Tellus
SERIES A
DYNAMIC
METEOROLOGY
AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2011), 63A, 263–282
Printed in Singapore. All rights reserved

Operational assimilation of GPS slant path delay measurements into the MM5 4DVAR system

By HANS-STEFAN BAUER[1,*], VOLKER WULFMEYER[1], THOMAS SCHWITALLA[1], FLORIAN ZUS[2] and MATTHIAS GRZESCHIK[3], [1]*Institute of Physics and Meteorology, University of Hohenheim, Garbenstrasse 30, 70599 Stuttgart, Germany;* [2]*Helmholtz-Center Potsdam, German Research Center for Geosciences, Telegrafenberg A17, 14473 Potsdam, Germany;* [3]*WESS - Water & Earth System Sciences, c/o University of Tübingen, Center for Applied Geosciences, Keplerstrasse 17, 72076 Tübingen, Germany*

(Manuscript received 8 January 2010; in final form 27 September 2010)

ABSTRACT

A forward operator for Global Positioning System (GPS) slant total delay (STD) data and its adjoint were implemented into the Mesoscale Model version 5 (MM5) 4DVAR system to investigate its impact on quantitative precipitation forecasting (QPF). An operational forecast system was set up providing two forecasts per day, one driven by ECMWF forecasts only and the other additionally by 4DVAR of GPS STD data.

The investigation of statistics for August 2007 demonstrated a positive impact on the representation of the water-vapour field and the diurnal cycle of precipitation in southwest Germany. The spread of observation-minus-model departures was strongly reduced during the first 6 h of the simulations. As compared to the control simulation, the averaged diurnal cycle of precipitation in the 4DVAR integration was closer to observations in spite of the limitations of the 4DVAR system. Especially promising is the almost complete removal of the spin-up at the beginning of the simulation.

This study not only demonstrates the potential of water vapour data assimilation for nowcasting and short-range QPF. It also suggests that improved GPS retrievals and extended networks are important to further improve the forecast performance and demonstrates the potential to apply observation operators for model verification.

1. Introduction

Impact studies with additional or new observations and state-of-the-art data assimilation methods are an important contribution for improving numerical weather prediction (Xiao and Sun, 2007; Richard et al., 2007) and climate models (Bauer and Wulfmeyer, 2009). The impact depends on the type of observation, its coverage and data quality, the assimilation system and the quality of model physics. Although forecasts of atmospheric variables such as temperature and wind have been improved in recent years, large deficiencies remain in quantitative precipitation forecasting (QPF) (e.g. Ebert et al., 2003; Rossa et al., 2008). Improving QPF is a challenging task as a complex process chain from initiation of convection to the organization, development and decay of clouds and precipitation needs to be simulated as correctly as possible.

Successful QPF has to be performed with a resolution that accurately represents orography and land use. Currently, this is only possible with limited-area models (LAMs) nested into global models. In this configuration, the predictive skill of LAMs becomes dependent on: (1) the performance of global models providing initial and boundary conditions, (2) LAM numerics, (3) LAM physics and its consistency with the physics of the driving global model, and (4) the predictability of the weather situation (e.g. Gleckler et al., 2008; Skamarock and Klemp, 2007; Küll and Bott, 2009; Hohenegger and Schär, 2007).

Current observations for the initialization of LAMs mainly stem from the global radiosonde network, aircraft measurements, satellite and surface data whose spatial coverage is too sparse to accurately initialize mesoscale models (WMO, 2009). Information concerning dynamics and humidity are lacking although they are closely related to the initiation of convection and the formation of clouds and precipitation (Weckwerth and Parsons, 2006; Dierer et al., 2009). Additional high-quality data sets in combination with advanced data assimilation systems on the mesoscale have the potential to reduce remaining uncertainties and imbalances between the mesoscale model and the driving analysis to improve the forecast quality from nowcasting to medium range. The operation of this combination is particularly challenging, as balance conditions applicable on the global scale, for example, geostrophy, become invalid. Furthermore, Gaussian error statistics of observations and model background

*Corresponding author.
e-mail: hans-stefan.bauer@uni-hohenheim.de
DOI: 10.1111/j.1600-0870.2010.00489.x

as well as linearity of model and observation operators are questionable.

A variety of data assimilation methods are applied operationally such as nudging (Stauffer and Seaman, 1994; Schraff and Hess, 2003), three-dimensional and four-dimensional variational analyses (3DVAR and 4DVAR) (Fischer et al., 2006; Rawlins et al., 2007; Saito et al., 2007) and ensemble Kalman filtering (EnKF) (e.g. Houtekamer et al., 2005; Houtekamer et al., 2009). From a theoretical point of view, 3DVAR, 4DVAR or EnKF-based techniques are superior to nudging even on the mesoscale, as they consider the error characteristics of model and observations more realistically and permit the incorporation of various, even indirect types of observations.

The combination of sophisticated data assimilation techniques with the application of new data sets demonstrated strong potential for improving QPF. Xue and Martin (2006a, b) applied a 1-km version of the Advanced Regional Prediction System with assimilated Mesonet surface data for a convective case during the International Water Vapor Project (IHOP_2002; Weckwerth et al., 2004) and derived a detailed conceptual model for convective initiation along the dryline. Wulfmeyer et al. (2006) assimilated high-resolution water vapour Differential Absorption Lidar (DIAL) data from the NASA LASE[1] system for one intense observation period (IOP) during IHOP_2002 into the MM5 4DVAR system (Ruggiero et al., 2002). They were able to improve the representation of the humidity structure in the model leading to an improved representation of the subsequent development of convection. Kawabata et al. (2007) applied a convection permitting model in combination with 4DVAR to simulate a severe precipitation event. They assimilated radar radial velocity, Global Positioning System (GPS) precipitable water, 2 m temperature as well as surface wind and were able to analyze the key processes leading to severe precipitation. Grzeschik et al. (2008) assimilated Raman lidar water vapor data from a network of three ground-based systems over Eastern Germany. Comparison with GPS stations in the region confirmed a significant improvement of the water vapor field in the assimilation window. Unfortunately, a detailed comparison of different data assimilation techniques studying the impact of exactly the same observations is lacking to date and most of these studies were based on single cases, which do not allow to draw general conclusions.

The application of GPS ground-based networks to derive integrated water vapour (IWV) has widely been demonstrated in recent years (e.g. Bengtsson et al., 2003; Hagemann et al., 2003). The combination of the data of several satellites over a longer time period results in the zenith total delay (ZTD). It is then related to the IWV content above the GPS station. To date, most of the ground-based GPS data assimilation experiments applied ZTD or retrieved IWV estimates. It was found that the assimilation of such data has a weak, but positive impact on Numerical Weather Prediction (NWP) of humidity and precipitation (e.g. Vedel and Huang, 2004; Faccani et al., 2005; Boniface et al., 2009).

Applying only the projection to ZTD, key information about the 3-D distribution of water vapour is lost. Therefore, new developments consider the direct assimilation of delay values to single satellites, so-called slant total delays (STDs) (MacDonald et al., 2002; Ha et al., 2003; Zus et al., 2008).

Ha et al. (2003) performed a series of observing system simulation experiments to assess the impact of slant wet delay (SWD), the part of STD which depends on the water vapour content along the ray path, on the short-range prediction of a squall line. SWD data from a hypothetical network of ground-based GPS receivers were assimilated into an earlier single-CPU version of the MM5 4DVAR system with a grid resolution of 27 km. They showed that the assimilation of SWD produced a considerably improved forecast of the squall line, both in terms of rainfall prediction and mesoscale frontal structure, as compared to forecasts without the assimilation of GPS data and forecasts where only IWV was assimilated. This improvement was mainly caused by the more accurately retrieved temperature and moisture structure in the convectively unstable region.

Järvinen et al. (2007) analyzed the assimilation of hypothetical and real STD observations in the framework of the High-Resolution Limited Area Model (HIRLAM) 3DVAR system with a grid resolution of 9 km. Their experiments with hypothetical data showed that the assimilation of STD observations resulted in reasonable specific humidity analysis increments and captures fine-scaled humidity structures in the hypothetical observations. Real STD observations were assimilated in an arbitrary single case and the resulting specific humidity analysis increments were found to have a similar structure to those obtained with a comparable radiosonde or ZTD observing network.

This study takes advantages of the infrastructure provided by the World Weather Research Program (WWRP) research and development project COPS[2] (Wulfmeyer et al., 2008) and the WWRP forecast demonstration project D-PHASE[3] (Rotach et al., 2009). It goes several steps beyond previous studies. An operational 4DVAR data assimilation system based on the MM5 model was set up for central Europe during the whole D-PHASE Operation Period to investigate statistically whether water-vapour information from a ground-based GPS network improves the representation of the 3-D water-vapour field and short- to medium-range QPF. Furthermore, it is studied what are the main elements of the forecast system limiting the spatio-temporal impact of GPS data assimilation on the predictive skill of the model.

[1] LASE: Lidar Atmospheric Sensing Experiment

[2] COPS: Convective and Orographically-induced Precipitation Study (http://www.cops2007.de

[3] D-PHASE: Demonstration of Pobabilistic Hydrological and Atmospheric Simulation of Flooding Events in the Alpine region (http://www.map.meteoswiss.ch/map-doc/dphase/dphase_info.htm)

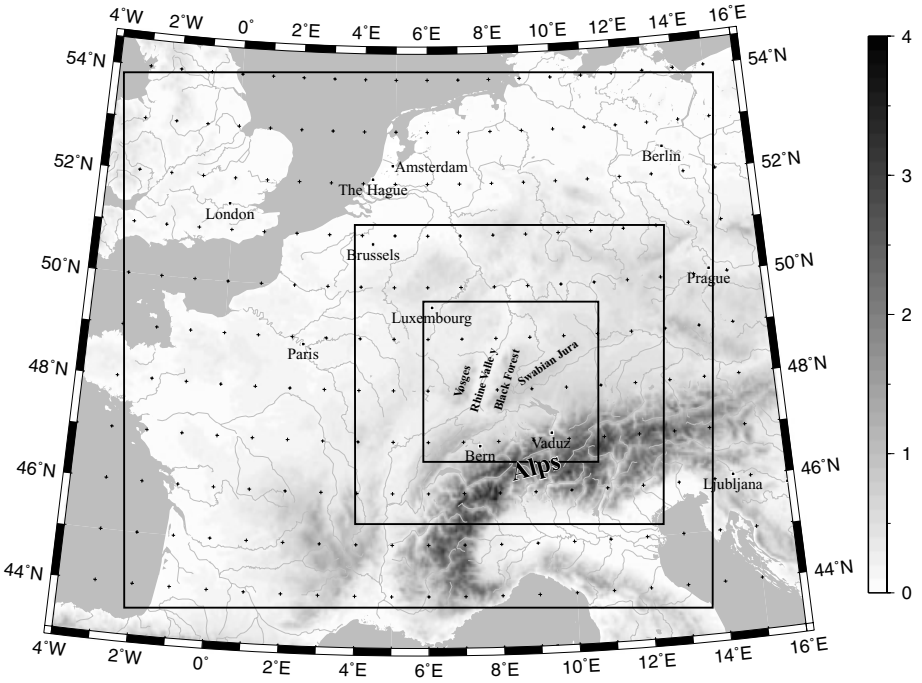

Fig. 1. Domain configuration of the operational system. The frames show the locations of the three domains. From large to small the domains were 64 × 70 grid points with 18 km horizontal resolution, 106 × 109 grid points with 6 km horizontal resolution, and 169 x 184 grid points with 2 km horizontal resolution. Assimilation was only performed in the outermost domain.

Particularly, we are addressing the following scientific questions:

- How good was the representation of the water vapour field in the driving European Centre for Medium Range Weather Forecasting (ECMWF) forecast?
- How large are the analysis increments with respect to the water vapour field, dynamics and temperature and do they result in an improvement of the initial water vapour field?
- What is the impact on the prediction of the water vapour field and on QPF in comparison to observations and other D-PHASE models?
- What are the processes limiting the spatio-temporal impact of the assimilation on the water vapour field?

The paper is structured as follows. In Section 2, the MM5 and the assimilation system are briefly described, GPS meteorology is introduced, a more detailed description of the GPS data and the derivation of the necessary forward operator is given. Section 3 describes the set-up of the operational system. In Section 4, the IMPACT and CONTROL runs are compared for August 2007. Thereby, a brief description of the observations used for validation is followed by a statistical investigation of the performance of 4DVAR of GPS STD data. In Section 5, possible causes that may deteriorate the performance of the assimilation system on different time scales and forecast lead times are discussed. Finally, Section 6 briefly summarizes the main results, conclusions and future plans are given.

2. MM5 4DVAR and assimilated observations

2.1. MM5

For the study, we selected the 5th generation Penn State/NCAR mesoscale model MM5 (Grell et al., 1995). It applies a terrain-following sigma vertical coordinate system with horizontal Arakawa-B grid staggering and has widely been used for applications ranging from regional climate simulations to high-resolution process studies on the km-scale (e.g. Das et al., 2003; Wu et al., 2005; Wulfmeyer et al., 2006; Grzeschik et al., 2008).

For the forecast system, a 3-domain, 2-way nested configuration with horizontal resolutions of 18, 6 and 2 km was configured with domain sizes of 64 × 70, 106 × 109 and 169 × 184 grid boxes (see Fig. 1). In the vertical, 36 levels up to the model top at 100 hPa were defined. For the meteorological initialization, the operational forecast of the ECMWF, providing data every 6 h, was selected. We replaced the default land use description from the US geological survey by the more realistic EU CORINE data set (Bossard et al., 2000). The soil characteristics are initialized from the global 5 min resolution United Nations FAO data set (FAO, 1995) converted to 17 soil categories.

MM5 supports several physical parameterization schemes for cloud microphysics, convection, radiation and for the boundary layer. For the model simulations from the two different initial states, the schemes summarized in the right column of Table 1 were selected, following the suggestions of a high-resolution

Table 1. Overview of the model set-up for the assimilation run and the model integrations

Physics	Assimilation	Forecasts
LW radiation	simple radiative cooling	RRTM (Mlawer et al., 1997)
SW radiation	simple radiative cooling	Dudhia (Dudhia, 1989)
Convection	Anthes-Kuo (Anthes, 1977)	Kain Fritsch 2 (Kain, 2004)
Microphysics	not simulated	Reisner2 (Reisner et al., 1998)
Boundary layer	MRF (Hong and Pan, 1996)	MRF (Hong and Pan, 1996)
Land surface	no soil model, fixed ground temperature	Dudhia 5-layer soil model

Note: Both simulations were operated with 18 km horizontal resolution.

modelling study by Schwitalla et al. (2008). They tested the performance of different combinations of parameterizations in MM5 for several convective cases from the year 2005 in south-western Germany applying horizontal resolutions down to 1 km. During the assimilation, a simplified set of parameterizations was chosen for which the required adjoint versions are available (see the middle column of Table 1).

2.2. Assimilation system

At the moment, 4DVAR, improved versions of the EnKF or hybrid combinations of both are the most promising candidates for future data assimilation on the global as well as on the mesoscale, since they allow to ingest high-resolution observations in their raw form at the exact time of observation (Kalnay et al., 2007a,b; Gustafsson, 2007; Wang et al., 2008a,b; Huang et al., 2009). Due to the strong flow dependence of the water vapour field, we selected the 4DVAR, provided by MM5 (Ruggiero et al., 2002). An EnKF system was not available.

Apart from the observations, a model simulation is needed to fit the observations in a least-square sense. Expecting Gaussian statistics, the minimum of the cost function J, measuring the distance between the model solution and observations, is determined from Bayes statistics:

$$J[\mathbf{x_0}] = (\mathbf{x}_0 - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_b) + \sum_{n=0}^{m} (H[\mathbf{x}_n] - \mathbf{y}_n)^T \mathbf{R}^{-1} (H[\mathbf{x}_n] - \mathbf{y}_n) \quad (1)$$

where $\mathbf{x}_0$ denotes the analysis vector, $\mathbf{x}_b$ is the forecast background vector and $\mathbf{y}_n$ denotes the observation vector at time n. The model forecast vector $\mathbf{x}_n$ at time n is predicted by the (non-linear) forecast model M, i.e. $\mathbf{x}_n = M \ldots M[\mathbf{x_0}]$. H represents the observation operator which transforms model variables into observation space. $\mathbf{R}$ denotes the observation error covariance matrix containing information about the errors of the observations and their correlations. $\mathbf{B}$ denotes the background error covariance matrix containing information about the errors in the background field and their correlations. The analysis vector of the MM5 4DVAR system consists of the horizontal wind components, the temperature, the water vapour mixing ratio, the pressure perturbation and the vertical velocity. Cloud water, cloud ice, surface variables and model boundaries are not included.

The MM5 4DVAR system only offers diagonal background error covariance matrices $\mathbf{B}$. This approximation has proven to work well in several studies conducted with the system (Zou et al., 1995; Xiao et al., 2000; Wulfmeyer et al., 2006; Grzeschik et al., 2008). This can be explained by the ability of the 4DVAR system to self-generate physically consistent structure functions during model integration due to the inclusion of the forecast model into the assimilation process.

For each control variable, the diagonal elements of $\mathbf{B}$ are specified by constructing the differences between a 15 min forecast and the initial values at each grid point. At each vertical level, the maximum value of the difference is retrieved and assigned to all grid points on that level resulting in a vertical profile of forecast errors valid for all model grid columns. The forecast errors are then squared to produce the diagonal elements of $\mathbf{B}$. The near-surface variances were in the order of 1–2 $\mathrm{g\,kg^{-1}}$ for water vapour mixing ratio and 1–2 K for temperature. Since these calculations were repeated for each assimilation, the values can differ considerably from day to day.

2.3. GPS data

GPS meteorology provides reliable, operational information of the water-vapour distribution under all weather conditions (Bender and Raabe, 2007; Bender et al., 2008, 2009). Its temporal resolution is currently 15 min for ZTD and 2.5 min for STDs. The spatial resolution is determined by the number of Global Navigation Satellite System satellites and ground stations that is expected to increase considerably in the near future.

In order to better coordinate the efforts in improving the observation of the water vapour field with the GPS methodology, several GPS observing networks were set up worldwide in recent years. In Europe, the EUMETNET GPS water vapour programme (E-GVAP)[4] was set up in 2005 to provide its partners with ZTD estimates for operational usage and to ensure that ground-based GPS near-real time data gradually meet the

[4] http://egvap.dmi.dk

requirements of operational NWP regarding data quality, homogeneity, stability and availability.

The Helmholz Center Potsdam, German Research Center for Geosciences (GFZ) is one of the data processing centres in Europe. During COPS, the GPS network, analysed in near-real time at the GFZ, consisted of about 200 receivers in Germany and neighbouring countries. The data analysis is performed on an hourly basis with a delay of about 30 min using the GFZ Earth Parameters and Orbit determination System (EPOS) software (Gendt et al., 1999) which is based on a least-square adjustment of undifferenced phase measurements.

The ZTD is a combined quantity based on all data processed from each station during 1 h. The available GPS observations are analysed in order to find an optimal parameter estimation for the ZTD data in 15 min steps. During 2006, the GFZ started to analyse individual STDs of a given station making use of the full temporal resolution of 2.5 min with an elevation cut-off angle of 7°. This was originally done for about 130 GPS stations leading to about 1 million STDs per day. During COPS, the GFZ was the only data provider in Europe which was able to provide STD data in near-real time. Therefore, we set our preference to the GFZ network. Data used in this study are based on the output of an operational version of EPOS during COPS in summer 2007 and is retrieved from 170 to 200 stations. To improve the data coverage in upwind regions west and southwest of the COPS region, the GFZ STD data was complemented by ZTD data from the French Satellite Geodetic Network (SGN).

STD data was validated with pointed water vapour radiometers (WVR) and showed promising results with a RMS of 1.3 mm in the IWV (Bender et al., 2008, 2009). The most promising approach to use STD data is its direct assimilation with accurate error estimates. This provides the most complete information content for the construction of the optimal water-vapour field according to the background, observational errors and model physics. To derive STD data, the smallest number of assumptions is necessary. For ZTD, the projection of the slants from different satellites to the zenith direction above the station location requires additional assumptions. Even more assumptions as well as the surface pressure and the column-mean temperature are necessary to invert ZTD to IWV.

Although several studies were carried out to determine the error of STD data by means of intercomparisons with independent observations like WVR, it remained difficult to estimate the real error of GPS STDs (Bender et al., 2008). An initial comparison of GPS stations in Germany with the COSMO-EU model of the German Weather Service (DWD) gives a standard deviation of about 1 mm in the IWV (Dick et al., 2001) which converts to a standard deviation of about 6 mm in the ZTD. In the absence of better error estimates and on the basis of data provided by GFZ, we estimate the error of a STD with an elevation angle of ε_r at the ground based GPS receiver to be of the order of a few millimeters. Since less error-prone assumptions are applied during the derivation of STD data, our observation error σ_r is estimated to be

$$\sigma_r(\varepsilon_r) = \frac{d}{sin(\varepsilon_r)}, \tag{2}$$

where $d = 1$ mm.

Although it is likely that STD measurements are correlated, we assume for simplicity neither any correlation of STD data of the same receiver nor any correlation of STD data of different receivers. To justify this, the data is thinned. This is explained in more detail in Section 3. Without correlations, the observation error covariance matrix **R** is diagonal.

2.4. Forward operator

The introduction of a new observation type into variational assimilation systems requires the development of the observation operator H, the tangent-linear operator $\vec{\nabla} H$, its adjoint operator $(\vec{\nabla} H)^T$ and the estimation of the observation-error covariance matrix **R** (see Ide et al., 1997 for a complete list of commonly applied notations). H is the projection from model space to observation space. A STD is an integrated quantity and can be determined with the following equation:

$$\mathrm{STD} = \int_0^s \left(k_1 \frac{p}{T} + k_2 \frac{e}{T^2}\right) ds. \tag{3}$$

Here p denotes the total atmospheric pressure, T is the temperature and e denotes the water vapour pressure. The empirical constants are given by $k_1 = 77.6\,\mathrm{K\,hPa^{-1}}$ and $k_2 = 3.73 \times 10^5\,\mathrm{K^2\,hPa^{-1}}$ (Wickert and Gendt, 2006). The first term of the integral is known as the hydrostatic delay and the second part as the SWD. The integration is carried out along the geometric path from the ground-based GPS receiver to the satellite in question. For a discussion on the accuracy of the line of sight assumption the reader is referred to Eresmaa et al. (2007). In general, the bending of the ray path must be taken into account for elevation angles smaller than 20° at the GPS receiver. Therefore, only STD data with elevation angles larger than 30° at the GPS receiver are assimilated in this study.

For a mathematical description of the numerical algorithm as well as the derivation of the tangent linear and the adjoint, the reader is referred to the online supplement provided on the *Tellus* Webpage (link at the end of the paper). The trapezoid rule is applied to evaluate the integral numerically in model grid point space. The geographical location (longitude, latitude and height above sea level) of the supporting points and the weighted line elements are determined along the geometric path from the GPS receiver to the model top. The supporting points are determined by calculating the intersection of the geometric path and the MM5 terrain following sigma half levels. The refractivity at a supporting point is constructed using a tri-directional interpolation of the model variables of adjacent model grid points. The reconstruction consists of a vertical interpolation followed by a horizontal interpolation using low-order piecewise polynomials. Finally, at each supporting point the integrand is

calculated and the summation is carried out. The STD above the model top STD_{top} is determined with the Saastamoinen model (Saastamonien, 1972):

$$STD_{top} = \frac{c}{\sin(\epsilon_{top})} \frac{p_{top}}{f(\theta_{top}, h_{top})}, \quad (4)$$

where $c = 2.2768 \times 10^{-5} [m\ Pa^{-1}]$, ϵ denotes the elevation angle and $f(\theta, h) = 1 - 0.00266 \cos(2\theta) - 0.0000028 \frac{1}{m} h$ accounts for the variation of the gravitational acceleration with latitude θ and height h. The subscript top indicates that the quantities refer to the model top. For a discussion on the accuracy of the Saastamoinen model refer to Cucurull et al. (2000).

The numerical integration below the model top and the assumptions applied above (spherical symmetry of the atmosphere) suggested that the error of the forward model is roughly proportional to the cosecant (the reciprocal of the sine) of the elevation angle at the GPS receiver.

3. Set-up of operational GPS STD 4DVAR system during D-PHASE

The forecasts were driven by the 00 UTC operational forecast of ECMWF. The observations for the assimilation window 00-03 UTC were downloaded from the FTP servers of the GFZ and the MetOffice. Then, the MM5 preprocessing system was started to prepare the GPS observations for assimilation. To reduce temporal and spatial correlations, the data were thinned so that only observations every 30 min (15 and 45 min past the hour) enter the assimilation. To reduce spatial correlations, it was ensured that not more than one STD has the same starting and ending grid box. Furthermore, stations located more than 50 m below or above the model topography were permanently blacklisted. The remaining data set consisted of 900–1200 slants in the 3 h assimilation window 00-03 UTC. No routine cross checks for temporal and spatial consistency of the STD data were performed. The ZTDs provided by the UK Met Office were treated as STD data with an elevation angle of 90°. Figure 2 shows the GPS stations available in the model domain (upper left panel), the coverage of the complete GPS STD data set (upper right panel) and the coverage of the remaining GPS observations after data thinning (lower panel).

From the initial conditions created from the ECMWF forecast, the 24 h control simulation (CONTROL) was started. With the pre-processed observations available, the 4DVAR with GPS STD data was initiated for the assimilation window 00-03 UTC. Since 4DVAR is a time-consuming assimilation method and the assumed Gaussian probability distributions are more and more violated at higher resolution, in this study, the operational assimilation was performed in the coarse 18 km domain. During the first forward simulation of the 4DVAR, additional filtering of the STD data based on the relative departure (observation-background) in the SWD was carried out to reject questionable data and to avoid that the minimization may not work correctly due to large differences between model background and measurements. GPS stations whose relative departure in SWD deviated by more than 30% from the background value were permanently blacklisted during the assimilation. The quality control was done in terms of SWD, since we expect that pressure p and temperature T are reasonably represented by the model.

The minimization of the cost function was cut off after 20 iterations to restrict the running time of the 4DVAR. In most cases where the minimization worked properly, the gradient norm was reduced by 3–5 orders of magnitude. The calculated increment was applied to modify the initial state provided by the ECMWF analysis. The lateral boundaries during the assimilation window were not touched by the 4DVAR system. From the modified initial state another 24 h model integration, the IMPACT experiment, was started. Both simulations were performed in the 3-domain 2-way-nested configuration described above using the most recent MM5 version 3.7.4 and most sophisticated physics (see the right column of Table 1). In this publication, we focus on the investigation of low-resolution statistics. Another systematic investigation of the statistical performance of MM5 and other models participating during D-PHASE is given in Bauer et al. (2010). High-resolution case study simulations are subject of future papers.

4. Results

4.1. Methodology and selected observations

The main objective of our forecast system, producing one CONTROL and one 4DVAR forecast (IMPACT) during COPS and D-PHASE, is to investigate the statistical influence of GPS STD/ZTD observations on the representation of the water vapour field and the forecast of precipitation. In the following, we restrict the discussion to the month of August 2007, since we tuned our system until mid of July. Later in September, October and November precipitation totals well below average were observed.

As an example illustrating the influence of the assimilation, Fig. 3 shows the differences IMPACT-CONTROL of the 850 hPa water vapour field for two different forecast lead times. It is seen that large changes of up to 2 g kg^{-1} are included by the assimilation and that they are still present after more than 5 h of model integration.

GPS STD data from the German GFZ network and ZTD data from the French SGN network are applied to calculate 'model-GPS' departures. The corresponding model values are calculated with the GPS STD forward operator developed for the assimilation system. In addition, the model results are validated with the gridded REGNIE (Regionalisierung von Niederschlagsdaten) precipitation data set of the German Meteorological Service (DWD). REGNIE is routinely created on a 1×1 km grid from about 1000 stations over Germany. The interpolation

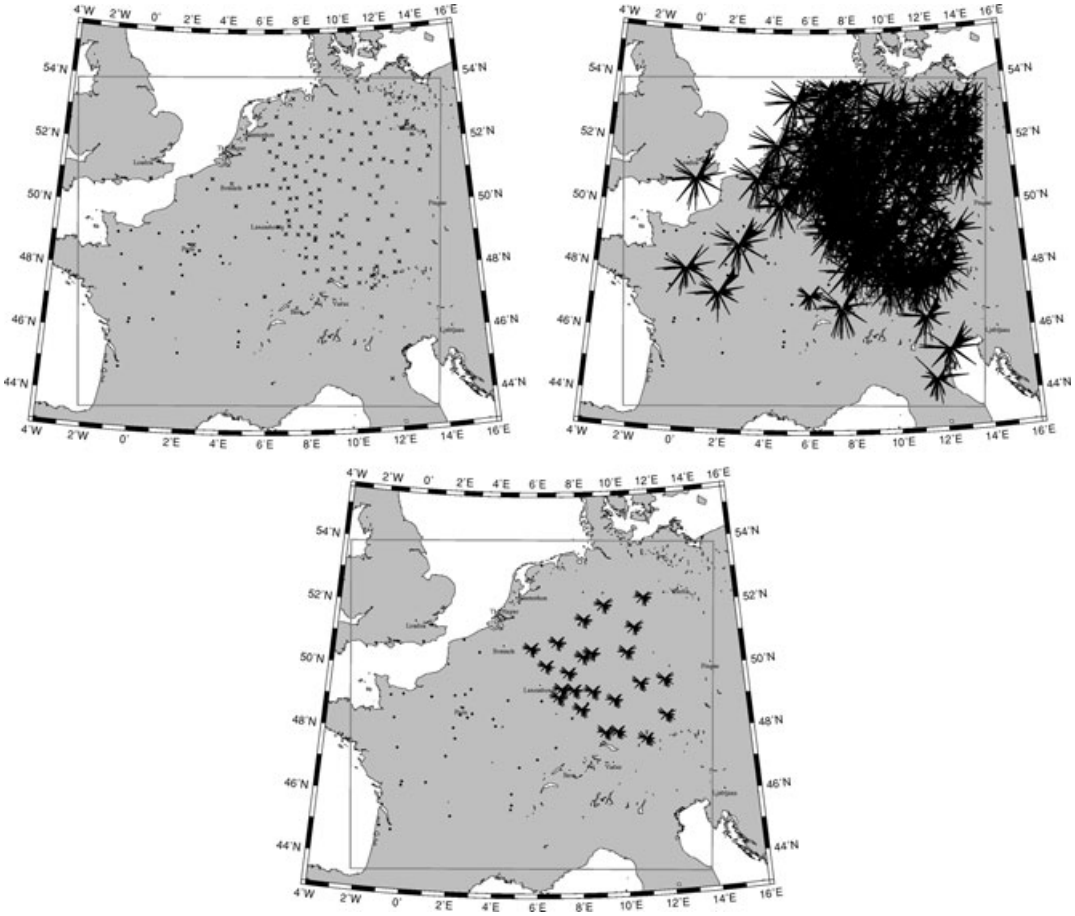

Fig. 2. GPS network applied for the assimilation. The dots on the top left panel mark the French GPS receiver stations and the crosses the German GFZ stations. The top right panel shows the area covered by the full GPS data set available neglecting information from elevation angles smaller than 30°. The bottom panel shows the data actually fed into the system after data thinning. Each day 900–1200 slants were assimilated in the assimilation window 00-03 UTC.

algorithm is based on the optimal interpolation approach of Wiener (1949) also considering the station elevation and exposition. For each station, the precipitation accumulated from 6 to 6 UTC of the following day is measured and then summed up to the total precipitation of August 2007.

The investigations are exclusively done with the coarse 18 km resolution domain to avoid artefacts caused by the nesting (e.g fronts that dry out from the outer to the inner domain due to the change in resolution or model physics). They would make the interpretation of the performance of the 4DVAR system more difficult. To avoid the feedback from the inner domain and to be able to compare modelled precipitation with observed precipitation from 6 to 6 UTC of the following day, we re-ran the 18 km CONTROL and IMPACT simulations for 30 h. To compare the model with the REGNIE data set, the observations are interpolated to the model grid.

The monthly averaged diurnal cycle of precipitation in Germany is validated with hourly reporting precipitation stations downloaded from the webpage of the General Observations Period (Crewell et al., 2008). In this data set, hourly reporting DWD precipitation stations are complemented by measurements from regional flood forecasting centres. The final data set contains up to 950 stations in total. The comparison is carried out by interpolating the hourly model precipitation to the station location using bilinear interpolation.

4.2. Comparison of GPS and model data at initial time

In this subsection, we focus on the first time step that can be compared, namely 0015 UTC where GPS data are available for the first time. So far, only GPS radio occultation data are assimilated by ECMWF. Preparations to also apply ZTD data in the ECMWF 4DVAR system were started recently after studies investigating the influence of the data on the forecast system. Therefore, GPS STD and ZTD can be applied as an independent verification of the ECMWF water vapour field. This comparison is also the necessary first step when thinking about bias

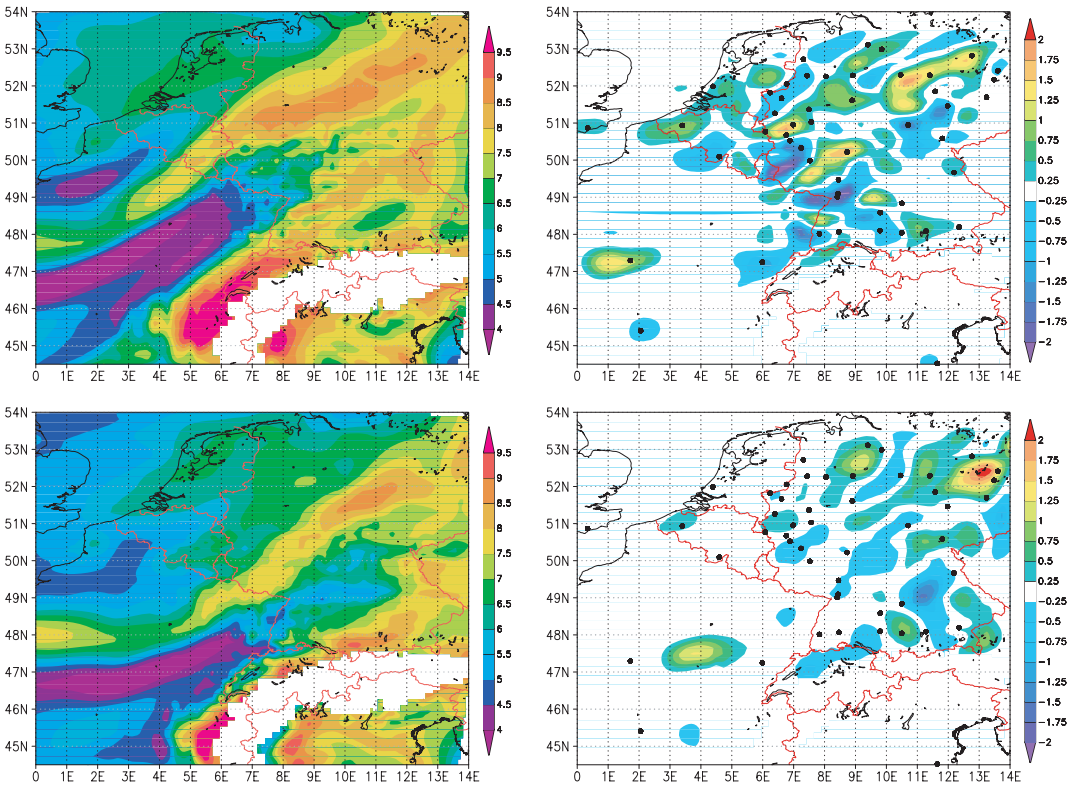

Fig. 3. Water vapor mixing ratio [g kg^{-1}] at 850 hPa for two different time steps on 27 July 2007. Left column: Water vapour field in the 18 km simulation initialized by 4DVAR. Right column: Difference IMPACT-CONTROL simulations. The upper row shows the situation at 01:30 UTC in the middle of the assimilation window. The lower row shows the situation at 05:30 UTC 2.5 h after the end of the assimilation window. Black dots in the difference plots mark the locations of the GPS stations.

correction of the data before assimilation. Furthermore, the data will be applied to verify whether the 4DVAR is able to improve the initial ECMWF water vapour field in MM5.

Normalized model minus observation departures of STDs are calculated with the following equation:

$$STD_{nbias} = \frac{2 \cdot \sum_{i=1}^{n} \frac{STD_{mod} - STD_{obs}}{STD_{mod} + STD_{obs}}}{n} \tag{5}$$

STD_{mod} represents either the ECMWF analysis interpolated to the MM5 model grid (CONTROL) or the model field after the correction by 4DVAR (IMPACT). The model STDs are derived with the developed forward operator. The normalization is necessary to eliminate the influence of the varying elevation angle in the STD data. The left panel of Fig. 4 shows the normalized departure from the German GFZ network consisting of STD data from about 50 stations. The right panel shows the departure from the French SGN network consisting of ZTD data from about 40 stations. These station numbers remain after data thinning and orographic as well as background filtering. This separation between the networks is necessary since ZTD and STD data cannot be compared directly. Furthermore, different software packages are applied by the two processing centres resulting in different delays calculated from the same raw data.

The solid line in both plots represents the normalized difference CONTROL-GPS (either GFZ or SGN network). At most of the GPS receiver stations, the ECMWF initial water vapour field interpolated to the MM5 model grid deviates by only ±0.1%, which corresponds to approximately 2.5 mm in zenith direction. This suggests that the ECMWF model accurately predicts the IWV content. Interestingly, a notable difference appears between the two networks. Whereas the deviations of CONTROL are evenly distributed around zero for the GFZ network, leading to an almost unbiased Gaussian distribution, the majority of the departures to the French SGN network are positive, resulting in a positive bias of CONTROL as compared to the SGN network.

The dot-dashed line in both plots represents the situation after the assimilation of GPS data from both networks (IMPACT). The 4DVAR worked properly in reducing the spread. For almost all GPS stations the differences IMPACT-GPS are reduced by up to 50% as compared to the CONTROL-GPS departures. This is true for both networks. However, as compared to the SGN network, again, a positive bias remains.

To investigate the spatial differences between CONTROL and GPS data, the normalized departures were plotted on a map

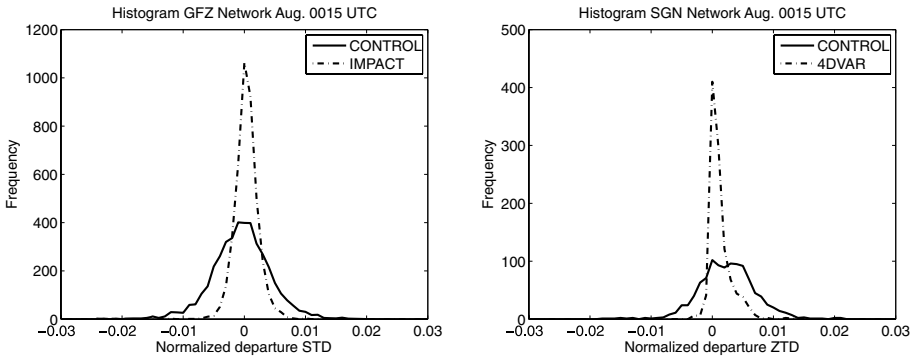

Fig. 4. Histogram of the normalized departures of model-GPS STD [%] (BIAS) at 00:15 UTC. The solid line shows the departure CONTROL-GPS and the dotted line IMPACT-GPS. The left panel is for the German GFZ network and the right panel for the French SGN network. The simulations from both initial states were performed with MM5 version 3.7.4 and sophisticated model physics.

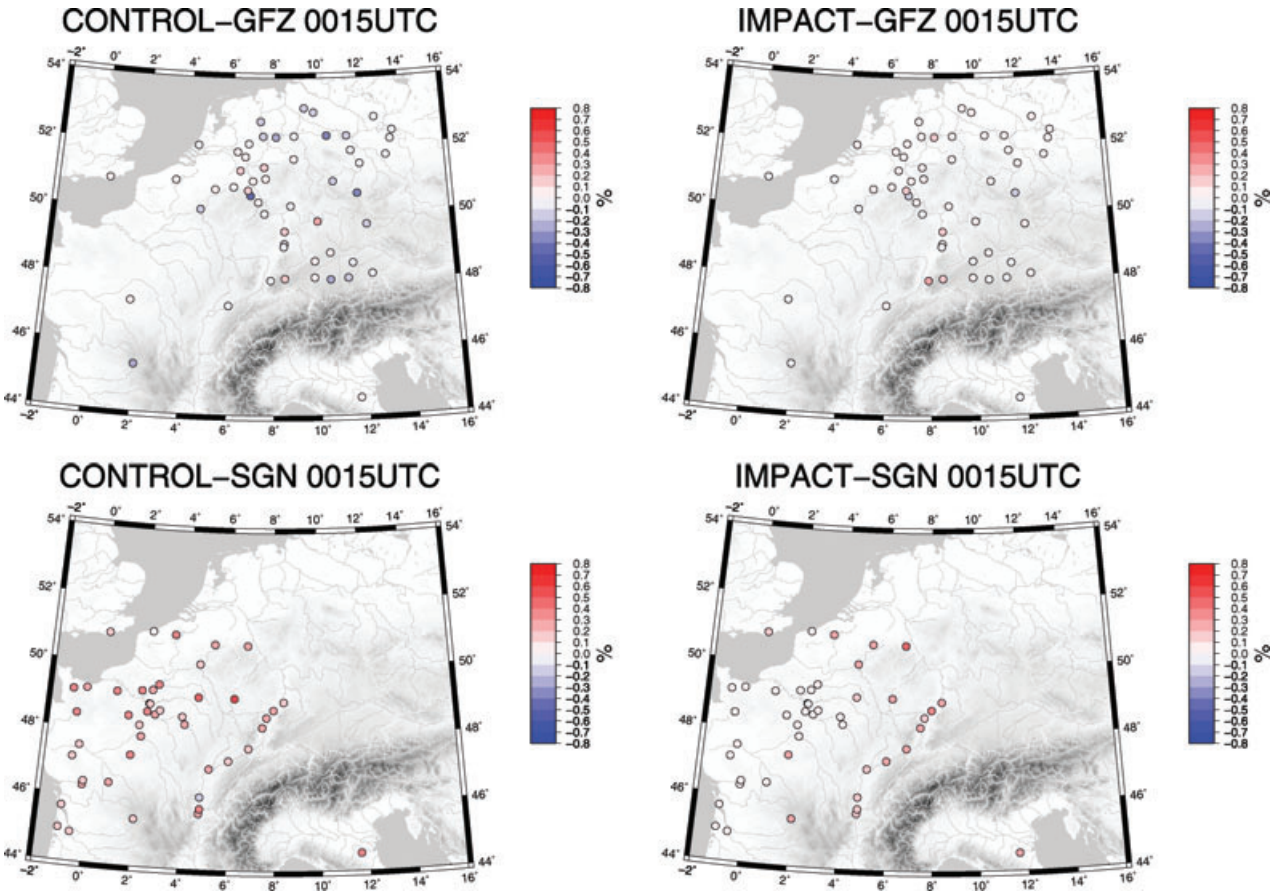

Fig. 5. Averaged and normalized model-GPS departures [%] (BIAS) at 00:15 UTC for the STD observations of the German GFZ network (top row) and the French SGN network (ZTD's) (bottom row). 'Model' represents either the ECMWF forecast interpolated to the MM5 grid (CONTROL) (left column) or the same field modified by 4DVAR of GPS (IMPACT) from both networks (right column).

averaging STD_{nbias} of all 0015 UTC time steps of August 2007 for each station, shown in Fig. 5. The dots on the maps represent the locations of the GPS stations. They are coloured with the averaged value of the normalized departures of August 2007 as calculated by eq. (5). As also visible in Fig. 4, a more even distribution around zero is seen when comparing with the German GFZ network (upper left panel). For most GPS stations an averaged departure of less than ±0.1% (approximately 2.5 mm in zenith direction) is seen suggesting that the spatial distribution of humidity provided by ECMWF is close to what

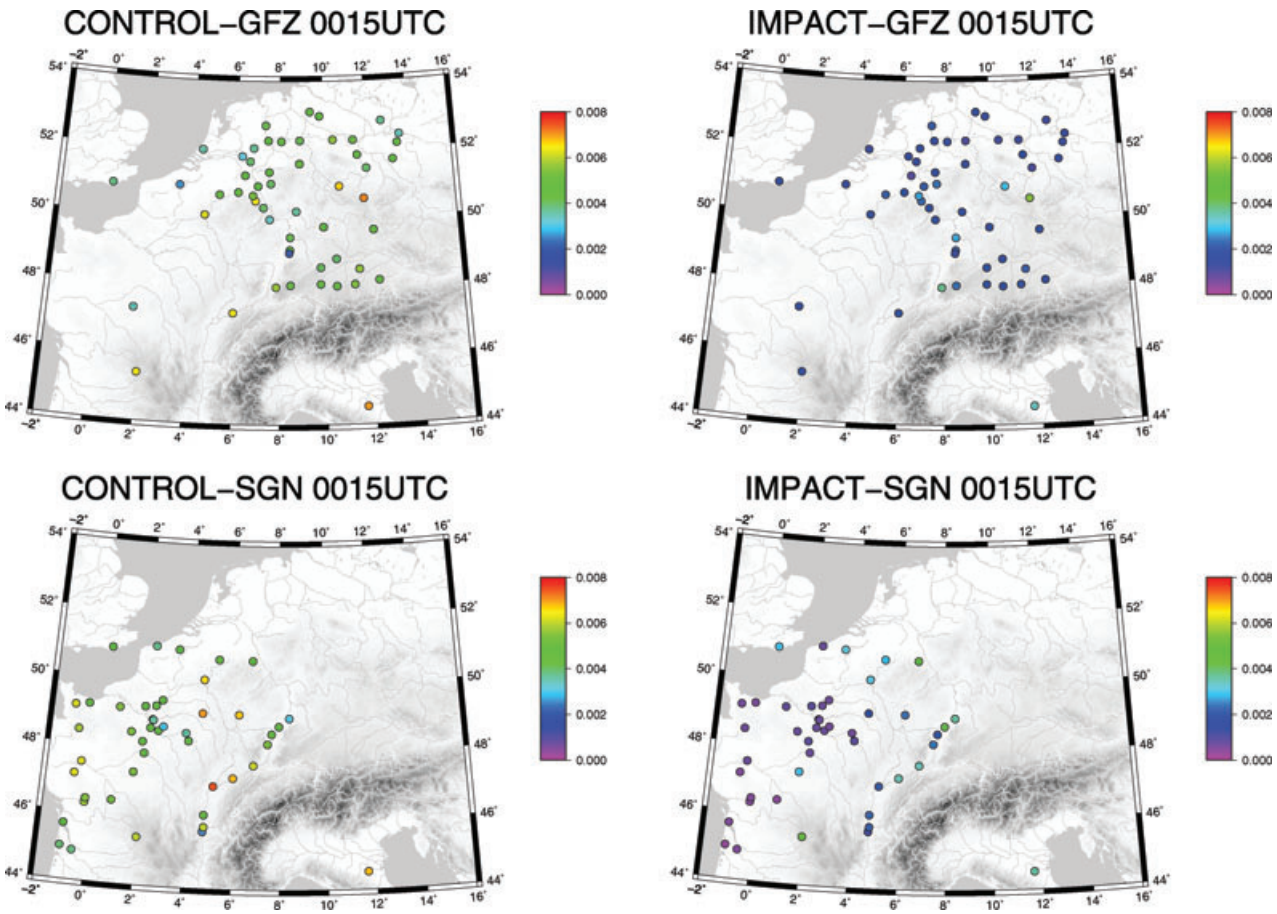

Fig. 6. Same as Fig. 5 but for the variance (RMS error) [m] instead of the bias.

was observed by the GPS network. However, larger departures of up to 0.3% occur at single stations (one outlier with −0.5% occurred in the low mountain region southwest of Cologne). The majority of stations revealing larger differences between models and observations are either located in river valleys or at the boundary of orography, suggesting that the model resolution of 18 km and the correspondingly interpolated orography is an important contributor to the differences.

The lower left panel of Fig. 5 shows the averaged normalized departure of CONTROL from the French SGN network (ZTDs). Here, with the exception of one station in the Rhone valley, a positive bias is detected for all stations corresponding to the results given on the right panel of Fig. 4. This confirms that the positive bias is not caused by single outliers but is rather due to a systematic difference between the ECMWF analysis and SGN data.

The right column of Fig. 5 shows the same comparisons after 4DVAR of GPS data (IMPACT) for the GFZ network (top) and the SGN network (bottom). The general properties of the 4DVAR to reduce the deviation is clearly seen. Again, the distribution of the IMPACT-GFZ departures are evenly distributed around zero, while a positive bias remains in the IMPACT-SGN departures. Furthermore, spatial differences in the influence of the 4DVAR are seen. Stations located in regions with steep topography, for example, in mountains or river valleys show a reduced or even detrimental influence of the assimilation. A good example is the station 'Freiburg' on the western side of the southern Black Forest. Whereas CONTROL shows a slightly negative bias, IMPACT changes this to an even larger positive bias. Another example is the line of GPS stations in the Rhine valley where hardly any influence of the 4DVAR is seen. On the other hand, for almost any station in France or northern Germany the biases are clearly reduced by the assimilation. Possible causes for the reduced performance of the 4DVAR system in mountainous terrain are discussed in the next section.

Apart from the normalized bias, the normalized RMS error was investigated. It was calculated with the following equation

$$STD_{nrms} = \sqrt{\frac{\sum_{i=1}^{n} \left(\frac{STD_{mod} - STD_{obs}}{STD_{mod} + STD_{obs}} \times 2 \right)^2}{n}} \tag{6}$$

Again, the normalization is necessary to compensate the influence of the different elevation angles. Figure 6 compares the averaged normalized RMS error of CONTROL (left column)

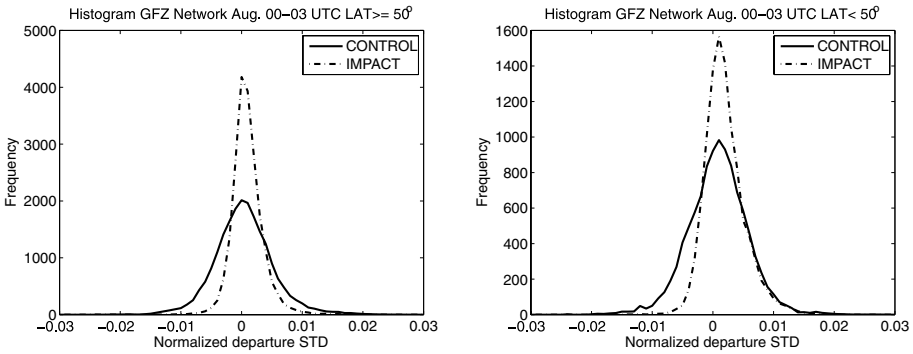

Fig. 7. Histogram of the normalized departures of model-GPS STD for the assimilation window 00-03 UTC. The solid line shows the departure CONTROL-GPS and the dotted line IMPACT-GPS. Both plots are for the GFZ network. The left panel collects all stations at or north of 50°N and the right panel all stations south of 50°N.

and IMPACT (right column) as compared to the GFZ (top row) and SGN networks (bottom row). The considerable reduction of the RMS error by 4DVAR is clearly seen. As for the bias, exceptions where the RMS errors are not or only slightly reduced are mostly related to mountain or river valley locations. This supports the assumption that the performance of the 4DVAR is reduced in regions of steep topography.

4.3. Influence of GPS 4DVAR in the assimilation window

For our forecast system, we selected the assimilation window to be 00-03 UTC each day. This selection has two major advantages. At 00 UTC and 12 UTC the largest amount of observations is assimilated into the ECMWF 4DVAR system leading to the best possible background state. Furthermore, the absence of incoming solar radiation reduces errors due to incorrect simulation of land-surface exchange processes in the 4DVAR system.

To further investigate the influence of orography on the performance of the assimilation, we coarsely separate stations in flat and orographic terrain by distinguishing between GPS stations located north of 50°N and south of 50°N. Figure 7 shows the normalized bias for the 3 h assimilation window for the GFZ network. This comparison is only shown for the GFZ STD network due to the much larger amount of STD data. For the stations south of 50°N (right panel), a positive bias in the IMPACT-GFZ departure remains although the differences are clearly reduced. This shift of the maximum of the distribution to the right does not occur for the stations north of 50°N. In addition, it is seen that the differences are generally not as strongly reduced as it is the case at 00:15 UTC (estimated 30% instead of 50%), suggesting a reduction of the influence of 4DVAR during the assimilation window.

Figure 8 shows the spatial distribution of the normalized bias for the GPS stations, the two networks and the assimilation window 00-03 UTC. The same tendency as for the first time step is seen when comparing the departures of the CONTROL simulation from the GFZ and SGN networks. Whereas the values from CONTROL are evenly distributed around the observations of the GFZ network (upper left panel), they are systematically too moist as compared to the SGN network (lower left panel). When comparing the IMPACT results from the assimilation window (Fig. 8, right column) with the first time step (Fig. 5, right column), interesting differences occur. Although the bias is still reduced from CONTROL to IMPACT, the reduction is not as strong as for the first time step. This is especially true for stations in mountainous terrain (as e.g. in the Rhine valley), confirming the reduced performance of the 4DVAR in orographic terrain.

4.4. Influence of GPS 4DVAR over the forecast length

To investigate the performance of the 4DVAR system over the forecast length, apart from the water vapour comparisons shown above, also the statistical representation of precipitation for August 2007 is analysed. Figure 9 compares the normalized departures CONTROL-OBS (solid line) and IMPACT-OBS (dashed line). Although reduced by the longer analysis period, the same signal as in the preceding comparisons is seen. Whereas the departures of the CONTROL simulation occur evenly distributed around zero, a slight shift to positive values is visible in the IMPACT simulation. This implies that on average the 4DVAR includes a slight positive bias. The previously mentioned large reduction of the normalized departures in the 4DVAR simulation has more or less vanished, illustrating that the 4DVAR looses its influence over Europe during the 24 h model integration when only GPS data from Germany and France are assimilated. Reasons for this are discussed in the next section.

Whereas the plots shown so far analysed the bias and RMS for selected time periods, it is also necessary to investigate their temporal evolution during the forecast. The spatial averages of normalized bias and RMS error over time are shown in Fig. 10. Both, CONTROL and IMPACT are too moist during the first hours of the simulation. Comparing the two simulations, it is seen that a small positive bias is included by the 4DVAR. It is largest during the first hours of the model integration and steadily reduced later on. Interestingly, the bias changes its sign from positive to negative after 13 h. In the second half of the

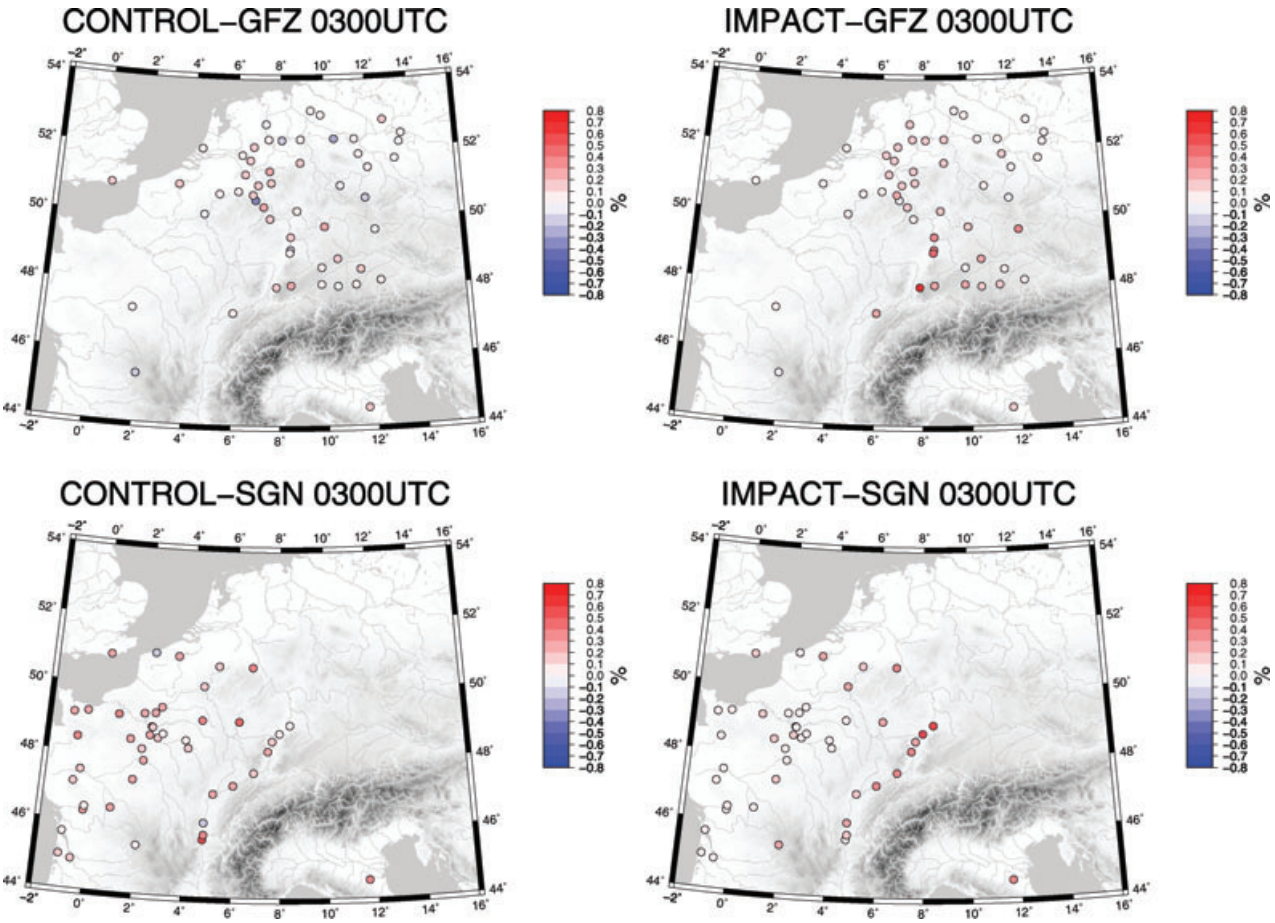

Fig. 8. Same as Fig. 5 but for the assimilation window 00-03 UTC instead of the time step 00:15 UTC.

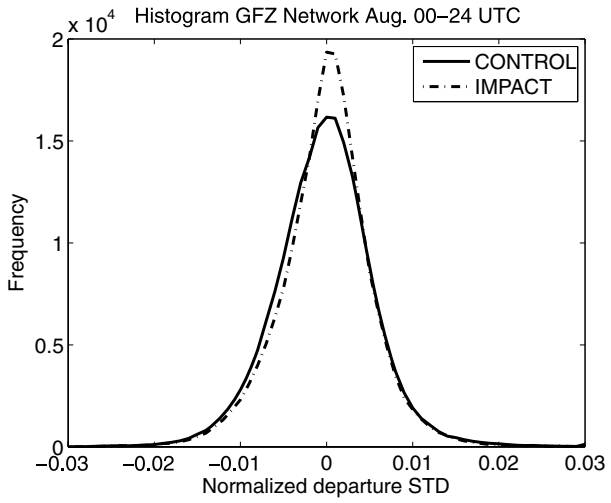

Fig. 9. Same as Fig. 5 but for the 24 h forecast instead of the time step 00:15 UTC.

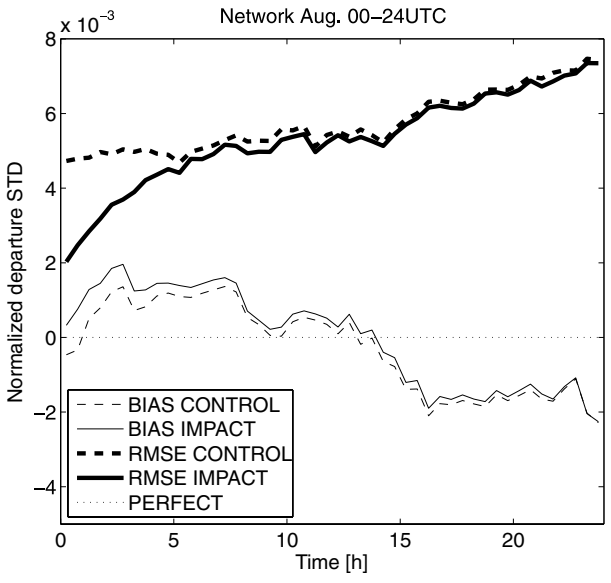

Fig. 10. Temporal evolution of BIAS and RMS of GPS STD during the 24 h forecasts of CONTROL and IMPACT forecasts. Data included are the whole GFZ network.

simulation, the bias of IMPACT is smaller than that of the CONTROL forecast. The influence of 4DVAR is clearly seen in the temporal evolution of the RMSE. At initial time it is reduced by up to 50% in the IMPACT simulation and remains smaller than for the CONTROL simulation during the whole 24 h forecast. It is important to note that the difference between CONTROL and IMPACT is steadily reduced during the model integration and that its absolute value is constantly rising.

Figure 11 compares the 24 h precipitation sum of August 2007 for the REGNIE observation (top panel), the CONTROL simulation (middle panel) and the IMPACT simulation (lower panel). Both model integrations overestimate precipitation over southwestern Germany and the differences between the two simulations are clearly smaller than the differences to the REGNIE observations. While the precipitation maximum in the Alpine foreland is captured by the simulations, the observed maximum in western Germany is underestimated. MM5 simulates a large maximum shifted to the southwest to upwind slopes of the mountains. This so-called windward-lee effect is a typical error occurring when a convection parameterization is applied (Schwitalla et al., 2008) and is further amplified by the coarse resolution orography. Whereas the structure of the distribution of precipitation is more or less identical in CONTROL and IMPACT, the latter slightly reduces the amount of precipitation in southwestern Germany and eastern France, for example, the maxima over the Vosges mountains and the southern Black Forest. On the other hand, the northern part of the maximum over western Germany is simulated slightly stronger. Both changes are small but in better accordance with the REGNIE observation.

Since such comparisons neglect the changes during day, it is essential to also investigate the representation of the averaged diurnal cycle to judge the temporal evolution of precipitation in the domain. This is shown in Fig. 12. The CONTROL simulation underestimates the amount of precipitation in the first 9 h of the forecast. This is due to the performed cold starts from dry conditions each day. Since the set-up of an assimilation cycle is not an option in the MM5 4DVAR system, the simulation needs time to develop clouds and precipitation. It is promising in view of future nowcasting applications that the spin-up is almost completely removed by the 4DVAR. The IMPACT simulations even tend to slightly overestimate precipitation in the first 3 h of the model integrations. Although both simulations, CONTROL and IMPACT, overestimate the precipitation from 0900 UTC to the end of the simulations and show a late afternoon maximum that was not observed, the averaged diurnal cycle of the IMPACT simulations is closer to the observation than that of the CONTROL forecasts over the whole 24 h.

Figure 12 also includes the averaged diurnal cycles of the 7 km resolution models of MeteoSwiss (COSMO-7) and the DWD (COSMO-EU) that were operated during COPS and D-PHASE.

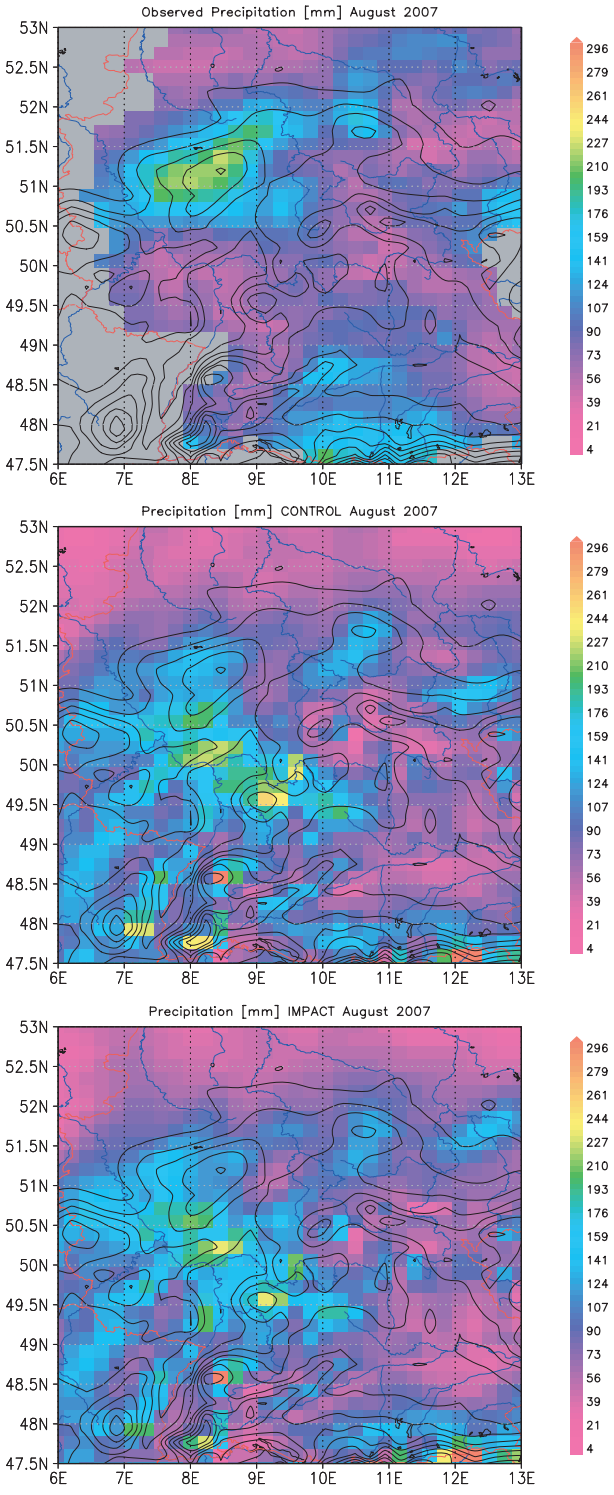

Fig. 11. Accumulated precipitation [mm] for August 2007 at 18 km horizontal resolution for DWD REGNIE (top panel) CONTROL (middle panel) and IMPACT (bottom panel). Contour lines in the three panels represent the orography at 18 km resolution.

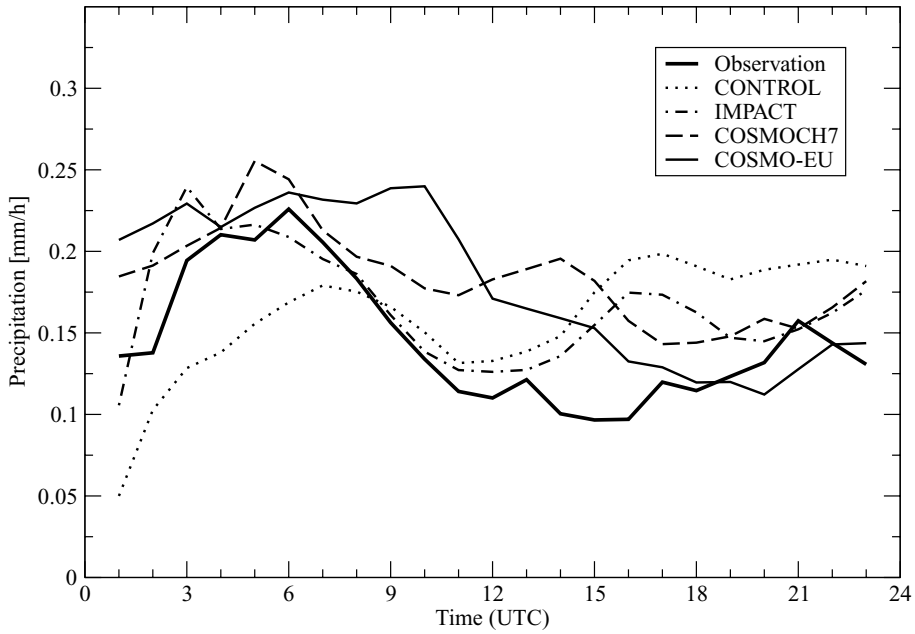

Fig. 12. Averaged diurnal cycle of total precipitation of CONTROL (18 km), IMPACT (18 km), COSMOCH-7 (7 km), COSMO-EU (7 km) and observations. The model grid box values are interpolated to the observation locations using bilinear interpolation.

We know that the models are too different to be compared in detail. This is not our intention. The only aim of this comparison is to provide a benchmark for the MM5 simulations allowing to put their performance into context with other operational systems. The COSMO models overestimate precipitation during the first hours of the simulation. Between 09 and 15 UTC, the observed minimum is better represented by MM5, whereas in the late afternoon, the COSMO models are closer to the observations than both MM5 simulations. This confirms that the forecast performance of MM5 is comparable with other operational systems.

Finally, the discussion is quantified with the calculation of different scores for the CONTROL and IMPACT simulations. Figure 13 shows the probability of detection (POD), the false alarm ratio (FAR) and the true skill score (TSS) for different precipitation thresholds ranging from 50 to 200 mm per month. The POD measures the fraction of observed events that were correctly forecasted. A perfect forecast has a POD of 1. The higher the POD the better the localization of precipitation in the model. The FAR measures the fraction of false alarms in comparison to all forecasts. A perfect forecast has a FAR of 0. The TSS measures the accuracy of the forecast in predicting the correct precipitation category relative to random chance.

The scores were calculated for the 24 h precipitation sum compared to the REGNIE product, neglecting the influence of the assimilation on the diurnal cycle. Therefore, it is not surprising that the assimilation has only a small impact. Interestingly, some of the scores were especially improved for larger precipitation thresholds, whereas no changes or even lower scores occurred for lower precipitation thresholds. An example is the 150 mm threshold, where the FAR is improved by 5%, the POD by 20% and the TSS by even 60%.

5. Discussion

In this section, the causes that may have a detrimental influence on the performance of the 4DVAR system for different time scales and forecast lead times are discussed.

At initial time, errors in the model and assimilation system as well as errors in the selected observations may limit the performance of the 4DVAR. Usually, the inclusion of heterogeneous observations generates high-frequency inertia-gravity waves ('meteorological noise'), caused for example by imbalances between model and observations or model and observation errors, reducing the overall performance of the assimilation system. In advanced versions of 4DVAR, as for example in the ECMWF forecast system or in the Weather Research and Forecasting (WRF) model, gravity-wave filtering is an important task and a major research topic (Gauthier and Thepaut, 2001; Huang et al., 2009).

The presented results revealed different performances of the 4DVAR system over flat and orographically structured terrain. This may either be caused by the coarse representation of the orography in the 18 km model grid or by errors in the model or assimilation system. Since we neglect the GPS data that vary considerably from the model background and stations more than 50 m above or below the model topography, the difference between model background and observations is reduced and the

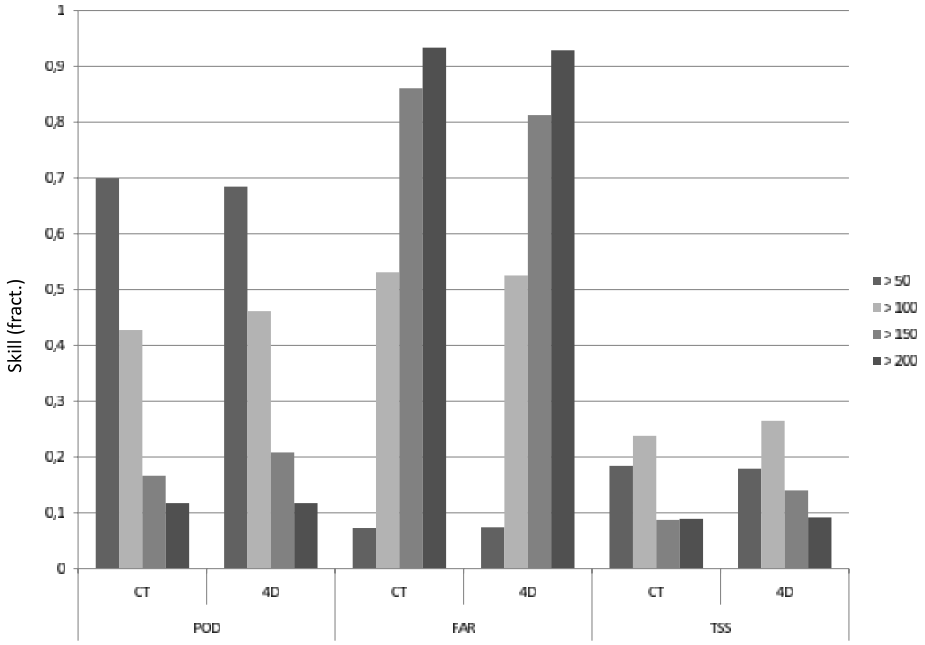

Fig. 13. Selection of forecast scores for different precipitation thresholds [mm month^{-1}]. POD, Probability Of Detection; FAR, False Alarm Ratio; TSS, True Skill Score.

excitation of gravity waves is minimized. Therefore, errors in the model physics and the assimilation system are the most probable causes for the performance differences.

Observation-minus-background statistics and sensitivity studies with changed model physics during the operational phase of D-PHASE revealed that the errors are caused by the necessity to change the model physics from the assimilation to the free forecasts. This introduces a wet bias in the first forecast hours in mountainous terrain (Zus et al., 2008). We had to use the simple Anthes-Kuo convection scheme (Anthes, 1977) at a resolution of 18 km for the assimilation run, because the MM5 4DVAR system does not offer adjoints for more sophisticated alternatives. Furthermore, this scheme is known to produce excessive convective rainfall (Grell et al., 1995). In the MM5 4DVAR system, diffusion is calculated along terrain-following sigma-levels. The more accurate z-diffusion scheme (Zängl, 2002) is not available in the MM5 version that serves as the basis for the 4DVAR system. The sigma diffusion scheme leads to significant errors over mountainous terrain, particularly for atmospheric variables with a strong vertical gradient (Zängl, 2002). Temperature diffusion along sigma surfaces tends to cool down the air in valleys and heat it above mountains, resulting in artificial upward motion above the mountains facilitating the triggering of the sensitive Anthes-Kuo convection scheme. Diffusion of the water vapour mixing ratio along sigma surfaces tends to dry out the air in valleys and moisten the atmosphere above mountains (Zängl, 2002; Zus et al., 2008). Both effects, together with the Anthes-Kuo scheme, are responsible for the wet bias created by the 4DVAR in orographic terrain during the first 3 h of the forecast. The removal of moisture from the atmosphere by the excessive precipitation is recognized by the 4DVAR system which changes the water vapour field to moister values. In the free forecasts, a more realistic convection scheme (Kain, 2004) and diffusion along z-levels are applied. Since the 4DVAR adjusted the water vapour field to the older model version with simplified physics, this results in overestimated precipitation in the first hours of the simulation to compensate for the too moist lower troposphere.

In the assimilation window, additional processes contribute to the further reduction of the performance of the 4DVAR system. Most of the reduction can be explained by the change of the model version from 3.4 for the assimilation to 3.7 for the forecasts. The latter is not in balance, since the 4DVAR adjusted the initial state for the older model version 3.4 and the simplified physics applied in the 4DVAR system. Nevertheless, we decided to change the model version, since the physics parameterizations are more advanced and many errors have been corrected from version 3.4 to version 3.7.

Later in the forecast, the development of the atmosphere strongly depends on the external forcing across the lateral boundaries. Since this forcing is not influenced by the assimilation of data in the MM5 domain, a damping effect on the corrections implied by 4DVAR is not surprising. The air advected into the domain origins from the global ECMWF model. Although ECMWF operates a sophisticated 4DVAR system, their forecasts run with different horizontal and vertical resolutions and apply different descriptions of dynamics and physics. This leads to inconsistencies degrading the performance of the

4DVAR system. Furthermore, the ambient flow transports the air influenced by the assimilation out of the domain, lowering the impact of the assimilation.

Another issue is that we only assimilate GPS STD data. The model representation of the dynamics is questionable at 18 km resolution in orographic terrain. Therefore, improvements in the representation of the water vapour field can be counteracted by a wrong representation of the dynamics within short time. For future activities, it is therefore vital to assimilate water vapour data from GPS and radar radial velocity data simultaneously.

Comparing the 24 h model integration, further processes that may have a detrimental influence on the performance of the assimilation are discussed. The atmospheric water vapour content is strongly influenced by the description of the land surface, namely the land use as well as the soil characteristics. The latter determine what amount of water can be stored in the soil and how fast it can be transported through the soil layers. The land use type, on the other hand, controls the fluxes of latent and sensible heat providing heat and moisture to the boundary layer. A too strong evaporation at the surface, combined with a too strong transport or a too large amount of water below the surface, would result in a too moist lower troposphere and therefore a wrong feedback to the boundary layer and microphysics parameterizations.

Although we improved the representation of the land use with the CORINE data set, the influence remains small since the forecast performance only benefits from a better land-use description when the fluxes of sensible and latent heat are improved. Another problem is that our optimized description of land use is not accompanied by a more realistic description of the soil characteristics. Since the implemented soil data set has an even coarser resolution than the land-use data set, we expect a wrong description of the soil moisture influence on the atmosphere, resulting in a further reduction in the performance of the assimilation system due to wrongly represented surface fluxes.

Another issue is that we did not initialize the soil moisture with observations. Instead, initial values from ECMWF were applied to determine the temperature and the moisture in the top soil layer. The ECMWF soil moisture is the result of an analysis system based on optimal interpolation of proxy observations of 2 m temperature and relative humidity. Soil moisture needs a long time to reach a balance to the lower tropospheric humidity field and the strength of the soil moisture forcing on the lower troposphere is governed by the strength of the land-atmosphere exchange which varies with the synoptic situation. A wrong soil moisture initialization will therefore destroy the lower tropospheric humidity field. On the other hand, a sophisticated initialization of the soil moisture field is not yet possible for a large region since no operational network of soil moisture sensors exists. New satellite sensors like the ESA Earth Explorer mission 'Soil Moisture Ocean Salinity' (SMOS; Kerr et al., 2001), launched recently, are expected to improve the situation.

6. Summary and Conclusions

Although considerable progress has been achieved in forecasting mesoscale and small-scale atmospheric phenomena, there is still potential for improving numerical weather prediction by more accurately obtaining the correct initial condition by advanced data assimilation. During D-PHASE from June 1 to November 30, IPM ran an operational assimilation and forecasting system based on the MM5 4DVAR. This paper discussed the performance of the system in a statistical sense and identified several processes that had a detrimental influence on the assimilation acting on different time scales. Summing up the scientific questions posed in the introduction, the following conclusions are drawn from our analyses.

(1) How good is the representation of the water vapour field in the driving ECMWF forecast?

The comparison of the ECMWF analysis interpolated to the MM5 model grid and GPS STD data at the first time step when the observations were available demonstrated that ECMWF provides a good representation of the water vapour field at 18 km resolution as seen from the GPS perspective. For almost all stations, the normalized model-observation departures of STD did not exceed $\pm 0.1\%$, corresponding to 2.5 mm projected to ZTD. However, relatively large normalized RMS values of up to $\pm 0.8\%$ (corresponding to 2 cm projected to ZTD) indicated that the small-scale variability of the water vapour field was not fully captured by the ECMWF analysis. This is not surprising, as meso-gamma scale variability is not resolved but does exist due to land-surface heterogeneity and orography.

(2) How large are the analysis increments with respect to water vapour, temperature and wind. Do they result in an improvement of the initial water vapour field?

The comparisons before and after the assimilation of GPS data showed that the 4DVAR system worked properly in reducing model-observation departures of GPS STD. Whereas the bias was only slightly reduced, a large reduction of up to 50% at initial time occurred for the variance. This corresponds to changes in the mixing ratio q of up to $2\,\mathrm{g\,kg^{-1}}$ in the boundary layer. The comparisons also revealed differences between the two applied GPS networks. This discrepancy may be caused either by systematic differences between STD (GFZ) and ZTD (SGN) measurements or point to different biases occurring in the applied processing systems. This motivates the verification of the retrieval algorithms with data from NWP models since they provide a complete data set and allow the calculation of all retrieved products (Bender et al., 2008). Furthermore, bias removal should be included before data from different processing centres is assimilated. Analysis increments were also found for wind and temperature (not shown). These were typically of the order of 0.2 K and $0.3\,\mathrm{m\,s^{-1}}$ in the atmospheric boundary layer in orographic terrain and below 0.1 K and $0.1\,\mathrm{m\,s^{-1}}$ in flat terrain. During the course of the forecast, the effect of GPS

STD data was gradually reduced due to the influence of several processes acting on different time scales.

(3) What is the impact on the prediction of the water vapour field and on QPF in comparison to observations and other D-PHASE models?

The performance of the 4DVAR system is different for stations located in flat and orographically structured terrain. In mountainous regions or river valleys, a wet bias was introduced by the assimilation caused by the inaccurate physics we had to apply during the assimilation.

The comparison of the horizontal distribution for the August 2007 precipitation sum with the gridded REGNIE observation of DWD revealed an almost neutral impact of the assimilation with improvement in some regions and deterioration in others. Nevertheless, forecast scores calculated from the 24 h precipitation sum in comparison with the DWD REGNIE product reveal improvements, especially for larger precipitation thresholds.

Comparisons with hourly reporting precipitation stations, on the other hand, showed that the averaged diurnal cycle was clearly improved by the assimilation of GPS STD data. The spin-up occurring in the CONTROL simulation was almost completely removed by the assimilation and the representation of the diurnal cycle was better than that of the CONTROL simulation for the whole 24 h model integration. This is true in spite of the limitations of the 4DVAR system. The reduction of the spin-up is an important issue since precise flood forecasts with hydrological models need short-range forecasts of the spatial and temporal distribution of precipitation as accurately as possible. These comparisons indicate that it is not enough to study averaged spatial distributions as it is usually done when validating NWP and climate model simulations. Such comparisons can completely hide improvements in the representation of the diurnal cycle due to the cancellation of time-dependent systematic errors. Additionally performed comparisons with the COSMO models of MeteoSwiss and DWD confirmed that our system is comparable in performance with other operational NWP systems.

(4) What are the processes limiting the spatio-temporal impact of the assimilation on the water vapour field?

We identified several processes that limit the performance of the assimilation system on different time scales. The result is that the difference between CONTROL and IMPACT is steadily reduced and the RMS error is constantly rising during the model integration.

At initial time, the performance of the 4DVAR system is reduced by (1) inconsistencies in the physics and dynamics of the driving ECMWF model and MM5, (2) the inclusion of observations which are not in balance with model physics, (3) errors in the model physics and the assimilation system, and 4) only assimilating GPS STD observations without the consideration of additional wind observations representing the dynamical evolution.

The inconsistency between the assimilation and the subsequent forecasts is further increased by the change of the model version from the assimilation (MM5 3.4) to the subsequent model integration (MM5 3.7.4) as well as by the transport of air not affected by 4DVAR into the domain and a simultaneous transport of air influenced by 4DVAR out of the domain. In addition, we hypothesize that despite the applied soil moisture analysis, the data provided by ECMWF is not in balance with the rest of the MM5 model physics. Together with the coarse soil data set, this results in a wrong soil moisture feedback to the lower troposphere counteracting adjustments by 4DVAR during the course of the simulation.

Although the 4DVAR was able to almost completely remove the spin-up occurring in the CONTROL simulation, the forecast performance is quickly degraded with lead time after the end of the assimilation window, leaving room for improvements of the system. An important extension is to include more observations into the assimilation process. In addition to the water vapor field, at least an adjustment of the model dynamics is necessary to avoid that the reasonable correction of the water vapour field is quickly destroyed by a wrong representation of the wind field. Adding additional observations may further improve the performance of the system due to a stronger influence of the assimilation on the model. Here, also the improved density of the GPS network and improved retrievals shall be taken into account. In the meantime, the E-GVAP network increased to more than 1200 stations and the German GFZ network has grown to more than 350 stations.

It is necessary to improve the model physics in the assimilation system to avoid imbalances between the assimilation and the forecast. The goal must be to use the same representation of physics for both, complemented by an additional filter that deals with the initial inconsistencies between the observations and the model (e.g. Wee and Kuo, 2004). In case of variational assimilation, the development of adjoints for complicated parameterizations is challenging and time consuming. Therefore, thorough comparisons of 3DVAR with accurate physics, 4DVAR with optimized physics and ensemble-based Kalman Filter techniques are required in a data assimilation testbed to identify the best suited data assimilation method for future applications especially on the convection-permitting scale (WMO, 2009).

Another important point is to choose the model domain as large as possible to reduce the influence of the lateral forcing on the region of interest as much as possible. This requires that the whole development of the major components determining the synoptic situation has to take place within the high-resolution domain.

In the meantime, we apply the WRF system for high-resolution assimilation and process studies for selected COPS IOPs (Schwitalla et al., 2010). WRF supports variational as well as Kalman filter based methods and the systems are more advanced than the MM5 system we applied in this study. With the assimilation of numerous different observations, the best

possible initial state is derived for an European domain with 3.6 km horizontal resolution. Kain et al. (2008) compared different horizontal resolutions for a large US domain and found 4 km to be the best compromise between realism of the results and computational efforts. For detailed process studies and comparison with COPS observations and synergetically derived quantities at COPS supersites, higher resolution sub-domains shall then be nested into this European domain.

7. Acknowledgments

We thank Galina Dick and Michael Bender from the Helmholz Center Potsdam, German Research Center for Geosciences (GFZ) for their tireless efforts to provide us with GPS STD data during COPS and for their assistance during the development of the forward operator. E-GVAP and Dave Offiler from the MetOffice are acknowledged for the provision of GPS ZTD data from French stations during COPS. We thank the German Research Foundation for funding the priority program DFG SPP1167 into which this project is embedded. DWD and ECMWF are acknowledged for the permission to use the most recent ECMWF forecast data and the necessary computing time on the ECMWF computer system to perform our forecasts. Furthermore, Carsten Maass and Paul Dando from the ECMWF user support are thanked for their help in setting up the operational system. The FZK Garmisch is acknowledged for the provision of the EU CORINE data set. Finally, the two anonymous reviewers are acknowledged for their helpful comments that considerably improved the paper.

References

Anthes, R. A. 1977. A cumulus parameterization scheme utilizing a one-dimensional cloud model. *Mon. Wea. Rev.* **105**, 270–286.

Bauer, H.-S., Weusthoff, T., Dorninger, M., Wulfmeyer, V., Schwitalla, T., Gorgas, T., Arpagaus, M. and Warrach-Sagi, K. 2010. Predictive skill of a subset of the D-PHASE multi-model ensemble in the COPS region. *Quart. J. Roy. Meteorol. Soc*. (COPS Special Issue), in press.

Bauer, H. S. and Wulfmeyer, V. 2009. Validation od components of the water cycle in the ECHAM4 general circulation model based on the newtonian relaxation technique: a case study of an intense winter cyclone. *Meteorol. Atmos. Phys.* **104**, 135–162.

Bender, M., Dick, G., Wickert, J., Ramatschi, M., Ge, M. and co-authors. 2009. Estimates of the information provided by GPS slant data observed in germany regarding tomographic applications. *J. Geophys. Res.* **114**, 6303–6313.

Bender, M., Dick, G., Wickert, J., Schmidt, T., Song, S. and co-authors. 2008. Validation of GPS slant delays using water vapour radiometers and weather models. *Meteorol. Zeitschrift* **17**(6), 807–812.

Bender, M. and Raabe, A. 2007. Preconditions to ground-based GPS water vapour tomography. *Ann. Geophysicae* **25**, 1727–1734.

Bengtsson, L., Robinson, G., Anthes, R. A., Aonashi, K., Dodson, A. and co-authors. 2003. The use of GPS measurements for water vapour determination. *Bull. Amer. Meteor. Soc*. **84**, 1249–1258.

Boniface, K., Ducroq, V., Jaubert, G., Yan, X., Brousseau, P. and co-authors. 2009. Impact of high-resolution data assimilation of GPS zenith delay on mediterranean heavy rainfall forecasting. *Ann. Geophys.* **27**, 2739–2753.

Bossard, M., Feranec, J. and Otahel, J. 2000. CORINE land cover technical guide, Addendum 2000. Technical report, EEA, Techincal Report 40, Copenhagen.

Crewell, S., Mech, M., Reinhardt, T., Selbach, C., Betz, H.-D. and co-authors. 2008. The general observation period 2007 within the priority program on quantitative precipitation forecasting: concept and first results. *Meteorol. Zeitschrift* **17**(6), 849–866, doi:10.1127/0941-2948/2008/0336.

Cucurull, L., Navascues, B., Ruffini, G., Elósegui, P., Rius, A. and co-authors. 2000. The use of GPS to validate NWP systems: the HIRLAM model. *J. Atmos. Oceanic Technol.* **17**, 773–787.

Das, S., Singh, S. V., Rajagopal, E. N. and Gall, R. 2003. Mesoscale modelling for mountain weather forecasting over the himalayas. *Bull. Amer. Meteor. Soc*. **84**, 1237–1244.

Dick, G., Gendt, G. and Reigber, C. 2001. First experience with near real-time water vapor estimation in a german GPS network. *J. Athmos. and Solar-Terrestrial Physics* **63**, 1295–1304.

Dierer, S., Arpagaus, M., Seifert, A., Avgoustoglou, R., Dumitrache, R. and co-authors. 2009. Deficiencies in quantitative precipitation forecasts: sensitivity studies using the COSMO model. *Meteorol. Zeitschrift* **18**, 631–645, doi:10.1127/0941-2948/2009/0420.

Dudhia, J. 1989. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.* **46**, 3077–3107.

Ebert, E. E., Dammrath, U., Wergen, W. and Baldwin, M. E. 2003. The WGNE assessment of short-term quantitative precipitation forecasts. *Bull. Amer. Meteor. Soc.* **84**(4), 481–492.

Eresmaa, R., Järvinen, H., Niemelä, S. and Salonen, K. 2007. Azimuthal asymmetry in ground-based GPS slant delay observations and their NWP model counterparts. *Atmos. Chem. Phys. Discuss.* **7**, 3179–3202.

Faccani, C., R., F., Pacione, R., Paolucci, T., Vespe, V. and co-authors. 2005. Impact of a high density GPS network on the operational forecast. *Adv. Geosci.* **2**, 73–79.

FAO 1995. Digital soil map of the world and derived soil properties, *Technical Report 1*, FAO, Rome.

Fischer, C., Montmerle, T., Berre, L., Auger, L. and Stefanescu, S. E. 2006. An overview of the variational assimilation in the ALADIN/FRANCE NWP system. *Quart. J. Roy. Meteorol. Soc.* **131**, 3477–3492.

Gauthier, P. and Thepaut, J.-N. 2001. Impact of the digital filter as a weak constraint in the pre-operational 4DVAR assimilation system of Meteo-France. *Mon. Wea. Rev.* **129**, 2089–2102.

Gendt, G., Dick, G. and Söhne, W. 1999. GFZ Analysis Center of IGS- Annual report 1998. *IGS 1998 Technical Reports* pp. 79–87. Jet Propulsion Laboratory, Pasadena.

Gleckler, P. J., Taylor, K. E. and Doutriaux, C. 2008. Performance metrics for climate models.. *J. Geophys. Res.* **113**, D06104, doi:10.1029/2007JD008972.

Grell, G. A., Dudhia, J. and Stauffer, R. R. 1995. A description of the fith-generation penn state/NCAR mesoscale model (MM5), *NCAR Technical Note TN-398+STR*, NCAR, P.O. Box 3000, Boulder, CO, 80307. 122 pp.

Grzeschik, M., Bauer, H.-S., Wulfweyer, V., Engelbart, D., Wandinger, U. and co-authors. 2008. Four-dimensional analysis of water-vapor Raman lidar data and their impact on mesoscale forecasts. *J. Atmos. Oceanic Technol.* **25**, 1437–1453.

Gustafsson, N. 2007. Discussion on '4D-Var or EnKF?' *Tellus* **59A**(5), 774–777.

Ha, S., Kuo, Y.-H., Guo, Y.-R. and Lim, G.-H. 2003. Variational assimilation of slant-path delay measurements from a hypothetical ground-based GPS network. Part I: comparison with precipitable water assimilation. *Mon. Wea. Rev.* **131**, 2635–2655.

Hagemann, S., Bengtsson, L. and Gendt, G. 2003. On the determination of atmospheric water vapor from GPS measurements. *J. Geophys. Res.* **108**(D21), 4678–4691, doi:10.1029/2002JD003235.

Hohenegger, C. and Schär, C. 2007. Predictability and error growth dynamics on cloud-resolving models. *J. Atmos. Sci.* **64**, 4467–4478.

Hong, S.-Y. and Pan, H.-L. 1996. Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Wea. Rev.* **124**, 2322–2339.

Houtekamer, P. L., Mitchell, H. L. and Deng, X. 2009. Model error representation in an operational ensemble Kalman filter. *Mon. Wea. Rev.* **137**(7), 2126–2143.

Houtekamer, P. L., Mitchell, H. L., Pellerin, G., Buehner, M., Charron, M. and co-authors. 2005. Atmospheric data assimilation with the ensemble Kalman filter: results with real observations. *Mon. Wea. Rev.* **133**, 604–620.

Huang, X.-Y., Q., X., Barker, D. M., Zhang, X., Michalakes, J. and co-authors. 2009. Four-dimensional variational data assimilation for WRF: formulation and preliminary results. *Mon. Wea. Rev.* **137**, 299–314.

Ide, K., Courtier, P., Ghil, M. and Lorenc, A. C. 1997. Unified notation for data assimilation. operational, sequential and variational. *J. Meteorol. Soc. Japan* **75**, 181–189.

Järvinen, H., Reima, E., Vedel, H., Kirsti, S., Sami, N. and co-authors. 2007. A variational data assimilation system for ground-based GPS slant delays. *Quart. J. Roy. Meteorol. Soc.* **133**, 969–980.

Kain, J. S. 2004. The Kain-Fritsch convective parameterization: an update. *J. Appl. Meteorol.* **43**, 170–181.

Kain, J. S., Weiss, S. J., Bright, D. R., Baldwin, M. E., Levit, J. J. and co-authors. 2008. Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecast.* **23**, 931–952.

Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C. and Ballabrera-Poy, J. 2007a. 4D-Var or ensemble Kalman filter?. *Tellus* **59A**(5), 758–773.

Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C. and Ballabrera-Poy, J. 2007b. Response to the discussion on "4D-Var or EnKF?" by Nils Gustafsson. *Tellus* **59A**(5), 778–780.

Kawabata, T., Seko, H., Saito, K., Kuroda, T., Tamiya, K. and co-authors. 2007. An assimilation and forecasting experiment of the Nerima heavy rainfall with a cloud-resolving nonhydrostatic 4-dimensional variational data assimilation system. *J. Meteorol. Soc. Japan* **85**, 255–276.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Martinuzzi, J., Font, J. and co-authors. 2001. Soil moisture retrieval from space: the Soil Moisture OceanSalinity (SMOS) mission. *IEEE Trans Geosci Remote Sens* **39**(8), 1729–1735.

Küll, V. and Bott, A. 2009. Application of the hybrid convection parameterization scheme HYMACS to different meteorological situations. *Atmos. Res.* **94**, 743–753.

MacDonald, A., Xie, Y. and Ware, R. 2002. Diagnosis of three dimensional water vapor using slant observations from a GPS network. *Mon. Wea. Rev.* **130**, 386–397.

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

Rawlins, F., Ballard, S. P., Bovis, K. J., Clayton, A. M., li, D. and co-authors. 2007. The Met Office global four-dimensional variational data assimilation scheme. *Quart. J. Roy. Meteorol. Soc.* **133**, 347–362.

Reisner, J., Rasmussen, J. and Bruintjes, R. T. 1998. Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. Roy. Meteorol. Soc.* **124B**, 1071–1107.

Richard, E., Buzzi, A. and Zängl, G. 2007. Quantitative precipitation forecasting in the Alps: the advances achieved by the Mesoscale Alpine Programme. *Quart. J. Roy. Meteorol. Soc.* **133**, 831–846.

Rossa, A., Nurmi, P. and Ebert, E. 2008. Precipitation: Advances in Measurement, Estimation and Prediction edition, Springer, Berlin Heidelberg, 419–452.

Rotach, M. W., Arpagaus, M., Dorninger, M., Hegg, C., Montani, A. and co-authors. 2009. MAP D-PHASE: real-time demonstration of weather forecast quality in the Alpine region. *Bull. Amer. Meteor. Soc.* **90**, 1321–1336.

Ruggiero, F. H., Michalakes, J. G., Nehrkorn, T., Modica, G. M., Cerniglia, M. and co-authors. 2002. Performance of the scalable mesoscale four dimensional variational assimilation system, *Technical report*, Department of Defense High Performance Computing Modernization Program. UGC 2002, Austin, Texas, U.S.

Saastamoinen, J. 1972. Atmospheric correction for troposphere and stratosphere in radio ranging of satellites. In *The Use of Artificial Satellites for Geodesy* Volume 15 eds S. W. Henriksen, A. Mancini and B. H. Chovitz). Washington DC American Geophysical Union Geophysical Monograph Series, 247–251.

Saito, K., Ishida, J., Aranami, K., Hara, T., Segawa, T. and co-authors. 2007. Nonhydrostatic atmospheric models and operational development at JMA. *J. Meteorol. Soc. Japan* **85B**, 271–304.

Schraff, C. and Hess, R. 2003. A description of the non-hydrostatic regional model LM. Part III: data assimilation. *LM Documentation* **3**. 85 p.

Schwitalla, T., Bauer, H.-S., Wulfmeyer, V. and Aoshima, F. 2010. High-resolution simulation over central Europe: assimilation experiments with WRF 3DVAR during COPS IOP9c. *Quart. J. Roy. Meteorol. Soc.* (COPS Special Issue), in press.

Schwitalla, T., Bauer, H.-S., Wulfmeyer, V. and Zängl, G. 2008. Systematic errors of QPF in low mountain regions. *Meteorol. Zeitschrift* **17**, 903–917.

Skamarock, W. C. and Klemp, J. B. 2007. A time-split nonhydrostatic atmospheric model for research and NWP applications. *J. Comp. Phys.* 3465–3485. Special issue on environmental modeling.

Stauffer, D. R. and Seaman, N. L. 1994. Multiscale four dimensional data assimilation. *J. Appl. Meteorol.* **33**, 416–434.

Vedel, H. and Huang, X.-Y. 2004. Impact of ground based GPS data on numerical weather prediction. *J. Meteorol. Soc. Japan* **82**(1B), 459–472.

Wang, X., Barker, D. A., Snyder, C. and Hamill, T. A. 2008a. A hybrid ETKF-3DVAR data assimilation scheme for the WRF model.

Part I: observing system simulation experiment. *Mon. Wea. Rev.* **136**, 5116–5131.

Wang, X., Barker, D. A., Snyder, C. and Hamill, T. A. 2008b. A hybrid ETKF-3DVAR data assimilation scheme for the WRF model. Part II: real observation experiments. *Mon. Wea. Rev.* **136**, 5132–5147.

Weckwerth, T. M. and Parsons, D. B. 2006. A review of convection initiation and motivation for IHOP_2002. *Mon. Wea. Rev.* **134**, 5–22.

Weckwerth, T. M., Parsons, D. B., Koch, S. E., Moore, J. A., LeMone, M. A. and co-authors. 2004. An overview of the international h2o project (IHOP_2002) and some preliminary highlights. *Bull. Amer. Meteor. Soc.* **85**, 253–277.

Wee, T.-K. and Kuo, Y.-H. 2004. Impact of a digital filter as a weak constraint in MM5 4DVAR: an observing system simulation experiment. *Mon. Wea. Rev.* **132**, 543–559.

Wickert, J. and Gendt, G. 2006. Fernerkundung der Erdatmosphäre mit GPS. *Promet* **32**(3/4), 176–184.

Wiener, N. 1949. *Extrapolation, Interpolation and Smoothing of Stationary Time Series Wiener*, N. John Wiley, New York.

WMO 2009. WWRP Strategic Plan 2009-2017. available online: http://www.wmo.int/pages/prog/arep/wwrp/new/documents/final_WWRP_SP_6_oct.pdf.

Wu, W., Lynch, A. H. and Rivers, A. 2005. Estimating the uncertainty in a regional climate model related to initial and lateral boundary conditions. J. Climate **18**, 917–933.

Wulfmeyer, V., Bauer, H.-S., Grzeschik, M., Behrendt, A., Vandenberghe, F. and co-authors. 2006. Four-dimensional variational assimilation of water vapor differential absorption lidar data: the first case study within IHOP_2002. *Mon. Wea. Rev.* **134**(1), 209–230.

Wulfmeyer, V., Behrendt, A., Bauer, H.-S., Kottmeier, C., Corsmeier, U. and co-authors. 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.

Xiao, Q. and Sun, J. 2007. Multiple radar data assimilation and short-range QPF applied to a squall line observed during IHOP_2002. *Mon. Wea. Rev.* **135**, 3381–3404.

Xiao, Q., Zou, X. and Wang, B. 2000. Initialization and simulation of a land-falling hurricane using a variational bogus data assimilation scheme. *Mon. Wea. Rev.* **218**, 2252–2269.

Xue, M. and Martin, W. J. 2006a. A high-resolution modeling study of the 24 may 2002 case during IHOP. Part I: numerical simulation and general evolution of the dryline and convection. *Mon. Wea. Rev.* **134**(1), 149–171.

Xue, M. and Martin, W. J. 2006b. A high-resolution modeling study of the 24 may 2002 case during IHOP. Part II: horizontal convective rolls and convective initiation. *Mon. Wea. Rev.* **134**(1), 172–191.

Zängl, G. 2002. An improved method for computing horizontal diffusion in a sigma-coordinate model and its application to simulations over mountainous topography. *Mon. Wea. Rev.* **130**, 1423–1432.

Zou, X., Kuo, Y.-H. and Guo, Y.-R. 1995. Assimilation of atmospheric radio refractivity using a nonhydrostatic adjoint model. *Mon. Wea. Rev.* **123**, 2229–2249.

Zus, F., Grzeschik, M., Bauer, H.-S., Wulfmeyer, V., Dick, G. and co-authors. 2008. Development and optimization of the IPM MM5 GPS slant path 4DVAR system. *Meteorol. Zeitschrift* **17**, 867–885.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Numerical algorithm, tangent linear and adjoint.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.