2.5 km) qui permet de s'affranchir de la paramétrisation de la convection profonde et sa paramétrisation microphysique relativement élaborée pour un modèle opérationnel constituent des facteurs d'amélioration de la prévision des systèmes convectifs et des précipitations associées. Mais ils ne garantissent pas nécessairement le succès de la prévision car celle-ci dépend aussi des conditions initiales. AROME a ainsi également été doté d'un système d'assimilation de résolution horizontale 2.5 km pour lui fournir des conditions initiales à fine échelle.

L'assimilation de données pour ces systèmes de prévision à mesoéchelle requiert

tout particulièrement des observations à haute fréquence spatiale et temporelle, et en particulier de vapeur d'eau. La vapeur d'eau présente en effet une plus forte variabilité spatiale et temporelle que d'autres paramètres comme la température ou la pression. Aussi, Ducrocq et al. (2002) ont montré que les prévisions à haute résolution de systèmes convectifs étaient particulièrement sensibles à la description du champ initial d'humidité. Les mesures de vapeur d'eau troposphérique sont cependant encore insuffisantes en termes de résolutions spatiale et temporelle. Les radiosondages qui permettent d'avoir une observation de la distribution verticale de la vapeur d'eau dans la troposphère ont une faible résolution spatiale et temporelle (seulement 7 radiosondages à 00 et 12 UTC pour la France métropolitaine). Si les satellites géostationnaires offrent des observations à haute résolution temporelle et spatiale, celles-ci ne décrivent pas l'humidité dans les basses couches, et on ne sait les assimiler à l'heure actuelle qu'en dehors des nuages. Le sondeur IASI embarqué sur le satellite défilant METOP constitue sans aucun doute une avancée avec son nombre important de canaux vapeur d'eau, incluant des canaux sensibles dans les basses couches. Cependant sa fréquence de revisite d'un même site est sous-optimale pour une assimilation de données à mesoéchelle.

Dans ce contexte pauvre en observations à mesoéchelle d'humidité, la possibilité d'exploiter pour des applications météorologiques, les données des réseaux GPS en pleine expansion pour les besoins de la géolocalisation, est apparue très intéressante à explorer. Il y a maintenant plus d'une quinzaine d'années que le GPS a été proposé comme instrument de mesures du contenu en vapeur d'eau troposphérique (Bevis et al., 1992). Les signaux émis par la constellation de satellites GPS subissent en effet un retard lorsqu'ils traversent l'atmosphère qui dépend des caractéristiques en pression, température et vapeur d'eau des couches de l'atmosphère traversées. Ce retard est exprimé comme un délai total au zénith (Zenith Total Delay, ZTD). La collecte des données GPS à l'échelle européenne, leur traitement et leur diffusion en temps quasi-réel réel vers les services météorologiques pour une utilisation dans les systèmes opérationnels de prévision numérique du temps se sont par la suite organisés, tout d'abord au travers de l'action COST 716 et des projets MAGIC et TOUGH, puis depuis avril 2005 au sein du programme E-GVAP d'EUMETNET avec une évolution significative depuis ses débuts en termes de qualité et de nombre des données collectées.

Plusieurs travaux ont étudié par le passé l'impact de l'assimilation des données GPS de ZTD sur la prévision des précipitations et ont généralement montré un impact neutre à faiblement positif. Ces études ont cependant concerné des modèles de prévision de résolution inférieure à la dizaine de kilomètres et utilisant une paramétrisation de la convection profonde. L'objectif de ma thèse était d'examiner l'impact de l'assimilation des données GPS pour les modèles de prévision à fine échelle (2-3 km) permettant une résolution explicite de la convection profonde. Mes travaux de thèse ont aussi contribué au développement de l'assimilation des données GPS de ZTD dans les systèmes d'assimilation ALADIN et AROME. Ils ont été organisés en deux parties :

– *Assimilation des données GPS de ZTD avec le 3DVAR ALADIN (9.5 km) et*

évaluation de l'impact sur la prévision à haute résolution (2.5 km) du modèle Meso-NH. Cet impact a été examiné plus particulièrement pour l'épisode précipitant du 5 au 9 septembre 2005 qui a affecté la région de l'Hérault et du Gard.

– *Assimilation des données GPS de ZTD avec le 3DVAR AROME (2.5 km) et évaluation de l'impact sur la prévision AROME.* Cet impact a été examiné plus particulièrement pour les évènements précipitants du 18 au 20 juillet 2007, période qui correspond à la POI9 de la campagne de mesures COPS sur les Vosges et la Forêt Noire.

Ce manuscrit de thèse est organisé en cinq chapitres. Le premier chapitre décrit la relation entre les observations de délais GPS et la réfractivité atmosphérique, ainsi que le principe de l'estimation des délais zénithaux. Le second chapitre dresse un état de l'art des travaux sur l'assimilation de données GPS. Ensuite, une description technique des systèmes d'assimilation de Météo-France (3DVAR/ALADIN, 3DVAR/AROME) utilisés au cours de mon travail de thèse est présentée au chapitre 3. Le chapitre 4 présente les expériences d'assimilation de données GPS de ZTD réalisées avec le 3DVAR/ALADIN (9.5 km) et l'évaluation de l'impact sur la prévision Meso-NH des évènements de pluie intense des 5 au 9 septembre 2005. Les résultats d'une étude de sensibilité sur la formulation de l'opérateur d'observation sont aussi décrits. Le chapitre 5 est consacré aux expériences d'assimilation de données GPS de ZTD réalisées avec le 3DVAR/AROME (2.5 km) et à l'impact sur la prévision AROME de la POI9 de la campagne COPS. Une étude de sensibilité pour évaluer l'apport des données des récepteurs GPS déployés spécifiquement pendant la campagne a aussi été conduite. Les conclusions et perspectives à ce travail de thèse sont ensuite présentées.

Chapitre 1

Les observations GPS

Résumé

Ce chapitre commence par présenter la relation entre le système de positionnement GPS et la réfractivité de l'atmosphère. L'intégration de la réfractivité le long du chemin de l'onde électromagnétique, appelé le retard troposphérique, est la principale source d'erreur pour le positionnement par GPS. Ce délai induit par la réfractivité de l'atmosphère est fonction des caractéristiques en température, pression et vapeur d'eau rencontrées par l'onde électromagnétique entre le satellite émetteur et le récepteur GPS au sol. Ce délai oblique est exprimé dans les méthodes de traitement GPS à partir du délai au zénith (ZTD pour Zenith Total Delay), qui est un des paramètres estimés en même temps que la position par minimisation à partir des observations GPS.

Connaissant la pression et la température à la localisation du récepteur, on peut obtenir une estimation du contenu intégré en vapeur d'eau à partir du délai zénithal. Cette mesure du contenu intégré en vapeur d'eau a une précision comparable à celle des radiosondages ou des radiomètres, inférieure à 10 mm. L'intérêt de la communauté météorologique pour les données GPS (délais zénithaux et contenus intégrés en vapeur d'eau) a donc tout naturellement augmenté au cours des dernières années avec le développement de leur utilisation dans des applications météorologiques (validation, assimilation de données) et de l'accès à ces données en temps quasi-réel pour les applications opérationnelles.

1.1 Relationship between the tropospheric delay and the Global Positioning System (GPS)

1.1.1 The Global Positioning System

The Global Positioning System (GPS) is a satellite-based navigation system. As of March 2008, GPS system is constituted of 31 satellites, which orbit around the earth at about 20000 km above the earth surface. The satellites were launched into 6 orbital planes 60 degrees apart from each other. The orbits are arranged so that at least four satellites are always within line of sight from almost everywhere on Earth's surface. GPS works in any weather conditions, anywhere in the world and 24 hours a day.

1.1.2 Brief positioning theory and formulation

GPS receivers on the ground can receive the signal transmitted by the GPS satellites once they are above the horizon. All GPS satellites transmit C/A code at the same carrier frequencies : L1 (1575.42MHz) and L2 (1227.6MHz). Two types of observables can be recorded by the GPS system (Parkinson and Spiker Jr, 1996). They are the code pseudorange measurement and the carrier phase measurement. Both observables have a length unit. Code measurements detect how much the C/A code sequence has changed, from which the actual travelling time of the signal can be obtained. Multiplying this time by the speed of light c_0 yields the code pseudorange measurement. Phase measurement detects the phase change of the L1 carrier signal from its emission by the satellite till its reception by the receiver. Multiplying this phase change with the wavelength yields the phase measurement.

Both pseudorange and phase measurement are not the true geometrical distance between the receiver and satellite since the signal is not travelling with the speed of light in the atmosphere. An equation can be established between the true geometrical distance and the observed false distance (the two observables) by compensating the excess path caused by the error sources.

The true geometrical distance G_r^sbetween the satellite and receiver can be expressed using receiver and satellite position in XYZ WGS-84 coordinates :

$$G_r^s = \sqrt{(X_r - X_s)^2 + (Y_r - Y_s)^2 + (Z_r - Z_s)^2} = c_0 t_0 \tag{1.1}$$

t_0 :	time of signal travelling from satellite to receiver in a vacuum condition
c_0 :	the speed of light in vacuum
X_r, Y_r, Z_r :	GPS receiver position
X_s, Y_s, Z_s :	satellite position

The relationship between the receiver position and the GPS observables can be established using code pseudorange measurement or carrier phase measurement :

1) Using code pseudorange measurement

$$P_r^s = G_r^s + c_0(dt_r - dt^s) + \Delta^T + \Delta^I \tag{1.2}$$

2) Using carrier phase measurement

$$\lambda\phi + \lambda\mathbf{N} = G_r^s + c_0(dt_r - dt^s) + \Delta^T - \Delta^I \tag{1.3}$$

$$\lambda\phi = G_r^s + c_0(dt_r - dt^s) + \Delta^T - \Delta^I - \lambda\mathbf{N} \tag{1.4}$$

– P_r^s is the code pseudorange measurement, in length unit. Calculated by $P_r^s = c_0 t$, where c_0 is the speed of light in vacuum and t is the actual travelling time of the signal from satellite to the receiver. t is obtained by the following method : the receiver becomes aware of the satellite ID after decoding the ephemeris message from the incoming satellite signal. Then the receiver is able to reproduce the same C/A code series as that satellite according to its unique ID. A correlation algorithm is then done between the receiver generated code and the code coming from the satellite. The difference between the two codes are considered as the travelling time of the signal from the satellite to receiver.

– G_r^s is the geometrical distance between the GPS receiver and the satellite. It can be expressed as a function of the receiver position (Xr,Yr,Zr) and satellite position(Xs,Ys,Zs) in the WGS-84 coordinate system(1.1).

– dt_r and dt^s are the receiver and satellite clock errors. Satellite and receiver clocks are used as the device to generate the code signal sequence. If there is an error from these two clocks, the travelling time t that we obtained by comparing the two code sequences will not be precise. Therefore clock error also need to be corrected. The satellite clock is calibrated regularly by the ground control segments (Kaplan, 1996) and the atomic clock that equips on the GPS satellite, is of high precision and stability. The problem is more originated from the receiver clock. The receiver clock error dt_r is treated as the fourth unknown parameter together with the 3 coordinates of receiver position when solving equation 1.2. The precision of code measurement positioning depends on the width of the C/A code (300 m).

– λ is the wavelength of the carrier wave, in vacuum condition.

– ϕ is the phase measurement detected by the GPS receiver. It is in fact only the non-integer part of the total phase change of the travelling carrier signal-L1 from satellite to the receiver.

– N is the integer part of the total phase change called cyclic ambiguity. It will only be resolved together with other unknown parameters during the data processing. Note that the correct estimation of N is a precondition to achieve the

high precision advantage from phase measurement

- Δ^I is the atmospheric delay caused by ionosphere effects.

- Δ^T is the atmospheric delay caused by troposphere.

If the receiver can observe 4 or more satellites at the same time, 4 or more equations can be established with pseudorange or phase measurements, from which the positions of receiver and its clock offset can be solved (4 unknowns).

1.1.3 Positioning error source - atmospheric delays

From the right side of the equation 1.4, we can see that there are two major error sources due to the atmospheric effects for the GPS positioning. They are caused by the extra time that the signal used to travel when it is slowed down by the substances in ionosphere and troposphere of the earth atmosphere.

Equation 1.2 can be rewritten as :

$$\begin{aligned} c_0 t &= c_0 t_0 + c_0(dt_r - dt^s) + \Delta^T + \Delta^I && (1.5)\\ \Delta^T + \Delta^I &= c_0[t - t_0 - (dt_r - dt^s)] && (1.6)\\ &= c_0 \Delta t_{atm} + c_0(dt_r - dt^s) && (1.7)\\ &= c_0(\Delta t_{iono} + \Delta t_{trop}) + c_0(dt_r - dt^s) && (1.8) \end{aligned}$$

$c_0(dt_r - dt^s)$ is the satelite and receiver clock offset error and this term can be eliminated in equation 1.8 by double difference technique or estimation of extra parameters. Atmospheric delay originally means the extra time (Δt_{atm}) that signal used to travel through the atmosphere compared when it travels in vacuum. This extra time can be divided into two parts : Δt_{iono}, extra time induced by the ionospheric effects and Δt_{trop}, extra time induced by the tropospheric effects. But in practical use, this extra time has been converted into length unit by multiplying with c_0, the speed of light.

1) Ionospheric effects

$$\Delta^I = c_0 \Delta t_{iono} \quad (1.9)$$

Due to the existence of free electrons in the ionosphere, the index of refraction does not equal 1 and the value varies according to the season, time of day, solar activity, location, etc. This caused the slow down of group velocity of the satellite signal compared with the velocity when it travels in vacuum. But ionospheric effects increase the phase velocity of the satellite signal. That is why for the phase measurement in equation 1.4, the sign of Δ^I is negative while in the code measurement equation it is positive. However, with the dispersive nature of ionosphere (speed is dependent on the signal frequency), we can use a linear combination of the phase measurement from both carrier frequencies L1 and L2 to eliminate the ionospheric effects.

2) Tropospheric delay

$$\Delta^T = c_0 \Delta t_{trop} \tag{1.10}$$

This delay is due to the variation of the index of refraction in the troposphere. Δt_{trop} is the excess time the signal travelled due to the slowing down of speed and the increase of the path length (bending effect) compared with the time of travelling in vacuum t_0. But because of the non-dispersive nature of troposphere, one cannot perform the same correction using dual frequency phase measurement as with ionosphere. Other methods have to be used in order to deal with this error source. In the following sections, the issue concerning the modelling and estimating of the troposphere delay is addressed.

1.2 Formulation of tropospheric delay using refractivity

1.2.1 Relationship between tropospheric delay and atmospheric refractivity index

The phase speed of an electromagnetic (EM) wave slows down due to the interaction between the electromagnetic wave and the bounded electrons in the gas molecules in the troposphere. In other words, the refractive index does not equal 1 in the troposphere. Note that as the water molecule is polar, it produces a stronger effect of slowing down than other gases in the troposphere. Therefore the degree of refraction depends on the density of normal gas molecules and water vapour molecules. Moreover, the ray path is no longer a straight line between the satellite and the receiver, but bended to a curved line which is caused by the inhomogeneous refractive index value along the path (Fig. 1.1). This phenomenon is called refraction. The effect of slowing down and bending of the EM wave together created an extra travelling time (Δt_{trop}) of the signal between the satellite and the receiver compared with when it is in vacuum. Multiplying the extra time with speed of light, we obtain an overall excess path length that the signal travelled between the satellite and receiver compared with the distance when the signal travels in vacuum (index of refraction = 1). This excess path length is called tropospheric delay and is well illustrated by Equation 1.10. But it is also possible to estimate this delay using the physical characteristics of troposphere.

According to the Fermat law, the electromagnetic wave chooses a trajectory to travel from point A to B, which requires a minimum travelling time. The distance that the EM wave travelled between two points is defined as :

$$ds = cdt \tag{1.11}$$

c : speed of the EM wave in the medium
dt : travelling time

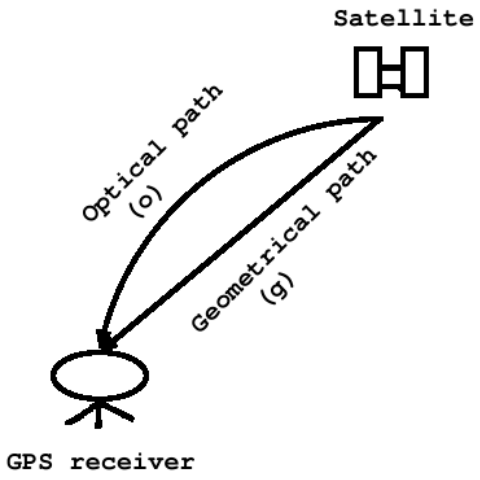

FIG. 1.1: Optical and geometrical path of signal

Let both sides of equation 1.11 multiplied by c_0/c, then perform the integration along the signal path o, it yields :

$$\int_o c_0 dt = \int_o \frac{c}{c_0} ds \tag{1.12}$$

$$c_0 t = \int_o n ds \tag{1.13}$$

with :

t :	the actual travelling time of signal between point A and B. It can be detected by a GPS receiver
$n = c_0/c$:	the index of refraction in the medium
o :	the optical path of the signal, which is the actual trajectory of the travelling signal in the atmosphere (see Fig. 1.1)

The integration of the refractive index n along the optical path o is defined as the electrical path length L_e, in length unit. According to equation 1.13 and the definition of L_e :

$$L_e = \int_o n ds = c_o t \tag{1.14}$$

When the signal travels through a vacuum, the electrical path length is in fact the geometrical path length :

$$L_g = \int_g n_{vac} ds = \int_g 1 ds = G_r^s = c_0 t_0 \tag{1.15}$$

L_g : geometrical path length
g : geometric path of the signal (straight line between two points)
G_r^s : total geometric path length
n_{vac} : refractive index in vacuum, which equals 1

From the electrical path length point of view, the tropospheric delay Δ^T can be defined as the difference between the electrical path length (L_e) in the presence of tropospheric refraction and in the vacuum condition.

$$\begin{aligned} \Delta^T &= L_e - L_g && (1.16)\\ &= \int_o n ds - \int_g 1 ds && (1.17)\\ &= \int_o (n-1)\, ds + \left(\int_o 1 ds - \int_g 1 ds \right) && (1.18) \end{aligned}$$

- $\int_o (n-1)\, ds$ is the excess path length due to the slowing down of signal caused by the tropospheric refractivity.
- $\left(\int_o 1 ds - \int_g 1 ds \right)$ is the excess path length due to the bending effect caused by the tropospheric refractivity. In figure 1.1, this excess path length is illustrated as (o-g). This bending part can be neglected since it is quite small compared with the total tropospheric delay. It is only important in the case of very low elevation angle.

If the second term of 1.18 is neglected, tropospheric delay can be written as :

$$\Delta^T = \int_o (n-1)\, ds = 10^{-6} \int_o N ds \qquad (1.19)$$

with $N = (n-1) \times 10^6$, the refractivity of the troposphere.

According to equation 1.14, 1.15, and assume that troposhere is the only source to delay the signal($t - t_0 = \Delta t_{trop}$), then equation 1.19 can be re-written as :

$$\begin{aligned} \Delta^T &= L_e - L_g && (1.20)\\ &= c_0 t - c_0 t_0 && (1.21)\\ &= c_0 (t - t_0) && (1.22)\\ &= c_0 \Delta t_{trop} && (1.23) \end{aligned}$$

notice, this Δ^T defined by the difference of electrical path length, equals the tropospheric delay $c_0 \Delta t_{trop}$ as in the GPS observation measurement equations (Eq. 1.10). It implies that the tropospheric delay in the GPS observation equation 1.4 can be

modelled if we know the refractivity of troposphere.

The tropospheric model based on refractivity is quite essential for providing an a priori value of Δ^T in order to solve all the parameters in equation 1.4, using parameter estimation methods.

1.2.2 Zenithal Tropospheric Delay

According to equation 1.19, knowledge of N is needed in order to model the tropospheric delay. In the microwave band, N can be expressed as a function of pressure, temperature, and water vapour pressure. The refraction due to the liquid and solid water is often neglected. The physical reason for that is the relatively low polarizability per unit mass of water molecules in condensed form, relative to vapour, at the GPS frequency. Under extremely adverse conditions of heavy rain and cloud along the signal path, the liquid water contribution to wet refractivity of GPS is at the most 10% of the contribution from water vapour (Mannucci et al., 2004). The significant contributions are very concentrated in the heart of severe thunderstorms (Brenot et al., 2006).

Thayer (1974) has proposed the following expression for N :

$$N = k_1 \frac{P_d}{T} Z_d^{-1} + k_2 \frac{e}{T} Z_w^{-1} + k_3 \frac{e}{T^2} Z_w^{-1} \tag{1.24}$$

with

P_d :	pressure of dry air in [hPa]
e :	pressure of water vapour in [hPa]
T :	temperature in [K]
Z_d, Z_w :	compressibility factors of dry air and water vapour
k_1, k_2, k_3 :	empirical coefficients

Different sets of coefficients k_1, k_2, k_3 have been proposed; the most used one was proposed by Bevis et al. (1994) with : $k_1 = 77.60$ K/hPa, $k_2 = 69.4$ K/hPa, $k_3 = 370100 K^2$/hPa. Also, it is generally assumed that gases are incompressible ($Z_d = 1, Z_w = 1$).

The first term in equation 1.24 is the refractivity due to the induced dipole moment of the dry air. The second term is the refractivity caused by the induced dipole moment of the water molecules in water vapour and the third term is caused by the effect of orientation of the permanent dipole moment of water molecules in water vapour.

The zenith total delay - ZTD - is defined as the tropospheric delay Δ^T for a path at the vertical of the receiver at height z_r :

$$ZTD = 10^{-6}\int_{z_r}^{\infty} N dz \tag{1.25}$$

$$= 10^{-6}(\int_{z_r}^{\infty} k_1\frac{P_d}{T}dz + \int_{z_r}^{\infty} k_2\frac{e}{T}dz + \int_{z_r}^{\infty} k_3\frac{e}{T^2}dz) \tag{1.26}$$

Given the following relations :

$$P = \rho_h R_d T_v \tag{1.27}$$

$$P_d = \rho_d R_d T \tag{1.28}$$

$$e = \rho_v R_v T \tag{1.29}$$

$$P = P_d + e \tag{1.30}$$

with
R_d, R_v : specific molar gas constants for dry air and water vapour .
ρ_d, ρ_v, ρ_h : density of dry air, water vapour and moist air.
T_v : virtual temperature

Equation 1.26 becomes :

$$ZTD = 10^{-6}\int_{z_r}^{\infty} (k_1\rho_d R_d + k_2\rho_v R_v + k_3\frac{e}{T^2})dz \tag{1.31}$$

$$= 10^{-6}\int_{z_r}^{\infty} (k_1 R_d\rho_h - k_1 R_d\rho_v + k_2\rho_v R_v + k_3\frac{e}{T^2})dz \tag{1.32}$$

$$= 10^{-6}\int_{z_r}^{\infty} k_1\frac{P}{T_v}dz + 10^{-6}\int_{z_r}^{\infty} [(k_2 - k_1\frac{R_d}{R_v})\frac{e}{T} + k_3\frac{e}{T^2}]dz \tag{1.33}$$

$$= 10^{-6}\int_{z_r}^{\infty} k_1\frac{P}{T_v}dz + 10^{-6}\int_{z_r}^{\infty} k_2^{'}\frac{e}{T} + k_3\frac{e}{T^2}dz \tag{1.34}$$

The first term of equation 1.34 is called the Zenith Hydrostatic Delay (ZHD) which is caused by the total air density of the atmosphere. The second and third term together are called the Zenith Wet Delay (ZWD) and is caused by the density of water vapour in the atmosphere.

1.3 Estimation of tropospheric delay using GPS observations

1.3.1 Principles of tropospheric delay estimation

Though equation 1.34 gives quite good description of ZTD, we cannot account on using it directly to correct the tropospheric delays in the GPS measurement equation 1.4 during the practical GPS data processing due to the lack of meteorological data. Instead, ZTD is treated as an unknown parameter in GPS measurement equation 1.4 and is estimated together with all the other unknown parameters in this equation.

The least square estimation technique can be used when the estimation problem is over-determined. Over-determined means there are more observations than required to estimate the parameters. This is the case of GPS observations since a receiver can always observe more than 4 satellites at anytime of the day.

The estimation methods solve all the unknown parameters in equation 1.4 using the GPS phase measurement and a priori values provided for all the unknown parameters in the slant direction. A priori SHD (slant hydrostatic delay) is obtained by converting a priori ZHD value to different zenith angels using Neil dry mappping function. Those unknown parameters include the station coordinates, the tropospheric delay and the phase ambiguity. All these parameters are then adjusted iteratively using least square method until the residual becomes stable and minimized. Residual is defined as the difference between the GPS observables (the phase measurement) and the modelled observables. To perform the adjustment of each parameter, a stochastic model which is defined by an appropriate covariance matrix indicating the uncertainty of the parameters is needed. The minimized residual contains delays caused by the not modelled or badly modelled parameters like liquid water. At last, the estimated ZTD will be the average of all ZTD from different zenith angels.

From the estimation procedure and 1.4, we could see that the following two aspects effect the precision of the ZTD estimated value : 1. the precision of the phase observation measurement and 2. the precision of the modelled value used during the least square estimation. For the precision of phase measueremnt, the main error source include satellite and receiver clock error, multipath effect, phase center offsets for receiver and satellite antenna. The main error source for the modelled value comes from precision of mapping functions, earth orientation parameters(ocean loading, geodetic coordinates), satellite orbit, ionospheric effects.

Several softwares are available nowadays to perform the parameter estimation using the GPS raw observations collected from the receivers in order to obtain high precision station position and atmospheric parameters. The most used ones are GAMIT, GIPSY and BERNESE. The basic idea behind those three software is more or less the same as we have described above in the ZTD estimation method. PPP (precise point positioning) or network approach can be choosed acoording to the user's preference. PPP uses only the observation from one single GPS receiver while network approach uses many observations which forms some baselines. More information about those softwares can be found in the their official website : GAMIT : www-gpsg.mit.edu/ simon/gtgk/ GIPSY :gipsy.jpl.nasa.gov/orms/goa BERNESE :www.bernese.unibe.ch/

For the detailed strategy for ZTD estimation, such as the network configuration, parameter weighting and modelling of delay, length of the treating period and sliding window, see

1.3.2 A priori values of tropospheric delays for parameter estimations

In order to perform the parameter estimation, an a priori value of tropospheric delay between the satellite and the receiver is needed. This delay, called total slant delay (STD), can be expressed as a function of ZHD and ZWD under the help of mapping functions that project the values of ZHD and ZWD from zenith direction to the slant direction.

$$\Delta_{apr}^{T} = STD_{apr} = ZHD_{apr} m_h(\varepsilon) + ZWD_{apr} m_w(\varepsilon) \tag{1.35}$$

Δ_{apr}^{T} : the a priori value for tropospheric delay
$m_h(\varepsilon)$ and $m_w(\varepsilon)$: the mapping function for hydrostatic and wet zenith delay, respectively.

Mapping functions of Niell (1996) have been the most widely used in GPS parameter estimating. Under the assumption that the STD is assumed to be azimuthally symmetric, the Niell mapping functions can be written as a function of the elevation angle ε :

$$m_h(\varepsilon) = \frac{1 + \frac{a_h}{1+\frac{b_h}{1+c_h}}}{\sin\varepsilon + \frac{a_h}{\sin\varepsilon + \frac{b_h}{sin\varepsilon + c_h}}} \tag{1.36}$$

$$m_w(\varepsilon) = \frac{1 + \frac{a_w}{1+\frac{b_w}{1+c_w}}}{\sin\varepsilon + \frac{a_w}{\sin\varepsilon + \frac{b_w}{sin\varepsilon + c_w}}} \tag{1.37}$$

ε : the angle between the slant delay path and the horizon
a_h, b_h, c_h : coefficients for hydrostatic delay mapping function
a_w, b_w, c_w : coefficients for wet delay mapping function

Equation 1.36 and 1.37 showed that $m_h(\varepsilon)$ and $m_w(\varepsilon)$ have the same form, only the coefficients are different. These coefficients vary in function of latitude, altitude and day of the year.

The a priori value of ZHD and ZWD in equation 1.35 generally followed Saastamoinen (1972). A hydrostatic delay model was derived from the first term of equation 1.34 under the assumption that the dry atmosphere is in hydrostatic equilibrium.

$$ZHD_{priori} = 10^{-6} \frac{k_1 R_d P_s}{g_m} \tag{1.38}$$

$$= \frac{0.002277 P_s}{f(\lambda, H)} \tag{1.39}$$

$$f(\lambda, H) = 1 - 0.0026(\cos 2\lambda) - 0.000279H \tag{1.40}$$

P_s : surface pressure in [hPa]
λ : latitude of the GPS receiver
H : GPS receiver height from sea level in [km]
g_m : mean gravity acceleration in $[m/s^2]$

The advantage of this formulation is that only a surface pressure value is needed instead of the vertical distribution of the pressure above the receiver. This makes the calculation of hydrostatic delay much simpler and practical. The main error in this model comes therefore from the surface pressure value. Originally, the surface pressure value used for the a priori estimation was the standard surface pressure value (1013.15 hPa). Nowadays, the surface pressure values used for the a priori ZTD come from atmospheric model climatologies or from Numerical Weather Prediction analyses of the day. If in-situ meteorological measurements are available near the GPS receiver and are used to compute the a priori ZTD, then the precision of this estimation can be at millimeter level (if the atmosphere is hydrostatic).

The following equation is generally used in order to provide a priori value for ZWD.

$$ZWD_{priori} = \frac{0.002277(\frac{155}{T_s} + 0.05)e}{f(\lambda, H)} \tag{1.41}$$

with T_s : surface temperature in [K]

1.3.3 From tropospheric delay to Integrated Water Vapour

Once ZTD is estimated by the GPS processing, we can estimate ZWD by subtracting ZHD calculated with equation 1.39 using the in-situ surface pressure measurement at the receiver location :

$$ZWD = ZTD_{estimated} - ZHD_{saastamoinen} \tag{1.42}$$

Methods have been developed to convert ZWD into a more comprehensive meteorological parameter, *i.e.* the Integrated Water Vapour (IWV). IWV maybe also expressed as precipitable water in (mm). Knowing the ZWD value, IWV above a GPS receiver is obtained through a conversion factor κ which is proportional to the atmosphere temperature profile (Askne and Nordius, 1987), referred to as κ_a hereafter :

$$IWV = \kappa \cdot ZWD \tag{1.43}$$

$$\kappa_a \approx \frac{10^8}{R_w(\frac{k_3}{T_m} + k_2')}) \quad \text{with} \quad T_m = \frac{\int_L \frac{e}{T} dz}{\int_L \frac{e}{T^2} dz} \tag{1.44}$$

But in practice, it would require atmospheric temperature profiles at each GPS receiver. Several authors have therefore proposed approximation of κ using surface temperature measurements instead of the vertical temperature profile. Among the different proposals, Emardson and Derks (2000) have developed a conversion factor

κ which depends on the site and region, by performing statistical regressions using radio-sounding data. For the Mediterranean region's climate conditions, for instance, κ is given by :

$$\kappa \approx \frac{10^3}{6.324 - 0.0177(T_S - 289.76) + 0.000075(T_S - 289.76)^2} \quad (1.45)$$

1.4 Advantages, precision and applications of GPS observations for atmospheric water vapour monitoring and meteorological applications

1.4.1 Advantages of GPS water vapour measurements

Besides GPS water vapour measurements, water vapour observing systems include radiosondes, surface-based microwave radiometers, satellite-based microwave remote sensing, infrared radiometers and research-based aircraft measurements. Each system has its own advantages and limitations. Radiosondes provide information about the vertical distribution of water vapour in the atmosphere, but they are only launched twice per day. Surface-based radiometers have high temporal resolution, but with a high operational cost and their biggest default is that they do not function well in all weather conditions, especially when it is raining. Satellites provide global coverage, but they have difficulty in observing the lower part of atmosphere where water vapour concentrates. Finally infrared radiometers are only capable of providing data under clear sky conditions. Research aircraft-based measurements are expensive. Some commercial aircrafts started to install airborne instruments to measure the water vapour. But this method can only perform the measurement along the flight path and during the take off and landing near the major airports. VLBI (Very Long Baseline Interferometry) is a technique developed within the geodetic community and it can also measure the water vapour with a good precision. But it is very expensive to make such measurements.

Advantages of the GPS technique compared with the existing water vapour measuring methods are : It is not affected by the rain and clouds thanks to its carefully chosen carrier wave frequency. It is inexpensive, accurate, reliable and is capable of being widely deployed. But GPS is unable to provide vertical profile of water vapour measurements. Only tomography based on high resolution GPS networks can supply an information about the vertical distribution of water vapour.

1.4.2 Precision of GPS water vapour measurements

Dixon and Wolf (1990) were the first to suggest using GPS techniques to estimate the water vapour in the atmosphere. Based on the method of Bevis et al. (1992), several experiments were set up during 1993-1994 which proved the possibility to monitor the total water vapour amount above the ground-based GPS station in a continuous manner, with a precision of a few mm and a time resolution less than 1h (Rocken et al., 1993, 1995). Comparison between the GPS obtained total water vapour amount and the water vapour radiometer showed a good match.

Encouraged by the above results, various comparison between the GPS derived water vapour data and other water vapour measuring instruments such as radiometer and radiosondes have been carried out to prove the coherence of these observations methods (Kuo et al., 1993; Rocken et al., 1993; Businger et al., 1996). At the same time, several efforts of improving the techniques like using the improved mapping function (Niell et al., 2001) and correction of antenna phase center (Fang et al., 1998) have lead to a better agreement of these observation methods. The common conclusion drawn is that GPS derived water vapour data have the same level of accuracy as radiosondes and microwave radiometers (Elgered et al., 1997; Emardson et al., 1998; Duan et al., 1996) with a precision better than 10mm.

1.4.3 Meteorological applications of GPS water vapour measurements

After the accuracy of GPS observations being evaluated and proved by comparing with conventional methods, the potential and feasibility of using GPS observations in meteorological applications have been examined. Comparison of GPS observations with the NWP model data and measurements from radar and satellite have been carried out. Yang et al. (1999), Cucurull et al. (2000), Bock et al. (2004) and Bock et al. (2005) have suggested that GPS can be a useful tool for the validation of the model forecast as the radiosonde data, but with even higher temporal and spatial resolution. Moreover, it can be used as a quality control tool for radiosounding observations or satellite calibration and validation.

In the recent years, with the development of data assimilation techniques, GPS data is no longer only a validation or calibration tool for the meteorology models or satellite data, but also can be used directly to improve the model initial conditions. Research has been thus centered on the variability and the error statistics of GPS observations in order to prepare the GPS data assimilation in NWP models. Haase et al. (2003) have worked on estimation of the error statistics necessary for future assimilation of GPS ZTD data in numerical models. A continuous time series of GPS ZTD data over two years of duration have been compared with the equivalent values calculated from radiosonde data. The standard deviation of the difference of the two data sets is about 12mm and the bias about 7mm. The same comparison has been done between the GPS ZTD data observations and the HIRLAM estimates with a 18mm of standard deviation found. They have pointed out that the standard deviation has a

strong dependence on the amount of humidity which follows an annual variation with higher values during the summer time. They suggested that the better agreement with radiosonde data than HIRLAM estimates indicate that the NWP models could benefit from the additional information provided by GPS data if assimilated.

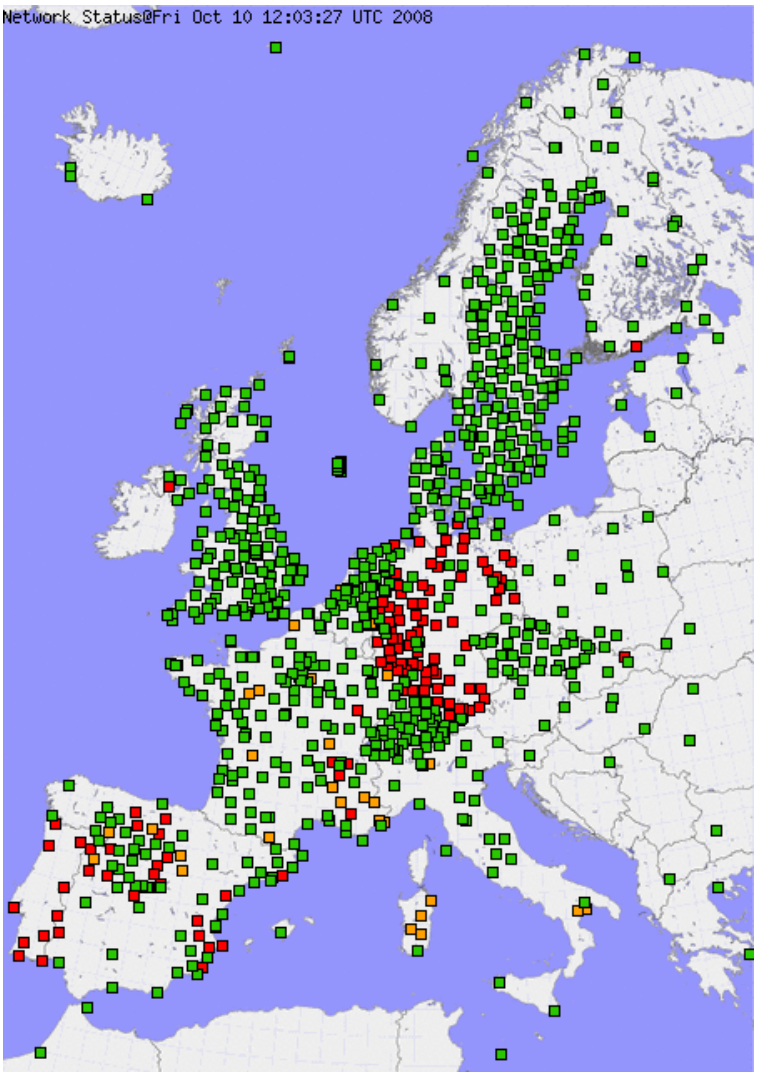

FIG. 1.2: E-GVAP network status on 10/10/2008

Efforts to organise the GPS water vapour network for serving the atmospheric research have been carried out worldwide. The United States, Japan, Germany, Sweden, Korea have built up their own continuous operational GPS networks. A ground-based GPS real-time network was set up by NOAA (National Oceanic and Atmospheric Administration) Forecast System Laboratory. Its main research goal is to study the impact using the high temporal resolution of GPS derived total water vapour data in weather forecast. It has more than 200 stations by now. In 2008, NOAA is planning to extend this network into an operational ground-based GPS water vapour network covering the whole United States, with 400 main stations and more than 600 expan-

ded stations. In Japan, the high-density Japanese geodetic GPS network (GEONET) contains more than 1000 GPS receivers and is used for water vapour monitoring. Those networks have been served as solid bases for developing the GPS applications in meteorology, especially in data assimilation.

With the potential of GPS derived ZTD or IWV data being exploited in the meteorological research and operational applications, a few projects have been organized in Europe with the aim to bring together different GPS observation network providers in order to perform certain research tasks under coordination. Such projects are MAGIC, COST716, TOUGH, E-GVAP. MAGIC (Haase et al., 1997, 2001) was the first effort in Europe to demonstrate the use of GPS IWV data in meteorological applications. COST716 was the first action on the exploitation of real-time GPS observations in operational meteorological application. The TOUGH project has developed the methods enabling optimal use of GPS data from existing European GPS stations in numerical weather prediction models. The impact of such data upon the skill of weather forecasts were assessed. E-GVAP pushes the results of the former 3 projects one step further towards the operational meteorological applications such as collecting and providing near-real time GPS observations for data assimilation in the meteorological centres. At the same time, E-GVAP tries to feed back the meteorological data to the geodetic community. This leads to a mutual benefits for both the meteorological and geodetic community.

In practice, E-GVAP aims to bring together the GPS networks which belong to the individual companies, institutes or the states. A GPS processing center is the place where observations from one or more GPS network is received and then processed into different products, such as ZTD and IWV. There are several GPS data processing centres (Tab. 2.1) now in Europe mainly distinguished by the country. Different processing strategies are used by the centres. Those strategies include : constrain used, configuration of the network shape, software used, iteration procedure, IGS satellite orbits, averaging time (sampling time), modelling differences.

Under the framework of E-GVAP, GPS water vapour observations from more than 400 GPS sites (Fig. 1.2) processed by different centres are being sent to a common ftp-server at the UK Met Office within 1h45 after the valid time of the observation. When the data are received they are converted from the COST716 ASCII format into BUFR format, then the BUFR format files are distributed on the GTS (Global Telecommunication System). They will be then distributed to meteorological centres where data assimilation of GPS observations is performed for operational or research purposes.

In France, the research observation service RENAG (http ://renag.fr/) aims at gathering GPS observations from operational and research networks over France for geodetic and atmospheric research such as the earth crust deformation and water vapour variation. It includes some regional research GPS networks with high spatial resolution, such as the one originally initiated under the framework of the research observation service OHM-CV (http ://www.lthe.hmg.inpg.fr/OHM-CV/) in south of France. Denser networks dedicated to GPS tomography purposes (Champollion et al.,

TAB. 1.1: The different processing centres of E-GVAP - 3 December 2008, http ://egvap.dmi.dk/

Center	Software version	Observation Frequency (minutes)
SGN1	BERNESE 5.0	15
NGAA	GIPSY	5
KNM1	BERNESE V5.0 abs-pcv	15
GOP	BERNESE V5.0	60
IESG	BERNESE V5.0	15
ROB	BERNESE V5.0	15
GFZ	EPOS P.V2	15
IEEC	GIPSY	10
ASI	GIPSY NET	15
LPT	BERNESE V5.0	60
SGN	BERNESE 5.0	15
LPTR	GPSNET 2.0	5
BKG	BERNESE V5.0	60
METO	BERNESE V5.0	15
KNMI	BERNESE V5.0 abs-pcv	15

2005) are also setting up during field experiments such as the one deployed for the COPS field campaign during summer 2007.

Chapitre 2

Etat de l'art de l'assimilation de données GPS

Résumé

Ce chapitre commence par présenter les différentes méthodes d'assimilation qui ont été utilisées dans les études antérieures pour assimiler les données GPS. Les premières études ont impliqué des méthodes d'assimilation par relaxation newtonienne (ou nudging) ou par interpolation optimale, et les données assimilées étaient les contenus intégrés en vapeur d'eau déduits des délais zénithaux totaux GPS. Avec le développement des méthodes d'assimilation de données variationelles, l'assimilation des délais zénithaux totaux au lieu du contenu intégré en vapeur d'eau est devenue plus répandue.

Les études antérieures montrent que l'assimilation des données GPS a un impact principalement sur le champ d'humidité analysé. Un impact positif sur la prévision des forts cumuls de précipitation est généralement trouvé, alors que l'impact est plus mitigé sur la prévision des précipitations plus faibles. Le bénéfice de l'assimilation est plus évident dans les situations où la distribution d'humidité connaît une évolution temporelle importante. Plusieurs études ont aussi montré un bénéfice mutuel à assimiler les observations GPS avec d'autres types de données (observations de surface, profileurs de vent, etc). Un réseau GPS plus dense, ou des stations GPS situées en amont du flux affectant la région augmentent l'impact de l'assimilation des données GPS sur les prévisions.

Ces études ont ainsi démontré le potentiel des observations GPS de ZTD pour améliorer la prévision des fortes précipitations réalisées par des modèles à méso-échelle (résolution inférieure à 6 km). Dans un contexte d'augmentation de la densité des réseaux GPS et du développement de modèles atmosphériques aux échelles kilométriques, le travail de thèse présenté ici avait pour objet de développer et d'évaluer l'impact de l'assimilation variationelle de données de ZTD sur la prévision à fine échelle de systèmes convectifs. Ces travaux constituent une première puisqu'aucune étude antérieure n'a utilisé une résolution aussi fine. A ces échelles, la convection n'est plus paramétrée mais résolue explicitement. Comme l'a montré l'étude de Ducrocq et al. (2002), des réponses différentes peuvent être obtenues à partir des mêmes condi-

tions initiales entre un modèle utilisant une convection paramétrée et un autre une convection profonde résolue et constitue une des motivations de ce travail de thèse. Notre démarche guidée par ces considérations et aussi par l'état d'avancement du nouveau modèle AROME a été constituées de deux étapes :

Dans une première partie, nous avons réalisé des expériences d'assimilation avec le système 3DVAR ALADIN de Météo-France à une résolution de 9.5 km de résolution et testé son impact sur la prévision à fine échelle (2.4 km) de MESO-NH

Puis, l'assimilation des données de ZTD a été développée et évaluée dans le nouveau système de prévision numérique AROME de Météo-France à 2.5 km de résolution.

2.1 Introduction

Numerical Weather Prediction (NWP) is basically an initial condition problem : given the starting state of the atmosphere such as pressure, temperature, wind and humidity, the evolution of those variables follows the time integration of a set of non-linear partial differential equations (Holton et al., 1992). There are several factors deciding the accuracy of this integration results : accuracy of the initial field, numerical representation of atmospheric dynamics, the physical parametrization and accuracy of the lateral boundary conditions when it is a limited area model.

Data assimilation is the procedure for importing the information content of observations into the NWP system to produce the initial conditions while keeping the consistency with the model dynamics and physics. Assimilation can be defined as : an analysis technique in which the observed information is included into the model state (also called *background field* or *first-guess*) by taking advantage of consistency constraints with laws of time evolution and physical properties (Bouttier and Courtier, 2002). The output of data assimilation is an estimate of the atmospheric state, called *analysis*, which serves as the initial condition for the deterministic forecast.

One of the current challenges in NWP is the improvement of short-term cloud and precipitation forecasts. Water vapour is quite a crucial factor in achieving this goal since it plays a key role in several processes, such as cloud and rain formation, the absorption and emission of latent heat, solar and terrestrial radiative transfer. However, partly due to its high spatial and temporal variability, the water vapour field is one of the less well-described parameters in initial conditions of numerical weather prediction systems. Moreover, it is in fact one major source of uncertainties in short-range high-resolution forecasts of precipitation (Kuo et al., 1996). Ducrocq et al. (2002) have demonstrated the importance of initial humidity fields in improving kilometric scale forecast of heavy precipitation events. As mentioned in the previous chapter, GPS observations such as ZTD, ZWD, and IWV appeared therefore appealing to improve initial humidity fields at mesoscale. Data assimilation of those GPS observations have therefore been developed using various assimilation techniques to incorporate GPS water vapour information into the NWP models.

2.2 The different techniques used for GPS data assimilation

Several data assimilation methods are used in meteorology. The choice of which method to use is depending on the type of NWP models used and the computational resources available at each institution. Following we only introduce the brief principle of the assimilation techniques that have been used by the GPS data assimilation research in the past years. Methods that have been used in assimilating GPS observations into NWP models so far include : nudging, OI, 3DVAR and 4DVAR.

2.2.1 Nudging

Nudging is a method that gradually relaxes the model state towards observations or a grid-point analysis of observations. The influence of the observation to the grid points around is based on the distance between them. The weight and coefficients of the relaxation method are pre-fixed empirically. More detailed descriptions on the nudging technique can be found in Hoke and Anthes (1976) or Stauffer and Seaman (1990). Integrated water Vapour or Precipitable water measurements can not be used directly in the nudging data assimilation scheme since PW is not a model prognostic variable. Thus retrieval methods must be used to recover the water vapor mixing ratio from PW data.

2.2.2 Optimal Interpolation(OI)

The optimal interpolation (OI) is a statistical method and is based on the least square estimation theory in contrast to the nudging technique. The least-squares analysis is obtained as :

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - H[\mathbf{x}_b]) \tag{2.1}$$

with the gain (weighting) matrix defined as :

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \tag{2.2}$$

with
- $\mathbf{x}_a$: the analysis model state
- $\mathbf{x}_b$: the background model state
- $\mathbf{y}$: the vector of observations
- H : the observation operator
- $\mathbf{B}$: covariance matrix of the background errors
- $\mathbf{R}$: covariance matrix of the observation errors

However, the huge dimension of the least-squares analysis problem necessitates some simplifications in practical applications. Optimal interpolation (OI) is an algebraic simplification of the computation of the weight $\mathbf{K}$ in order to solve the least-squares analysis problem operationally. The fundamental hypothesis in OI is : For each model variable, only a few observations are needed for determining the *analysis increment* $(\mathbf{x}_a - \mathbf{x}_b)$. In practice, assuming that the background error covariances are small for large separation, only the observations in a limited geometrical domain around the model variable need to be selected.

The advantage of OI is its simplicity of implementation and relatively small cost if the observation selection is efficient. The shortcoming of OI is that due to the use of different sets of observations, noise could be produced in the analysis fields. Also the analysis increment is determined one grid point by one grid point sequentially. This method had been widely used in the NWP centers worldwide in the seventies.

Nowadays, most of the NWP systems have replaced OI by variational methods.

2.2.3 Three-dimensional variational data assimilation (3D-VAR)

It can be shown that the optimal least-squares estimator given by Eq.2.1 and 2.2 is equivalently obtained as a solution to the variational optimization problem :

$$x_a = ArgminJ \tag{2.3}$$

$$J(\mathbf{x}) = J_b(\mathbf{x}) + J_o(\mathbf{x}) \tag{2.4}$$

$$J_b(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) \tag{2.5}$$

$$J_o(\mathbf{x}) = \frac{1}{2}(\mathbf{y}-H(\mathbf{x}_b))^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{y}-H(\mathbf{x}_b)) \tag{2.6}$$

$J(\mathbf{x})$ is the cost function of the analysis, $J_b(\mathbf{x})$ is the background term and $J_o(\mathbf{x})$ is the observation term. The optimal estimate of the analysis is obtained by minimizing this cost function which is defined as the quadratic distance between the modelled state of the atmosphere $\mathbf{x}$ and the other resources of information, such as background $\mathbf{x}_b$ and observations $\mathbf{y}$; these quadratic distances being weighted by the covariance error matrix of background ($\mathbf{B}$) and observations ($\mathbf{R}$). x_a is the maximum likehood under the hypothesis that the probability density functions of background and observational errors are Gaussian.

The main difference which distinguishes the variational method with the OI method is that analysis is solved using a different numerical approach : Analysis in variational methods is found as an approximate solution which minimizes the global cost function $J(x)$. The minimization problem becomes to solve the analysis $\mathbf{x}_a$ which makes the gradient of J equal 0. This avoids the computation of the gain $\mathbf{K}$ in equation 2.2, which costs a lot of computation time. The analysis is sought iteratively by performing several evaluations of the cost function $J(x)$ and of its gradient :

$$\nabla J(\mathbf{x}) = 2\mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) - 2\mathbf{H}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}(\mathbf{y}-H(\mathbf{x}_b)) \tag{2.7}$$

In practice, the initial point of the minimization, called first-guess, is equal to the background $\mathbf{x}_b$.

The main advantage of the variational method compared with OI is that it takes into account all the observations in the domain at the same time to solve a global minimization problem (Parrish and Derber, 1992b). With such minimization approach, the assumptions made when solving with the OI method are not needed anymore (Parrish and Derber, 1992b; Courtier et al., 1998). Another very important feature is that it allows operationally the use of indirect observations which have a non-linear relationship with the model variables. For instance, those observations include as instance satellite radiances, radar reflectivity and GPS atmospheric delays for example. H in equation 2.6 is called the observation operator and it plays a very important role in the variational data assimilation since it allows using indirect observations in the data assimilation system. It calculates the model equivalent value of a certain

observation type using the model background variables $\mathbf{x}_b$ at each observation point. The departure between the observation value and the model equivalent value is called the *first-guess departure* $\mathbf{d}$ (or the *innovation vector*), denoted in the equation of variational methods as $\mathbf{d} = y - H[\mathbf{x}_b]$, where H is the non-linear forward observation operator.

In order to perform the minimization, this non-linear forward operator has to be linearized around the background $\mathbf{x}_b$, thus we need to write a tangent-linear version of the observation operator, which is noted as $\mathbf{H}$ in equation 2.7. Also, in real applications, incremental formulations of 3DVAR can be used in order to reduce the cost of solving the variational problem. This is performed by reducing the resolution of the increments in the minimization.

Since the beginning of 90', OI has been replaced gradually by variational methods in the NWP centers worldwide. The details of the 3DVAR technical aspects in the operational ALADIN and AROME NWP systems of Météo-France will be presented in Chapter 3.

The observation error covariance matrix R

In one-dimensional case, the observation error covariance matrix $\mathbf{R}$ is reduced to become the variance of the observation errors at each observation location for each observation type. For the three dimensional case, the observation errors are assumed to be un-correlated in space, which means the measurement from location A is independent from the measurement from location B. With this assumption, the observation error covariance matrix R turns out to be diagonal and only the observation error standard deviations σ_o need to be specified. In theory, this error should include the instrumental, representativeness and forward observation operator errors. Proper evaluation of these errors is one of the difficulty of variational methods. $\mathbf{R}$ is also used in the observation quality control procedure to decide whether or not to reject an observation.

The background error covariance matrix B

One of the major difficulties in 3DVAR method is to design the matrix $\mathbf{B}$ that properly describes the background error covariances for all pairs of model variables. $\mathbf{B}$ is a very critical part in 3DVAR techniques because it decides the structure of the analysis increments. In another way, it means the quality of the analysis in 3DVAR mainly relies on whether $\mathbf{B}$ can well represent the atmospheric situation in that region and at that time. While $\mathbf{R}$ is usually assumed to be diagonal, this is not the case for matrix $\mathbf{B}$. The non-diagonal elements in $\mathbf{B}$ are very important for deciding the analysis increments, as they determine the spatial spreading of the corrections of the background state due to the observations in the region around that observation. Besides that, $\mathbf{B}$ also contains the information that determines the spreading of observational information concerning one variable to other type of variables not

directly related to the observations. For instance, the correction issued from the observation of temperature will also lead to corrections of pressure and wind. Thus, $\mathbf{B}$ is important in keeping the dynamic consistency or balance between the model variables.

In 3DVAR, the common way to estimate $\mathbf{B}$ does not depend on the observation at all as in OI ; instead, it is evaluated solely from the difference of forecasts at different validation time or starting from different initial conditions. The common methods to estimate the $\mathbf{B}$ matrix are NMC method (Parrish and Derber, 1992a) and ensemble methods (Berre et al., 2006).

2.2.3.1 4DVAR

4DVAR is a generalization of the 3DVAR assimilation method by taking into account the time dimension. Observations which are not close to the analysis time can therefore be included in the assimilation process with 4DVAR assimilation and more observations are potentially used. 4DVAR has the same cost function form as 3DVAR, except that the observation operators include also the forward meteorological model M which is used to compare the model state (produced by $M_i(\mathbf{x}_0)$) and the observations $\mathbf{y}_i$ at the appropriate time i :

$$J(x) = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_b)^T\mathbf{B}^{-1}(\mathbf{x}_0 - \mathbf{x}_b) + \frac{1}{2}(H(M_i(x_0)) - \mathbf{y}_i)^T\mathbf{R}^{-1}(H(M_i(x_0)) - \mathbf{y}_i) \quad (2.8)$$

$\mathbf{B}$ ($\mathbf{R}$) is still the background (observation) error covariance matrix as used in 3DVAR.

The disadvantage of 4DVAR is its requirement of large computational time as it needs several time integrations of the model over the 4Dvar time window. It also needs the tangent linear version of the numerical model in the minimization of the cost function, which can be a problem for the highly non-linear physical parametrizations used in some atmospheric models. In real time applications, 4DVAR requires the assimilation to wait for the observations over the whole time window whereas 3Dvar (or OI) can assimilate the observations just after they are available.

4DVAR uses a flow-dependent background error covariance matrix for estimating the atmospheric state. This is an advantage compared with 3DVAR, since a statistically estimated background error covariance is pre-settled and used in 3DVAR. And such pre-settled values cannot reveal the changing state of the background fields.

Presently, 4DVAR is only used by a few very large meteorological services such as ECMWF, Météo-France, UK Met office, Japan Meteorological Agency and Canadian Meteorological Center, and mainly for large scale NWP systems.

2.3 Assimilation of GPS IWV observations using nudging and OI methods

With the growing interest in GPS for meteorological applications, studies have been carried out in investigating the impact of ground GPS data assimilation on improving the analysis quality and forecast skill for different weather conditions. The experiments that have been conducted, mainly focused on testing the impacts on the moisture field and rainfall prediction. Various methods have been developed in order to assimilate those GPS observations into different atmospheric models.

Before the use of variational methods in the meteorological centres, nudging and OI were used as the first efforts to introduce the GPS derived water vapor information into the NWP models. Nudging can only assimilate the observations which are the model variables. In this case, IWV needs to be converted into the model humidity such as vapour mixing ratio or specific humidity. Theoretically OI can assimilate indirect observations under the help of observation operator, but in reality, it is hard to use indirect observations such as ZTD or satellite radiances due to the computation costs with OI algorithms.

Some preliminary but essential conclusions about IWV assimilation experiments on case studies have been drawn under the work of Kuo et al. (1996), Falvey and Beavan (2002), Guerova et al. (2004), Pacione et al. (2001), Tomassini et al. (2002), Nakamura et al. (2004) using nudging or OI assimilation techniques. Those conclusions have been serving as a good guidance for the research carried out later and they include : *i)* overall positive impact in the rainfall prediction with IWV data ; *ii)* importance of upwind moist information coming from the upwind GPS stations for the improvement of rainfall precipitation skills (Falvey and Beavan, 2002; Tomassini et al., 2002; Nakamura et al., 2004) ; and *iii)* better results obtained when wind and temperature data are assimilated together with IWV observations (Pacione et al., 2001; Falvey and Beavan, 2002).

IWV data assimilation has been also carried out continuously for a long time in U.S.A. at NOAA. IWV observations from a wide coverage of GPS network in North America have been assimilated into the research version of Rapid Updated Cycle (RUC) NWP model. The technique used to assimilate precipitable water observations into RUC is an OI-based columnar adjustment (Kuo et al., 1993; Smith et al., 2000). The basic idea is to compare the model value of IWV, obtained by integrating vertically the model specific humidity value at the GPS location, to the observed GPS IWV value. The ratio between the two is used to scale the vertical profiles of specific humidity keeping their shapes so that the integrated value is equal to the analyzed IWV. The verification of five years of experiment with 3-h RUC at 60-km resolution against the radiosonde data revealed a steady increase in the positive impact in short-range relative humidity forecasts below 500hPa (Gutman et al., 2004; Smith et al., 2007). This improvement on relative humidity forecasts goes up with the increasing of GPS stations during the 5-year study period. Gutman et al. (2004) suggested future

increases in the network density over United States would result in further forecast improvements. Also this paper pointed out that the greatest improvement in relative humidity forecasts is found in lower levels (below 700hPa) and during cold season where moisture change is dominated by synoptic scale weather change. This proved again the former research conclusion by Smith et al. (2000) that the impact of IWV is more substantial when the atmosphere is more active and undergoing a rapid change of water vapor distribution.

Further work has been done by Smith et al. (2007) with a more recent version of the RUC at 20-km resolution over 3-month trial period. The strongest impact on precipitation forecast is within 6-h forecast range, however small but positive improvements are still observed up to 12 hours. Evident impact is found for relative humidity under 850hPa. In the severe convective case that they studied, the inclusion of IWV data improved the forecasts of relative humidity and CAPE (convective available potential energy), which is an important predictor of severe storm potential. The degree of IWV impact is found depending on the season, time of day and geographic location. The GPS IWV data have a greater impact during the cooler seasons and during nighttime. Indeed, during the night or cold season, when the vertical mixing is less strong than daytime or warm weather, the conventional surface observations are not enough representative of the moisture situation aloft and thus GPS observations are more useful to correct the low to mid-level troposphere.

2.4 Assimilation of GPS observations using variational methods

Theoretically OI can assimilate the observations, which is not directly the model variables. But those observations are only limited in those who have a linear relationship with the model variables due to the numerical implementation limitations. Recently, variational assimilation methods have been implemented in the NWP systems. One reason why the variational method is considered to be superior (Rabier et al., 1993) is that it allows for non-linear relations between observations and the model variables to be retrieved with a relatively smaller computational cost. The introducing of the variational methods has therefore started a new era that not only IWV but also ZTD observations which is a non-linear function of model temperature, pressure and humidity, can be assimilated.

2.4.1 Assimilation of IWV observations

Guo et al. (2000) was among the first to assimilate IWV using a four dimensional variational assimilation method. The system used in this study is 10-km resolution MM5/4DVAR system, which is based on the Pennsylvania State University-National Center for Atmospheric Research non-hydrostatic mesoscale model. They have asses-

sed the relative importance of different mesoscale observations such as profiler winds, surface virtual temperature and GPS IWV observations. They have found out when only IWV is assimilated, there is only a relatively small influence on the recovery of the moisture structure. Indeed, impact lies only in the prediction of convective rainfall. In order to achieve both positive impacts on rainfall prediction and moisture structure, profiler wind data need to be assimilated together with GPS IWV observations.

Then Koizumi and Sato (2004) have used the 4DVAR data assimilation associated with the 10-km resolution mesoscale model (MSM) of Japan Meteorological Agency to investigate the impact brought by combination of GPS IWV data and satellite water vapor data (TMI). GPS IWV data from the Japanese geodetic GPS network (GEONET) provided high density of water vapor information inland. The TMI precipitable water data covers mainly the ocean region. The combination of precipitable water vapor data both inland and from the ocean yields quite good improvements in the rainfall prediction. However, no experiment has been done in their study using only the GPS IWV data, due to the expensive computational cost of their data assimilation experiment.

This gap was soon filled up by the research work of Nakamura et al. (2004) using the same 4DVAR system. However, neutral impacts on the rainfall forecasts were obtained. The authors incriminated improper deformation of the vertical profiles of water vapor which affect the vertical static stability of the atmosphere, one of the key factors to control cumulus convection and cloud formation. Thus they suggested that the vertical profile of background error covariance of relative humidity might be not well adapted for assimilation of IWV in their assimilation system. Guo et al. (2005) suggested also that the problem encountered with the moisture structure during IWV assimilation could be due to a badly defined $\mathbf{B}$ matrix, since $\mathbf{B}$ is the key factor in variational methods method to define the structure of analysis increment. Using an improved $\mathbf{B}$ matrix improves in their case the convection forecast.

2.4.2 Assimilation of ZTD observations

At the same time, the direct assimilation of ZTD observations started to draw the interests of research. The direct assimilation of ZTD can reduce the error originated from converting ZTD to IWV as detailed in Brenot et al. (2006). With the implementation of variational methods worldwide in the meteorological centers, potential of ZTD assimilation especially on the rainfall prediction has been and is still being evaluated for different weather situations.

Several experiments have been carried out using relatively low model resolution by Poli et al. (2007), Macpherson (2007) and Vedel and Huang (2004a).

Poli et al. (2007) has conducted experiments of ZTD data assimilation using the Météo-France 4DVAR/ARPEGE global model assimilation system with a horizontal resolution of 40-km. Impact of ZTD assimilation for a long time period of spring, sum-

mer and winter 2005 has been evaluated. All three trials suggested a positive impact of the ZTD data in helping constrain the synoptic circulation in 1-4 day forecasts. Positive impact on precipitation and an improved skill for 12-36 hour accumulated rainfall were found.

Vedel and Huang (2004a) have performed their study under a horizontal resolution of 0.45/0.15 degrees[1] respectively with 3DVAR/HIRLAM. ZTD data from 117 ground-based GPS stations belonging to the COST ACTION 716 network were assimilated. They have found neutral impact on the analysis of parameters such as geopotential, wind, temperature and humidity but positive impact for rainfall forecasts. An improved skill was found in the prediction of strong precipitation (>5mm/12h) when GPS observations were assimilated. This result together with the similar results obtained by Vedel and Huang (2003) and others strengthened the conclusion that GPS ZTD observations improve the prediction of strong precipitation. They have also demonstrated that this improvement found in forecasting high precipitation was not due to a bias of the ZTD observations which would lead to increase all the precipitation.

Macpherson (2007) have found a mitigated impact on precipitation forecasts using 15-km resolution 3DVAR system of the EC (Environment Canada) regional analysis and forecast model. The impact is however more positive than negative for forecasts of large precipitation amounts during the summer period. The importance of adding collocated surface weather data is mentioned in achieving the positive impact on precipitation forecasts in central U.S region. The study also showed the impacts for the winter cases are generally smaller because of the low precipitable water over North America at that time, but some positive impacts are found for the precipitation forecast. Together with the results of Gutman et al. (2004), it reinforces the conclusion that GPS data assimilation does a good job under rapidly changing moisture conditions. Also, the location of GPS stations in the upwind region has been found to have a positive impact on the forecast.

To conclude, the assimilation of GPS ZTD observations in coarse resolution NWP systems generally yields positive impact on the precipitation forecast, especially for high precipitation and for the short-range (12-36 hours). The amplitudes of the impact depend on the assimilation methods and the density and size of the GPS network. Recently, exploiting the potential of GPS data is more focused on high-resolution NWP systems which are better designed to make use of the high temporal and spatial resolution of the current GPS networks.

The work of De Pondeca and Zou (2001) is among one of the first attempts to assimilate GPS ZTD observations into the non-hydrostatic MM5 model using 4DVAR data assimilation. The impact of assimilating GPS ZTD observations from a dense GPS network was examined for a frontal system. A three-domain nested configuration was used, with 54, 18 and 6km horizontal resolution, respectively. The 4DVAR data assimilation was only performed within the most inner domain (6km). A small but beneficial impact on the short-range precipitation forecast has been found. The simultaneous assimilation of GPS ZTD and wind profiler observations were found to

[1] 0.15 degrees is about 15km in the middle latitude

be mutually improving the forecast. Peng and Zou (2004) using the same 4DVAR system at 6-km found for their part that the assimilation of GPS ZTD observations together with precipitation data improves the forecast of a winter storm compared to those obtained with the assimilation of only GPS ZTD data or only the precipitation observations.

Based on the former conclusion over the potential of ZTD on the mesoscale phenomenon prediction, Cucurull et al. (2004) have focused their study on a snowstorm over the western Mediterranean Sea. The non-hydrostatic 3DVAR/MM5 system has been used. They have conducted two parallel experiments to investigate the impact of the assimilation. GPS ZTD data are assimilated either into the low-resolution model domain (54 km) or into the 3-nested model domains, with 6-km horizontal resolution for the most inner domain. GPS ZTD data provided by 23 stations within the COST716 project are used in those experiments. They found out that the benefits of the assimilation of the GPS observations are quickly lost if the assimilation is only conducted in the coarser domain. They also found out that the maximum impact is achieved when ZTD observations are assimilated together with mesonet surface observations. The decrease of the root mean square (rms) error is found to be larger for the specific humidity as expected with a decrease of 18% of the rms.

As for the case of GPS IWV data assimilation, the performance of GPS ZTD data assimilation is also a function of the density of ZTD observations. Faccani et al. (2005) used ZTD observations from a very high-density GPS network over South Italy. They conducted daily operational data assimilation experiments using 3DVAR/MM5 at 9 km from winter 2003 to spring 2004. The results suggested that the assimilation of high density GPS observations improves the results even though it is not performed in a very high resolution system.

2.5 Synthesis

In this chapter, the various methods used to incorporate the GPS observations into the initial conditions of the NWP models have been described. The former studies involved OI or nudging methods to assimilate IWV. With the development of the variational data assimilation methods, assimilation of GPS ZTD observations instead of IWV becomes common.

Ground based GPS water vapor observations generally have an impact on the analysed moisture fields. Most of the studies showed a positive impact on precipitation forecasts for high precipitation events, whereas it is found generally neutral for weak rainfall. The benefit is more obvious for situations with rapid change in moisture distribution. Several studies have also shown mutual benefit when GPS observations are assimilated with other types of observations (profiler winds, precipitation). Also, the increasing of GPS station density and location of GPS stations in the upwind region have been found to increase the impact on the forecast.

The past research has demonstrated the potential of GPS ZTD observations in improving the heavy rainfall forecast, especially for deep convection cases. With the rapid increase of GPS network and the development of mesoscale numerical models down to kilometric resolution, we must take advantage of those developments. Thus in this study we use the 3DVAR assimilation systems available at Météo-France which allow the direct assimilation of ZTD observations, avoiding the error produced during the conversion to IWV. Our main goal is to investigate the impact of GPS data assimilation on kilometric scale forecasts. So far, no assimilation experiments of ZTD observations have been done under such high resolution. Moreover, the convection is not parametrized as in the previously quoted studies but explicitly resolved by the model. Ducrocq et al. (2002) had pointed out different responses to the same initial conditions between parametrized and explicitly resolved convection.

Based on the former research results, we choose also to assimilate all the observations assimilated by the NWP systems used in this study such as surface observations, wind profiler data, satellite radiances, sounding observations, etc together with ZTD observations. GPS observations from a regional campaign are also used to achieve a higher GPS observation density. In order to test the maximum potential of ZTD observations, convective rainfall events are also chosen as the case studies.

We have carried out our research in two steps :

- First, we perform the ZTD assimilation using the Météo-France 3DVAR/ALADIN system with a 9.5-km resolution. The analysis obtained will be used as the initial and lateral boundary conditions to carry out a simulation using the kilometric scale, non-hydrostatic model - MesoNH. The benefit of using different initial conditions (with and without assimilation of GPS ZTD data at 9.5 km) is worth evaluating for the high-resolution quantitative precipitation forecast. This is particularly relevant for operational meteorological centers that can not run operationally on their computers full high-resolution limited-area data assimilation systems but only limited-area high resolution numerical weather prediction models, using the analyses produced by other meteorological centers as initial conditions.
- Then, we perform the ZTD assimilation directly in the new 3DVAR/AROME assimilation system of Météo-France, running at 2.5 km. It is for the first time that GPS ZTD assimilation cycle experiments are performed in such a high-resolution (2.5-km), non-hydrostatic data assimilation system. AROME uses a physical package from the MesoNH one and dynamics of the non-hydrostatic version of the ALADIN model. Better modelled orography and explicitly resolved convection benefiting from this high resolution model can bring the following advantages concerning the assimilation of GPS ZTD data : First, with such highly resolved model orography, one can reduce the errors introduced during the vertical interpolation of model variables from the model surface to the real GPS station height. Second, high resolution, non-hydrostatic models provide better representation of deep convection processes. Improved first-guess and model equivalent ZTD are therefore expected, and thus weaker first-guess departures which helps the assimilation process.

TAB. 2.1: A summary of the past research on ZTD data assimilation

Author	De Pondeca and Zou (2004)	Cucurull et al. (2004)	Vedel and Huang (2004)	Our study
Assimilation model	Non-hydrostatic MM5/4DVAR	Non-hydrostatic MM5/3DVAR	HIRLAM/3DVAR	ALADIN/3DVAR and AROME/3DVAR
Resolution	6km	6km	15km	2.5km
Convection	deep convection parameterized	deep convection parameterized	deep convection parameterized	explicitely resolved
Results	Scores improved when precipitation observation is also assmilated	18% reduction in the RMS of specific humidity, maximun impact reached when surface observation is also assimilated	Positive impact over the precipitation forecast	?

Chapitre 3

Description des systèmes d'assimilation 3DVAR/ALADIN et 3DVAR/AROME

Résumé

Ce chapitre présente les modèles et systèmes d'assimilation utilisées au cours de la thèse : le système d'assimilation 3DVAR ALADIN, le modèle MESO-NH et le système d'assimilation 3DVAR AROME. Les différentes étapes nécessaires à l'assimilation des données GPS sont aussi détaillées : pré-sélection des observations, débiaisage, opérateur d'observation, estimation des erreurs d'observations.

Les similarités et différences entre ces différents systèmes sont discutées. Les systèmes d'assimilation 3DVAR/ALADIN et 3DVAR/AROME partagent une bonne partie de leur code (flux de données, opérateurs d'observation directs et adjoints,etc). Les principales différences résident dans la mise en oeuvre de ces systèmes : spécification des matrices de covariances d'erreur de l'ébauche et de variances d'erreur d'observations, nature de l'ébauche, fréquence du cycle d'assimilation,... Le modèle AROME et la version du modèle MesoNH utilisées dans cette étude utilisent une résolution similaire et les mêmes paramétrisations physiques. Les différences principales résident dans les schémas numériques des modèles.

3.1 The 3DVAR/ALADIN assimilation system

3.1.1 The ALADIN model

ALADIN (Aire Limitée Adaptation dynamique Développement InterNational) is a spectral limited-area version of the global ARPEGE model in Météo-France. It is used operationally over France, eastern and central Europe as well as some North African countries. The domain which is run at Météo-France (called ALADIN/France), covers Western Europe (Fig. 3.1) with a 9.5km horizontal resolution on a conformal Lambert projection grid. The version used in this study is the one which was operational at Météo-France in 2006 and 2007. The vertical discretisation is done according a following-terrain pressure hybrid coordinate with 41 vertical levels.

The physical parametrization package included the ISBA soil-vegetation-atmosphere transfer parametrization (Noilhan and Planton, 1989), a first order closure turbulence scheme based on (Louis, 1979), the RRTM longwave radiation scheme (Mlawer et al., 1997) and an improved version of the Fouquart's shortwave radiation scheme, an improved version of the Bougeault's mass-flux convection scheme (Bougeault, 1985) and a large-scale prognostic cloud and precipitation scheme (Lopez, 2002), a gravity wave drag parametrization. Prognostic variables are the wind components of the horizontal wind V , the temperature T, the specific humidity q and the logarithm of surface pressure log P_s. The prognostic equations are solved using semi-implicit and semi-lagrangian numerical schemes.

As all limited area models, ALADIN needs lateral boundary conditions which are provided by the ARPEGE model forecasts each 3 hours ; a Davies relaxation (Davies, 1976) is applied on the outermost points in order to smooth the transition between the ARPEGE lateral boundary conditions and the ALADIN solution.

3.1.2 The 3DVAR/ALADIN system

Since July 2005, the ALADIN/France model is running operationally at Météo-France starting from analyses obtained with a 3DVAR data assimilation scheme. The initial conditions are non longer provided by the ARPEGE analyses, only the lateral boundary conditions are still supplied by the ARPEGE forecasts. Figure 3.2 is an illustration of the 3DVAR ALADIN/France 6-hour assimilation cycle running each day. 42-h to 54-h short-range ALADIN forecasts are issued from each 6-hourly ALADIN analyses.

The ALADIN 3DVAR data assimilation software relies highly on the ARPEGE/IFS data assimilation system (common observation operators, data flow, incremental formulation, etc)

3.1.2.1 Incremental formulation used in the 3DVAR/ALADIN

Incremental formulation of 3DVAR (Courtier et al., 1994) is used in the 3DVAR/ALADIN system in order to reduce the computational cost during the minimization process. Incremental formulation has basically the same form as classic 3DVAR formulation

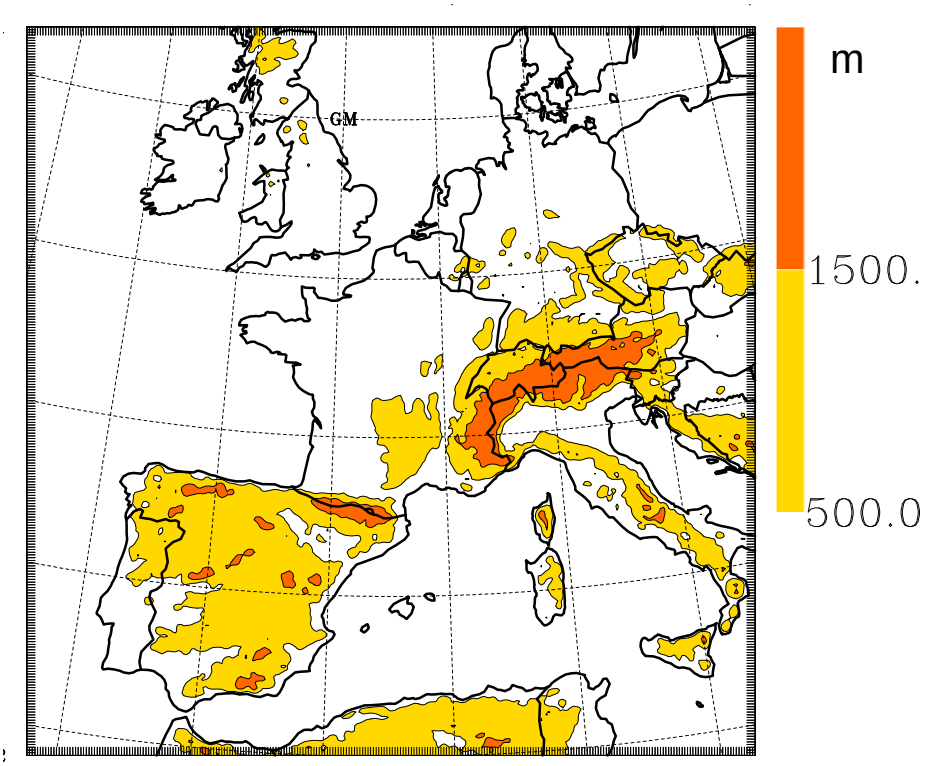

FIG. 3.1: The orography (m) of the ALADIN/France domain

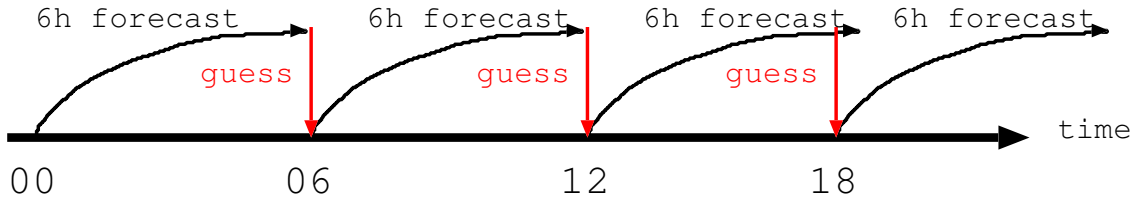

FIG. 3.2: 3DVAR/ALADIN assimilation cycle

as in equation 2.4 except that the correction $\delta\mathbf{x}$ is used as control variable instead of

$\mathbf{x}$:

$$J(\delta\mathbf{x}) = J_b(\delta\mathbf{x}) + J_o(\delta\mathbf{x}) \tag{3.1}$$

$$J_b(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^t\mathbf{B}^{-1}\delta\mathbf{x} \tag{3.2}$$

$$J_o(\delta\mathbf{x}) = \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - d)^t\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - d) \tag{3.3}$$

$$d = y - \boldsymbol{H}(\mathbf{x}_b) \tag{3.4}$$

$J(\delta\mathbf{x})$:	cost function
$\delta\mathbf{x}$:	analysis increment
$\mathbf{H}$:	linearized observation operator
$\boldsymbol{H}$:	direct observation operator
$\mathbf{B}$:	background error covariance matrix
$\mathbf{R}$:	observation error covariance matrix
$\mathbf{x}_b$:	background model state (first guess)
y :	observation vector
d :	departure

d is the departure between the observations and the background model state, interpolated to the observation space. $\boldsymbol{H}$ contains both spatial vertical and horizontal interpolations, and physical inversions or conversions if needed. $\boldsymbol{H}$ can be strongly non-linear. $\mathbf{H}$ is the linear approximation of $\boldsymbol{H}$ in the vicinity of the background $\mathbf{x}_b$.

The minimization of cost function is done under a low model resolution. The obtained increments $\delta\mathbf{x}$ is then changed from low model resolution to high model resolution and added up to the model background to obtain the new model state in high resolution. To find $\delta\mathbf{x}$ that minimizes $J(\delta\mathbf{x})$, the gradient of cost function $J(\delta\mathbf{x})$ is derived by differentiation with respect to $\delta\mathbf{x}$. At the minimum, the gradient vanishes :

$$\nabla_{\delta\mathbf{x}}J = \mathbf{B}^{-1}\delta\mathbf{x} - \mathbf{H}^t\mathbf{R}^{-1}(d - \mathbf{H}\delta\mathbf{x}) = 0 \tag{3.5}$$

The minimization of cost function 3.1 is done by solving equation 3.5 iteratively using quasi-Newton method. This is done by first calculating the gradient in the observation space using the tangent linear code of forward observation operator $\boldsymbol{H}$, denoted as $\mathbf{H}$. Then, gradient in the model space can be obtained by projecting gradient in observation space to the model space, under the help of adjoint code of $\mathbf{H}$, denoted as $\mathbf{H}^t$. Details of code flow can be found in Fischer (2005). At the end of minimization process, $\delta\mathbf{x}$ is obtained and the analysis state deduced by :

$$\mathbf{x}_a = \mathbf{x}_b + \delta\mathbf{x} \tag{3.6}$$

The analysis $\mathbf{x}_a$ will be served as the initial conditions to start the 6h ALADIN forecast.

3.1.2.2 Background covariance matrix B in the 3DVAR/ALADIN

In 3DVAR/ALADIN, the control variables used are : vorticity (ζ), divergence (η), temperature (T), logarithm of the surface pressure (P_s) and specific humidity (q).

More precisely, the control variables are a subspace of the model variables and are stated in fact as the correction to the background, *i.e.* $\delta\mathbf{x} = \mathbf{x}_a - \mathbf{x}_b$. Moreover, in order to avoid the direct calculation of inverse of matrix $\mathbf{B}$ in equations 3.2 and 3.5, a base change of the control variable is carried out with :

$$\chi = \mathbf{B}^{-1}\delta\mathbf{x} \tag{3.7}$$

χ is the actual control variable used during the minimization.

To express the balances that exist in the atmosphere among the control variables, a multivariate formalism has been proposed for ALADIN by Berre (2000). This formalism separates balanced and unbalanced components, based on scale-dependent statistical regressions and with an extra balance tuned for specific humidity. Following are the statistical relationship between the forecast errors of each control variable :

$$\begin{aligned}
\zeta &= \zeta && (3.8)\\
\eta &= MH\zeta + \eta_u && (3.9)\\
(T, P_s) &= NH\zeta + P\eta_u + (T, P_s)_u && (3.10)\\
q &= QH\zeta + R\eta_u + S(T, P_s)_u + q_u && (3.11)\\
& && (3.12)
\end{aligned}$$

M, N, P, Q, R :	vertical balance operators relating spectral vertical profiles of predictors and predictants
u :	subscript that stands for the unbalanced fields of control variables. Unbalanced fields equal total fields minus the balanced fields.

Using this multivariate formalism, the background error covariance matrix $\mathbf{B}$ is then obtained from an ensemble method. The ensemble-based covariances are sampled from an ensemble of ALADIN 6-h forecasts with initial conditions coming from an ensemble of perturbed ARPEGE analyses (Ştefănescu et al., 2006). Finer scale structure functions than those of the global ARPEGE model are therefore obtained from this specific background error matrix. Figure 3.3 shows the horizontal correlation length for the unbalanced part of the humidity near the surface and in mid-troposphere.

3.1.2.3 Observations assimilated in the 3DVAR/ALADIN

In ALADIN, the same observations as in ARPEGE are used. In addition, the METEOSAT/SEVIRI radiances and surface observations are assimilated (*see* Montmerle et al. (2007)). The observations are : radiosondes, surface stations, buoy, ships, and aircraft measurements, wind profilers, horizontal winds from atmospheric motion vectors (AMV), and the QuickSCAT scatterometer, AMSU-A radiances from NOAA-15, NOAA-16 and AQUA, AMSU-B radiances from NOAA-16 and NOAA 17, High-Resolution Infrared Sounder (HIRS) radiances from NOAA-17. For METEOSAT/SEVIRI, the IR 8.7, 10.8 and 12 μm channels are used only over sea and

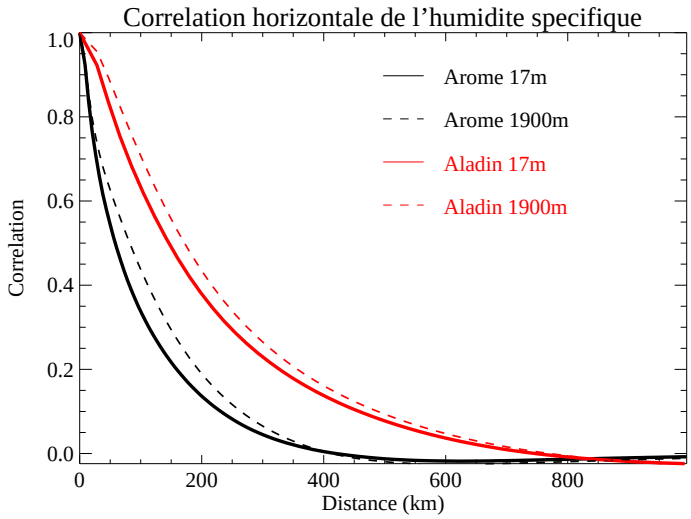

FIG. 3.3: Horizontal correlation of the unbalanced humidity q_u used in the $\mathbf{B}$ matrix of 3DVAR/ALADIN (red) and of 3DVAR/AROME (black)

in clear-sky conditions. The two water vapour 6.2 and 7.3 μm chanels are used in clear-sky conditions and over low clouds.

3.1.2.4 Screening

The observations which arrived in real time in the operational meteorological databases are not always exact. These wrong observations should be eliminated prior to enter in the assimilation scheme in order to not introduce wrong corrections of the first-guess. Also, in some places, we can have too many observations compared to the model resolution. The observation datasets should be thinned. These both actions are done within the screening. If gross errors can be suppressed by checking the observations alone, less evident errors are detected by comparison of the observations to the model forecast. Screening consists in running the direct forecast model and to call the observation operators to compute the model equivalent observation, followed then by the first guess check, the quality control and the thinning of the observations. The background quality control rejects the observations with such first-guess departure d :

$$d = y - H(\mathbf{x}_b) > (n(\sigma_o^2 + \sigma_b^2))^{\frac{1}{2}} \tag{3.13}$$

σ_o^2 : variance of observation error
σ_b^2 : variance of background error
n : predefined multiple, 4 is used in 3DVAR/ALADIN.

3.2 Assimilation of GPS ZTD observations in the 3DVAR/ALADIN

3.2.1 Set up of H and R for the GPS ZTD assimilation

3.2.1.1 The observation operator H

The ZTD observation operator computes the model equivalent ZTD by performing the vertically integration of tropospheric refractivity above the GPS station as in equation 1.26. The observation operator we implemented in the 3DVAR ALADIN system follows Brenot et al. (2006)'s study.

As for other observation operators, the vertical profiles of model variables (here Temperature and Vapour mixing ratio, surface pressure and model terrain) are first interpolated at the observation horizontal location. Then the vertical integration of equation 1.26 is performed as follows :

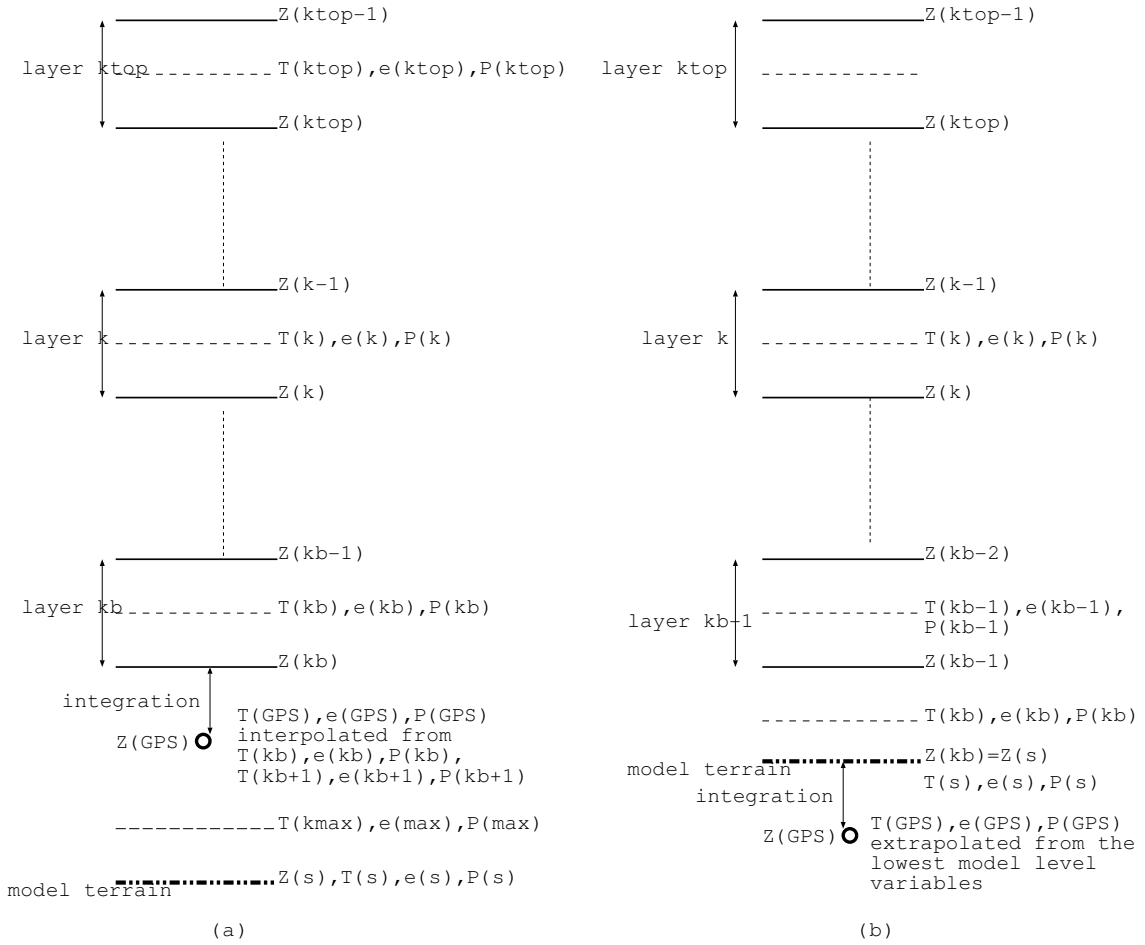

FIG. 3.4: Vertical integration of the ZTD observation operator

First, the integration of equation 1.26 is done from the top of the model level

(k_{top}) to level k_b. k_b is defined as the model level that is just above the layer in which the GPS station is located. This first integration can be expressed as :

$$ZTD_1 = 10^{-6}\sum_{k_b}^{k_{top}}(k_1\frac{p(k)}{T_v(k)}(z(k-1)-z(k))+10^{-6}\sum_{k_b}^{k_{top}}(k_2^{'}\frac{e(k)}{T(k)}+k_3\frac{e(k)}{T(k)^2})(z(k-1)-z(k)) \tag{3.14}$$

Then the second part of the integration depends on two situations :

- **The GPS station height is above the model terrain** (Fig. 3.4a)
  In this case, model temperature, specific humidity and pressure are interpolated to the height of the station using the values from the model level just above and below the station height.

- **The GPS station height is below the model terrain** (Fig. 3.4b)
  In this case, extrapolation is done in order to obtain the model variables below the model terrain. This extrapolation routine follows the same assumption and algorithms used for the other conventional observations of temperature, humidity and pressure in the ECMWF and Météo-France data assimilation systems, trying to preserve the boundary layer properties. Specific humidity is assumed to be constant during the extrapolation, which means it will have the same value as the specific humidity from the level just above the model terrain. Algorithm used to estimate the pressure at the sea level is used to obtain the pressure at the height of GPS station. Finally, the temperature is extrapolated using :

$$T = T_s[1 + \alpha \ln\frac{P}{P_s} + \frac{1}{2}(\alpha \ln\frac{P}{P_s})^2 + \frac{1}{6}(\alpha \ln\frac{P}{P_s})^3] \tag{3.15}$$

T_s : surface temperature at the observation location
P_s : surface pressure at the observation location
P : pressure at the GPS station height
α : $\alpha = 0.0065Km^{-1}R_d/g$, $R_d = 287.0586J/(kmol.K)$ is the specific molar gas constants for dry air, g is the gravity acceleration

Then, for both cases, the second part of integration is performed following equation 3.14 from the station height z_{GPS} to z_{kb} and then added to ZTD_1 to obtain the total ZTD within the model domain.

Finally, the contribution above the top of the model ΔZTD_{top} is also added to the total ZTD according to Saastamoinen (1972) :

$$\Delta ZTD_{TOP} = 10^{-6}\frac{k_1 R_d P_{TOP}}{g_{TOP}} \tag{3.16}$$

P_{TOP} : pressure at the top of the model
g_{TOP} : acceleration of gravity at the top of the model

The tangent linear model of the observation operator described above, as well as the adjoint of this tangent linear model, have been developed. These models were validated for an ensemble of random perturbations around a standard atmosphere, following Poli et al. (2007).

3.2.1.2 Observation covariance matrix R

Like most of the other observations assimilated in the 3DVAR/ALADIN system, GPS ZTD observations are assumed to have uncorrelated errors. Thus the diagonal of the observation error covariance matrix R for GPS ZTD observation is the variance σ_o^2 of the observation error. In our study with the 3DVAR/ALADIN system, σ_o has been chosen constant (10 mm) for all the GPS stations.

3.2.2 Pre-processing of GPS observations

3.2.2.1 A priori station-center selection

Observations from a GPS station can be post-processed by several centers and therefore various ZTD solutions can be available for the same station. In the assimilation, we need only one ZTD observation per station. Therefore we need to select a ZTD solution among the others prior the assimilation. Also, when we started the work on GPS data assimilation at Météo-France, GPS ZTD solutions provided by E-GVAP centers were sometimes very different for a same station or of poor quality. All these motivated Poli et al. (2007) to develop for the ARPEGE data assimilation a pre-selection of stations and centers that provide *a priori* good quality data (with respect to the assimilation criteria). In the present study, this pre-selection has been adapted to the ALADIN model. The selection is done based on the time-series of first-guess departures during a period of 15-30 days before the assimilation period. Hereafter, we call this period the selecting period. We remind that first-guess departures are defined as the difference between ZTD observations and model equivalent ZTD calculated from the first-guess, denoted by $[y - H(x_b)]$. Each station can be processed by several centers, thus the selection is based on the first-guess departure time-series of each station-center pair, instead of station.

The selection follows the procedure below :

- **Step 1 : Reject the station-center pairs if their first-guess departures does not follow a Gaussian distribution**
  This is needed under the fundamental hypothesis of the variational methods that the probability density functions of background and observational errors must be Gaussian. The chi-square test used by Poli et al. (2007) has been replaced by the non-parametric Kolmogorov-Smirnov test (K-S test) for this selection.

- **Step 2 :Reject the station-center pairs which have not sufficient time coverage of data during the whole period**
  This step rejects the station-center pairs, which have too much missing data. It assures a homogeneous data set during all the assimilation period. Moreover, it has been found that GPS ZTD data obtained following a long data gap are more likely resulted from sub-optimal solutions during the first few hours of calculation in the GPS ZTD retrieving process at each center. This is because other parameters such as station coordinates, ambiguity N have to be established again after a long gap.

- **Step 3 :Reject the station-center pairs that have too large height difference with the model terrain**
  Due to the complexity of the boundary layer characteristics, the interpolation or extrapolation of model variables from the model terrain to the station height works fine only within a short height difference. The height difference limit is set to be 150 meters in 3DVAR/ALADIN for GPS ZTD observations.

- **Step 4 : Among the station-center pairs that passed the first 3 criteria, select only one center for each station**
  As explained in Chapter 1, each center's solution has different error characteristics depending on the treatment strategies used. Therefore it is better to use ZTD data from only one unique center for each station in order to keep the same error characteristics during the assimilation period. Also, bias corrections can benefit from using solution from one unique center for each station since bias is likely to be constant for the same center-station pair. The selection of one center is done by taking the center, which has the smallest standard deviation of the first-guess departure series among all the centers for each station.

- **Step 5 : Horizontal thinning**
  Horizontal thinning is performed so as to have a coherence between the model resolution and observation resolution. This is done in order to avoid the representativeness error. Since the resolution of 3DVAR/ALADIN is 9.5km, the minimum distance between the observations should be no less than 9.5km. Thus, some stations are removed when they are too much concentrated in one area.

The thresholds used in this pre-selection for 3DVAR ALADIN have been chosen after examining the results after each step and the visual inspection of the time series of observations as well as Q-Q plots for the data of the whole month of September 2005. A Q-Q plot ("Q" stands for quantile) is a graphical method for diagnosing differences between the probability distribution of the sample and the theoritical distribution (here the Gaussian distribution).

Finally, a list is obtained and only the observations from those station-centers in this list can enter the screening procedure in the beginning of the assimilation experiment. We call this list a white-list as opposed to the well known black-list in data assimilation that indicates the observations that should not be assimilated. Note that

a few more quality control will be done during the screening procedure, that is to say, the final observations entered in the minimization is less than what is allowed in the white-list.

3.2.2.2 Bias correction

Unbiased errors is one of the fundamental hypotheses of data assimilation techniques based on statistical methods. In practice, unbiased errors means that the average of first-guess departure should be zero. A bias cannot be included in the first guess departure since it will create a bias in the analysis increment as well. If a bias can be estimated, it must be removed from the first-guess departures.

In 3DVAR/ALADIN for the GPS observations, the bias is calculated as the average of first-guess departures of GPS ZTD observations during the first 10 days of the selecting period for each station-centers on the list. At the end, the bias will be subtracted from each observation just before it enters the screening procedure of the assimilation experiments. Some sensitivity tests to the length of the period to compute the bias have been carried out by Hakam (2006). He found that the time average of the first-guess departure do not evolve anymore significantly after 10 days.

3.2.3 Synthesis of the data flow of GPS ZTD assimilation

To sum-up, the data flow of the assimilation of GPS ZTD observations consists of :

- First, the pre-selection of station-centers, the bias and the observation error variance estimation are performed during the selecting period. As the ZTD values have a seasonal cycle (weaker value in winter, larger in summer), this period should be chosen within the season corresponding to the assimilation cycle carried out afterward. This requires to have 6-hour forecast (first guess) available over the period to compute the first-guess departures.
- Then, the assimilation of GPS ZTD observations consists in performing the screening and then the minimization of the cost function each 6 hours during an intermittent assimilation cycle (Fig.3.2). It produces each 6 hours an analysis which can be used as initial conditions to run the NWP models and produce forecasts.

3.3 The MesoNH model

MesoNH is a non-hydrostatic mesoscale model developed for research purposes (Lafore et al., 1998). The prognostic variables are the three components of the wind, the potential temperature, the turbulent kinetic energy and the mixing ratio of six water species (vapor,cloud water, rain water, primary ice, graupel and snow).

The configuration used in our study is quite similar to the one used by Nuissier et al. (2008) or Lebeaupin Brossier et al. (2008). It consists of two nested grids inter-

acting with each other according to a two-way interactive grid-nesting method (Clark and Farley, 1984; Stein et al., 2000). The coarser grid provides the lateral boundary conditions to the finer grid, while the variables of the coarser grid are relaxed with a short relaxation time toward the finer grid's value on the overlapping area. The horizontal grid intervals are set to 9.5 km and 2.4 km respectively. Therefore the outer MesoNH domain has the same resolution as the 3DVAR ALADIN analysis and the inner domain has almost the same resolution as the AROME model described below. The same physical parametrisations as in AROME are also used in the 2.4 km domain. A bulk microphysical scheme (Caniaux et al., 1994) governs the prognostic equations of the six water species. No deep convection parametrization is used, assuming that at 2.4 km it is explicitly resolved. The turbulence parametrization is based on a 1.5-order closure (Cuxart et al., 2000), using a mixing length based on Bougeault and Lacarrère (1989). The radiative scheme is the RRTM parametrization (Mlawer et al., 1997). The surface energy exchanges are parametrized according to four different schemes depending on the surface types (nature surfaces, urban areas, ocean, lake) included in the grid mesh. The natural land surfaces are handled by the Interactions Soil-Biosphere-Atmosphere (ISBA) scheme (Noilhan and Mahfouf, 1996), whereas energy exchanges over urban surfaces are parametrized according to the Town Energy Balance (TEB) scheme (Masson, 2000).

3.4 The 3DVAR/AROME assimilation system

3.4.1 The AROME model

Since end of 2008, AROME is the new operational kilometric-scale model of Météo-France. We used during this study the pre-operational version of AROME which was running in 2008 over the France domain (Fig. 3.5).

AROME is the acronym of "Applications of Research to Operations at Mesoscale". AROME is a non-hydrostatic model as Meso-NH, but uses more efficient numerical schemes (semi-Lagrangian, semi-implicit time schemes) inherited from the adiabatic part of the ALADIN code. The main difference with the current version of ALADIN being the physical parametrisation package, which is coming from the MesoNH model (*see* sec. 3.3). A main difference with MesoNH is also that AROME has its own mesoscale data assimilation.

3.4.2 The 3DVAR/AROME data assimilation

The 3DVAR/AROME assimilation system shares most of its code with the 3DVAR/ALADIN system. The two components of the wind, temperature, specific humidity and surface pressure are the control variables and are analyzed at the model resolution (2.4km) while the other model fields (TKE, non-hydrostatic and microphysical fields) are cycled from the previous AROME guess. The frequency of the assimilation cycle is also more rapid than the ALADIN one, with a 3-hour window (Fig. 3.6). From the 00, 06, 12 and 18 UTC AROME analyses, 30-h forecasts are issued operationnaly each day

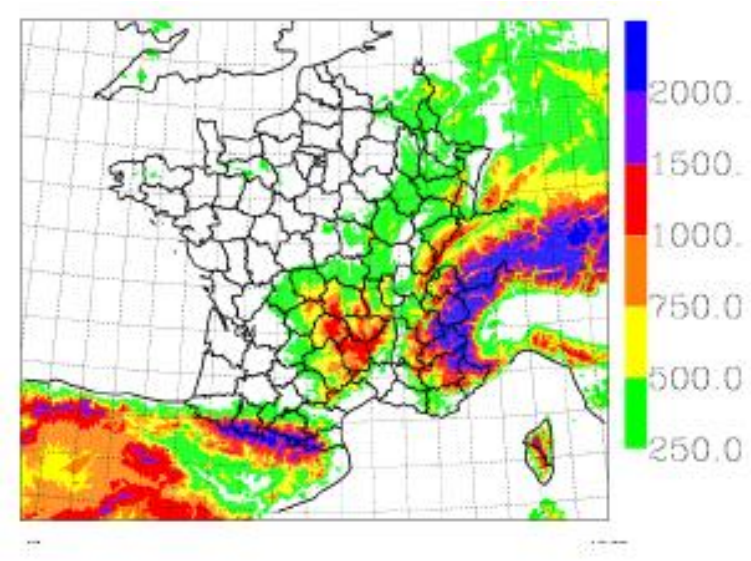

FIG. 3.5: The orography (m) of the AROME/France domain.

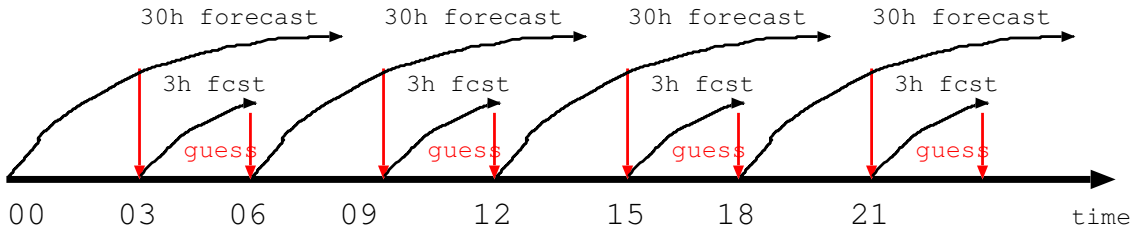

FIG. 3.6: The 3DVAR/AROME assimilation cycle

using the ALADIN/France forecasts as lateral boundary conditions.

In terms of observations, the 3DVAR AROME system used in this study assimilates the same types of observations as the ones operationally used in the ALADIN/France model, including observations from radio-soundings, screen-level stations, buoy, ship and aircraft measurements, wind profilers, horizontal winds from atmospheric motion vectors (AMVs) and the Quickscatt scatterometers, IR radiances from ATOVS and SEVIRI.

The background error covariance matrix shares also the same multivariate formulation as in ALADIN/France. The background error statistics have been calculated using an ensemble-based method (Berre *et al.*, 2006), with a six member ensemble of AROME forecasts carried out over two 15-day periods; the initial and lateral boundary conditions were provided by an ensemble of perturbed ARPEGE-ALADIN/France assimilation ensemble (stefanescu *et al.*, 2006). Compared to the ALADIN/France ones, background-error standard deviations used in AROME are increased and horizontal correlation lengthscales are much shorter (reduced by a factor 2 for low-level humidity for example, *see* Fig. 3.3) : analysis increments are therefore much stronger and narrower in AROME than in ALADIN.

For all the other aspects such as estimating of **R**, set up of GPS ZTD observation operator **H** and the pre-selection procedure, they are defined and estimated in a quite same way as in the 3DVAR/ALADIN.

3.5 Synthesis

To sum-up, the 3DVAR/ALADIN et 3DVAR/AROME data assimilation systems share almost the same code (data flow, forward obervation operators, adjoint operators, etc). The differences lie mainly in the way the systems are configured : background error covariance matrix **B**, observation variance error matrix **R**, type of first-guess, time window of the data assimilation cycle, ...

The AROME model and the MesoNH model version used here have a similar resolution and share the same physical parameterizations. The main differences between the two models concern the dynamic part of the model with more efficient numerical schemes for AROME.

Chapitre 4

Impact de l'assimilation de données GPS avec le 3DVAR/ALADIN

Résumé

Ce chapitre décrit les résultats de l'étude de l'impact de l'assimilation des données GPS avec le 3DVAR ALADIN sur la prévision à haute résolution MesoNH d'un épisode de pluie intense Méditerranéen. Il se présente sous la forme d'un article accepté au Journal of Geophysical Research, précédé d'une étude de sensibilité à la formulation de l'opérateur d'observation.

Le cas sélectionné pour cette étude est l'épisode de pluie intense du 5 au 9 septembre 2005 qui a affecté principalement le Gard et l'Hérault. Les données GPS assimilées proviennent du réseau européen E-GVAP et des stations de l'OHM-CV. Un cycle d'assimilation ALADIN à la résolution de 9.5 km a été effectué pour tout le mois de septembre. Les analyses ALADIN ont ensuite été utilisées comme conditions initiales et aux limites pour réaliser des prévisions MesoNH à 2.5 km de résolution de l'évènement du 5 au 9 septembre 2005. Deux runs (à 00UTC et 12 UTC) par jour avec une prévision jusqu'à 18 heures d'échéance ont été réalisés sur la totalité de l'évènement.

Les résultats de cette étude peuvent être résumés ainsi :

- Les expériences où une seule observation est assimilée nous ont permis de montrer la faible sensibilité à la formulation des deux opérateurs d'observation testés et de confirmer que l'assimilation des observations de ZTD modifiait principalement le rapport de mélange en vapeur d'eau dans la basse et moyenne troposphère dans un rayon d'environ 70 km autour de la station GPS.
- Lorsque toutes les observations de ZTD sont assimilées avec les autres types d'observations, l'impact de l'assimilation des observations de ZTD a été trouvé neutre en général sur l'analyse et l'ébauche (prévision à 6 heures d'échéance d'ALADIN), seul un faible impact positif sur l'humidité spécifique de l'ébauche a été trouvé par comparaison aux observations de radiosondages.
- Lorsque les analyses issues du cycle d'assimilation ALADIN sont utilisées comme conditions initiales aux prévisions MesoNH à 2.5 km, l'impact sur la prévision MesoNH des cumuls de précipitation, évalué à la fois qualitativement et quan-

titativement à l'aide de scores, a été trouvé faiblement positif.

4.1 Impact of the observation operator

In fact, two observation operators were available in the assimilation software : one issued from the Poli et al. (2007) study with ARPEGE (called hereafter POLI07) and an other one (called hereafter BREN06) developped according to the results of Brenot et al. (2006) which had evaluated many expressions to calculate the model equivalent ZTD. In the previous chapter, the BREN06 observation operator was fully described. We presents here only the main differences between the two observation operators.

POLI07 integrates from the bottom to the top of the model the following equation, using the total pressure P, the temperature T and the partial pressure of water vapour e of the atmospheric model column :

$$ZTD = \int (k_1 \frac{P}{T} + k_3 \frac{e}{T^2}) dz \tag{4.1}$$

According to Smith and Weintraub (1953)
k_1 : $0.776 \cdot 10^{-6}$ Pa^{-1}·K
k_3 : $3730 \cdot 10^{-6}$ Pa^{-1}·K^2

whereas the equation for BREN06 is the following :

$$ZTD = \int (k_1 \frac{P}{T_v} + k'_2 \frac{e}{T} + k'_3 \frac{e}{T^2}) dz + \Delta ZTD_{TOP} \tag{4.2}$$

k_2 : $0.704 \cdot 10^{-6}$ Pa^{-1}·K
k'_3 : $3739 \cdot 10^{-6}$ Pa^{-1}·K^2
k'_2 : $k'_2 = k_2 - k_1 \frac{R_d}{R_v}$ (Bevis et al., 1994)

BREN06 have evaluated different sets of (k_1, k_2, k_3) coefficients proposed in the literature for ZTD values computed from high-resolution (2.4km) non-hydrostatic atmospheric simulated fields. The results showed that there are no significant differences in the evaluation of ZTD between most of the coefficient sets (with a mean bias of ZTD less than 2 mm), except for the set using the two-coefficient formula as in POLI07, which presents a mean ZTD bias reaching nearly 12 mm.

A ZTD contribution above the model top level ΔZTD_{TOP} is also added in BREN06 (see Eq.(3.16) in chapter 3). POLI07 did not consider this correction in their observation operator and left this to be handled by their bias correction scheme. Poli et al. (2007) evaluated to 2.3 mm this contribution above the ALADIN/ARPEGE model top at 1 hPa for the mid-latitudes.

In order to perform the integration of the ZTD operator inside the model, POLI07 and BREN06 follow nearly the same implementation, except when the GPS station is located below the model terrain. In that case, extrapolation of pressure, temperature and humidity values from the model lowest level down to the GPS station altitude is needed. In that extrapolation procedure, the temperature is considered as constant in POLI07 while BREN06 assumes a constant temperature gradient.

In order to study the impact of how the model equivalent ZTD is calculated in the assimilation system, we perform analyses where the 3DVAR ALADIN system only assimilates one zenith delay observation using the two different observation operators available, as their tangent-linear and adjoint codes, in the assimilation software.

Figure 4.1 shows results of the assimilation of a single observation of ZTD at Aix-en-Provence of value 2.4586 m ; the model equivalent ZTD computed from the first guess with BREN06 and POLI07 observation operators departs from the observation by 14.34 mm and 18.6 mm, respectively. Note that 2.3 mm (out of the 18.6 mm difference) would normally be corrected by a bias correction (as in POLI07) but we neglect this effect here. After the minimization of the cost function in the assimilation, the model equivalent ZTD analysis differs from the observation of only 1.47 mm for BREN06 and of 1.9 mm for POLI07, showing a good behaviour of the assimilation with both operators. The main impact of assimilating ZTD data is seen on the humidity field. There is almost no impact on the pressure or temperature field. The horizontal influence range of the observation is in agreement with the horizontal correlation length of the background-error covariance matrix which is about 70 km in the middle troposphere for specific humidity with the 3DVAR/ALADIN. We can see in Figure 4.1cd that the assimilation experiment using POLI07 yields an analysis increment of humidity in the low to mid- troposphere more intense than in the assimilation experiment using the observation operator from BREN06. The larger analysis increments imply more corrections to the humidity field by the analysis.

To summarize, in the present example, the impact on the analysis of the formulation of the observation operator is not so much ; it can lead to about 2% of difference in the final analysis increment of humidity, without however modifying the vertical distribution of the humidity analysis increment.

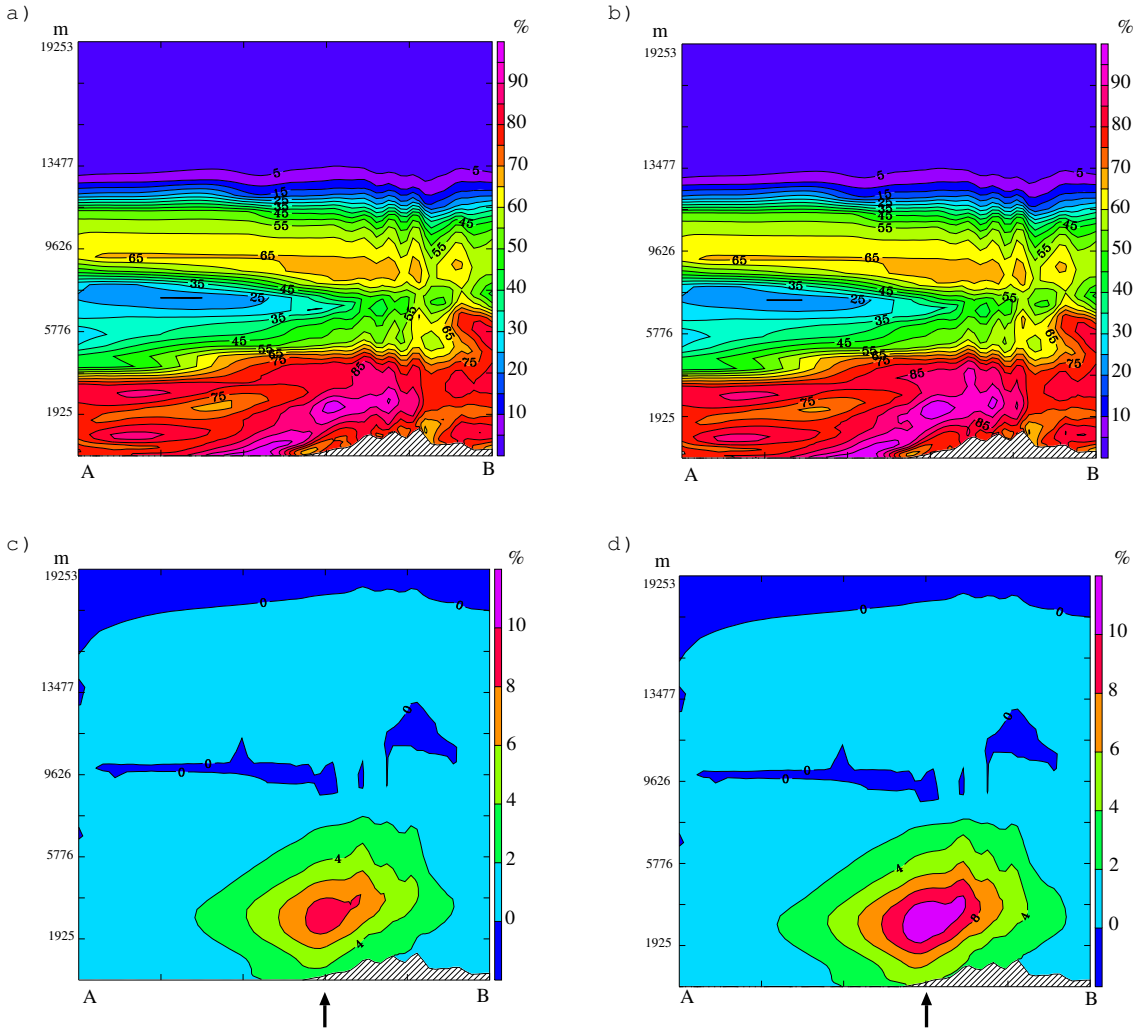

FIG. 4.1: Vertical cross section along a South-North 670-km line of analyses and analysis increments of relative humidity, obtained by assimilation of a single ZTD observation using the BREN06 (ac) or the POLI07 (bd) observation operator. Panels ab display the relative humidity from the 3DVAR analysis and panels cd the analysis increment of relative humidity. The vertical arrows in panels cd indicate the location of the ZTD observation.

4.2 Impact of GPS Zenith Delay Assimilation on Convective Scale Prediction of Mediterranean Heavy Rainfall

The following article is accepted for publication in the Journal of Geophysical Research. Yan, X., V. Ducrocq, P. Poli, M. Hakam, G. Jaubert, and A. Walpersdorf (2009), Impact of GPS zenith delay assimilation on convective-scale prediction of Mediterranean heavy rainfall, J. Geophys. Res., 114, D03104, doi :10.1029/2008JD011036

Chapitre 5

Impact de l'assimilation de données GPS avec le 3DVAR/AROME

Résumé

Ce chapitre décrit les résultats de l'étude de l'impact de l'assimilation de données GPS avec le 3DVAR/AROME sur la prévision AROME de la POI9 de la campagne COPS. Il se présente sous la forme d'un article préparé pour une soumission au Quaterly Journal of the Meteorological Society. Il est précédé d'une présentation générale de la campagne COPS. Une étude de sensibilité à la formulation de l'opérateur d'observation conclut ce chapitre.

Pendant la campagne COPS, un réseau dense de stations GPS a été déployé dans la région des Vosges et de la Forêt Noire. Les données de ces stations ainsi que celles du réseau opérationnel E-GVAP ont été assimilées avec le 3DVAR/AROME. Les analyses à 00UTC issues du cycle d'assimilation AROME ont servi de conditions initiales à la prévision AROME pour les 3 jours de la POI9 de COPS (18-20 juillet 2007). Cette POI est caractérisée par des systèmes convectifs précipitants qui se forment sur la France et balayent ensuite la région COPS.

Un impact positif sur la prévision AROME est clairement mis en évidence lorsque les données GPS sont assimilées. L'impact est particulièrement significatif pour la journée du 19 juillet (POI9b).

Deux expériences de sensibilité ont ensuite été conduites. Dans une première expérience, seules les données GPS du réseau opérationnel E-GVAP sont assimilées. Les résultats de cette expérience montrent qu'un impact postif est déjà obtenu lorsqu'uniquement les données du réseau E-GVAP sont assimilées. Ce résultat est très encourageant par rapport à une assimilation en opérationnel des données de ZTD dans AROME. Une seconde expérience a évalué l'impact de la formulation de l'opérateur d'observations. Un impact très faiblement positif en faveur de la formulation de l'opérateur selon Brenot et al. (2006) est trouvé.

5.1 The COPS field campaign

5.1.1 Motivation

Precipitation is one of the most important meteorological variables, which is highly demanded by different end-users such as hydrological, agriculture and disaster warning systems, etc. Nowadays, the demand of quantitative precipitation forecasts (QPF) from those end-users is increasing. However, the quality of QPF especially for convective rainfall is quite poor and has remained still a problem in the NWP research and operational community worldwide. Among all the difficulties, which have been encountered during the improvement of QPF skills, the convective rainfall induced by the complexity of the terrain is one of the most crucial. Many efforts have been carried out to address this problem.

Quantitative Precipitation Forecast in low-mountain regions is a real challenge for the NWP community. The forcing mechanisms which trigger the convection initiation are not well understood and also not well described in the NWP models. For example, one of the main systematic errors in mesoscale model precipitation forecasting is the windward/lee effects. On the windward side the mesoscale models generally overestimate precipitation whereas on the lee side of the mountain, they underestimate the rain.

In order to progress in QPF with mesoscale NWP models, under the organization of the German Research Foundation (DFG), a program called "quantitative precipitation prediction" (PQP) has been established in 2004 with 23 QPF related research projects. COPS is one of those QPF projects within the framework of the PQP program to provide extensive observations for validating mesoscale models and improving the representation of the key processes, and finally, to lead to the improvement of QPF skills. COPS contains mainly an observation field campaign and the research carried out after the campaign using the data obtained. Institutions from 8 countries in Europe have been involved and collaborate in this project. Interaction between the instrumentation groups and the modelers at NWP centers or research institutes have made the COPS project being recognized also as part of the research and development project of WWRP (World Weather Research Program of WMO).

As indicated in its name, COPS (Convective and Orographically-induced precipitation Study) is concentrated on understanding of the precipitation processes, especially on the convection initiation in low-mountain regions which is expected to be determined by a complex balance between forcing mechanisms due to land-surface processes and orography, as well as to mesoscale and large-scale conditions. This leads to the choice of the campaign site being located in the bordering regions between France and Germany. The main goal of COPS was "to advance in the quality of forecasts of orographically induced convective precipitation by four-dimensional observations and modeling of its life cycle for identifying the physical processes responsible for deficiencies in QPF over low-mountain regions" (Wulfmeyer et al., 2008).

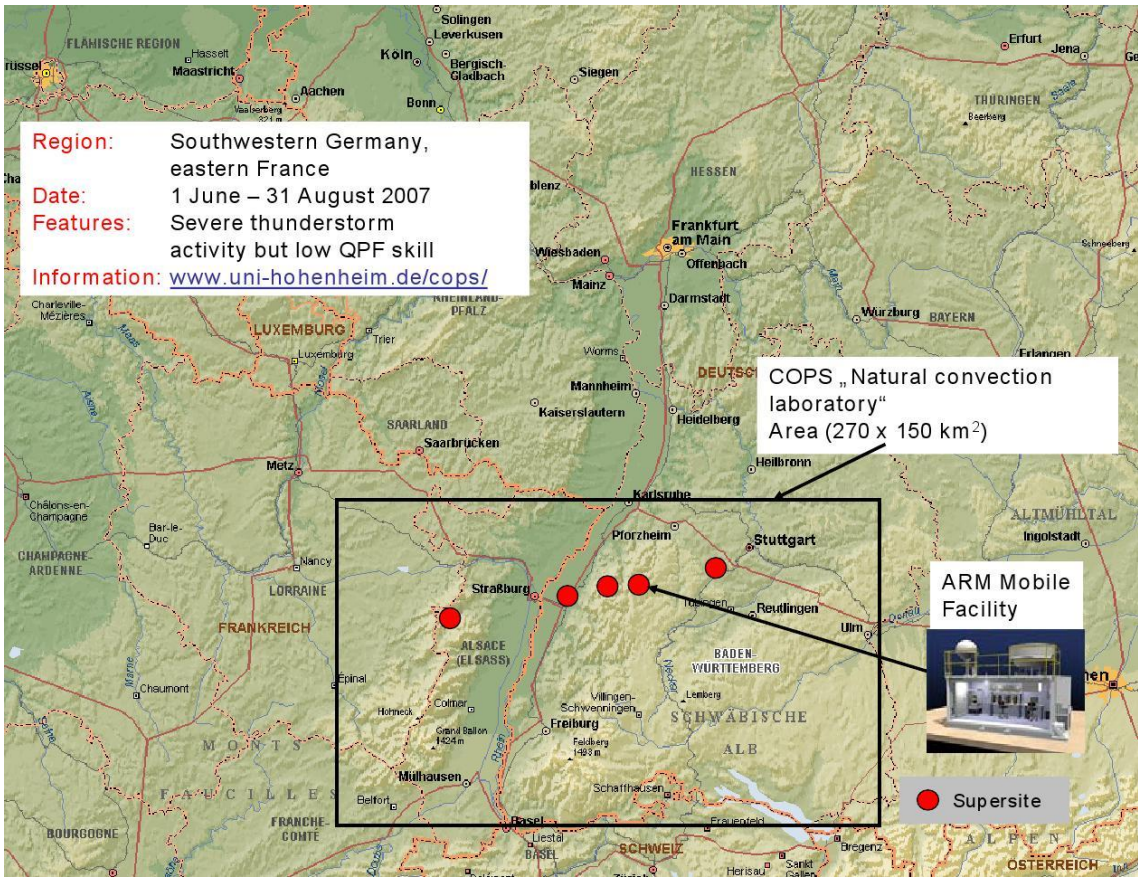

FIG. 5.1: The COPS field campaign domain with the five supersites

The region illustrated by Figure 5.1 (called COPS domain hereafter) is about 250 km in west-east direction and 170 km in south-north direction. From west to east, this mountainous region covers the French Vosges mountain area, then the low Rhine valley, and after the rapid rise of the Black Forest region in southwest of Germany. The climate characteristics of this region are governed by a westerly flow in winter with frontal rainfall and convective precipitation in summer caused by local instability of the atmosphere. This region experiences precipitation 50% of the time during the summer. Large deficiencies in QPF skill have been found in this region and during the summer time in cases of convective rainfall. Thus the campaign has been chosen to be performed from June to August, 2007.

5.1.2 Observations in the COPS campaign

Five "supersites" are organized along a transect from the Vosges mountains, the Rhine valley, the Hornisgrinde Mountain, the Murg valley, and on the lee side close to Stuttgart (Fig. 5.1). The instrumentation at each supersite was chosen in order to document the full life cycle of convective precipitation, from the preconvective environment to the development of precipitation. All of the supersites were equipped with soil moisture sensors, turbulence or energy balance and radiosonde stations, GPS, and surface meteorology instrumentation. Thirteen lidar systems measuring water-vapour or Doppler wind or temperature were deployed on 2 airborne and 11 ground-based platforms. Furthermore, seven microwave radiometers, two Fourier transform infrared spectrometers, three cloud radars, eight precipitation radars (in addition to the existing operational network of the national weather services), and five microrain radars were operated.

Regarding the GPS observations, a first network has been set up in order to obtain the regional distribution of water vapor for the entire COPS domain with a resolution of 50*50 km^2. Secondly, a densification of the Vosges, Rhine valley, and Black Forest has been performed to provide a quasi-continuous observation of the water vapor around and between the supersites (*see* sec.5.2 for a more comprehensive description of the network).

A suite of airborne platforms performed large-scale observations as well as observations between supersites. The German DLR Falcon with a water-vapor lidar and a Doppler lidar as well as dropsondes, and the French SAFIRE Falcon 20 with water-vapor lidar and dropsondes had the mission to sample the preconvective environment around the COPS region. The German DO 128 aircraft provided meteorological and turbulence measurements within the COPS region. The UK FAAM BAe 146, the SAFIRE FALCON20, and the Partenavia aircraft brought cloud microphysical measurements.

These research means have been deployed during 18 IOPs (sometimes of several days) from 1^{st} June to 31 August 2007.

5.1.3 Role and goal of data assimilation in the COPS campaign

The COPS field campaign has also incorporated real-time or post-analysis data assimilation studies in order to demonstrate the potential impact of new observing systems used in this campaign. The measuring systems which directly provide observations of NWP model prognostic variables or provide observations with suitable forward operators are considered. Different data assimilation schemes such as nudging, 3DVAR, and 4DVAR are compared as well as different mesoscale models.

The observations assimilated during or after the COPS campaign include : Water vapour, wind, temperature lidar, ground-based GPS observations, radiosondes, Doppler radial velocity and satellites remote sensing data.

One major goal of the data assimilation is to help in separating the errors due to the initial conditions to those due to the model parametrizations. This can be achieved by studying and comparing the impacts of different types of observation systems. If data assimilation of a certain observation systems lead to a significant improvement of the forecast, it is possible that the forecast quality is mainly limited by errors in initial conditions. Our study on the assimilation of GPS ZTD from both the COPS campaign and the E-GVAP network contributes to this general goal.

5.2 Benefit of GPS Zenith Delay Assimilation on high-resolution Quantitative Precipitation Forecast of the COPS cases IOP9

The following article has been prepared for submission to Quat. J. Roy. Meteor. Soc.

Benefit of GPS Zenith Delay Assimilation on high-resolution Quantitative Precipitation Forecast of the COPS cases IOP9

X. Yan[(1)],V. Ducrocq[(1)], G. Jaubert[(1)], P. Brousseau[(1)], P. Poli[(1,3)], C. Champollion[(2)], K. Boniface[(2)]

(1) GAME-CNRM, CNRS & Météo-France, 42 Avenue Coriolis, 31057 TOULOUSE Cedex 1, France

(2) UM2/CNRS,Place E.Bataillon, Montpellier, France

(3) ECMWF, Shinfield Park, RG2 9AX, United Kingdom

abstract

The benefit of assimilating high spatial and temporal resolution GPS ZTD observations within a high-resolution NWP system is evaluated during IOP9 of the COPS campaign. During the international COPS field experiment which took place from June to August 2007, a high density GPS network was deployed in the Vosges mountain region in eastern France and the Black forest region in southwestern Germany. GPS ZTD observations collected during the COPS field campaign, in addition to the European operational GPS data, are assimilated into the AROME/3DVAR system. We investigate the impact of assimilating these GPS ZTD observations on the forecast of convective systems that travelled over Eastern France during the 3-day COPS IOP9. The results show a clearly positive impact in terms of short-range quantitative precipitation forecast when GPS ZTD observations are assimilated. The impact is most important for the IOP9b day. As most of the improvement is gained by assimilating GPS ZTD observations from the European operational GPS network, this study opens up encouraging horizons for the assimilation of GPS ZTD in the new high-resolution operational NWP systems.

Keywords: GPS zenith delays; quantitative precipitation forecast; mesoscale data assimilation

1 Introduction

During the last decades, almost no significant improvement on Quantitative Precipitation Forecast (QPF) has been achieved, especially for convective precipitation. One reason for such QPF deficiency is that deep convection parametrization used in Numerical Weather Prediction (NWP) systems do not reproduce correctly the convective systems (Wulfmeyer *et al.*, 2008). Several previous studies have shown that the response in terms of quantitative precipitation forecast (QPF) of non-hydrostatic convection-resolved models are different from those of coarser models with parametrized convection using the same initial conditions (*e.g.* Ducrocq *et al.*, 2002; Richard *et al.*, 2007). Ducrocq *et al.* (2002) showed that increasing the model resolution down to 2-3km allowing explicit resolved deep convection improves the forecast of convective systems. This improvement is however subject to a relevant description of the mesoscale initial conditions and especially to the initial humidity field. But due to the limitation of water vapour measurements and the highly variable nature of the moisture in the troposphere, the improvement of the moisture field remains still a hot research topic for the NWP centers worldwide. Assimilation of GPS Zenith Total Delay (ZTD) observation has emerged to become one of the most promising method to address this moisture problem during the past 10 years (De Pondeca and Zou, 2001; Vedel and Huang, 2004; Poli *et al.*, 2007, among others)

Further studies (Cucurull *et al.*, 2004; Vedel *et al.*, 2004; Peng and Zou, 2004) have been carried out in assimilating high spatial and temporal resolution GPS ZTD observations into mesoscale numerical weather prediction systems. In general, those former studies suggest that better results can be obtained by increasing the model resolution. We have also conducted a former research on assimilating GPS ZTD observations into the 9.5-km ALADIN-France/3DVAR system and used its analyses as the initial conditions to start a 2.5-km non-hydrostatic model MesoNH runs (Yan *et al.*, 2008). Slightly positive results have been found by assimilating GPS ZTD data for a heavy convective precipitation event in Northwestern Mediterranean.

Now, high-resolution data assimilation systems based on non-hydrostatic resolving convection models and with rapid update cycle have been developed and become to be operational in several meteorological services. This is the case of the Météo-France AROME system which is now running in a pre-operational mode over a domain cov-

ering France (Bouttier, 2007). Both the model forecast and data assimilation are performed at 2.5 km resolution and the 3DVAR scheme assimilates observations every 3 hours. With such high horizontal resolution, the model orography is closer to the true terrain, thus a more precise equivalent model ZTD can be computed. The short-range high-resolution forecast which serves as the first guess for the data assimilation is also expected to be better, allowing equivalent model ZTD closer to the observations. This is especially true for convective rainfall events. The aim of our study is to evaluate the benefit of assimilating high spatial and temporal resolution GPS ZTD observations within a high-resolution NWP system. The pre-operational AROME system is used for that purpose. Our study is based on the COPS field campaign where a high density GPS network was deployed in the Vosges mountain region in eastern France and the Black forest region in southwestern Germany. GPS ZTD observations collected during the COPS field campaign, in addition to the European EGVAP GPS data, are assimilated into the AROME/3DVAR system. We investigate the impact of assimilating these GPS ZTD observations on the forecast of convective systems that travelled over Eastern France during the 3-day COPS IOP9.

The outline of the paper is as follows. Section 2 presents the data and the case study from the COPS field campaign. Then section 3 describes the AROME NWP system and the assimilation experiments performed in the study. Section 4 discusses the impact of assimilating GPS ZTD data on the AROME forecast of IOP9. The conclusions follow in section 5.

2 Data and case study from the COPS field campaign

2.1 The COPS GPS network

The international field campaign COPS (Convective and Orographically-induced Precipitation Study, Wulfmeyer *et al.*, 2008) took place from June 1 to August 31 2007 in southwestern Germany and eastern France which includes the Vosges mountain and the Black forest (see Fig. 1 for geography of the region). This region of low mountains frequently experiences severe thunderstorm activity during the summer while

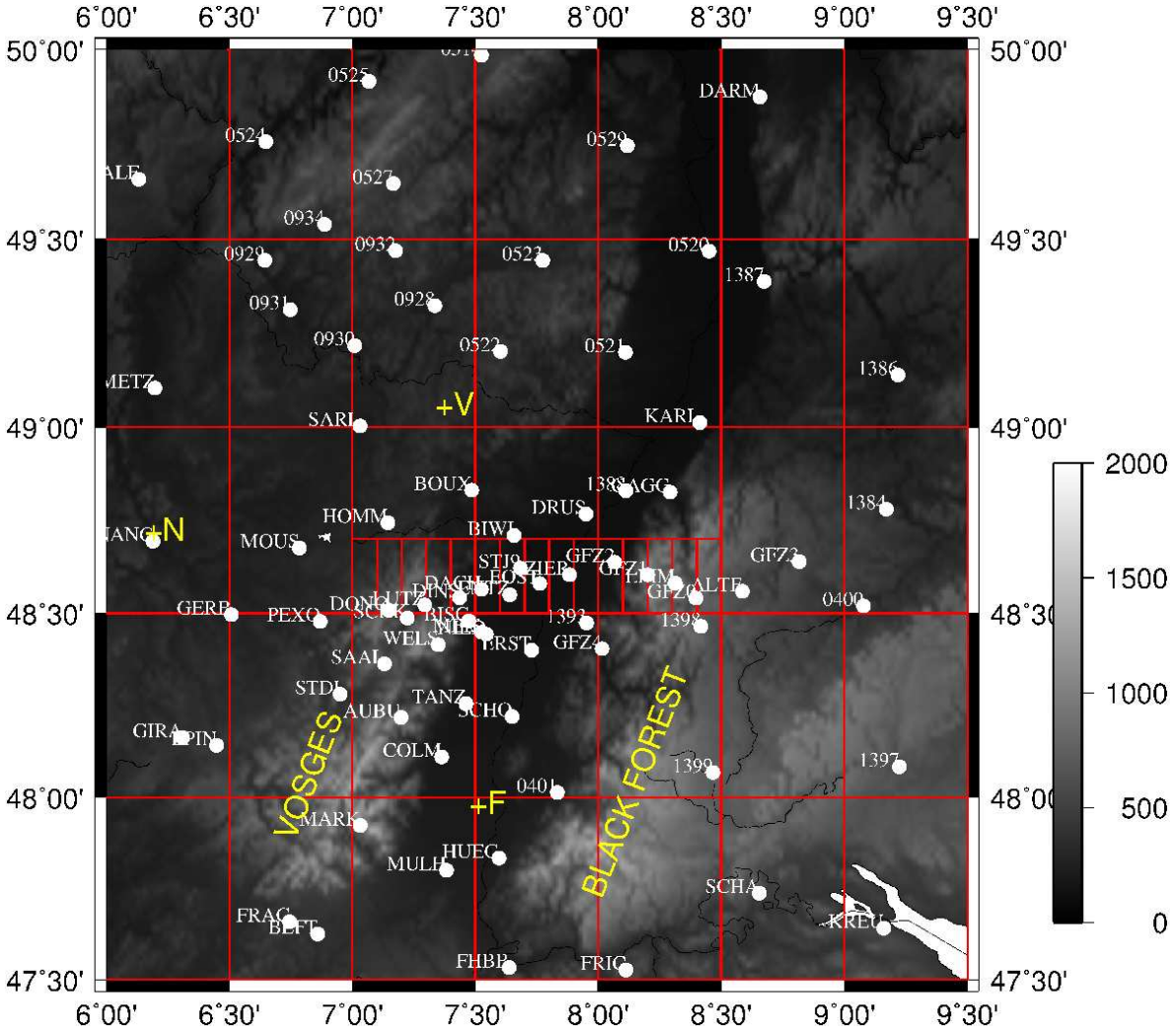

Figure 1: The GPS COPS network with the orography (m) of the region. The location of the raingauge stations of Fessenheim (F) and Vomunster(V) and of the Nancy radiosounding (N) are also indicated.

the numerical weather prediction skill is especially weak in those cases. The goal of this campaign was to provide new 4-dimensional datasets for identifying the reasons of deficiencies in quantitative precipitation forecast for orographically-induced convective rainfall events. The observing strategy was based on densifying of existing networks (additional GPS receivers, radiosoundings, soil moisture mesonet, etc) and deployment of research instruments (lidars, Doppler X-band radars, etc.). The instruments were concentrated over 5 "supersites" almost aligned along a transect from the Vosges mountains to Stuttgart. As regards more specifically the GPS network covering the COPS region, it is composed of (Fig. 1):

- 20 permanent stations coming from EOST (Ecole et Observatoire des Sciences de la Terre, Strasbourg), RGP (Réseau GPS Permanent), Orpheon and INRA (Institut National de la Recherche Agronomique of Nancy) networks over France,
- 29 permanent stations from SAPOS over Germany,
- 22 temporary INSU/CNRS (Institut National des Sciences de l'Univers / Centre National de la recherche scientifique) owned stations mounted during the 3-month COPS campaign period over France,
- 5 temporary GFZ (German Research Centre for Geosciences) owned stations over Germany.

In total, there were 76 permanent or temporary GPS stations present in the COPS region during the campaign. The spatial resolution is about 50*50km^2 over the total domain, with an increase of the density along the "supersites" transect and in the Rhine valley. The raw GPS observations have been processed to issue 60 minutes accurate GPS ZTD and horizontal delay gradients solutions, using the GAMIT software - release 10.33 (King and Bock, 2008). During a first run, GPS station coordinates are estimated. These are re-used for a second run during which atmospheric parameters are estimated: ZTDs and horizontal delay gradients (Davis *et al.*, 1985; Chen and Herring, 1997). The gradients represent the anisotropy of the water vapor field around the GPS site. GPS data are processed using an elevation cut-off angle of 5° and the Vienna mapping functions (Boehm *et al.*, 2006). A priory ZHD (Zenith

Hydrostatic Delay) values used in the processing are provided by the ECMWF (European Centre for Medium-Range Weather Forecasts) analyses. A sliding windows strategy have been applied to minimize the edge effect of the processing. For more details on GPS data processing we refer to Champollion *et al.* (2004) and Brenot *et al.* (2006).

2.2 IOP9

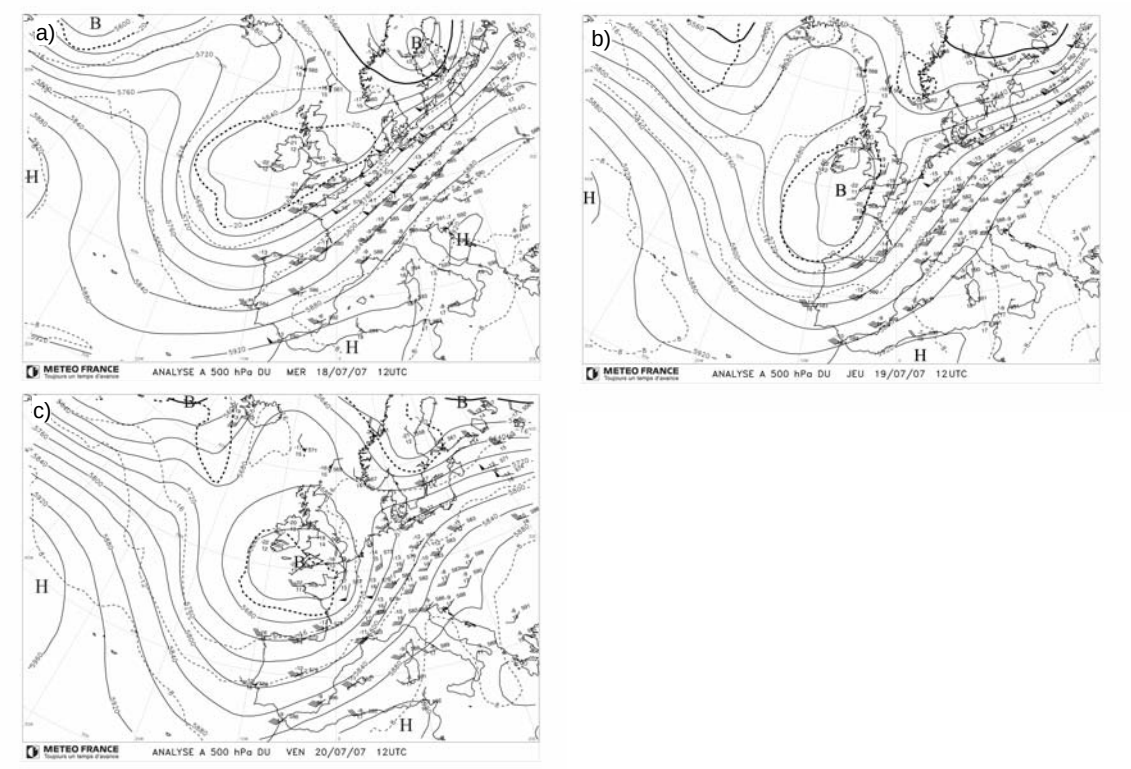

Figure 2: Maps of 500 hPa geopotential heights (solid lines) and temperatures (dashed lines) valid at 12 UTC on : a) 18 July 2007, b) 19 July 2007, and c) 20 July 2007.

The impact of assimilating GPS ZTD data is examined for the 3-day IOP9 (18-20 July 2007). The COPS goal for IOP9 was the study of the development of a frontal zone elongated from southwest to northeast and its influence on the deep convection. On 18 July, a large upper-level trough extending from Scandinavia to the Azores

generated a rapid southwesterly flow over France and Germany (Fig. 2a). A surface frontal zone is embedded within that region with hot and moist air ahead of it in the eastern part of France and Germany. Associated with the frontal zone, successive convective systems developed in the warm air over France and then affected the COPS region during the early morning (up to 1000 UTC). Later, isolated convective cells formed in the lee of the Vosges during the afternoon (1700-2000UTC). The upper-level trough progressively took a North-South orientation during the two following days (Fig. 2bc) and a low-pressure cut-off formed over the British Isles. Deep convection developed again over France within the frontal zone during the night from 18 to 19 July and the convective systems, transported by the southwesterly flow, concerned the COPS region during almost all the day the 19^{th}. Among the convective systems that travelled over the region, a major mesoscale convective system affected the COPS region from about 0600 UTC to 1200 UTC, 19 July. On 20 July, the cold front associated with the low surface pressure center over the British isles affected the COPS region during the morning. Precipitation organised in a bow-like structure that took the shape of a squall line ranging from the Netherlands to Germany. Severe precipitation with flooding took place in Germany associated with this system.

3 Assimilation experiments

3.1 The AROME NWP system

The system used in this study is the pre-operational AROME suite over France that runs daily in Météo-France (Seity *et al.*, 2008). This model is now used operationally since December 2008. A large part of the AROME software is common with the ALADIN and ARPEGE models from Météo-France and the IFS system from ECMWF. AROME uses a non-hydrostatic extension of the limited-area NWP model ALADIN for the adiabatic equations (Bubnovà *et al.*, 1995; Benard, 2004), whereby the vertical coordinate becomes an hydrostatic pressure, and two prognostic variables are added : vertical velocity and departure of the pressure from hydrostatism. The AROME physical package is extracted from the physical parameterizations of the Meso-NH research model (Lafore *et al.*, 1998): the prognostic equations of the six water species (vapour, cloud water, rain water, primary ice, graupel and snow) are governed by a

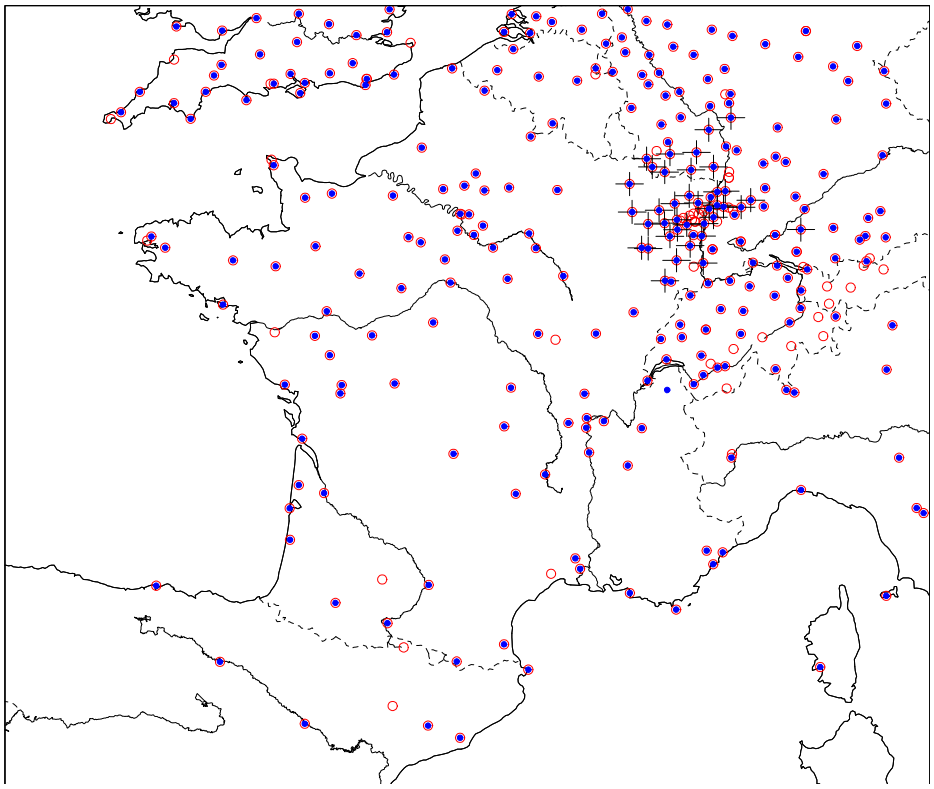

Figure 3: Location of the ground-based 316 GPS stations (circles)from the E-GVAP and COPS networks, plotted over the AROME domain. The 277 stations selected by the pre-processing and used after in the data assimilation are marked with a black bullet. The stations removed for the OPR experiments are marked also by a cross.

bulk microphysical scheme (Caniaux *et al.*, 1994) and the deep convection is resolved explicitly. The radiative scheme is the RRTM parameterization (Mlawer *et al.*, 1998) and the turbulence scheme is from Cuxart *et al.* (2000). The surface energy exchanges are parametrized according to four different schemes depending on the surface types (nature surfaces, urban areas, ocean, lake) included in the grid mesh. The natural land surfaces are handled by the Interactions Soil-Biosphere-Atmosphere (ISBA) scheme (Noilhan and Mahfouf, 1996), whereas energy exchanges over urban surfaces are parametrized according to the Town Energy Balance (TEB) scheme (Masson, 2000). The AROME domain covers mainly France (Fig. 3) at 2.5-km resolution and with 41 vertical levels. Lateral boundary conditions are provided by hourly ALADIN forecasts.

The AROME 3DVAR assimilation system is based on the ALADIN 3DVAR system (Fischer *et al.*, 2005). The data assimilation system uses a rapid forward intermittent assimilation cycle with a 3-hourly data analysis frequency. The two components of the wind, temperature, specific humidity and surface pressure are analyzed at the model resolution (2.5km) while the other model fields (turbulent kinetic energy, non-hydrostatic and microphysical fields) are cycled from the previous 3-hour AROME forecast. In terms of observations, the system assimilates the same types of observations as the ones operationally used in the ALADIN-France model (Montmerle *et al.*, 2007; Yan *et al.*, 2008), including observations from radio-soundings, screen-level stations, buoy, ship and aircraft measurements, wind profilers, horizontal winds from atmospheric motion vectors (AMVs) and the Quickscatt scatterometers, IR radiances from ATOVS and SEVIRI. The background error covariance matrix shares also the same multivariate formulation as in ALADIN-France (Berre, 2000). The background error statistics have been calculated using an ensemble-based method (Berre *et al.*, 2006), with a six member ensemble of AROME forecasts carried out over two 15-day periods; the initial and lateral boundary conditions were provided by an ensemble of perturbed ARPEGE/ALADIN-France assimilation ensemble (Stefanescu *et al.*, 2006). Compared to the ALADIN-France ones, background-error standard deviations used in AROME are increased and horizontal correlation lenghtscales are much shorter (*i.e.* reduced by a factor 2 for low-level humidity for example): analysis increments are therefore stronger and narrower.

3.2 GPS observations and pre-processing

In our study, we use two GPS ZTD datasets. The first one is from the European GPS station network whose data are already assimilated in the ARPEGE and ALADIN systems (Poli *et al.*, 2007). The European scale collection of ground based GPS delay measurements in near-real time is performed through the EUMETNET GPS water vapor program (E-GVAP, http://egvap.dmi.dk/). The second dataset comes from the COPS campaign as described in section 2.1. In total, there are 316 GPS stations gathering the two datasets within the AROME domain (Fig. 3). Note that each station can be treated by several centers at the same time which can result in more than one GPS ZTD solutions per station. Therefore 648 pairs of station-center within the AROME domain are finally available.

Pre-processing of GPS ZTD observations is needed in order to use for each station only the GPS ZTD observation processed by a selected center which provides the most reasonable solutions. The reader is referred to Poli *et al.* (2007) and Yan *et al.* (2008) for discussion about the necessity of such a selection. Spatial thinning, observation error assignment and bias correction are also performed at the same time. The pre-processing follows the one used for the GPS ZTD assimilation in the ALADIN system (Yan *et al.*, 2008). The pre-processing is based on the examination of the statistical behavior of departures for all the 648 station-center pairs during the period 23 July - 06 August 2007. "Departure" is defined here as the difference between the GPS ZTD observation and the model equivalent ZTD computed from the first-guess. The observation operator for GPS ZTD used in the AROME 3DVAR data assimilation system is based on the vertical integration of the refractivity expression as decribed in Poli *et al.* (2007). Spatial thinning is conducted with the distance between the stations no less than 10km. The standard deviation of the first 10-day departure time series is used as the bias for each selected station-center. The observation error assigned to each station varies also, depending on the standard deviation of the departure. Figure 3 shows the stations selected by the pre-processing that will be used for the assimilation experiment in our study. Note that some additional quality control by comparison to the first-guess will be applied later during the assimilation experiment to reject individual observations with large errors. For each 3-hourly analysis, we assimilate only the GPS observation closest to the analysis

time and within ± 1.5 h around the analysis time.

3.3 The data assimilation experiments

Three AROME assimilation cycle experiments are set up in order to investigate the impacts *i)* of the assimilation of GPS ZTD observations and *ii*) of the assimilation of the additional GPS ZTD data collected during the COPS field campaign. A first experiment, called hereafter REF, assimilates the observations used by the pre-operational AROME system, excluding all the GPS observations. Note that also are excluded conventional observations, such as those from radiosoundings, wind profilers and AMDAR, issued by the COPS campaign but not available usually in the operational databases. The reason for doing that is to identify the impact from the GPS ZTD observations in an operational context. The second assimilation cycle experiment, called hereafter COP, assimilates in addition to the observations assimilated in REF, all the GPS ZTD observations including the operational E-GVAP ones and those from the COPS campaign. The two assimilation cycles start at 2100UTC, 5 July 2007 and end at 2100UTC, 20 July 2007. The pre-processing described in section 3.2 has been carried out on an other 15-day period from 23 July 2007 to 6 August 2007.

Then, in order to evaluate the impact of assimilating the high-spatial resolution GPS ZTD dataset from the COPS campaign, a third assimilation cycle experiment, called hereafter OPR, assimilates in addition to the observations assimilated in REF, only the GPS ZTD observations from the operational European network. The GPS ZTD observations from the COPS campaign are excluded. OPR starts at 0000UTC 15 July 2007 using the first-guess from REF at the same time and ends at 0000UTC 20 July 2007.

Then, the AROME model is run for 30 hours starting from the analyses at 0000UTC on 18, 19 and 20 of July 2007, respectively, and for each of the data assimilation experiment in order to study the impact of the GPS ZTD assimilation on the forecast of IOP9 convective systems.

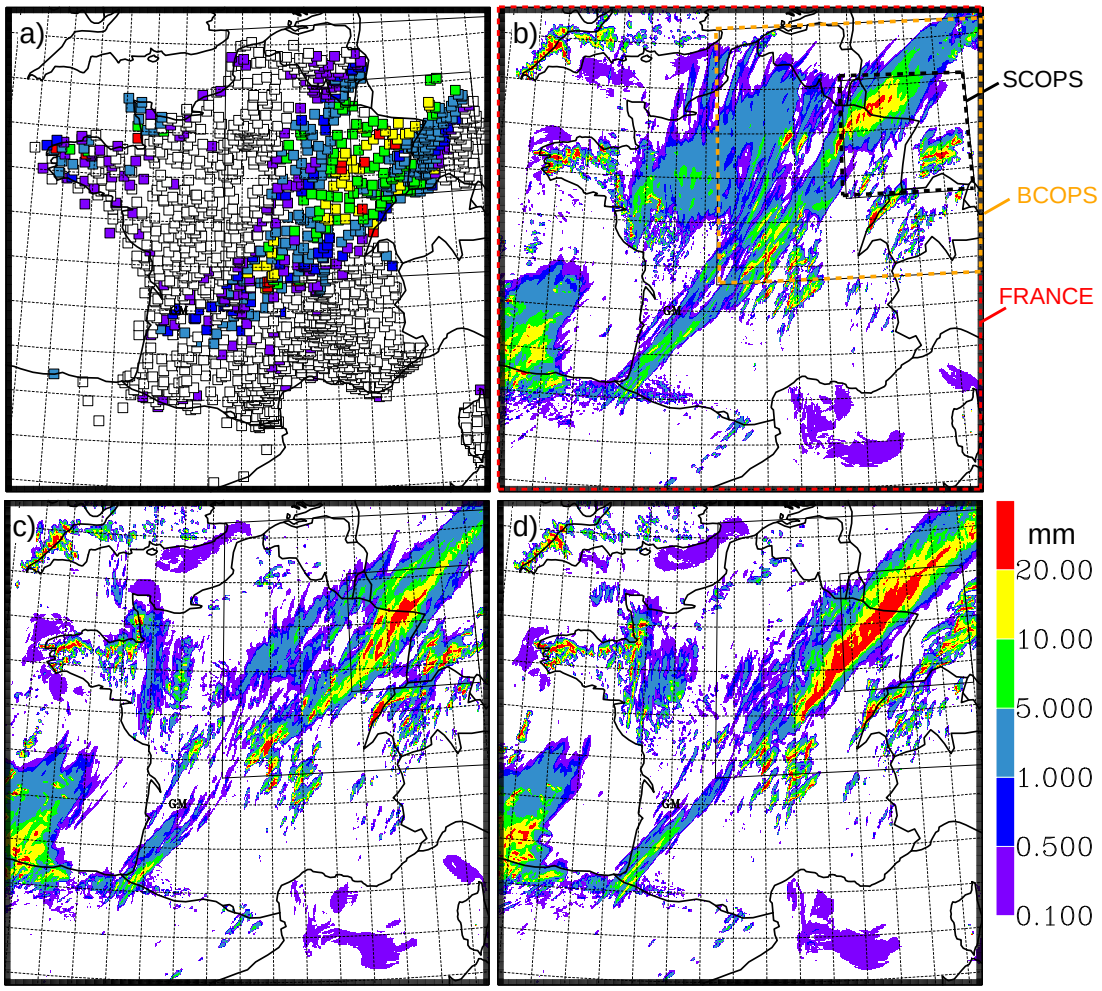

Figure 4: 12-h accumulated precipitation (mm) from 03 UTC to 15 UTC, 19 July 2007 from : a) raingauge observation; b) the REF AROME forecast; c) the COP AROME forecast; d) the OPR AROME forecast. The AROME forecasts start all at 00 UTC, 19 July 2007. In panel b, are delineated the domains used for computing the scores displayed in Fig. 5, 6, 4.1,4.1)

.

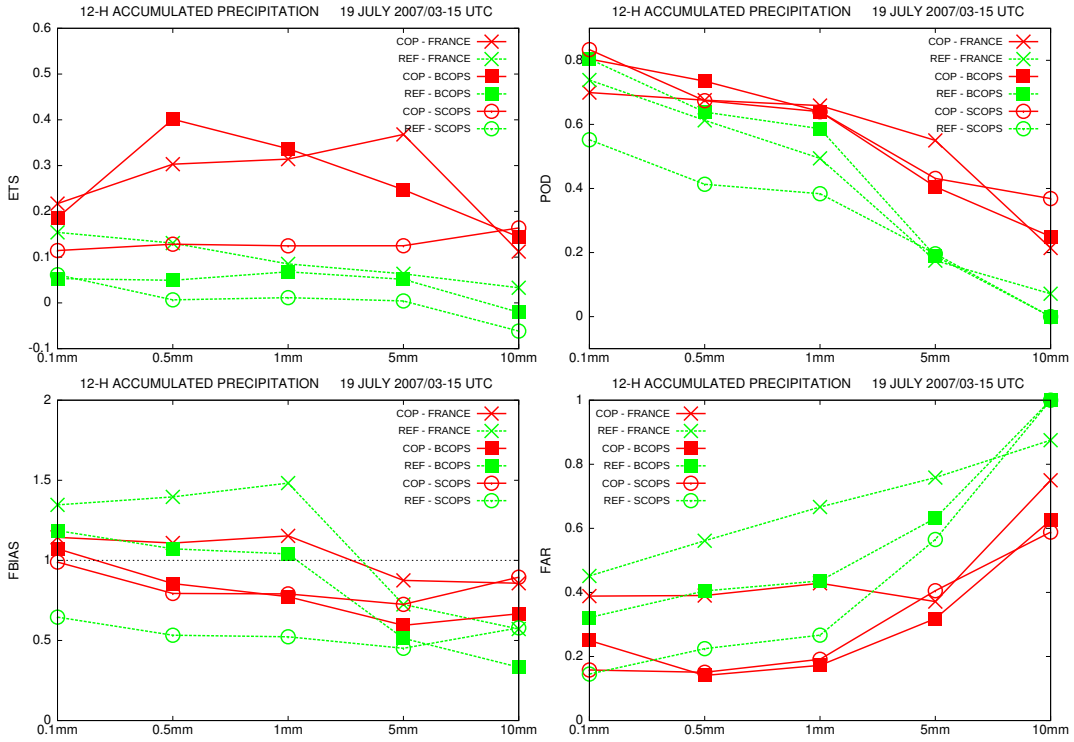

Figure 5: Quantitative scores against raingauge observations from 03 UTC to 15 UTC, 19 July 2007 for the 12-h accumulated precipitation AROME forecast from REF (dashed green lines) and COP (solid red lines) runs : a) Equitable Threat Score (ETS); b) Probability Of Detection (POD); c) Frequency Bias (FBIAS) and d) False Alarm Rate (FAR) are displayed for the 0.1 mm, 0.5 mm, 1 mm, 5 mm and 10 mm thresholds. The scores have been computed over three domains shown in Fig. 4: the FRANCE (cross), BCOPS (square) and SCOPS (circle) domains.

4 Results and discussion

4.1 Impact on the AROME forecast of assimilating GPS ZTD

The largest differences in quantitative precipitation forecast are found for IOP9b (19 July). Figure 4 shows the 12-h accumulated precipitation associated with convective systems observed during the morning of 19 July, together with the corresponding precipitation forecast from the REF and COP experiments (Fig. 4bc) starting from the AROME analyses at 00 UTC, 19 July 2007. Clearly, assimilating GPS ZTD observations improves QPF over all the AROME domain. The weak precipitation over northern France in REF is advisedly removed in COP, as well as the precipitation over western France is better simulated in COP. Concerning the COPS region, the high precipitation rainband is shifted eastward in COP, leading to a better agreement with the observations. Scores (Fig. 5) confirm that COP has a better skill than REF for IOP9b. The Equitable Threat Score (ETS), the Frequency Bias (FBIAS), the Probability Of Detection (POD) and the False Alarm Ratio (FAR) have been computed for thresholds 0.1mm, 0.5mm, 1mm, 5mm and 10mm and for three geographical domains (shown in Fig.4b), called France, BCOPS and SCOPS respectively. The France domain corresponds to the whole AROME domain, whereas the BCOPS domain focuses on precipitating systems that have affected Eastern France during IOP9. The SCOPS domain is designed to assess more specifically the local impact of assimilating high-spatial resolution GPS ZTD data over the COPS region. Whatever the domain, the scores show an overall improvement when GPS ZTD observations are assimilated. ETS are actually found better for COP than REF whatever the thresholds and the domain considered (Fig. 5). Lower FAR and higher POD are obtained for COP, except for the 0.1 mm threshold over the France domain for POD. Also FBIAS values are generally closer to one for COP than for REF.

Figure 6 shows ETS for the two other days and for the same forecast ranges which encompass most of the observed precipitating activity of IOP9a and IOP9c. The thresholds for computing the scores have been adapted to the precipitation amount observed each day. On July 18, ETS values indicate that COP performs better than REF, except for the 0.1 mm threshold over the France domain for POD. ETS values are globally also better for COP than for REF for the France and SCOPS domains

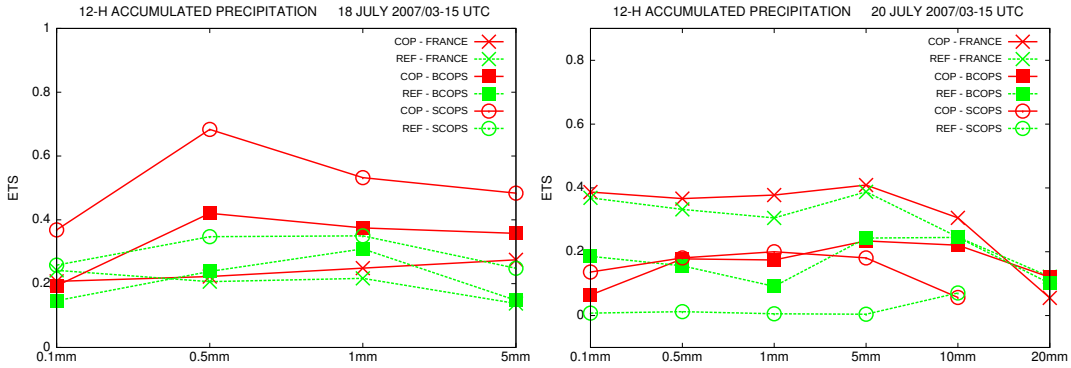

Figure 6: Equitable Threat Scores for the 12-h accumulated precipitation AROME forecast from REF (dashed green lines) and COP (solid red lines) runs against raingauge observations : a) from 03 UTC to 15 UTC, 18 July 2007 and b) from 03 UTC to 15 UTC, 20 July 2007. *See* Fig. 4 for location of the FRANCE, BCOPS and SCOPS domains on which ETS is computed. The scores are displayed for the 0.1 mm, 0.5 mm, 1 mm, 5 mm thresholds in (a) and for the 0.1 mm, 0.5 mm, 1 mm, 5 mm, 10 and 20 mm thresholds in (b).

on July 20, whereas results are more mitigated for the BCOPS domain. Globally, the COP AROME forecasts are thus also better than the REF ones on 18 and 20 July, even though the differences between the two sets are not as large as on 19 July.

Finally, a big picture of the skill of the quantitative precipitation forecast is provided by Fig. 8. Scores have been computed for the 24-h accumulated rainfall covering the whole IOP9 (i.e. 00h-24h forecast for 18, 19 and 20 July) and over the whole AROME domain (*i.e.* the France domain). Even though the differences are weak, the COP experiment produces weaker false alarm rate and higher probability of detection.

Not only the precipitation totals are improved when ZTD GPS data are assimilated but also the timing of the precipitating events. Figure 8 displays the temporal evolution of the 3-hour cumulative precipitation for two stations located in the COPS region (see Fig.1 for location of the raingauge stations). Fessenheim, the southern station, has been affected mainly by the precipitating systems during the morning of July 18, whereas Volmunster has registered the high precipitation associated with the precipitating systems on July 19 and experienced weaker precipitation the day after. Clearly, the timing of the precipitating systems is better in COP than in REF.

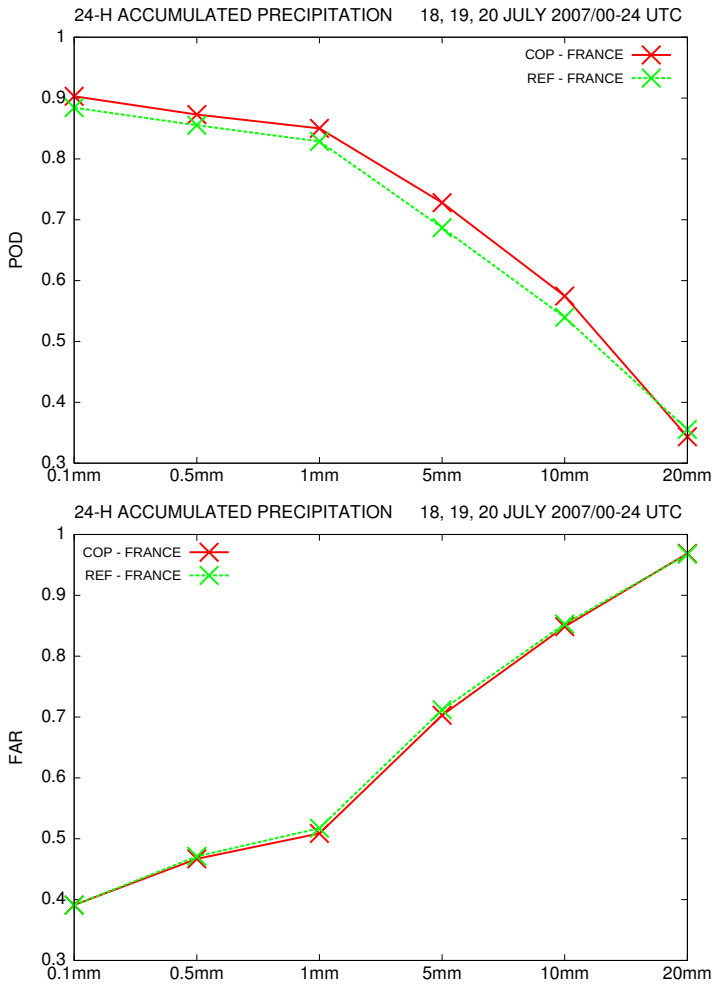

Figure 7: Quantitative scores against raingauge observations for the 24-h accumulated precipitation for all the AROME REF (dashed green lines) and COP (solid red lines) forecast runs: a) Probability Of Detection (POD)and b) Frequency Bias (FBIAS) are displayed for the 0.1 mm, 0.5 mm, 1 mm, 5 mm, 10 mm and 20 mm thresholds.

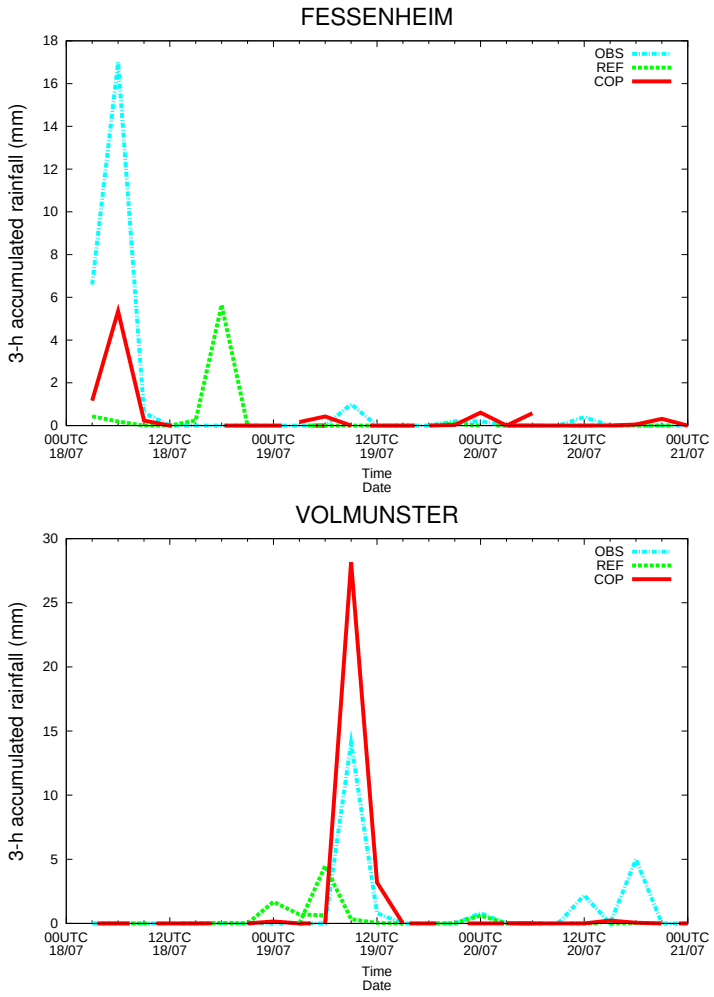

Figure 8: Time series of the 3-hour accumulated rainfall (mm) for the Fessenheim and Vomunster stations from 00 UTC, 18 July 2007 to 00 UTC, 21 July 2007. All the 30-h forecast REF (dashed green line) and COP (solid red line) runs are displayed together with the observations (dot-dashed line).

The precipitation peaks on July 18 and 19 occur at the same time in COP and in the observations whereas in REF the peaks are anticipated or delayed.

As the 06 UTC and 18 UTC radiosounding observations from the Nancy site over the COPS region (*see* Fig. 1 for location of the Nancy sounding) are not used in the assimilation cycle, they can be exploited to further evaluate the impact of assimilating GPS ZTD data. Figure 10 provides a comparison between the 18 UTC Nancy sounding and the 18-h forecast from the REF and COP AROME runs starting at 00UTC, on 18, 19 and 20 July respectively. Overall, the COP runs perform better than the REF ones. The differences are remarkable on 18 July, with a better simulation of the boundary layer in COP. This comparison shows also that the impact of the assimilation of the GPS ZTD observations on the forecast is not limited to only few hours, but it can last beyond 18 hours.

4.2 Impact of COPS GPS ZTD data

In order to assess the impact of assimilating additional high-spatial GPS ZTD over the COPS region, the assimilation experiment OPR does not use the GPS ZTD data from the COPS field campaign. The 40 stations from the COPS field campaign selected by the pre-processing (Fig. 3) are therefore excluded of the data assimilation cycle of OPR. Figure 4 which shows the 12-h accumulated precipitation forecast on IOP9b for REF, COP and OPR experiments allows us to assess the impact of the COPS GPS ZTD observations. First of all, the precipitation forecast from OPR is closer to the COP one than the REF one. Assimilating the European network GPS data (*i.e.* OPR experiment) helps thus a lot to remove the precipitation over northern France as well as to better locate the frontal convective rainband. The assimilation of additional COPS GPS ZTD data in the COPS region (*i.e.* COP experiment) leads to reduce the intensity of the frontal convective rainband. In general, COP gives much better QPF scores than OPR for threshold less than 10mm for all the 3 IOP9 days, as illustrated in Fig. 9 for IOP9b. However, for the thresholds greater than 10mm, COP seems to degrade the QPF scores. This is explained by a tendency of underestimating the heavy rainfall in the COP experiment for all the 3 days, and especially on 19 July.

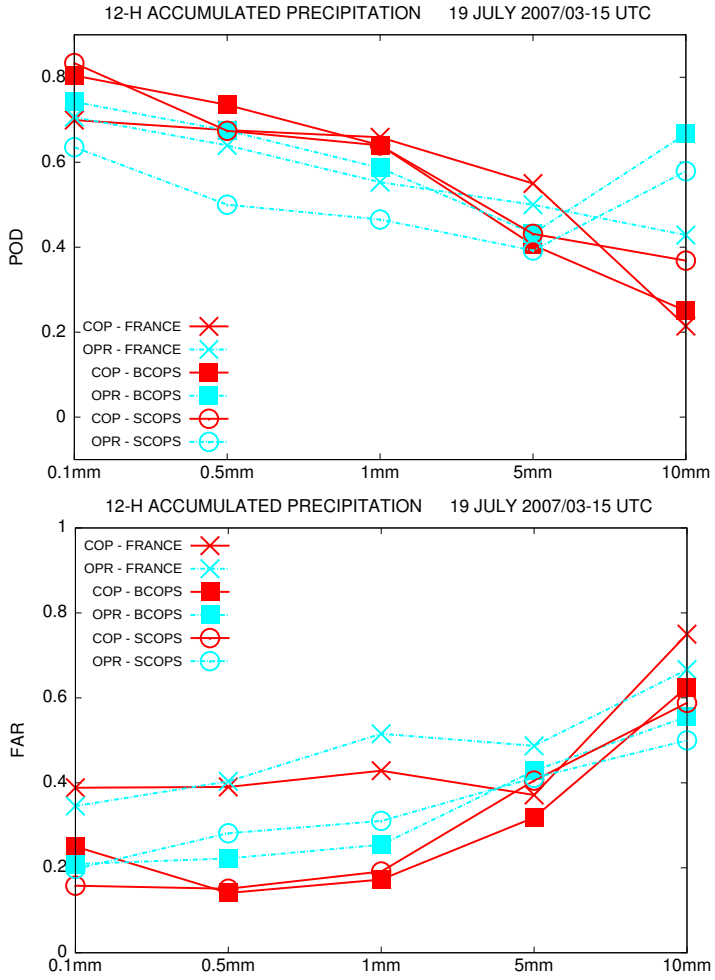

Figure 9: Probability Of Detection (POD)and False Alarm Rate (FAR) for the 12-h accumulated precipitation AROME forecast from COP (solid red lines) and OPR (dashed blue lines) runs against raingauge observations from 03 UTC to 15 UTC, 19 July 2007 over the FRANCE, BCOPS and SCOPS domains (see Fig. 4 for location of these domains). Results are displayed for the 0.1 mm, 0.5 mm, 1 mm, 5 mm and 10 mm thresholds.

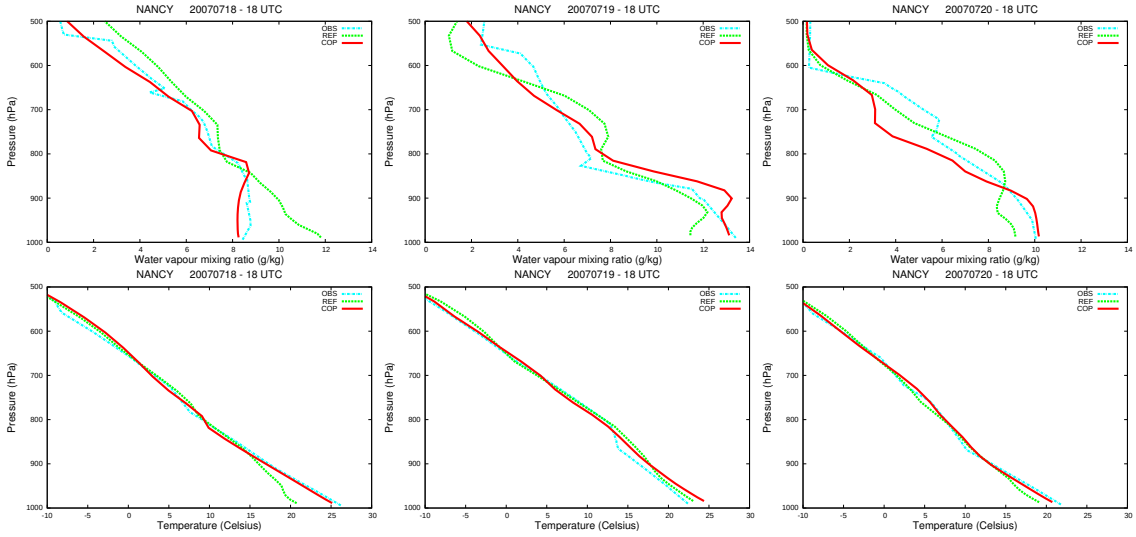

Figure 10: Vertical profiles of temperature (° C, top row) and water vapour mixing ratio (g/kg, bottom row) from the Nancy radiosounding observations (blue dot-dashed curves) at 18 UTC, on 18 (left), 19 (middle) and 20 (right) July 2007 and from the 18-h forecast REF (green dashed curves) and COP (red solid curves) runs. *See* Fig.1 for location of the Nancy sounding.

4.3 Discussions

In order to investigate the reason of the positive impact of GPS ZTD assimilation on the precipitation forecast, the initial conditions of the 30-h AROME forecast runs are now examined for IOP9b which shows the largest impact. The differences for the analyses valid at 00 UTC, 19 July are compared in Fig. 11b between the COP and the REF experiments and Fig. 11d between the COP and the OPR experiments for the Integrated Water Vapour (IWV). The differences between the first-guess used for these analyses are also shown Fig. 11a between the COP and the REF experiments and Fig. 11c between the COP and the OPR experiments. The first-guess are here the 3-hour AROME forecasts issued from the previous analysis of the assimilation cycles. IWV is considered here as it is closely related to ZTD and it is also a relevant parameter for quantifying the moisture in the troposphere available for precipitation formation.

First of all, Fig. 11 shows that the differences between the COP and the REF analyses are larger than between the COP and OPR analyses. The same comment applies to the first-guess. It agrees with weaker differences found for the precipitation forecast between OPR and COP than between COP and REF. Also, a large part of the differences found between the analyses can be attributed to the differences between the first-guess. Therefore the impact on the short-range precipitation forecast on 19 July cannot be only imputed to the assimilation of GPS ZTD observations at 00 UTC, 19 July but it also results from the assimilation of GPS ZTD in the 3-hourly update assimilation cycle prior July 19.

Both in the first-guess and in the analysis, IWV is increased in COP with respect to REF in the region of the convective frontal band in the eastern France and soutwestern Germany (Fig. 11ab) and can explain a better location of the rainy band in COP. Also, in the southwestern part of the domain shown in Fig. 11ab, IWV is decreased in COP. The precipitation that develops over that region and then are transported northward in REF is thus not allowed to form in COP due to weaker available water vapour in that region.

The comparison of the COP and OPR analyses (Fig. 11d) shows that the OPR analysis is globally more humid in the COPS region where the COPS stations haven't been assimilated and may explain that the frontal convective rainband is more intense

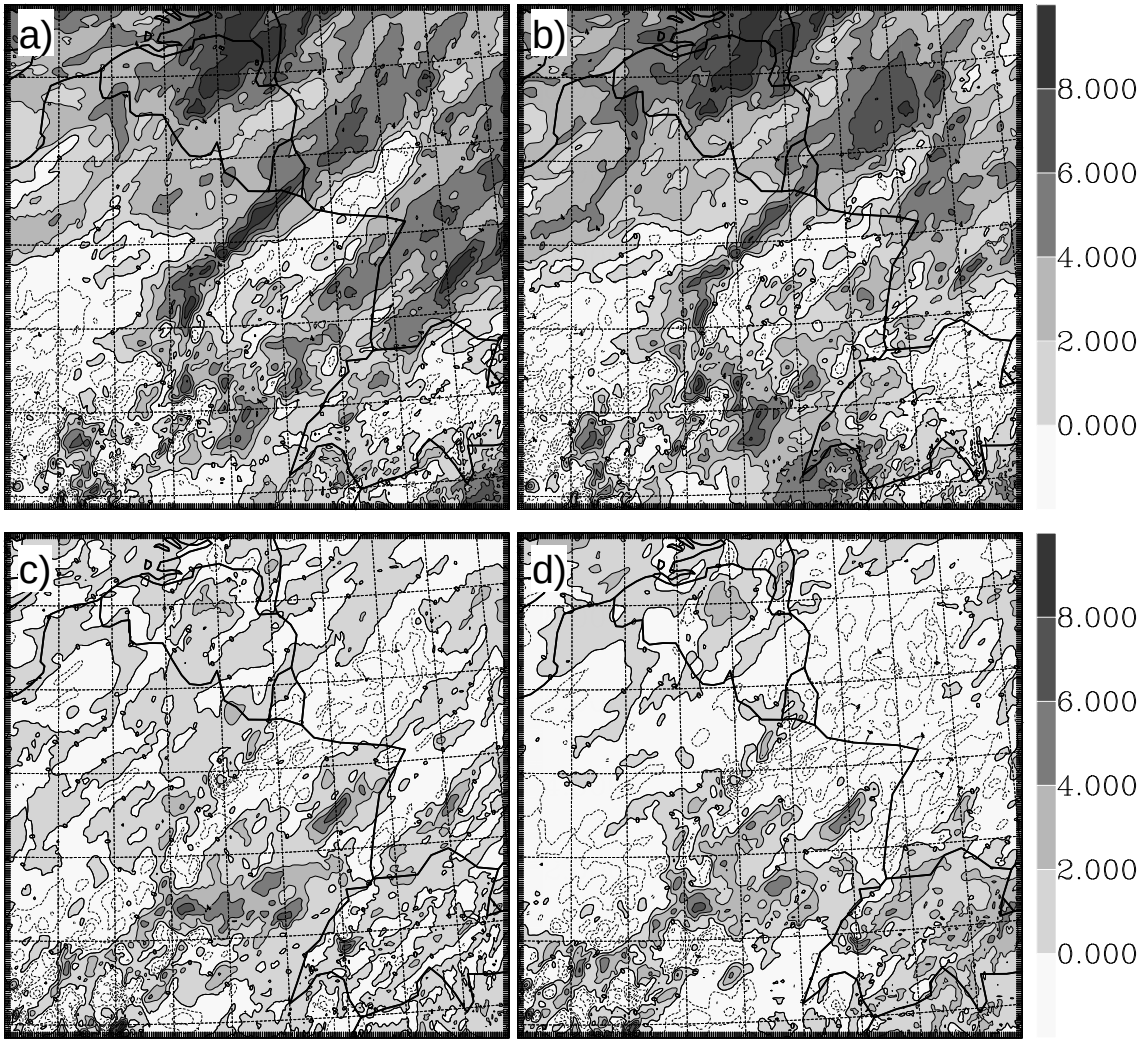

Figure 11: Integrated Water Vapour differences (mm) valid at 00 UTC, 19 July 2007 between : a) COP and REF first guess; b) COP and REF analyses; c) COP and OPR first-guess; d) COP and REF analyses.

in OPR. The differences between OPR and COP are not limited to the COPS region where the COPS GPS ZTD observations have been removed. They result from the previous analyses and propagation of the modifications within the 3-h forecast during the assimilation cycle.

5 Conclusion

The impact of assimilating GPS ZTD observations into a 2.4-km high-resolution convection-resolved model has been assessed through a case study of convective systems occurred on 18-20 July 2007 (IOP9) during the COPS field campaign. Three assimilation cycle experiments were conducted using the 3DVAR/AROME assimilation system: one called REF where no GPS ZTD observations were assimilated at all, one called COP where both GPS ZTD observations from the operational European E-GVAP network and the COPS-dedicated network were assimilated, and a last one called OPR where only GPS ZTD observations from the operational network were assimilated. In order to assess the impact on the precipitation forecast of the IOP9, nine 30-h AROME forecast runs have been carried out from the AROME analyses issued from the three assimilation cycles at 00 UTC, 18, 19 and 20 July 2007, respectively.

Clear improvements in terms of QPF are found when the AROME model starts from the analyses of the assimilation experiments where GPS ZTD observations are assimilated. Note that the QPF for high precipitation rate, which always remains a major problem in the forecast of convective systems, improves as well. The role of the additional GPS ZTD observations from the COPS field campaign over Eastern France and southwestern Germany has also been investigated. The impact of assimilating these additional observations on the analyses and precipitation forecasts is less significant than the one gained by assimilating the GPS ZTD observations from the European network. Whereas the assimilation of the additional COPS GPS ZTD observations improves the forecast of weak precipitation, it slightly deteriorates the scores for heavy precipitation. The differences in the analyses used as initial conditions for the 30-h AROME forecast runs cannot be simply explained by the assimilation of the GPS ZTD observations at the time of the analysis but result from the successive analysis/3-h forecast steps carrying out during the rapid update cycle which produce differences in the first-guess used for the analysis.

This impact brought by the addition of GPS ZTD observations is much more obvious than the results from the previous experiment (Yan *et al.* 2008) where assimilation of GPS ZTD observations was done at 9.5-km resolution. Even though the meteorological situations are not the same, better model terrain and better first-guess benefited from the 2.4-km high resolution model AROME allow to exploit better and more GPS ZTD observations in the assimilation, and thus to achieve better precipitation forecasts. As most of the improvement is gained by assimilating GPS ZTD observations from the European operational GPS network, our study opens up encouraging horizons for the assimilation of GPS ZTD in the new high-resolution operational NWP systems. One advantage of GPS ZTD observations with respect to other types of observation related to water vapour is that it provides mesoscale information to assimilate whatever the weather conditions. For instance, METEOSAT satellite radiances are assimilated only outside cloudy areas and therefore tend to bias the analyses toward dry conditions. The assimilation of GPS ZTD doesn't bring such bias as it can both remove moisture outside observed clouds or moisten in cloudy and rainy regions.

Acknowledgements: The authors would like to thank the E-GVAP project for their efforts in making European ground-based GPS data available in near real time. The COPS campaign and this study benefited from support from the ANR (Agence Nationale pour la Recherche) and the CNRS/INSU (Institut National des Sciences de l'Univers). The authors would like also to thank Michael Kunz for providing the raingauge data over Germany.

References

[1] Bénard P. 2004. Aladin/AROME dynamical core, status and possible extension to IFS. *In ECMWF Seminar proceeding, Sept. 2004*, available from http://www.ecmwf.int/publications/library in Nov. 2004, or by post from ECMWF, Shinfield Park, Reading RG29AX, Royaume Unis.

[2] Berre L. 2000. Estimation of synoptic and mesoscale forecast error covariances in a limited area model. Mon.Wea. Rev., 128, pp. 644-667.

[3] Berre L, Stefanescu SE, Belo Pereira M. 2006. The representation of the analysis in three error simulation techniques. *Tellus* **58A:**196—209.

[4] Boehm J, Werl B, Schuh H. 2006. Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data. *J. Geophys. Res.*, **111**, B02406.

[5] Bouttier F. 2007. Arome, avenir de la prévision régionale. *La Météorologie,* **58**:12-20 (in French).

[6] Bubnová R, Hello G, Bénard P, Geleyn JF. 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. Wea. Rev.*, **123:**515-535.

[7] Brenot H, Ducrocq V, Walpersdorf A, Champollion C, Caumont O. 2006. GPS Zenith Delay Sensitivity evaluated from High-Resolution NWP Simulations of the 8-9th September 2002 Flash-Flood over Southeastern France. *J. Geophys. Res.* **111**:D15105, doi:10.1029/2004JD005726.

[8] Caniaux G, Redelsperger J-L, Lafore J-P. 1994. A numerical study of the stratiform region of a fast-moving squall line. *J. Atmos. Sci.* **51:**2046-2074.

[9] Champollion C., Masson F., Van Baelen J., Walpersdorf A., Chéry J. and Doerflinger E., 2004: GPS monitoring of the tropospheric water vapor distribut ion and variation during the 9 September 2002 torrential precipitation episode in the Cévennes (southern France) J. Geophys. Res., Vol. 109, No. D24

[10] Chen G, Herring TA. 1997. Effects of atmospheric azimuthal asymmetry on the analysis of space geodetic data. *J. Geophys. Res.*, **102(B9):**20489-20502.

[11] Cucurull L, Vandenberghe F, Barker F, Vilaclara E, Rius A. 2004. Three dimensional variational data assimilation of ground-based GPS ZTD and meteorological observations during the 14 December 2001 storm event over the Mediterranean area. *Mon. Wea. Rev.* **132:**749-763.

[12] Cuxart J, Bougeault P, Redelsperger J-L, 2000: A turbulence scheme allowing for mesoscale and large eddy simulations. *Q. J. R. Meteorol. Soc.* **126**: 1-30.

[13] Davis JL, Herring TH, Shapiro II, Rogers AEE, Elgered G. 1985. Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimatoin of baseline length. *Radio Sci.* **20:**1593-1607.

[14] De Pondeca M S F V, Zou X. 2001. A case study of the variational assimilation of GPS zenith total delay observations into a mesoscale model. *J. Appl. Meteor.* **40(1):** 559-1576.

[15] Ducrocq V, Ricard D, Lafore J-P, Orain F. 2002. Storm-Scale Numerical Rainfall Prediction For Five Precipitating events over France: On the Importance of the initial humidity field. *Weather and Forecasting* **17:** 1236-1256, doi:10.1175/1520-0434(2002)017.

[16] King RW, Bock Y. 2008. Documentation for the GAMIT GPS analysis software, release 10.33. Departement of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology-Scripps Institution of Oceanography University of California at San Diego.

[17] Fischer C, Montmerle T, Berre L, Auger L, Stefanescu SE. 2005. An overview of the variational assimilation in the ALADIN/FRANCE NWP system. *Quart. J. Roy. Meteor. Soc.* **131:**3477-3492.

[18] Lafore J-P, Stein J, Asencio N, Bougeault P, Ducrocq V, Duron J, Fischer C, Héreil P, Mascart P, Masson V, Pinty J-P, Redelsperger J-L, Richard E, Vil-Guerau de Arellano J. 1998. The Meso-NH Atmospheric Simulation System. Part I: adiabatic formulation and control simulations. Scientific objectives and experimental design. *Ann. Geophys.* **16:**90-109.

[19] Masson V. 2000. A physically-based scheme for the urban energy budget in atmospheric models. *Bound. Layer Meteor* **1994:**357-397.

[20] Montmerle T, Rabier F, Fischer C. 2007. Respective impact of polar orbiting and geostationary satellite observations in the Aladin/France NWP system. *Quart. J. Roy. Meteor. Soc.* **133:**655-671.

[21] Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, CLough SA. 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.* **102D:**16663-16682.

[22] Noilhan J, Mahfouf JF. 1996. The ISBA land surface parameterisation scheme. *Global and Planetary Change*, **13:**145-159.

[23] Peng S-Q, Zou X. 2004. Impact on Short-Range Precipitation Forecasts from Assimilation of Ground-Based GPS Zenith Total Delay and Rain Gauge Precipitation Observations. *J. Meteorol. Soc. Japan.* **82(1B):**491-506, doi:10.2151/jmsj.2004.491.

[24] Poli P, Moll P, Rabier F, Desroziers G, Chapnik B, Berre L, Healy SB, Andersson E, EI Guelai FZ. 2007. Forecast impact studies of zenith total delay data from European near real-time GPS stations in Meteo France 4DVAR. *J. Geophys.Res.* **112**,D06114,doi:10.1029/2006JD007430.

[25] Richard E, Buzzi A, Zängl G. 2007. Quantitative precipitation forecasting in the Alps: The advances achieved by the Mesoscale Alpine Programme, *Q. J. R. Meteorol. Soc.* **133(625):** 831-846, doi:10.1002/qj.65

[26] Seity Y, Brousseau P, Malardel S, Masson V, Bouttier F, Hello G. 2007. Status of AROME developments, *Aladin Newsletter*, **33:**40-47

[27] Stefanescu SE, Berre L, Belo Pereira M. 2006. The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area analysis. *Mon. Wea. Rev.* **134:**3456-3478.

[28] Thayer, G.D. (1974). *An improved equation for the radio refractive index of air.* Radio Science, 9(10), 803-807.

[29] edel H, Huang XY. 2004. Impact of ground based GPS data on numerical weather prediction, *J. Meteorol. Soc. Japan.* **82(1B):**459-472, doi:10.2151/jmsj.2004.459.

[30] Vedel H, Huang X-Y, Haase J, Ge M, Calais E. 2004. Impact of GPS Zenith Tropospheric Delay data on precipitation forecasts in Mediterranean France and Spain. *Geophys. Res. Lett.* **31**, L02102 doi:10.1029/2003GL017715.

[31] Wulfmeyer V *et al.* 2008. The convective and orographically-induced precipitation study: A research and development project of the World Weather Research Program for improving quantitative precipitation forecasting in low-mountain regions. *Bull. Amer. Meteor. Soc.* **89(10):** 1477–1486, doi:10.1175/2008BAMS2367.1

[32] Yan X, Ducrocq V, Poli P, Hakam M, Jaubert G, Walpersdorf A. 2008: Impact of GPS zenith delay assimilaiton on convective cale predication of Mediterranean heavy rainfall.*J. Geophys.Res.*, in revision.

5.3 Impact of the observation operator

As in chapter 4, we have carried out a sensitivity study to the observation operator. Whereas in the COP assimilation experiment, the POLI07 operator was used, we operated in a companion assimilation experiment the BREN06 observation operator. We call this experiment MIC. MIC assimilates as COP both the E-GVAP operational GPS observations and the GPS data collected during the COPS field campaign. The cycle assimilation experiments were run for 5 days, starting from 15, July to 20, July 2007. Then 30-h AROME forecasts were issued starting from 00UTC for day 18, 19 and 20 of July during the heavy precipitation events.

As for the COP and OPR experiments, the largest differences in terms of rainfall forecasts are found for IOP9b. Figure 5.2 displays the 12h accumulated precipitation from 03UTC to 15UTC, 19 July 2007 from the COP and the MIC AROME forecast. The two rainfall fields are quite similar. The differences between COP and MIC are smaller than between the COP and OPR differences or between the REF and COP experiments (see Fig. 4 of sec. 5.2). That means changing the observation operator has a smaller impact than not assimilating GPS observations or assimilating only a part of the GPS ZTD observations. Figure 5.3 which shows the scores computing on the same period confirms that the differences are small between the experiments. The scores ETS (Fig5.3a), FAR (Fig5.3c) and POD (Fig5.3d) all however indicate a positive impact of MIC compared with COP for thresholds greater than 5mm.

To conclude, the impact of the formulation of the observation operator is also found weak with the AROME 3DVAR assimilation system. Slightly better results are found when using the BRE06 formulation.

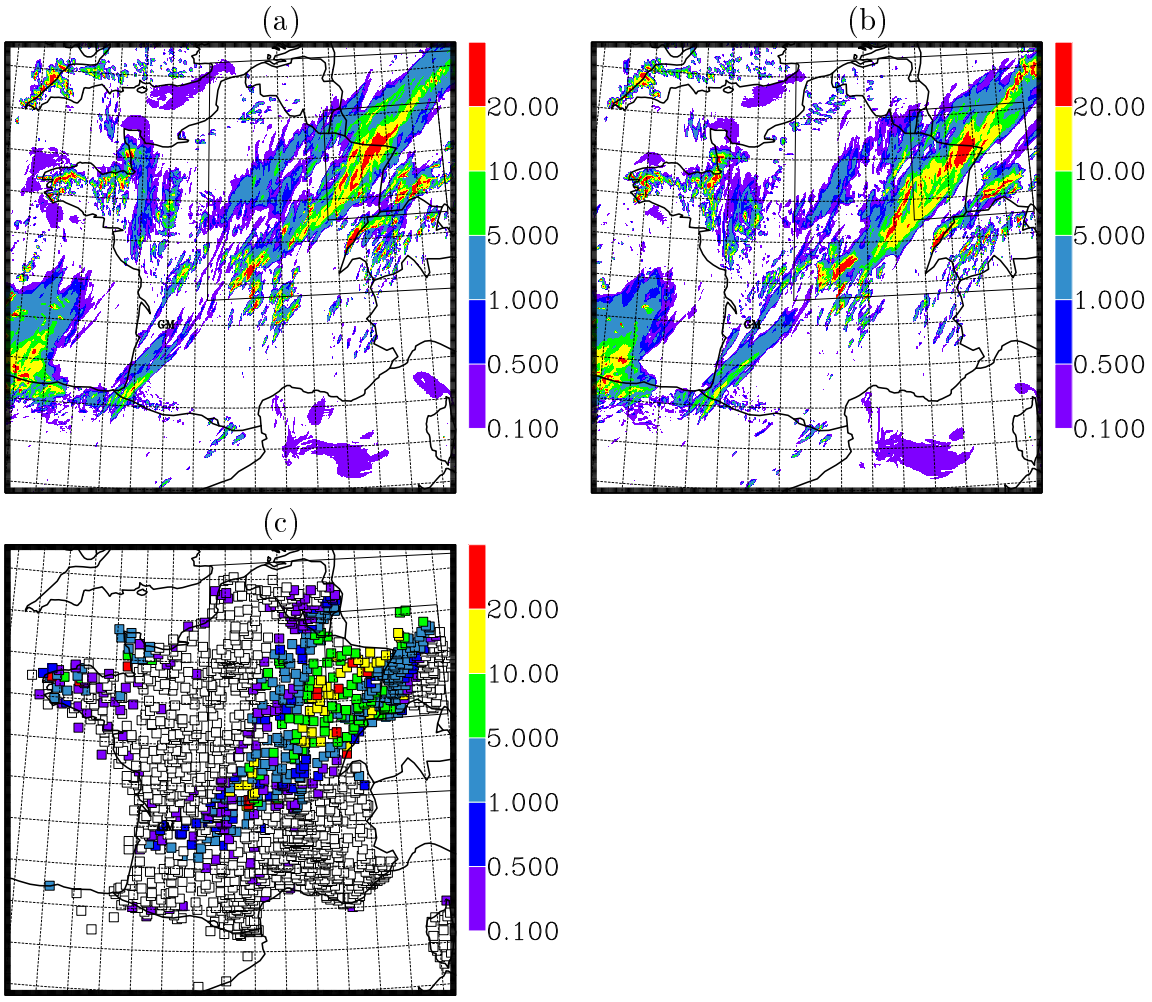

FIG. 5.2: 12-h accumulated precipitation (mm) from 03UTC to 15UTC, 19 July 2007 from : (a) the COP AROME forecast ; (b) the MIC AROME forecast ; (c) raingauge observations.

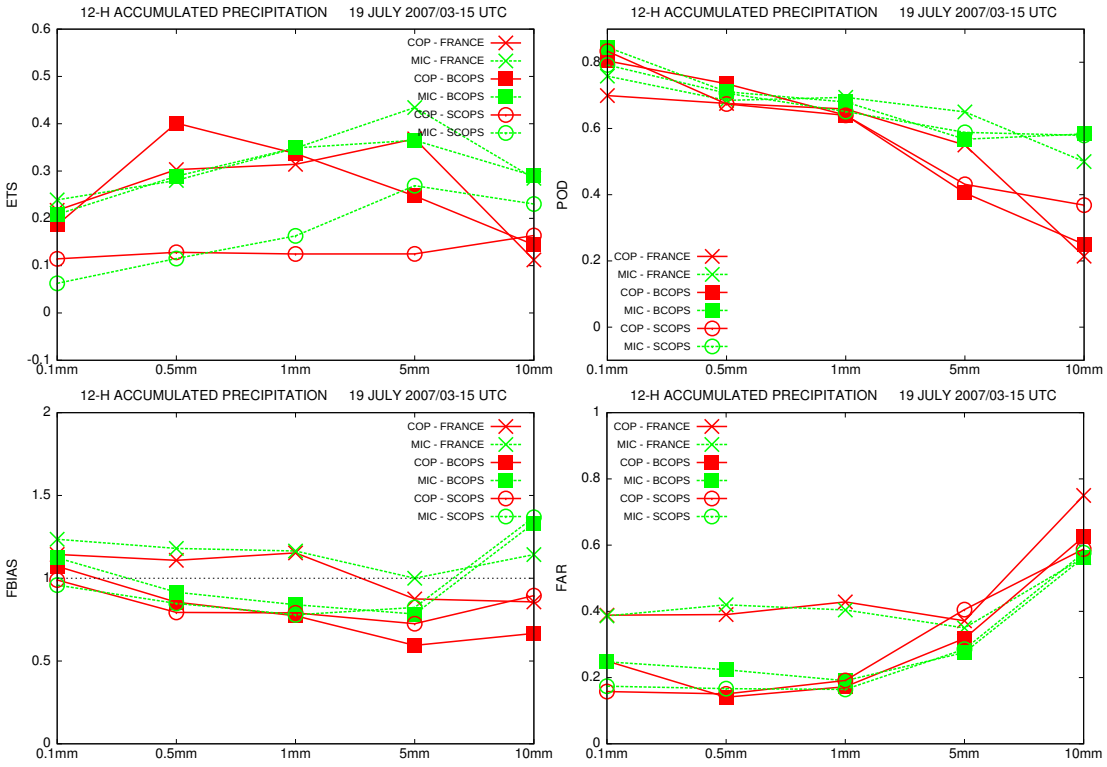

FIG. 5.3: QPF scores for the 12h accumulated rainfall from 03UTC to 15UTC, 19 July 2007. Idem Figure 5, sec.5.2 except for MIC and COP AROME forecasts.

Conclusion et perspectives

L'objectif de ce travail de thèse était d'évaluer l'impact de l'assimilation de données GPS à méso-échelle sur la prévision à haute résolution de systèmes convectifs. Ce travail, qui a aussi contribué au développement de l'assimilation de données GPS dans les systèmes de prévision numérique à méso-échelle de Météo-France, s'est inscrit dans la chronologie du développement et de la disponibilité de ces systèmes. Au démarrage de ce travail de thèse, l'assimilation 3DVAR/ALADIN était tout juste opérationnelle (sans assimilation de données GPS) et le nouveau modèle AROME encore en développement. Une première partie du travail de thèse a ainsi été consacrée à l'assimilation des données GPS de ZTD avec le 3DVAR/ALADIN (9.5 km). L'évaluation de l'impact sur la prévision à haute résolution (2.5 km) a été réalisée en initialisant le modèle MesoNH avec des analyses ALADIN ayant bénéficié ou non de l'assimilation des données de ZTD. La disponibilité de la version pré-opérationnelle du système de prévision numérique AROME dans la deuxième partie de ce travail nous a permis ensuite d'évaluer l'impact de l'assimilation des données GPS de ZTD lorsqu'à la fois le système d'assimilation et le modèle de prévision sont à haute résolution (2.5 km).

Dans la première partie, nous avons tout d'abord cherché à caractériser l'influence d'une observation de ZTD lorsque celle ci est analysée. Pour celà, nous avons réalisé des expériences où une seule observation de ZTD est assimilée avec le 3DVAR/ALADIN à un instant donné. Comme les études antérieures réalisées à plus grande échelle, nous avons trouvé que l'assimilation de ZTD modifie avant tout le champ de vapeur d'eau et très peu le champ de pression ou de température. En terme de répartition verticale de l'information contenue dans la valeur intégrée du ZTD, l'analyse modifie la basse et moyenne troposphère, avec un impact maximum vers 2-3 km d'altitude. L'assimilation d'observations de ZTD a donc pour effet principal d'apporter une information sur la vapeur d'eau dans la basse et moyenne troposphère, *i.e.* sur un paramètre et une région de l'atmosphère peu documentés par les autres systèmes opérationnels d'observation à méso-échelle.

Nous avons ensuite réalisé un cycle d'assimilation de durée un mois où toutes les observations de ZTD sont assimilées avec les autres types d'observations assimilés par le système de prévision numérique ALADIN-France. L'impact de l'assimilation des observations de ZTD a été trouvé neutre en général sur l'analyse et l'ébauche (prévision à 6 heures d'échéance d'ALADIN). Seul un faible impact positif sur l'humidité spécifique de l'ébauche a été trouvé par comparaison aux observations de radiosondages. Les analyses ALADIN ont ensuite été utilisées comme conditions initiales et aux limites pour réaliser des prévisions MesoNH à 2.5 km de résolution de l'évène-

ment du 5 au 9 septembre 2005 dont les fortes précipitations ont affecté le Gard et l'Hérault. Deux runs (à 00UTC et 12 UTC) par jour javec des prévisions jusqu'à 18 heures d'échéance ont été réalisée sur la totalité de l'évènement. L'impact sur la prévision MesoNH des cumuls de précipitation, évalué à la fois qualitativement et quantitativement à l'aide de scores, a été trouvé faiblement positif.

Avec le système d'assimilation AROME, nous avons examiné l'impact sur la prévision AROME de la POI9 (18-20 juillet 2007) de la campagne COPS. Cette POI est caractérisée par des systèmes convectifs précipitants qui se forment sur la France et balayent ensuite la région de la campagne COPS. Un cycle d'assimilation rapide a d'abord été réalisé sur une durée de 15 jours, puis des prévisions AROME de 30h d'échéance ont été réalisées à partir des analyses à 00 UTC pour les trois jours de la POI9. Un impact positif sur la prévision AROME des précipitations est clairement mis en évidence lorsque les données GPS sont assimilées. L'impact est particulièrement significatif pour la journée du 19 juillet (POI9b). Le calendrier de ce travail de thèse n'a pas permis d'exploiter complètement les nombreuses données potentiellement disponibles de la campagne COPS pour valider à la fois les analyses et les prévisions AROME. Une comparaison avec les données du LIDAR vapeur d'eau LEANDRE à bord de l'ATR42 français est notamment en cours et montre que les prévisions issues du cycle d'assimilation avec données GPS sont plus proches des observations vapeur d'eau LEANDRE.

Pour la campagne COPS, un réseau dense de stations GPS avait été déployé dans la région des Vosges et de la Forêt Noire, nous offrant l'opportunité de tester l'apport de l'assimilation des données d'un tel réseau dense en complément des stations du réseau européen E-GVAP. Les résultats de ces expériences d'assimilation montrent que la majorité du gain obtenu grâce à l'assimilation de données GPS est attribuable à l'assimilation des données du réseau opérationnel E-GVAP. Ce résultat est particulièrement important dans un contexte opérationnel. Même si la résolution du réseau E-GVAP peut apparaître faible par rapport à celle du système AROME, il apporte déjà une information à méso-échelle que le nouveau système d'assimilation AROME est capable d'exploiter pour améliorer l'analyse d'humidité dans la basse et moyenne troposphère.

Dans le cadre de ce travail de thèse, nous avons aussi implémenté et évalué un nouvel opérateur d'observation dans les systèmes d'assimilation 3DVAR/ALADIN et 3DVAR/AROME, en suivant les recommendations de la thèse de Brenot (2006). Les résultats obtenus avec cet opérateur ont été comparés à ceux obtenus avec l'opérateur d'observation original utilisé avec le système global 4DVAR/ARPEGE. Aussi bien les expériences d'assimilation à une observation que les expériences d'assimilation réalisées sur la POI9 de COPS montrent une faible dépendance à la formulation de l'opérateur d'observation. Un impact très faiblement positif en faveur de la formulation de l'opérateur selon (Brenot et al., 2006) est trouvé sur la POI9 de COPS avec le 3DVAR/AROME.

L'impact positif obtenu avec le 3DVAR/AROME sur la POI9 de COPS est plus

significatif que celui obtenu dans la première partie de la thèse avec le système 3DVAR/ALADIN pour l'épisode Méditerranéen de pluie intense. Une meilleure ébauche avec un relief modèle plus précis et une fréquence d'analyse plus importante dans le cas d'AROME peuvent expliquer ce plus fort impact. Mais ce n'est pas la seule explication possible : en effet, le cas d'étude et la configuration du réseau GPS sont aussi différents. Dans le cas de la POI9 de COPS, le flux humide de basses couches alimentant les systèmes peut être documenté par un ensemble de stations GPS terrestres en amont des systèmes, alors que dans le cas des systèmes méditerranéens de pluie intense, leur alimentation provient en grande partie des basses couches au dessus de la mer, dépourvues en stations GPS. Le déploiement de stations GPS sur navires ou bouées pourrait permettre de combler en partie le manque d'observations de vapeur d'eau en basse et moyenne troposphère au dessus de la mer. Une étude de faisabilité est actuellement conduite par la communauté GPS française dans le cadre du projet LEFE/VAPIMED, avec un déploiement envisagé dans le cadre de la future campagne de mesures HyMeX sur le cycle de l'eau en Méditerranée (http ://www.cnrm.meteo.fr/hymex/).

Pour conclure, ce travail de thèse a montré un impact globalement positif de l'assimilation des données GPS de ZTD sur la prévision à haute résolution de systèmes convectifs précipitants, et a constitué un résultat encourageant pour une assimilation de ces données en opérationnel. Le modèle de prévision numérique ALADIN-France assimile désormais les données du réseau E-GVAP. Le nouveau système de prévision numérique AROME, déclaré opérationnel en décembre 2008, les assimile également. Comme perspectives d'amélioration de l'assimilation opérationnelle de données GPS avec les systèmes de prévision numérique à mesoéchelle ALADIN et AROME, nous pouvons citer :

- Une augmentation de la densité spatiale du réseau E-GVAP dans les zones terrestres actuellement pauvres en observations de ZTD : Afrique du Nord, et certaines régions d'Espagne, d'Italie et de France. Il existe de nombreux capteurs GPS à ce jour mais ce sont l'accès à leurs données en temps quasi-réel et leur traitement par un des centres d'analyse qui font défaut généralement.
- Le développement de l'observation GPS en mer si les expérimentations mentionnées ci-dessous sont concluantes.
- L'assimilation des délais obliques à la place des délais zénithaux. Dans ce cas une information sur la distribution tri-dimensionnelle de la vapeur d'eau autour de la station GPS pourrait être exploitée et non plus seulement une valeur intégrée à la verticale de la station. Ceci suppose toutefois que les délais obliques fournis par les méthodes de traitement GPS ne soient pas "re-construits" à partir du délai zenithal total mais soient au plus proche de l'observation du retard entre le satellite et la station GPS.

Bibliographie

J. Askne and H. Nordius. Estimation of tropospheric delay for microwaves from surface weather data. *Radio science*, 22(3) :379–386, 1987. Cité page 24.

L. Berre. Estimation of Synoptic and Mesoscale Forecast Error Covariances in a Limited-Area Model. *Mon. Wea. Rev.*, 128(3) :644–667, 2000. Cité page 48.

L. Berre, S.E. Stefanescu, and M. Belo Pereira. The representation of the analysis in three error simulation techniques. *Tellus*, 58A : :196—209, 2006. Cité page 36.

M. Bevis, S. Businger, T.A. Herring, C. Rocken, R.A. Anthes, and R.H. Ware. GPS Meteorology : Remote Sensing of Atmospheric Water Vapor Using the Global Positioning System. *J. Geophys. Res.*, 97(D14), 1992. Cité pages 10 et 26.

M. Bevis, S. Businger, S. Chiswell, T.A. Herring, R.A. Anthes, C. Rocken, and R.H. Ware. GPS Meteorology : Mapping Zenith Wet Delays onto Precipitable Water. *Journal of Applied Meteorology*, 33(3) :379–386, 1994. Cité pages 20 et 61.

O. Bock, E. Doerflinger, F. Masson, A. Walpersdorf, J. Van-Baelen, J. Tarniewicz, M. Troller, A. Somieski, A. Geiger, and B. Bürki. GPS water vapor project associated to the ESCOMPTE programme : description and first results of the field experiment. *Physics and Chemistry of the Earth*, 29(2-3) :149–157, 2004. Cité page 26.

O. Bock, C. Keil, E. Richard, C. Flamant, and M.N. Bouin. Validation of precipitable water from ECMWF model analyses with GPS and radiosonde data during the MAP SOP. *Quat. J. Roy. Meteor. Soc.*, 131 :3013–3036, 2005. Cité page 26.

P. Bougeault and P. Lacarrère. Parameterization of orography-induced turbulence in a meso-beta scale model. *Mon. Wea. Rev.*, 117 :1872–1890, 1989. Cité page 55.

Ph. Bougeault. A simple parameterization of the largescale effects of cumulus convection. *Mon. Wea. Rev.*, 113 :2108–2121, 1985. Cité page 45.

F. Bouttier and P. Courtier. Data assimilation concepts and methods March 1999. *Meteorological training course lecture series. ECMWF*, 2002. Cité page 32.

H. Brenot. *Potentiel de la mesure GPS sol pour l'étude des pluies intenses méditerranéennes*. PhD thesis, Université Joseph-Fourier - Grenoble I, 2006. Cité page 102.

H. Brenot, V. Ducrocq, A. Walpersdorf, C. Champollion, and O. Caumont. 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, 2006. Cité pages 20, 39, 50, 61, 65 et 102.

S. Businger, S.R. Chiswell, M. Bevis, J. Duan, R.A. Anthes, C. Rocken, R.H. Ware, M. Exner, T. VanHove, and F.S. Solheim. The Promise of GPS in Atmospheric Monitoring. *Bulletin of the American Meteorological Society*, 77(1) :5–18, 1996. Cité page 26.

G. Caniaux, J.L. Redelsperger, and J.P. Lafore. A Numerical Study of the Stratiform Region of a Fast-Moving Squall Line. Part I : General Description and Water and Heat Budgets. *J. Atmos. Sci.*, 51(14) :2046–2074, 1994. Cité page 55.

C. Champollion, F. Masson, M.N. Bouin, A. Walpersdorf, E. Doerflinger, O. Bock, and J. Van Baelen. GPS water vapour tomography : preliminary results from the ESCOMPTE field experiment. *Atmospheric Research*, 74(1-4) :253–274, 2005. Cité page 28.

T.L. Clark and RD Farley. Severe Downslope Windstorm Calculations in Two and Three Spatial Dimensions Using Anelastic Interactive Grid Nesting : A Possible Mechanism for Gustiness. *J. Atmos. Sci.*, 41(3) :329–350, 1984. Cité page 55.

P. Courtier, J.N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4D-Var, using an incremental approach. *Quarterly Journal of the Royal Meteorological Society*, 120(519) :1367–1387, 1994. Cité page 45.

P. Courtier, E. Andersson, W. Heckley, J. Pailleux, D. Vasiljevic, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher. The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I : Formulation. *Quat. J. Roy. Meteor. Soc.*, 124 :1783–1808, 1998. Cité page 34.

L. Cucurull, B. Navascues, G. Ruffini, P. Elósegui, A. Rius, and J. Vilà. The Use of GPS to Validate NWP Systems : The HIRLAM Model. *Journal of Atmospheric and Oceanic Technology*, 17(6) :773–787, 2000. Cité page 26.

L. Cucurull, F. Vandenberghe, D. Barker, E. Vilaclara, and A. Rius. Three-Dimensional Variational Data Assimilation of Ground-Based GPS ZTD and Meteorological Observations during the 14 December 2001 Storm Event over the Western Mediterranean Sea. *Mon. Wea. Rev.*, 132(3) :749–763, 2004. Cité page 41.

J. Cuxart, Ph. Bougeault, and J.L. Redelsperger. A turbulence scheme allowing for mesoscale and large-eddy simulations. *Quart. J. Roy. Meteor. Soc.*, 126 :1–30, 2000. Cité page 55.

H. Davies. A lateral boundary formulation for multi-level prediction models. *Quart. J. Roy. Meteor. Soc.*, 102 :405–418, 1976. Cité page 45.

M.S.F.V. De Pondeca and X. Zou. A case study of the variational assimilation of gps zenith delay observations into a mesoscale model. *Journal of Applied Meteorology*, 40(9) :1559–1576, 2001. Cité page 40.

TH Dixon and S.K. Wolf. Some tests of wet tropospheric calibration for the CASA Uno Global Positioning System experiment. *Geophysical Research Letters*, 17, 1990. Cité page 25.

J. Duan, M. Bevis, P. Fang, Y. Bock, S. Chiswell, S. Businger, C. Rocken, F. Solheim, T. van Hove, R. Ware, et al. GPS Meteorology : Direct Estimation of the Absolute Value of Precipitable Water. *Journal of Applied Meteorology*, 35(6) :830–838, 1996. Cité page 26.

V. Ducrocq, D. Ricard, J.P. Lafore, and F. Orain. Storm-Scale Numerical Rainfall Prediction for Five Precipitating Events over France : On the Importance of the Initial Humidity Field. *Weather and Forecasting*, 17(6) :1236–1256, 2002. Cité pages 10, 30, 32 et 42.

G. Elgered, J.M. Johansson, B.O. Rönnäng, and J.L. Davis. Measuring regional atmospheric water vapor using the Swedish permanent GPS network. *Geophysical Research Letters*, 24(21) :2663–2666, 1997. Cité page 26.

T.R. Emardson and H.J.P. Derks. On the relation between the wet delay and the integrated precipitable water vapour in the European atmosphere. *Meteorological Applications*, 7(01) :61–68, 2000. Cité page 24.

T.R. Emardson, G. Elgered, and J.M. Johannson. Three months of continuous monitoring of atmospheric water vapor with a network of Global Positioning System receivers. *J. Geophys. Res.*, 103(D2) :1807–1820, 1998. Cité page 26.

C. Faccani, R. Ferretti, R. Pacione, T. Paolucci, F. Vespe, and L. Cucurull. Impact of a high density GPS network on the operational forecast. *Advances in Geosciences*, 2 :73–79, 2005. Cité page 41.

M. Falvey and J. Beavan. The Impact of GPS Precipitable Water Assimilation on Mesoscale Model Retrievals of Orographic Rainfall during SALPEX'96. *Mon. Wea. Rev.*, 130(12) :2874–2888, 2002. Cité page 37.

P. Fang, M. Bevis, Y. Bock, S. Gutman, and D. Wolfe. GPS meteorology : Reducing systematic errors in geodetic estimates for zenith delay. *Geophysical Research Letters*, 25(19) :3583–3586, 1998. Cité page 26.

C. Fischer. The variational computations inside ARPEGE/ALADIN : cycle CY30. 2005. Cité page 47.

G. Guerova, J.M. Bettems, E. Brockmann, and C. Matzler. Assimilation of the GPS-derived integrated water vapour (IWV) in the MeteoSwiss numerical weather prediction model—-a first experiment. *Physics and Chemistry of the Earth*, 29(2-3) : 177–186, 2004. Cité page 37.

Y.R. Guo, Y.H. Kuo, J. Dudhia, D. Parsons, and C. Rocken. Four-Dimensional Variational Data Assimilation of Heterogeneous Mesoscale Observations for a Strong Convective Case. *Mon. Wea. Rev.*, 128(3) :619–643, 2000. Cité page 38.

Y.R. Guo, H. Kusaka, DM Barker, Y.H. Kuo, and A. Crook. Impact of Ground-based GPS PW and MM5-3DVar Background Error Statistics on Forecast of a Convective Case. *SOLA*, 1(0) :73–76, 2005. Cité page 39.

S.I. Gutman, S.R. SAHM, S.G. BENJAMIN, B.E. SCHWARTZ, K.L. HOLUB, J.Q. STEWART, and T.L. SMITH. Rapid Retrieval and Assimilation of Ground Based GPS Precipitable Water Observations at the NOAA Forecast Systems Laboratory : Impact on Weather Forecasts. *J. Meteor. Soc. Japan*, 82(1B) :351–360, 2004. Cité pages 37 et 40.

J. Haase, A. Rius, E. Calais, I. Colomina, and F. Vespe. Meteorological Applications of Global Positioning System Integrated Column Water Vapor Measurements in the Western Mediterranean (MAGIC). *A Proposal for the European Commission DGXII/D*, 1997. Cité page 28.

J. Haase, E. Calais, J. Talaya, A. Rius, F. Vespe, R. Santangelo, X.Y. Huang, JM Davila, M. Ge, L. Cucurull, et al. The contributions of the MAGIC project to the COST 716 objectives of assessing the operational potential of ground-based GPS meteorology on an international scale. *Physics and Chemistry of the Earth, Part A*, 26(6-8) :433–437, 2001. Cité page 28.

J. Haase, M. Ge, H. Vedel, and E. Calais. Accuracy and Variability of GPS Tropospheric Delay Measurements of Water Vapor in the Western Mediterranean. *Journal of Applied Meteorology*, 42(11) :1547–1568, 2003. Cité page 26.

M. Hakam. Assimilation de données gps-sol pour la prévision des pluies intenses en régions méditerranéennes. stage d'approfondissement des eleves ingenieurs de l'ecole nationale de la meteorologie, Ecole Nationale de la Météorologie, 2006. Cité page 54.

J.E. Hoke and R.A. Anthes. The Initialization of Numerical Models by a Dynamic-Initialization Technique. *Mon. Wea. Rev.*, 104(12) :1551–1556, 1976. Cité page 33.

J. Holton et al. Dynamic Meteorology, 1992. Cité page 32.

E. Kaplan. *Understanding GPS*. Mobile communications series. Artech House INC., 685 Canton Street, Norwood MA 02062, 1996. Cité page 15.

K. Koizumi and Y. Sato. Impact of GPS and TMI Precipitable Water Data on Mesoscale Numerical Weather Prediction Model Forecasts. *J. Meteor. Soc. Japan*, 82(1B) :453–457, 2004. Cité page 39.

Y.H. Kuo, Y.R. Guo, and E.R. Westwater. Assimilation of Precipitable Water Measurements into a Mesoscale Numerical Model. *Mon. Wea. Rev.*, 121(4) :1215–1238, 1993. Cité pages 26 et 37.

Y.H. Kuo, X. Zou, and Y.R. Guo. Variational Assimilation of Precipitable Water Using a Nonhydrostatic Mesoscale Adjoint Model. Part I : Moisture Retrieval and Sensitivity Experiments. *Mon. Wea. Rev.*, 124(1) :122–147, 1996. Cité pages 32 et 37.

JP Lafore, J. Stein, N. Asencio, P. Bougeault, V. Ducrocq, J. Duron, C. Fisher, P. Héreil, P. Mascart, V. Masson, et al. Vila-Guerau de Arellano, J. : 1998,'The Meso-NH Atmospheric Simulation System. Part I : Adiabatic Formulation and Control Simulation'. *Ann. Geophys*, 16 :90–109, 1998. Cité page 54.

C. Lebeaupin Brossier, V. Ducrocq, and H. Giordani. Sensitivity of three Mediterranean heavy rain events to two different sea surface fluxes parameterizations in high-resolution numerical modeling. *J. Geophys. Res.*, 113(D21), 2008. Cité page 54.

P. Lopez. Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. *Quart. J. Roy. Meteor. Soc.*, 128 :229–257, 2002. Cité page 45.

J-F. Louis. A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Lay. Meteorol.*, 17(2) :187–202, 1979. Cité page 45.

S.R. Macpherson. Impact of ground-based GPS observations on the Canadian Regional Analysis and Forecast System. *11th Symposium on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, 2007. Cité pages 39 et 40.

A.J. Mannucci, C.O. Ao, G.A. Hajj, B.A. Iijima, M. de la Torre-Juarez, T.K. Meehan, and T. Schroder. Towards new scientific observations from GPS occultations : advances in retrieval methods. 2004. Cité page 20.

V. Masson. A physically-based scheme for the urban energy budget in atmospheric models. *Bound. Layer Meteor.*, 1994 :357–397, 2000. Cité page 55.

E.J. Mlawer, S.J. Taubman, P.D. Brown, M.J. Iacono, and S.A. CLough. Radiative transfer for inhomogeneous atmospheres : Rrtm, a validated correlated-k model for the longwave. *J. Geophys. Res.*, 102D :16663–16682, 1997. Cité pages 45 et 55.

T. Montmerle, F. Rabier, and C. Fischer. Respective impact of polar orbiting and geostationary satellite observations in the aladin/france nwp system. *Quart. J. Roy. Meteor. Soc.*, 133 :655–671, 2007. Cité page 48.

H. Nakamura, K. KOIZUMI, and N. MANNOJI. Data Assimilation of GPS Precipitable Water Vapor into the JMA Mesoscale Numerical Weather Prediction Model and its Impact on Rainfall Forecasts. *J. Meteor. Soc. Japan*, 82(1B) :441–452, 2004. Cité pages 37 et 39.

AE Niell. Global mapping functions for the atmosphere delay at radio wavelengths. *J. Geophys. Res.*, 101 :3227–3246, 1996. Cité page 23.

AE Niell, AJ Coster, FS Solheim, VB Mendes, PC Toor, RB Langley, and CA Upham. Comparison of Measurements of Atmospheric Wet Delay by Radiosonde, Water Vapor Radiometer, GPS, and VLBI. *Journal of Atmospheric and Oceanic Technology*, 18(6) :830–850, 2001. Cité page 26.

J. Noilhan and J.F. Mahfouf. The isba land surface parameterization scheme. *Global and Plan. Change*, 13 :145–159, 1996. Cité page 55.

J. Noilhan and S. Planton. A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.*, 117 :536–549, 1989. Cité page 45.

O. Nuissier, V. Ducrocq, D. Ricard, C. Lebeaupin, and S. Anquetin. A numerical study of three catastrophic precipitating events over southern France. I : Numerical framework and synoptic ingredients. *Q. J. Roy. Meteor. Soc.*, 134(630) :111, 2008. Cité page 54.

R. Pacione, C. Sciarretta, C. Faccani, R. Ferretti, and F. Vespe. GPS PW assimilation into MM5 with the nudging technique. *Physics and Chemistry of the Earth, Part A*, 26(6-8) :481–485, 2001. Cité page 37.

B.W. Parkinson and J.J. Spiker Jr. Global Positioning System : Theory and Applications Volume I, American Institute of Aeronautics and Astronautics. *Inc., Washington, DC*, 1996. Cité page 14.

D.F. Parrish and J.C. Derber. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.*, 120(8) :1747–1763, 1992a. Cité page 36.

D.F. Parrish and J.C. Derber. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Mon. Wea. Rev.*, 120(8) :1747–1763, 1992b. Cité page 34.

S.Q. Peng and X. Zou. Impact on Short-Range Precipitation Forecasts from Assimilation of Ground-Based GPS Zenith Total Delay and Rain Gauge Precipitation Observations. *J.Meteor.Soc.Japan*, 82(1B) :491–506, 2004. Cité page 41.

P. Poli, P. Moll, F. Rabier, G. Desroziers, B. Chapnik, L. Berre, SB Healy, E. Andersson, and FZ El Guelai. Forecast impact studies of zenith total delay data from European near real-time GPS stations in Météo France 4DVAR. *J. Geophys. Res*, 112 :D06114, 2007. Cité pages 39, 52 et 61.

F. Rabier, P. Courtier, J. Pailleux, O. Talagrand, and D. Thepaut, J.-N. andVasiljevic. Comparison of four-dimensional variational assimilation with simplified sequential assimilation. In UK ECMWF, Reading, editor, *Proceedings of the ECMWF on Variational Assimilation, With Special Emphasis on Three-Dimensional Aspects*, page 271–326, 1993. Cité page 38.

C. Rocken, R. Ware, T. Van Hove, F. Solheim, C. Alber, J. Johnson, M. Bevis, and S. Businger. Sensing Atmospheric Water Vapor With the Global Positioning System. *Geophysical Research Letters*, 20 :2631–2631, 1993. Cité page 26.

C. Rocken, T.V. Hove, J. Johnson, F. Solheim, R. Ware, M. Bevis, S. Chiswell, and S. Businger. GPS/STORM—GPS Sensing of Atmospheric Water Vapor for Meteorology. *Journal of Atmospheric and Oceanic Technology*, 12(3) :468–478, 1995. Cité page 26.

J. Saastamoinen. Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. In the use of artificial satellites for geodesy. *Geophys. Monogr. Ser. edited by Henriksen SW et al*, 15 :247–251, 1972. Cité pages 23 et 51.

EK Smith and S. Weintraub. The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies. *Proceedings of the IRE*, 41(8) :1035–1037, 1953. Cité page 61.

T.L. Smith, S.G. Benjamin, B.E. Schwartz, and S.I. Gutman. Using GPS-IPW in a 4-D data assimilation system. *Earth Planets and Space*, 52(11) :921–926, 2000. Cité pages 37 et 38.

T.L. Smith, S.G. Benjamin, S.I. Gutman, and S. Sahm. Short-Range Forecast Impact from Assimilation of GPS-IPW Observations into the Rapid Update Cycle. *Mon. Wea. Rev.*, 135(8) :2914–2930, 2007. Cité pages 37 et 38.

D.R. Stauffer and N.L. Seaman. Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part I : Experiments with Synoptic-Scale Data. *Mon. Wea. Rev.*, 118(6) :1250–1277, 1990. Cité page 33.

S.E. Ştefănescu, L. Berre, and M.B. Pereira. The Evolution of Dispersion Spectra and the Evaluation of Model Differences in an Ensemble Estimation of Error Statistics for a Limited-Area Analysis. *Mon. Wea. Rev.*, 134(11) :3456–3478, 2006. Cité page 48.

J. Stein, E. Richard, JP Lafore, JP Pinty, N. Asencio, and S. Cosma. High-Resolution Non-Hydrostatic Simulations of Flash-Flood Episodes with Grid-Nesting and Ice-Phase Parameterization. *Meteorology and Atmospheric Physics*, 72(2) :203–221, 2000. Cité page 55.

D Thayer. An improved equation for the radio refractive index of air. *Radio Science*, 9 :803–807, 1974. Cité page 20.

M. Tomassini, G. Gendt, G. Dick, M. Ramatschi, and C. Schraff. Monitoring of Integrated Water Vapour from ground-based GPS observations and their assimilation in a limited-area NWP model. *Physics and Chemistry of the Earth*, 27(4-5) : 341–346, 2002. Cité page 37.

H. Vedel and X.Y. Huang. An NWP impact study with ground based GPS data. *Proceedings of the International Workshop on GPS Meteorology—GPS Meteorology : Ground-Based and Space-Borne Applications*, 2003. Cité page 40.

H. Vedel and X.Y. Huang. Impact of Ground Based GPS Data on Numerical Weather Prediction. *J. Meteor. Soc.J apan*, 82(1B) :459–472, 2004a. Cité pages 39 et 40.

X. Yang, B.H. Sass, G. Elgered, J.M. Johansson, and T.R. Emardson. A Comparison of Precipitable Water Vapor Estimates by an NWP Simulation and GPS Observations. *Journal of Applied Meteorology*, 38(7) :941–956, 1999. Cité page 26.

Annexe A

Article "Mesoscale GPS Zenith Delay assimilation during a Mediterranean heavy precipitation event"

This article, sollicited for the Med-Storm Prize winners of the 9th Plinius conference, has been published in Advances in Geosciences, vol. 17, 71-77, 2008.

Adv. Geosci., 17, 71–77, 2008
www.adv-geosci.net/17/71/2008/

Advances in Geosciences

Mesoscale GPS Zenith Delay assimilation during a Mediterranean heavy precipitation event

X. Yan[1], V. Ducrocq[1], P. Poli[1], G. Jaubert[1], and A. Walpersdorf[2]

[1]GAME-CNRM, CNRS & Météo-France, 42 Avenue Coriolis, 31057 Toulouse Cedex 1, France
[2]LGIT, Université Joseph Fourier, Maison des Géosciences, BP 53, 38041 Grenoble Cedex 9, France

Received: 31 March 2008 – Revised: 11 June 2008 – Accepted: 30 June 2008 – Published: 29 July 2008

Abstract. The impact of assimilating Zenith Total delay (ZTD) observations from a mesoscale ground-based GPS network over Western Europe is evaluated for the heavy precipitation event of 5–9 September 2005 over Southern France. The ZTD assimilation is performed using a three dimensional variational data assimilation system at the 9.5-km horizontal resolution. Then using as initial conditions the 3DVAR analyses with and without assimilation of ZTD, we perform 2.4-km non-hydrostatic MESO-NH simulations. The results of the fine-scale simulations indicate that assimilation of ZTD help to improve the forecast of the tropospheric water vapour content and the quantitative precipitation forecast. We have also assessed through single observation experiments the influence of the formulation of the observation operator which is used to compute the model equivalent ZTD.

1 Introduction

Tropospheric water vapour is highly variable in space and time and, besides, is a key-ingredient for the success of fine-scale heavy rainfall forecast (Ducrocq et al., 2002). Lack of high-resolution water vapour observations is one of considerable sources of inaccuracy in model analyses that are used as initial conditions of the numerical weather prediction models (Kuo et al., 1996). The Zenith Total Delays (ZTD) deduced from GPS measurements are attractive for providing tropospheric water vapour data in a context of an increasing number of ground-based GPS receivers.

In the present study, we examine the impact of mesoscale ZTD data assimilation on the fine scale (2.4-km) forecast of a Mediterranean heavy precipitation event. The impact is evaluated more specifically for the 5 to 9 September 2005 period, during which several precipitating systems affected Southeastern France. The cumulative surface rainfall during the whole period was over 300 mm over a significant part of the region, reaching locally more than 500 mm.

Section 2 presents the ZTD assimilation methodology using the 9.5-km 3DVAR ALADIN data assimilation system. Section 3 evaluates the impact of the formulation of the observation operator based on a single observation assimilation experiment. Then, Sect. 4 discusses the impact of assimilating GPS data on the high-resolution forecast of the 5–9 September 2005 rainy event, based on non-hydrostatic 2.4-km MESO-NH simulations. The conclusions follow in Sect. 5.

2 Description of the ZTD assimilation

2.1 The 9.5-km 3DVAR assimilation system

The assimilation of ZTD is performed with the ALADIN 3DVAR data assimilation scheme, which has been running operationally at Météo-France since July 2005 (Fischer et al., 2005; Montmerle et al., 2007). To evaluate the impact of assimilating ZTD GPS data, two sets of six-hourly forecast-analysis cycles are run from 1st to 10th of September 2005 with the 3DVAR assimilation system. For the first set, called hereafter CTRL, the observations usually included in the operational 3DVAR ALADIN are assimilated. The latter include observations from radio-soundings, screen-level stations, wind profilers, buoys, ships and aircraft. The following satellite data are also assimilated: horizontal winds from atmospheric motion vectors (AMVs) and the Quickscatt scatterometers, Advanced Microwave Sounding Unit (AMSU)-A and -B radiances and Atmospheric Administration (NOAA)-15, -16, -17 and the AQUA satellites, High-resolution Infrared Sounder (HIRS) radiances from NOAA-17 and clear

Correspondence to: X. Yan
(xin.yan@cnrm.meteo.fr)

Published by Copernicus Publications on behalf of the European Geosciences Union.

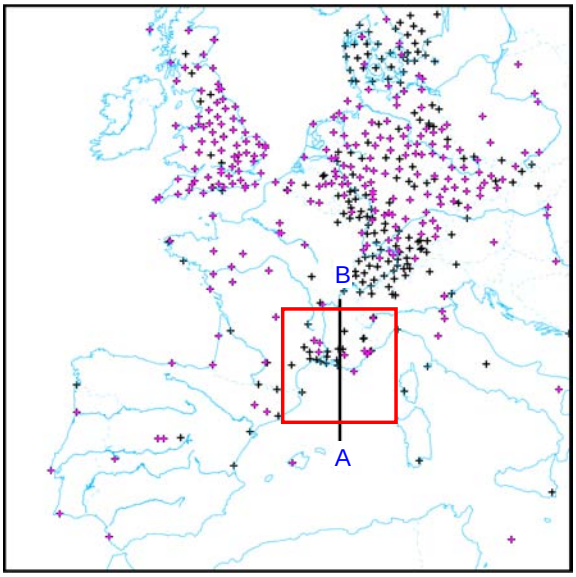

Fig. 1. Location in the 3DVAR ALADIN domain of the GPS stations of the E-GVAP and OHM-CV networks for September 2005. The pink-circled stations are those selected for data assimilation. The A–B line indicates the location of the cross-sections shown in Fig. 2. The 2.4-km MESO-NH domain is delineated by the red box.

SEVIRI radiances from the METEOSAT-8 satellite. The second set, called GPS, assimilates in addition the ZTD observations.

2.2 GPS ZTD data

In the GPS data assimilation cycle, we assimilate GPS data arising from the European GPS station network (E-GVAP, http://egvap.dmi.dk/) and from a 32-stations research network deployed within the Mediterranean Cévennes-Vivarais hydrometeorological Observatory (OHM-CV, Delrieu et al., 2005). For the studied period, data from more than 450 stations covering all the Western Europe were thus available (Fig. 1).

One GPS station can be processed by several analysis centers, leading to several ZTD time-series for a given station. For the assimilation, in order to have a uniform dataset per station, we select only one solution per station by applying the pre-processing developed by Poli et al. (2007) (called hereafter POLI07), slightly modified in order to take into account the higher resolution data assimilation system used in our study. The pre-processing selects pairs of station-center verifying that the first-guess departure follows a Gaussian distribution (The first-guess departure is defined as the difference between observed ZTD and model equivalent ZTD computed from the 6-hour ALADIN forecast which acts as the first-guess in the assimilation cycle). The difference between the station height and the model ground surface height should also be less than 150 m. Station-center pairs with large time availability, small standard deviation of the first-guess departure or the most Gaussian distribution are preferred. Within the assimilation window (±3 h around the analysis time), the GPS observation which is the closest to the analysis time is chosen. The pre-processing of GPS data retains at the end 262 stations out of 481 stations (Fig. 1).

To ensure that the observations meet the hypothesis of unbiased errors assumed in the assimilation scheme, the ZTD data are bias-corrected before the assimilation following the method used by POLI07. The bias is computed for each station-center pair based on a 15-day average of first-guess departure between 15 and 31 August 2005.

3 Sensitivity to the observation operator

To study the impact of how the model equivalent ZTD is calculated in the assimilation system, we perform analyses where the 3DVAR ALADIN system only assimilates one zenith delay observation using two different observation operators available, as their tangent-linear and adjoint codes, in the assimilation software. The first one has been used by POLI07 and the second one has been proposed by Brenot et al. (2006) (called hereafter BREN06) which had evaluated many expressions to calculate the model equivalent ZTD. POLI07 integrates from the bottom to the top of the model the following equation, using the total pressure P, the temperature T and the partial pressure of water vapor e of the atmospheric model column:

$$\mathrm{ZTD} = \int (k_1 \frac{P}{T} + k_3 \frac{e}{T^2}) dz \qquad (1)$$

with k_1=0.776 ·10^{-6} Pa^{-1}·K, k_3=3730 ·10^{-6} Pa^{-1}·K^2(Smith and Weintraub, 1953).

The equation for BREN06 is the following:

$$\mathrm{ZTD} = \int (k_1 \frac{P}{T_v} + k'_2 \frac{e}{T} + k'_3 \frac{e}{T^2}) dz + \Delta ZTD_{\mathrm{TOP}} \qquad (2)$$

with k_2=0.704·10^{-6} Pa^{-1}·K, k'_3= 3739·10^{-6} Pa^{-1}·K^2, and $k'_2 = k_2 - k_1 \frac{R_d}{R_v}$ (Bevis et al, 1994); T_v being the virtual temperature, R_d =287.0586 J/(kmol.K) and R_v =461.525 J/(kmol.K) the specific molar gas constants for dry air and water vapour, respectively.

BREN06 have evaluated different sets of (k_1, k_2, k_3) coefficients proposed in the literature for ZTD values computed from high-resolution (2.4 km) non-hydrostatic atmospheric simulated fields. The results showed that there is no significant differences in the evaluation of ZTD between most of the coefficient sets (with a mean bias of ZTD less than 2 mm), except for the set using the two-coefficient formula, which presents a mean ZTD bias reaching nearly 12 mm. A ZTD contribution above the model top level $\Delta ZTD_{\mathrm{TOP}}$ is

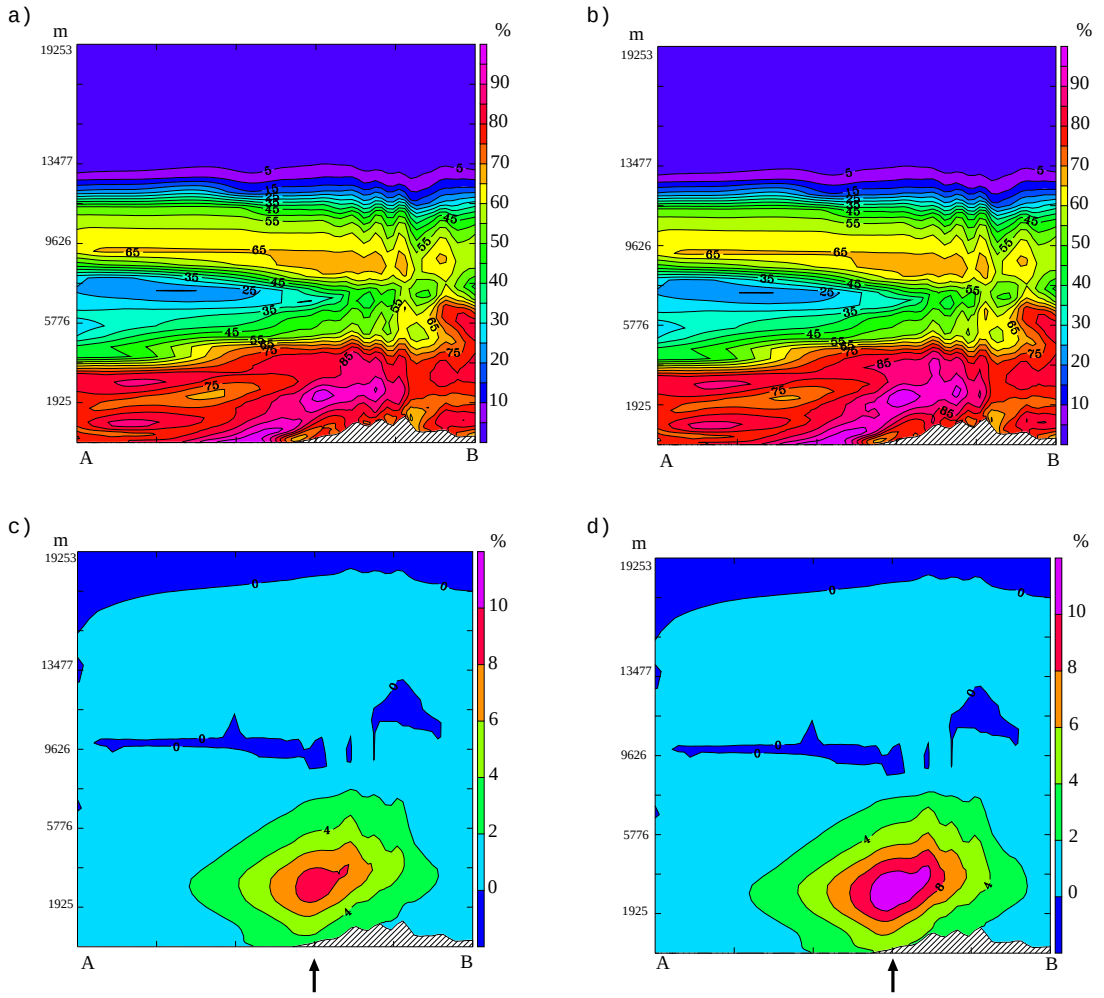

Fig. 2. Vertical cross section along the 670-km line A-B shown in Fig. 1 of analyses and analysis increments of relative humidity, obtained by assimilation of a single ZTD observation using the BREN06 (ac) or the POLI07 (bd) observation operator. Panels ab display the relative humidity from the 3DVAR analysis and panels cd the analysis increment of relative humidity (the analysis increment being defined as the difference between the analysis and the first-guess). The vertical arrows in panels cd indicate the location of the ZTD observation.

also added in BREN06. POLI07 did not consider this constant correction in their observation operator and left this to be handled by their bias correction scheme. ΔZTD_{top} in BREN06 is computed according to Saastamoinen (1972):

$$\Delta ZTD_{\text{TOP}} = 10^{-6} \frac{k_1 R_d P_{\text{TOP}}}{g_{\text{TOP}}} \quad (3)$$

with P_{TOP} and g_{TOP} being the pressure and the acceleration of gravity at the top of the model. POLI07 evaluated to 2.3 mm this constant contribution above the ALADIN/ARPEGE model top at 1 hPa.

In order to perform the integration of ZTD operator inside the model, POLI07 and BREN06 follow nearly the same implementation, except when the GPS station is located below the model terrain. In that case, extrapolation of pressure, temperature and humidity values from the model lowest level down to the GPS station altitude is needed. In that extrapolation procedure, the temperature is considered as constant in POLI07 while BREN06 assumes a constant temperature gradient.

Figure 2 shows results of the assimilation of a single observation of ZTD of value 2.4586 m; the model equivalent ZTD computed for the first guess with BREN06 and POLI07 observation operators departs from the observation by 14.34 mm and 18.6 mm, respectively. Note that 2.3 mm

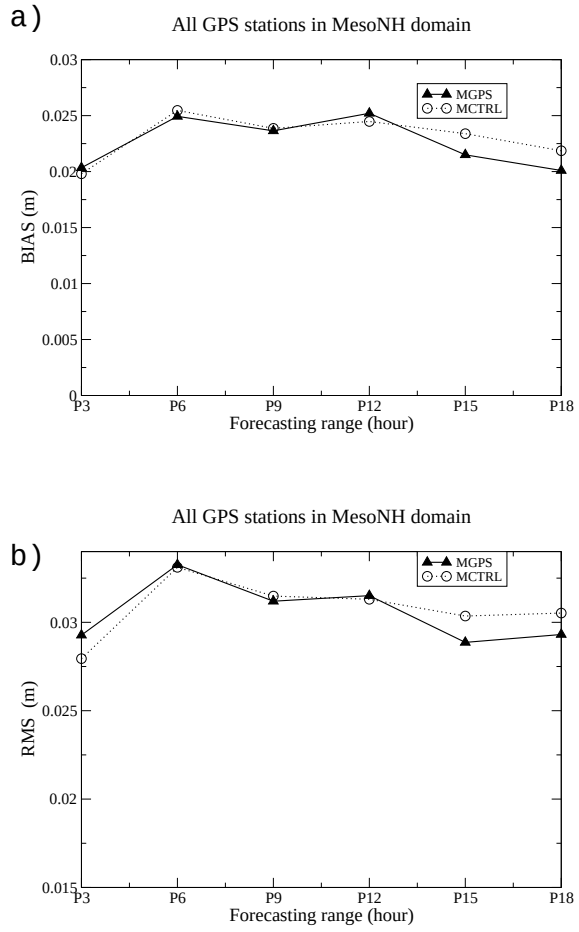

Fig. 3. BIAS and RMS for model equivalent ZTD as function of the forecast range for MGPS and MCTRL MESO-NH runs. Scores are computed against ZTD observations included in the MESO-NH domain (see red box in Fig. 1).The scores are computed gathering all the 3-hourly forecasts of all the MESO-NH 18-h duration runs from 12:00 UTC, 5 September 2005 to 18:00 UTC, 8 September 2005.

(out of the 18.6 mm difference) would normally be corrected by a bias correction (as in POLI07) but we neglect this effect here. After the minimization of the cost function in the assimilation, the model equivalent ZTD analysis differs from the observation of only 1.47 mm for BREN06 and of 1.9 mm for POLI07, showing a good behaviour of the assimilation with both operators. The main impact of assimilating ZTD data is seen on the humidity field. There is almost no impact on the pressure or temperature field. The horizontal influence range of the observation is in agreement with the horizontal correlation length of the background-error covariance matrix which is about 70 km in the middle troposphere for specific

Table 1. BIAS and RMS for ZTD simulated by GPS and CTRL MESO-NH experiments against observed ZTD. For the "all GPS stations", scores are computed gathering all the raw ZTD observations over the MESO-NH domain. For the "only assimilated GPS stations", scores are computed only for ZTD values used in the assimilation after removing of the station bias (see Sect. 2.1). The scores are computed gathering all the 3-hourly forecasts of all the MESO-NH 18-h duration runs from 12:00 UTC, 5 September 2005 to 18:00 UTC, 8 September 2005.

	all GPS Stations		only assimilated GPS Stations	
	MCTRL	MGPS	MCTRL	MGPS
BIAS	23.1 mm	22.6 mm	14 mm	13.4 mm
RMS	30.8 mm	30.6 mm	26.9 mm	26.2 mm

Table 2. BIAS and RMS for 12-h accumulated precipitation forecast from the MGPS and MCTRL MESO-NH experiments against observed precipitation. The scores are computing gathering all the 12-hourly forecasts (i.e. 0–12 h and 6–18 h forecasting ranges) of all the MESO-NH 00:00 UTC and 12:00 UTC runs covering the period from 12:00 UTC, 5 September 2005 to 18:00 UTC, 8 September 2005.

	MCTRL	MGPS
all GPS Stations		
BIAS	0.5 mm	0.3 mm
RMS	19.7 mm	20.2 mm

humidity. We can see in Fig. 2cd that the assimilation experiment using POLI07 yields an analysis increment of humidity in the low- to mid- troposphere more intense than in the assimilation experiment using the observation operator from BREN06. The larger analysis increments imply more corrections to the humidity field by the analysis. To summarize, in the present example, the use of different operators can lead to about 2% of difference in the final analysis increment of humidity, without however modifying the vertical distribution of the humidity analysis increment.

4 Impact on the fine scale forecasts

The impact of the ZTD assimilation on the convective scale forecast of the 5–9 September 2005 Mediterranean heavy precipitation episode has been assessed, using two-way grid-nesting MESO-NH domains at 9.5-km and 2.4-km, respectively. Ducrocq et al. (2002), Lebeaupin et al. (2006) and Nuissier et al. (2008) provided a comprehensive description of this MESO-NH model configuration, that was successfully used for simulating Mediterranean heavy precipitation events in these studies. The 3DVAR ALADIN analyses issued from the two assimilation cycles with and without ZTD data assimilation, as described in Sect. 2, are used as initial

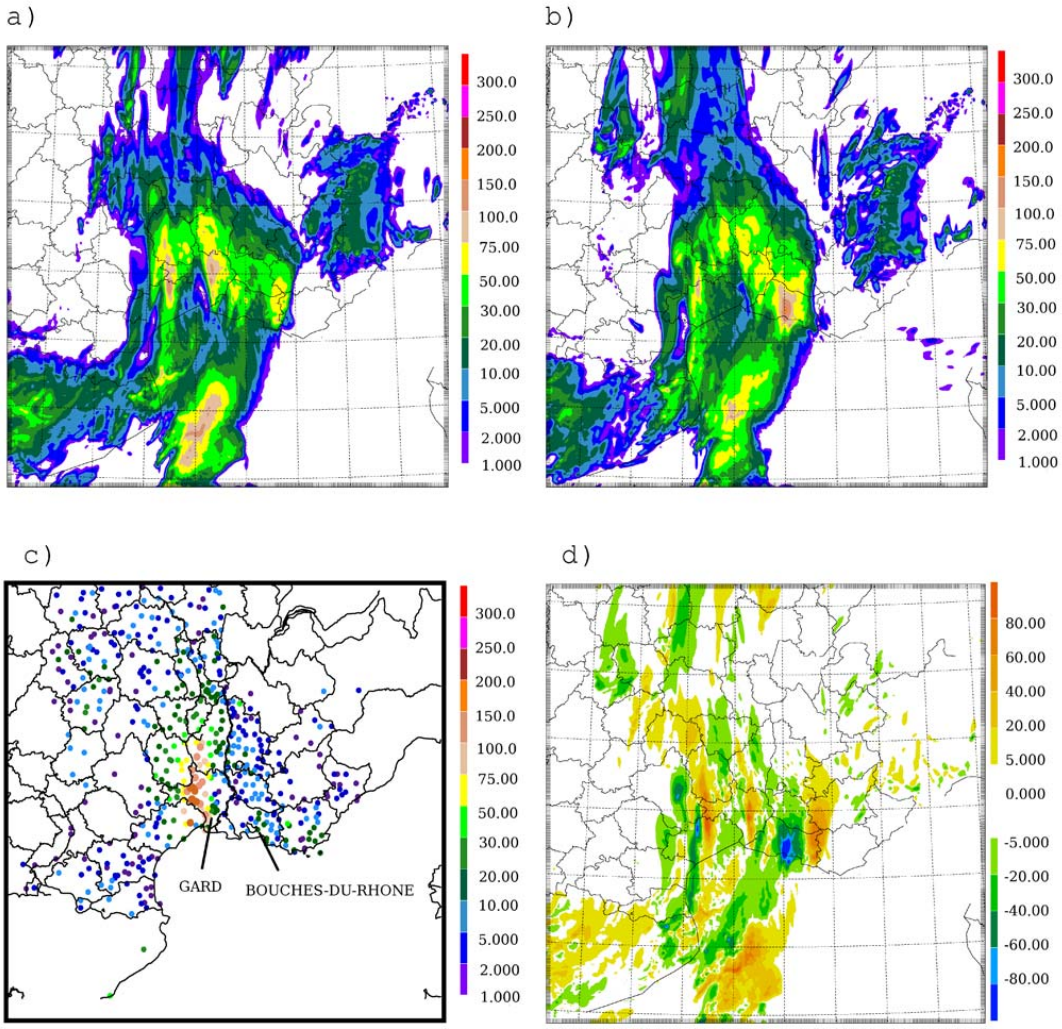

Fig. 4. 12-h accumulated precipitation (mm) from 00:00 UTC to 12:00 UTC, 6 September 2005 for: **(a)** MGPS MESO-NH run starting from the GPS ALADIN analysis at 00:00 UTC; **(b)** MCTRL MESO-NH run starting from the CTRL ALADIN analysis at 00:00 UTC; **(c)** raingauges; **(d)** differences between MGPS and MCTRL 12-h accumulated precipitation.

conditions to the MESO-NH simulations. Two 18-hour-duration MESO-NH runs were issued every day at 00:00 and 12:00 UTC from 12:00 UTC, 5 September to 00:00 UTC, 8 September 2005, covering the whole rainy period. The MESO-NH set using the 3DVAR ALADIN analyses with ZTD data assimilation is called MGPS hereafter, whereas the other one is called MCTRL. We focus here on results of the 2.4-km MESO-NH runs.

The model equivalent ZTD at the GPS stations have been computed for every 3-hourly forecasts from 3 to 18 h range issued from the two sets of MESO-NH runs. Figure 3 shows the bias and Root Mean Square error (RMS) computed against observed ZTD for all the GPS stations included in the 2.4-km MESO-NH domain (Fig. 1). Scores for MGPS runs show an improved forecast of ZTD for the longer ranges (15 and 18-h ranges) compared to MCTRL. BIAS and RMS are quite similar for the two MESO-NH sets for the very short-range (Fig.3). The BIAS and RMS computed over all the 3-hourly forecasts, whatever the forecasting ranges, lead to the same conclusion of slightly better results for the MGPS runs (Table 1, left columns). If the scores are computed only on the subset of GPS stations assimilated in the 3DVAR ALADIN system (Table 1, right columns), the MGPS runs are also found better. Note that the BIAS is weaker on that

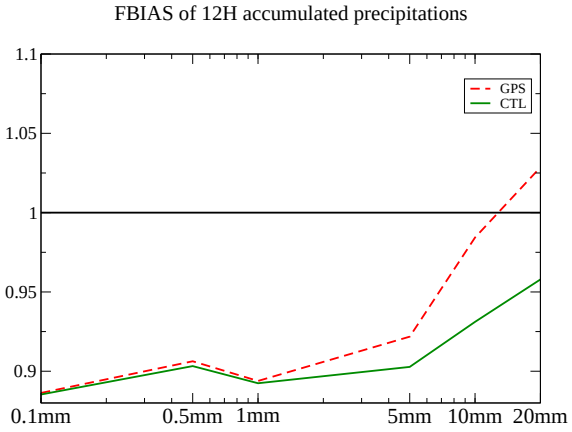

Fig. 5. FBIAS for 12-h accumulated precipitation forecast from the MGPS and MCTRL MESO-NH experiments against observed precipitation. FBIAS is computing gathering all the 12-hourly forecasts (i.e. 0–12 h and 6–18 h forecasting ranges) of all the MESO-NH 00:00 UTC and 12:00 UTC runs covering the period from 12:00 UTC, 5 September 2005 to 18:00 UTC, 8 September 2005.

subset than on the whole GPS data-set because the scores are computed on the biases-corrected ZTD values used for the 3DVAR assimilation. So, using as initial conditions an analysis in which ZTD data have been assimilated improves the forecast of zenith total delay and thus of tropospheric water vapour.

The impact of assimilating GPS on the high-resolution quantitative precipitation forecast is also found slightly positive. Figure 4 shows the 12-h accumulated precipitation results of MGPS and MCTRL runs issued from the corresponding GPS and CTRL Aladin analysis at 00:00 UTC, 6 September 2005. Compared with the raingauge data (Fig. 4c), we could see that MGPS run locates better the heavy precipitation region over the GARD department while MCTRL run misplaces the heavy precipitation area over the BOUCHES-DU-RHONE department. The differences reach more than 80 mm in some locations(Fig. 4d) between these two runs (MGPS-MCTRL). Table 2 shows the BIAS and RMS computed over every 12-h accumulated precipitation totals (i.e. 0 h–12 h and 6 h–18 h forecast ranges) for the MGPS and MCTRL runs. Scores are computed against raingauge observations in the 2.4-km MESO-NH domain, which counts about 220 raingauges. The MGPS set improves the BIAS whereas the RMS is slightly weaker for the MCTRL runs. Figure 5 shows the Frequency Bias (FBIAS, Bougeault (2003)) for precipitation events with observed 12-h accumulated rainfall above 0.1 mm, 0.5 mm, 1 mm, 5 mm, 10 mm and 20 mm. FBIAS indicates whether the experiment has a tendency to under-forecast (FBIAS$<$1) or to over-forecast (FBIAS$>$1) precipitation events, with a perfect score being 1. For all the thresholds, the MGPS experiment shows an improved skill compared to the MCTRL, especially for the higher precipitation events. FBIAS for the two experiments indicate a tendency to under-forecast the events for weak thresholds.

5 Conclusions

The impact of assimilating Zenith Total delay (ZTD) observations from a dense GPS network has been evaluated for the high-resolution forecast of the heavy precipitation event of 5–9 September 2005 over Mediterranean northwestern coasts. Using as initial conditions the analyses in which GPS ZTD data have been assimilated leads to a weak positive impact on the quantitative precipitation forecasts issued by the convective scale MESO-NH model. A better forecast of the model equivalent ZTD and thus of the tropospheric water vapour are also found.

The impact of the formulation of the observation operator has also been studied through a single observation experiment. Two slightly different observation operators showed that the assimilation of ZTD modifies mainly the low to mid-troposphere moisture. The main difference between the two operators lie in the amplitude of the relative humidity increments, only reflecting larger first-guess differences in one operator (some of these differences are explained by a constant model top contribution). Almost no difference is found for the vertical distribution of analysis increment of relative humidity in both cases.

A more comprehensive analysis of the 3DVAR ALADIN ZTD assimilation over the whole September 2005 month is the subject of a companion paper. Work is in progress to assimilate ZTD data directly at the convective scale using a 2.5-km data assimilation system and evaluate its impact on the quantitative precipitation forecast. A finer scale model is expected to reduce the height differences between the GPS station and the model orography, and therefore to improve the accuracy of the observation operator.

Acknowledgements. The authors would like to thank the E-GVAP project for their efforts in making European ground-based GPS data available in near real time. The authors are grateful to the constructive comments made by Romu Romero. The OHM-CV GPS network and this study benefited from support from the national research program LEFE, the Institut National des Sciences de l'Univers, and of the European Commission through the FP6/PREVIEW project.

Edited by: A. Mugnai
Reviewed by: R. Romero

References

Bevis, M., Businger, S., Chiswell, S., Herring, T. A., Anthes, R. A., Rocken, C., and Ware, R. H.: GPS Meteorology: Mapping Zenith Wet Delays onto Precipitable Water, J. Appl. Meteor., 33, 379–386,1994.

Bougeault, P.: The WGNE survey of verification methods for numerical prediction of weather elements and severe weather events, CAS/JSC WGNE Report, 18, WMO/TD-NO.1173, Appendix C, 1–11, 2003.

Brenot, H., Ducrocq, V., Walpersdorf, A., Champollion, C., and Caumont, O.: GPS Zenith Delay Sensitivity evaluated from High-Resolution NWP Simulations of the 8–9th September 2002 Flash-Flood over Southeastern France, J. Geophys. Res., 111, D15105, doi:10.1029/2004JD005726, 2006.

Delrieu G., Ducrocq, V., Gaume, E., Nicol, J., Payrastre, O., Yates, E., Kirstetter, P. E., Andrieu, H., Ayral, P.-A., Bouvier, C., Creutin, J.-D., Livet, M., Anquetin, S., Lang, M., Neppel, L., Obled, C., Parent-du-Chatelet, J., Saulnier, G.-M., Walpersdorf, A., and Wobrock, W.: The catastrophic flash-flood event of 8–9 September 2002 in the Gard region, France: a first case study for the Cévennes-Vivarais Mediterranean, Hydro-meteorological Observatory, J. Hydrometeorol., 6, 34–52, 2005.

Ducrocq, V., Ricard, D., Lafore, J. P., and Orain, F.: Storm-Scale Numerical Rainfall Prediction For Five Precipitating events over France: On the Importance of the initial humidity field, Weather Forecast, 17, 1236–1256, 2002.

Fischer, C., Montmerle, T., Berre, L., Auger, L., and Stefanescu, S. E.: An overview of the variational assimilation in the ALADIN/FRANCE NWP system, Q. J. Roy. Meteor. Soc., 131, 3477–3492, 2005.

Kuo, Y.-H., Zou, X., and Guo, Y.-R.: Variational assimilation of precipitable water using a non-hydrostatic mesoscale adjoint model. Part I: Moisture retrieval and sensitivity experiments, Mon. Weather. Rev., 124, 122–147, 1996.

Lebeaupin, C., Ducrocq, V., and Giordani, H.: Sensitivity of Mediterranean torrential rain events to the sea surface temperature based on high-resolution numerical forecsats, J. Geophys. Res., 111, D12110, doi:10.1029/2005JD006541, 2006.

Montmerle, T., Rabier, F., and Fischer, C.: Respective impact of polar orbiting and geostationary satellite observations in the Aladin/France NWP system, Q. J. Roy. Meteor. Soc., 133, 655–671, 2007.

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

Poli, P., Moll, P., Rabier, F., Desroziers, G., Chapnik, B., Berre, L., Healy, S. B., Andersson, E., and EI Guelai, F.-Z.: Forecast impact studies of zenith total delay data from European near real-time GPS stations in Meteo France 4DVAR, J. Geophys. Res., 112, D06114, doi:10.1029/2006JD007430, 2007.

Saastamoinen, J.: Introduction to practical computation of astronomical refraction, Bull. Geod., 106, 389–397, 1972.

Smith, E. and Weintraub, S.: The constants in the equation for atmospheric refractive index at radio frequencies, Proc. IRE, 1035–1037, 1953.

Glossary

AIUB	Astronomisches Institut der Universität Bern
ALADIN	Aire Limitée Adaptation dynamique Développement InterNational
AMV	Atmospheric Motion Vector
AROME	Applications de la Recherche à l'Opérationnel à Méso-Echelle
ARPEGE	Action de Recherche Petite Echelle et Grande Echelle
ATOVS	Advanced TIROS Operational Vertical Sounder
BKG	Bundesamt für Kartographie und Geodesie
BUFR	Binary Universal Form for the Representation of meteorological data
CAPE	Convective Available Pontential Energy
CNRM	Centre National de Recherches Météorologiques, Toulouse
CNRS	Centre National de la Recherche Scientifique
COST 716	Coopération europeénne dans les champs Scientifiques et Techniques de recherche, action 716
DLR	Deutches zentrum für Luft und Raumfahrt
DMI	Danish Meteorological Institute
ECMWF	the European Center for Medium range Weather Forecast
E-GVAP	Eumetnet - Global Water Vapour Project
FAAM	Facility for Airborne Atmospheric Measurement
GIPSY	GPS-Inferred Positioning SYstem
GPS	Global Positionning System
GTS	Greenwich Time Signal
HYMEX	Hydrological Cycle in the Mediterranean Experiment
HIRLAM	HIgh-Resolution Limited Area Modeling
IGN	Institut de Géographie National
IGS	International GPS Service for Geodynamics
IOPS	Intensive Observing Periods
ISBA	Interaction Soil Biosphere Atmosphere
ITRF	IERS Terrestrial Reference Frame
IWV	Integrated Water Vapor
JPL	Jet Propulsion Laboratory
LA	Laboratoire d'Aérologie, Toulouse
LGIT	Laboratoire de Geophysique Interne et Tectonophysique, Grenoble

LTHE	Laboratoire d'étude des Transferts en Hydrologie et Environnements, Grenoble
MCS	Mesoscale Convective System
MIT	Massachusetts Intitute of Technology
NCAR	National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
OHM-CV	Observatoire Hydrométéorologique Cévennes-Vivarais
PPP	Precise Point Positioning
QPF	Quantitative Precipitation Forecast
REGAL	REseau GPS permanent dans les ALpes occidentales
RGP	Réseau GPS Permanent
RINEX	Receiver INdependent EXchange format
RRTM	Rapid Radiative Transfer Model
RUC	Rapid Updated Cycle
STD	Slant Total Delay
TEB	Town Energy Balance
TKE	Turbulent Kinetic Energy
TOUGH	Targeting Optimal Use of GPS Humidity observations in meteorology
VLBI	Very Long Baseline Interferometry
WGS	World Geodetic System
WWRP	World Weather Research Program
ZHD	Zenith Hydrostatic Delay
ZTD	Zenith Total Delay
ZWD	Zenith Wet Delay

Assimilation of GPS data for the deep convection forecasts

Xin YAN

ABSTRACT :

The aim of this thesis work is to exploit the potential of ground based GPS observations for mesoscale data assimilation and high-resolution numerical weather prediction. Our main interest lies in the investigation of the impact of GPS observation assimilation in improving the forecast of convective scale weather phenomena such as convective storms. The data assimilation systems we have used are the Météo-France 3DVAR/ALADIN and 3DVAR/AROME systems. Two case studies were conducted with large numbers of GPS observations being assimilated. The cases selected are both convective rainfall events. For the first case (5-9 September 2005) with heavy precipitation over the French Mediterranean regions, the system we used is the 9.5km 3DVAR/ALADIN assimilation system. Analyses produced by one month of assimilation cycle experiments are used later as the initial and boundary conditions for starting the simulation of the 2.4km high resolution research model MesoNH. For the second case study (18-20 July 2007), high density GPS observations obtained from the COPS observations field campaign together with GPS observations coming from the European operational E-GVAP network are assimilated directly into the 3DVAR/AROME assimilation system with 2.5km horizontal resolution. Special attentions have been paid on the selection and pre-treatment of the GPS observations before they enter in the assimilation system. Such procedure can be viewed as a quality control for the observations. For both cases, results of twin experiments with and without assimilating GPS observations have suggested a positive impact on the prediction of heavy precipitation, the impact being more significant on the second case.

KEYWORDS : GPS ZTD Assimilation

AUTEUR : **Xin YAN**

TITRE : **Assimilation de données GPS pour la prévision de la convection profonde**

DIRECTEUR DE THESE : **Véronique DUCROCQ**

LIEU ET DATE DE SOUTENANCE : Centre National de Recherches Météorologiques, Météo-France, 10 Mars, 2009

RESUME :

Ce travail de thèse visait à exploiter le potentiel des observations GPS sol pour l'assimilation de données à mesoéchelle et la prévision numérique du temps à haute résolution. Nous avons examiné l'impact de l'assimilation des données GPS sur la prévision à l'échelle convective des systèmes orageux. Les systèmes d'assimilation utilisés sont les systèmes d'assimilation 3DVAR des modèles ALADIN et AROME de Météo-France. Deux cas d'étude ont été traités, avec pour chacun des cas un nombre important de données GPS assimilées. Les résultats des expériences d'assimilation et de prévision avec ou sans assimilation de données GPS montrent un impact positif sur la prévision des précipitations intenses de l'assimilation des données GPS ; l'impact est plus significatif sur le second cas d'étude.

MOTS CLES : GPS ZTD Assimilation

DISCIPLINE : Physique de l'atmosphère

INTITULE ET ADRESSE DU LABORATOIRE : Centre National de Recherches Météorologiques, Météo-France, 42 Avenue Gaspard Coriolis, 31057, TOULOUSE Cedex 1, France