Application of Global Positioning System slant path delay data for mesoscale model verification and four-dimensional variational assimilation

Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

Fakultät Naturwissenschaften
Universität Hohenheim

Institut für Physik und Meteorologie (IPM)

vorgelegt von
Florian Zus

aus Graz, Österreich
2009

Dekan: Prof. Dr. rer. nat. Heinz Breer

1. berichtende Person: Prof. Dr. rer. nat. Volker Wulfmeyer

2. berichtende Person: Prof. Dr. George Craig

Eingereicht am: 6.8.2009

Mündliche Prüfung am: 24.3.2010

Die vorliegende Arbeit wurde am 6.8.2009 von der Fakultät Naturwissenschaften der Universität Hohenheim als “Dissertation zur Erlangung des Doktorgrades der Naturwissenschaften” angenommen.

Abstract

Observation modeling is required in order to make use of slant path delay data, processed from ground-based Global Positioning System (GPS) measurements, for verification purposes and numerical weather prediction. A rigorous ray-tracing algorithm based on the Euler-Lagrange equation derived from Fermat's principle is developed to simulate the propagation of GPS radio signals in a mesoscale model. The ray-tracing algorithm is based on a finite difference scheme and allows the direct numerical simulation of GPS slant path delays. A statistical intercomparison of simulated and observed data for a large network of continuously operating ground-based receivers in Germany indicates a good agreement over the entire elevation range. The estimated RMS of the integrated water vapor is 1.3 mm.

The sensitivity of simulated slant path delays with respect to forecast model components of the Mesoscale Model 5 (MM5) is studied. It is demonstrated that the model minus observation statistics of the slant path delay data crucially depends on the parameterization of deep convection and the implementation of horizontal diffusion. The experiment indicates that slant path delay observations are potentially useful for the verification of forecast model components.

The development of tools for 4-dimensional variational assimilation (4DVAR) of slant path delay data in the framework of the MM5 system is described. A computational fast forward model and its adjoint were implemented into the message passing interface environment, such that assimilation of slant path delay data became routinely possible. A set of modifications to the existing non-linear, tangent-linear and adjoint forecast model is presented. They include an improvement of the horizontal diffusion scheme and the implementation of the Grell cumulus scheme in order to eliminate the observed systematic deviation in the model minus observation statistics of the slant path delay data and precipitation.

A single assimilation experiment was conducted and is analyzed in detail. The impact is area wide and reaches ± 1 gkg^{-1} in the water vapor mixing ratio, ± 0.5 K in the temperature and ± 0.5 ms^{-1} in the wind field. Slant path delay data of an independent network of receivers and radiosonde measurements confirm the results gained by 4DVAR. An assimilation experiment over a period of two weeks indicates that the impact in observation space of an independent network of receivers is positive. The experiment reveals, that, the impact in the humidity content lasts for about 12 h. Finally, an assimilation experiment over a period of one month was performed. The impact of the assimilation of slant path delay data indicates a positive impact on the precipitation forecast.

Zusammenfassung

Slant Path Delay-Daten, die aus bodengestützten Global Positioning System (GPS)-Messungen abgeleitet werden, können zur Verifikation von mesoskaligen Modellen und zur numerischen Wettervorhersage verwendet werden. Dazu ist die Entwicklung eines Vorwärtsmodells notwendig. Ein rigoroser Ray Tracing-Algorithmus basierend auf den Euler-Lagrange-Gleichungen abgeleitet aus Fermat's Prinzip wurde entwickelt, um die Ausbreitung von GPS Radiosignalen in einem mesoskaligen Modell zu simulieren. Der Ray Tracing-Algorithmus basiert auf einem Finite Differenzen-Verfahren und erlaubt die direkte numerische Simulation von GPS Slant Path Delay-Daten. Ein statistischer Vergleich simulierter und beobachteter Daten für ein kontinuierlich operierendes Netzwerk von bodengestützten Empfängern in Deutschland zeigt eine gute Übereinstimmung über den gesamten Elevationsbereich. Der abgeschätzte RMS im integrierten Wasserdampf ist 1.3 mm.

Die Sensitivität simulierter Slant Path Delay-Daten in Abhängigkeit von Modellkomponenten des mesoskaligen Modells 5 (MM5) wird untersucht. Es wird demonstriert, dass die Modell-minus-Beobachtungs-Statistik der Slant Path Delay-Daten entscheidend von der Konvektionsparameterisierung und der Implementierung der horizontalen Diffusion abhängt. Das Experiment zeigt, dass Slant Path Delay-Daten zur Verifikation von Modellkomponenten genützt werden können.

Die Entwicklung von Werkzeugen zur 4-dimensionalen variationellen Assimilation (4DVAR) von Slant Path Delay-Daten in das MM5 System wird beschrieben. Ein effizientes Vorwärtsmodell und das adjungierte Modell wurden in die Message Passing Interface-Umgebung implementiert, sodass eine routinemäßige Assimilation von Slant Path Delay-Daten möglich wurde. Ein Satz von Modifikationen des existierenden nicht-linearen, tangent-linearen und adjungierten Modells wird präsentiert. Ein verbessertes horizontales Diffusionschema wurde entwickelt und das Konvektionsschema von Grell wurde implementiert, um die systematische Abweichung in der Modell-minus-Beobachtungs-Statistik der Slant Path Delay-Daten und im Niederschlag zu eliminieren.

Ein einzelnes Assimilationsexperiment wurde durchgeführt und wird im Detail analysiert. Der Einfluss ist flächendeckend und erreicht ± 1 gkg^{-1} im Wasserdampf-Mischungsverhältnis, ± 0.5 K in der Temperatur und ± 0.5 ms^{-1} im Windfeld. Slant Path Delay-Daten eines unabhängigen Netzwerkes von Empfängern und Radiosondenmessungen bestätigen den positiven Einfluss der 4DVAR. Ein Assimilationsexperiment über eine Periode von zwei Wochen zeigt, dass der Einfluss im Beobachtungsraum eines unabhängigen Netzwerkes von Empfängern positiv ist. Das Experiment zeigt, dass der Einfluss im Wasserdampfgehalt 12 Stunden anhält. Ein Assimilationsexperiment über eine Periode von einem Monat wurde durchgeführt. Der Einfluss der Assimilation der Slant Path Delay Daten auf die Niederschlagsvorhersage ist positiv.

Contents

1 Introduction

The interest of atmospheric science in observations derived from the Global Positioning System (GPS) is caused by their sensitivity to water vapor. Water vapor is of key importance in weather prediction (Crook, 1996; Weckwerth, 2000) and climate research (Solomon et al., 2007). It is coupled to cloud and rain formation, and solar and terrestrial radiative transfer. Currently there is lack of knowledge about the actual humidity field, both due to a shortage of observations and a sub-optimal handling in data assimilation systems being used to estimate the state of the atmosphere. Knowledge of the state of the atmosphere is required to initialize numerical weather prediction (NWP) models, which provide the basis for modern weather forecasting. The availability of precise and continuous observations in data assimilation systems is crucial.

In variational assimilation, the model background field, which is typically a short-range model forecast from a previous analysis, is updated with observations in a statistically optimal way. Observations can be divided into conventional, in-situ observations, and non-conventional, remote sensing observations. The latter can be done in two different ways: passive and active. The various observation types have different characteristics, such as resolution, spatial and temporal coverage, which determines their ability to affect the analysis. Conventional observations, such as radiosonde observations, surface observations, ship and buoy observations or aircraft observations provide useful in-situ measurements. Radiosonde observations for example directly provide pressure, temperature, wind and water vapor information at specific locations. To date, radiosonde observations are one of the most important information sources for NWP models. Nevertheless, the networks of conventional observations are spatially and temporally sparse, e.g. radiosondes are launched typically only twice a day. On the other hand, remote sensing observations such as those from satellites have a very good spatial and temporal coverage. Although satellite data is slightly less accurate than conventional observations, their great advantage is their broad geographical coverage. Consequently there is a strong benefit from satellite observations and the influence of conventional observations becomes less critical (Persson and Grazzini, 2005).

However, remote sensing observations are usually not equal to model variables, and therefore observation modeling can be a difficult task. Observation modeling is required to determine the model equivalents of observations, i.e. it allows to simulate the observed quantity in the model. For in-situ measurements observation modeling is less of an undertaking. Typically it consists of interpolation from the gridded model field to the observation location, and conversion from model variables to the observed quantities. The implementation of conventional observations in a variational assimilation frame work is straight forward. For remote sensing measurements, observation modeling can be very complex.

Initially designed as a navigation aid for the military service, the number of high precision civilian applications of the GPS has grown with time and includes remote sensing of the atmosphere using ground-based (Bevis et al., 1992) and spaceborne techniques (Kursinski et al., 1996). The GPS space segment consists of 32 satellites distributed in 6 orbital planes about the globe. The satellites orbit the Earth at an altitude of 20200 km with a circulation time of about 12 h. Figure 1 provides a sketch of the satellite constellation. Each satellite continuously transmits radio signals at two frequencies, which are received by suitably equipped low earth orbit satellites and ground-based stations. The radio signals are changed in a characteristic way when they pass through the atmosphere. The phase delays, caused by the signal refraction in the neutral atmosphere, contain signatures of the atmospheric humidity field and can be considered as meteorological observations. The ability to use the GPS to perform accurate measurements related to the atmospheric humidity has led to the development of a promising new active remote sensing system.

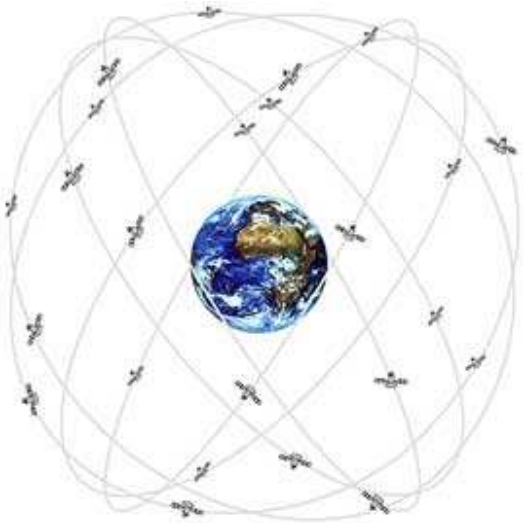


Figure 1: The GPS satellite orbits (URL http://www.kowoma.de/en/gps/orbits.htm). The inclination angle of the orbital planes towards the equator is 55°, the orbital planes are rotated in the equatorial plane by 60° against each other.

The first application of ground-based GPS meteorology involved the measurement of integrated water vapor in the atmosphere above receiver stations (Bevis et al., 1994). This technique has several advantages over conventional water vapor observing systems including low cost, high measurement accuracy, all weather operability, and long-term measurement stability. In addition, it requires no external calibration and operates unattended for long periods. One distinct advantage of GPS measurements compared to other satellite observations is that it provides data in all weather conditions, i.e. the measurements are not affected by clouds. As water vapor is often under-observed both in time and space during active weather, this capability of the GPS data is expected to improve

the skill of short range predictions of medium to heavy rainfall systems. To date, most of the ground-based data assimilation studies have been conducted using the so called zenith total delays or integrated water vapor estimates. It was found that their assimilation has a weak, but mainly positive impact on the NWP forecasts of precipitation (Vedel and Huang, 2004; Poli et al., 2007).

Processing techniques had to be developed to determine slant path delays, i.e. phase delays between ground-based receivers and each of a number of satellites in view (Alber et al., 2000; Braun et al., 2001). Compared with zenith total delays, slant path delays contain azimuthally asymmetric information and near-horizon slant path delays potentially contain atmospheric profile information (Sokolovskiy et al., 2001; Lowry et al., 2002). It may be possible to recover a significant amount of information about the horizontal and vertical structure of the atmosphere from slant path delays. However, the processing techniques developed so far are at an early stage. In addition, it has been recently recognized that accurate observation modeling, i.e. a ray-tracing algorithm, is required in order to completely utilize slant path delay observations (Eresmaa et al., 2008; Hobiger et al., 2008). Ray-tracing is necessary to simulate the propagation of the radio signals through the model atmosphere. So far, the proposed ray-tracing algorithms contain partly numerous simplifying assumptions. All of them can essentially be considered as being solvers for initial value problems. However, one is concerned with a boundary value problem, determined by the position of the satellite and the receiver, rather than with an initial value problem.

Partly due to a number of unsolved questions concerning processing of slant path delays and possibly due to the complexity concerning observation modeling, the assimilation experiments conducted so far have focused on the use of hypothetical observations (MacDonald et al., 2002; Ha et al., 2003). These case studies showed that the assimilation of hypothetical slant path delays have the potential to improve the retrieval of the humidity field compared to the assimilation of hypothetical zenith total delays. Järvinen et al. (2007) analyzed the data assimilation of hypothetical and real slant path delays for a limited number of ground based receivers. Since one of the benefits of slant path delays is their ability to capture asymmetric humidity structures present in the atmosphere (Eresmaa et al., 2007), the assimilation system was able to reproduce the asymmetric information content in the hypothetical observations. Real slant path delay observations were assimilated for an arbitrary single case and the resulting humidity analysis increment was found to have a similar structure as the humidity analysis increment obtained with a comparable radiosonde network.

While the results presented so far indicate that the measurements from ground-based networks are potentially valuable for short-range prediction of precipitation, it is important to recognize that the assimilation of observations related to water vapor in a convective and/or pre-convective environment is highly challenging. One major difficulty is the accuracy of model physics and numerics. In particular, the convection parameterization can

be regarded as an issue. Besides model physics, deficiencies in model numerics can be substantial (Zängl, 2002, 2004). Prior to any data assimilation efforts, systematic errors must be identified, quantified and removed. However, it is not trivial to separate the contribution of errors due to initial fields, errors in model physics and numerics, errors in observation modeling and errors inherent in the observations. A close inspection of a long time series of the model minus observation statistics and sensitivity analysis are one strategy to identify errors. Corresponding research is subject of several ongoing projects of the World Weather Research Program (WWRP) such as the Research and Development Project Convective and Orographically-induced Precipitation Study (COPS). This study provides a comprehensive set of observations for testing hypothesis on the improvement of Quantitative Precipitation Forecasting (QPF) in regions with complex terrain (Wulfmeyer et al., 2008).

Improvements of NWP can only be expected if progress is made on three research areas simultaneously: Handling of problems related to model numerics such as the treatment of the governing equations in complex terrain, the improvement of parameterization schemes and the optimization of initial conditions. All three research areas require the availability of reliable observations with good spatial and temporal coverage and sophisticated algorithms in observation modeling. This thesis is intended to make a step in this direction, taking advantage of slant path delay data provided by a large and continuously operating network of ground-based receivers.

Figure 2 shows a block diagram illustrating the research strategy. The first step consists of monitoring the observational data against their model equivalents. Therefore the background (a previous analysis) is mapped to observation space using the forward model. The monitoring can be used to study the sensitivity of simulated data with respect to forward model components and/or components in the processing of the observational data. The verification follows the monitoring and can be used to study the sensitivity of simulated data with respect to forecast model components. The strategy is to a) identify and b) to remove forecast model errors prior to any data assimilation effort by statistical comparisons of simulated and observed data. The last step consists of assimilating observational data. This step provides the analysis which is used for NWP. The feedback allows to study the sensitivity of the analysis with respect to forward/forecast model components. Following this resaerch strategy, this thesis is organized as follows:

In chapter 2, a brief description of three-dimensional variational assimilation (3DVAR) and four-dimensional variational assimilation (4DVAR) is provided. The basic steps to incorporate slant path delays into the MM5 4DVAR are described.

In chapter 3, the basic principles of GPS ground-based observation processing are provided. The phase pseudorange observation equation is introduced, the definition of the slant path delay is provided and the relation to atmospheric quantities is described. Information on current activities at GPS processing centers in Europe is provided and the

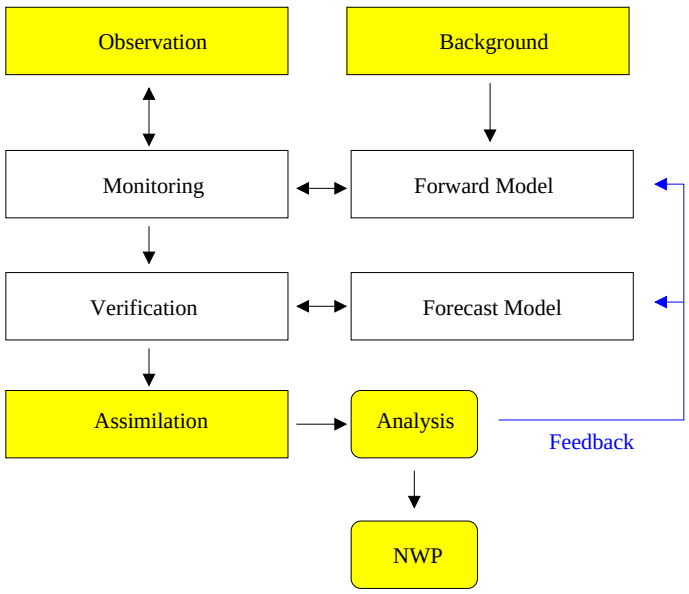


Figure 2: Block diagram illustrating the research strategy. The first step consists of monitoring, the second step includes verification and the final step is the assimilation.

GPS products processed at the German Research Center for Geosciences (GFZ) Potsdam are described.

In chapter 4, the basic equations determing the signal path form the satellite to the ground based receiver are derived from electromagnetic wave propagation in an inhomogeneous medium. A rigorous ray-tracing algorithm, based on a global approach, is described which allows the direct numerical simulation of the signal path and the determination of the slant path delay in a mesoscale model. As an alternative to the rigorous computational model for the slant path delay, a fast computational model is described. Potential applications of the ray tracing algorithm are discussed.

In chapter 5, the MM5 configuration and the GPS network used in this thesis are described. The proposed ray-tracing algorithm is validated using an alternative ray-tracing algorithm and an analytical solution. The main features of slant path delays are discussed, and the potential azimuthally asymmetric information content is estimated. A statistical intercomparison of simulated and observed slant path delays is performed. The sensitivity of simulated slant path delays with respect to forward model components is investigated.

In chapter 6, the sensitivity of simulated slant path delays with respect to two forecast model components, namely the horizontal diffusion scheme and the parameterization of deep convection, is investigated. A modified horizontal diffusion scheme is presented. A statistical intercomparison of simulated and observed slant path delays is performed.

In chapter 7, a description of basic principles of tangent linear and adjoint code construction is given. Tangent linear and adjoint code construction are illustrated for the

forward model and a forecast model component. The impact of a single slant path delay in variational assimilation is estimated analytically. The effect of an error in a forecast model component in 4DVAR is illustrated.

In chapter 8, a description of the implementation in the MM5 4DVAR Message Passing Interface (MPI) environment is provided. Problems encountered in model physics and model numerics are discussed and a set of modifications to the existing MM5 4DVAR system is presented. The forward model error and the observation error are estimated. A single case study is discussed in detail. The impact of the assimilation in observation and model space is analyzed. An assimilation experiment over a period of two weeks is performed and the impact in the humidity content is studied. An assimilation experiment over a period of one month is performed to study the performance on the precipitation forecast. The main conclusion and a summary of the results are provided at the end of this thesis.

2 Data assimilation

Data assimilation is an analysis technique in which information from observations is incorporated into the model state. The basic information that can be used to produce an analysis is a collection of observations $\mathbf{y}$. In most cases the analysis is under-determined because observations are sparse and only indirectly related to variables of the model state $\mathbf{x}$. In order to make it a well-posed problem, it is necessary to rely on some background information in the form of an a priori estimate of the model state, called the background state $\mathbf{x}_b$. To use observations in the analysis procedure it is necessary to be able to compare them with the model state. This is accomplished through the use of a projection from model space to observation space called the observation operator H. Neither the background state $\mathbf{x}_b$ nor the observations $\mathbf{y}$ are perfect which is accounted for through the background errors $\varepsilon_b = \mathbf{x}_b - \mathbf{x}_t$ and the observations errors $\varepsilon_o = \mathbf{y} - H[\mathbf{x}_t]$, where $\mathbf{x}_t$ denotes the true state.

2.1 Three-Dimensional Variational Assimilation (3DVAR)

The principle of the variational assimilation consists in solving for the most probable model state $\mathbf{x}_a$ defined both by the background state $\mathbf{x}_b$ and the observations $\mathbf{y}$. Bayes theorem states that (Bouttier and Courtier, 1999)

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y})} \quad \propto \quad p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x}) \tag{1}$$

where $p(\mathbf{x})$ denotes the a priori probability of the model state $\mathbf{x}$, $p(\mathbf{y})$ the a priori probability of the observations $\mathbf{y}$, $p(\mathbf{y}|\mathbf{x})$ the conditional probability of the observations $\mathbf{y}$ given the model state $\mathbf{x}$ and $p(\mathbf{x}|\mathbf{y})$ the conditional probability of the model state $\mathbf{x}$ given the observations $\mathbf{y}$. Assuming that both background errors and observation errors have the Gaussian statistics and are independent on each other, the conditional probability $p(\mathbf{x}|\mathbf{y})$ explicitly reads as

$$p(\mathbf{x}|\mathbf{y}) \quad \propto \quad \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T\, \mathbf{B}^{-1}\, (\mathbf{x}-\mathbf{x}_b) - \frac{1}{2}(\mathbf{y}-H[\mathbf{x}])^T\, \mathbf{R}^{-1}\, (\mathbf{y}-H[\mathbf{x}])\right) \tag{2}$$

where $\mathbf{R}$ is the observation error covariance matrix and $\mathbf{B}$ is the background error covariance matrix, determined through the covariances

$$\begin{aligned} \mathbf{R} &= \langle \varepsilon_o \, \varepsilon_o^T \rangle \\ \mathbf{B} &= \langle \varepsilon_b \, \varepsilon_b^T \rangle \end{aligned} \tag{3}$$

where the brackets $\langle\rangle$ denote the expectation value. The most probable model state is given by the minimum of the cost function

$$J[\mathbf{x}] = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \, \mathbf{B}^{-1} \, (\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(\mathbf{y}-H[\mathbf{x}])^T \, \mathbf{R}^{-1} \, (\mathbf{y}-H[\mathbf{x}]) \tag{4}$$

The determination of the minimum of the cost-function requires knowledge of its gradient

$$\nabla J[\mathbf{x}] = \mathbf{B}^{-1} \, (\mathbf{x}-\mathbf{x}_b) - \mathbf{H}[\mathbf{x}]^T \, \mathbf{R}^{-1} \, (\mathbf{y}-H[\mathbf{x}]) \tag{5}$$

Here $\mathbf{H}$ denotes the derivative of H with respect to $\mathbf{x}$. The most propable model state $\mathbf{x}_a$ is determined through

$$\nabla J[\mathbf{x}_a] = 0 \tag{6}$$

Equation (4) and (5) are the basis for 3DVAR.

2.2 Four-Dimensional Variational Assimilation (4DVAR)

4DVAR is a generalization of 3DVAR for observations that are distributed in time. The underlying equations are the same. The mapping from model to observation space is generalized to include the forecast model that will allow a comparison between the model state and the observations at the appropriate time. The minimum of the following cost function

$$J[\mathbf{x}] = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}\sum_{i=1}^{m} (\mathbf{y}_i - H_i[\mathbf{x}_i])^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i[\mathbf{x}_i]) \tag{7}$$

is determined where the subscript i denotes the time step. The model state $\mathbf{x}_i$ is determined from the initial state $\mathbf{x}$ through the forecast model M

$$\mathbf{x}_i = M_{0\to i}[\mathbf{x}] \tag{8}$$

The forecast model can be expressed as a sequence of forecast model steps. If $\mathbf{x}_0 = \mathbf{x}$ so that M_0 is the identity, then by denoting M_i the forecast model step form $i-1$ to i we have $\mathbf{x}_i = M_i[\mathbf{x}_{i-1}]$ and by recurrence

$$\mathbf{x}_i = M_i\, M_{i-1} ... M_1[\mathbf{x}] \tag{9}$$

The gradient of the cost function can be derived by rigorous application of the chain rule of differential calculus

$$\nabla J[\mathbf{x}] \;=\; \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) - \sum_{i=1}^{m} \mathbf{M}_1[\mathbf{x}]^T ... \mathbf{M}_i[\mathbf{x}_{i-1}]^T\, \mathbf{H}_i[\mathbf{x}_i]^T\, \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i[\mathbf{x}_i]) \tag{10}$$

Here $\mathbf{H}_i$ denotes the derivative of H_i with respect to $\mathbf{x}_i$ and $\mathbf{M}_i$ denotes the derivative of $\mathbf{x}_i$ with respect to $\mathbf{x}_{i-1}$. Equation (7) and (10) are the basis for 4DVAR.

2.3 The computation of the cost function and its gradient

The evaluation of the cost function and the gradient of the costfunction requires one forward integration and one suitably arranged backward integration. The forward integration consists of computing the forecast model state $\mathbf{x}_i$

$$\mathbf{x}_i = M_i\, M_{i-1} ... M_1[\mathbf{x}] \tag{11}$$

the normalized departure $\mathbf{d}_i$

$$\mathbf{d}_i = \mathbf{R}_i^{-1}(\mathbf{y}_i - H_i[\mathbf{x}_i]) \tag{12}$$

and the contribution to the cost function J_i

$$J_i = (\mathbf{y}_i - H_i[\mathbf{x}_i])^T\, \mathbf{d}_i \tag{13}$$

for $i = 1, ..., m$. Finally, the cost function is computed according to

$$J[\mathbf{x}] = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}\sum_{i=1}^{m} J_i \tag{14}$$

The backward integration consists of a recursive algorithm. Let $\delta\hat{\mathbf{x}}_i$ for $i = 0, ..., m$ denote a set of vectors where $\delta\hat{\mathbf{x}}_m = 0$. With

$$\delta\hat{\mathbf{x}}_{i-1} = \mathbf{M}_i[\mathbf{x}_{i-1}]^T \left(\delta\hat{\mathbf{x}}_i + \mathbf{H}_i[\mathbf{x}_i]^T \mathbf{d}_i\right) \tag{15}$$

for $i = m, ..., 1$ the gradient of the cost function can be written as

$$\nabla J[\mathbf{x}] \quad = \quad \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) - \delta\hat{\mathbf{x}}_0 \tag{16}$$

Due to the non-linearity of the forecast model and/or the forward model, the gradient of the cost function can be a highly non-linear function of the model state $\mathbf{x}$. The minimum of the cost-function must be determined iteratively, starting from a first guess. This is usually done using gradient descent algorithms or quasi-Newton methods. The limited memory quasi-Newton method L-BFGS (Liu and Nocedal, 1989) is widely used in minimization problems, because it is useful for optimization problems with a large number of unknowns and provides fast convergence. The L-BFGS is implemented in the MM5 4DVAR (Zou et al., 1997).

4DVAR has the following benefits over 3DVAR: (i) the observations are assimilated at appropriate time and (ii) the dynamics and physics of the forecast model are considered in assimilating data. As a result, observational data are optimally used in a meteorologically consistent way. However, it is important to recognize that 4DVAR works only under the assumption that the forecast model is close to perfect. Problems can be expected if forecast model errors are large (Bouttier and Courtier, 1999). For further discussion the reader is referred to section 7 of this thesis.

The principles of assimilation of slant path delays are the same as those of any other kind of observations. Three model components are required if a new observation type is to be incorporated into the MM5 4DVAR system: (i) the observation operator H (ii) the gradient of the observation operator $\mathbf{H}$, also referred to as the tangent linear observation operator and (iii) the transpose of the gradient of the observation operator $\mathbf{H}^T$, also referred to as the adjoint of the tangent linear observation operator. In addition, interfaces must be implemented to perform the relevant computations at the appropriate time step in the forward integration and the backward integration, following Equation (11) to Equation (16).

3 GPS observations

3.1 The phase pseudorange observation

In this section a brief description on the determination of ground-based GPS tropospheric delays from raw measurements is provided. Each GPS satellite continuously transmits radio signals at two frequencies $f_1 = 1575.42$ MHz and $f_2 = 1227.60$ MHz, corresponding to wavelengths of $L_1 = 19.0$ cm and $L_2 = 24.4$ cm respectively. The atmospheric effect on GPS signal propagation stems from the ionosphere on the one hand, and from the neutral atmosphere on the other. These two phenomena are known as ionospheric delay and tropospheric delay, respectively. The tropospheric delay is referred as the slant path delay. The phase pseudorange observation equation can be written as (Hofmann-Wellenhof et al., 1992)

$$\Theta = R + \lambda\,(\varphi_r - \varphi_s) + c\,(\delta_r - \delta_s) - \lambda N_i + S + I \tag{17}$$

where Θ is the phase pseudorange, i.e. the phase difference, in length units, between the received carrier wave and the referencing wave created by the receiver, R is the distance between the satellite and the receiver, λ is the wavelength of the carrier wave, φ_r and φ_s are the initial phases of the referencing wave and the carrier wave, respectively, c is the vacuum speed of light, δ_r and δ_s are the clock errors of the receiver and the satellite, respectively, N_i is the integer ambiguity, i.e. the number of full wavelengths between the satellite and the receiver, S denotes the slant path delay an I denotes the ionospheric delay. The ionospheric delay is dispersive and depends on the frequency of the radio signal. A linear combination of dual-frequency measurements (f_1 and f_2) allow to a large extent to eliminate the ionospheric delay. In the GPS observation processing the unknown parameters of the phase pseudorange observation equation, including the slant path delay, are solved using a least square adjustment.

3.2 The slant path delay

The slant path delay is proportional to the difference between the travel time of a signal from a satellite to a ground-based receiver and the travel time that would occur if there was no atmosphere affecting the signal propagation (Hofmann-Wellenhof et al., 1992). The slant path delay is expressed as the time difference multiplied by the vacuum speed of light and can be written as

$$S = \int_s n\,ds - \int_{s_0} ds_0 \tag{18}$$

where n is the atmospheric refractive index, s is the signal path through the atmosphere, and s_0 is the geometrical path, i.e. the hypothetical signal path that would occur without the atmosphere. The slant path delay can be written as

$$S = \int_s (n-1)\,ds + \int_s ds - \int_{s_0} ds_0 \tag{19}$$

The refractivity N is related to n through

$$N = 10^6 \cdot (n-1) \tag{20}$$

In the microwave range N is related to atmospheric quantities through

$$\begin{aligned} N &= N_d + N_w \\ &= k_1 \cdot \frac{P}{T} + k_2 \cdot \frac{e}{T^2} \end{aligned} \tag{21}$$

where P is the total pressure of air (hPa), e is the partial pressure of water vapour (hPa) and T is the temperature (K). The empirically-determined constants are given by $k_1 = 77.6\ \mathrm{K\,hPa^{-1}}$ and $k_2 = 3.73 \cdot 10^5\ \mathrm{K^2\,hPa^{-1}}$ (Ware et al., 1997). Even though N_d is usually referred to as the dry refractivity, the first term on the right hand side contains water vapor effects through pressure. It is referred to as dry to differentiate it from the wet refractivity N_w which includes the effects of water vapor only. The slant path delay can be written in terms of N_d and N_w as

$$S = \underbrace{10^{-6} \cdot \int_s N_d\,ds}_{S_d} + \underbrace{10^{-6} \cdot \int_s N_w\,ds}_{S_w} + \underbrace{\int_s ds - \int_{s_0} ds_0}_{S_g} \tag{22}$$

The first term on the right hand side is referred to as the slant dry delay S_d, the second term on the right hand side is referred to as the slant wet delay S_w and the third term on the right hand side is referred to as the slant geometric delay S_g. The slant wet delay contributes to about 5 percent to the slant path delay. In the zenith case, the slant path delay is referred to as Zenith Total Delay (ZTD). The magnitude of the ZTD is of the order of 2.5 m for a receiver located at mean sea level. Note that all three contributions to the slant path delay, i.e. S_d, S_w and S_g, depend on the total refractivity N through the signal path s. It is the dependency of the signal path s on the refractivity N which makes the simulation of the slant path delay in a mesoscale model not trivial. Ray-tracing is necessary to simulate the signal path from the satellite to the ground-based receiver.

3.3 Derived quantities

The slant wet delay S_w can be roughly related to Slant Water Vapor SWV according to

$$\text{SWV} = \Pi \cdot S_w \tag{23}$$

where the conversion parameter Π is a function of the weighted mean temperature of the atmosphere (Davis et al., 1985). Bevis et al. (1994) showed that $\Pi \sim 1/6$, which implies that 1 mm of integrated water vapor corresponds to a slant wet delay of about 6 mm. However, the actual value of Π can vary due to the dependency on the weighted mean temperature and it is therefore not recommended to use SWV for verification or assimilation purposes. In the zenith case, SWV is referred as Integrated Water Vapor IWV. Since IWV is closely related to Precipitable Water PW, ground-based GPS observations are tightly coupled to precipitation.

3.4 GPS observation processing

Several data processing centres in Europe have shown that the processing and delivery of tropospheric delay estimates is possible in near-real time. An overview of data processing centers providing ZTD estimates can be found at the EUMETNET GPS water vapour programme (E-GVAP) website (egvap.dmi.dk). E-GVAP was set up to provide its EUMETNET partners with ZTD estimates for operational usage. Currently data from more than 400 GPS receivers are being sent hourly to a common ftp-server at the UK Met Office. Figure 3 shows the GPS network status in Europe on July 14, 2009. The data are intended for operational and scientific purposes. Implementation of ground based GPS data in operational NWP has certain requirements regarding quality, homogeneity and stability. A key goal of E-GVAP is to gradually bring the ground based GPS near-real time data to meet those requirements.

The GFZ Potsdam is one of the data processing centers in Europe. The GPS network analyzed in near real-time at the GFZ consists presently of about 200 receivers in Germany and neighboring countries. The data analysis is performed on an hourly basis with a delay of about 30 minutes using the GFZ Earth Parameters and Orbit determination System (EPOS) software (Gendt et al., 1999a,b) which is based on a least squares adjustment of undifferenced phase measurements. The GPS data are processed in a sliding window mode with a window length of 12 hours and a forward step of one hour. The technique of parallel analysis of station clusters with the Precise Point Positioning (PPP) strategy (Zumberge et al., 1997) is implemented. The high quality orbits and clocks required by the PPP analysis are determined by further refining the International GNSS (Global

Navigation Satellite System) Service (IGS) routine products. All GPS data are taken with a sampling period of 2.5 minutes. The E-GVAP website shows the performance of the GFZ data analysis in relation to other processing centres and a NWP model.

The GFZ provides two products for atmospheric applications: ZTDs and slant path delays. The ZTD is a combined quantity based on all data taken during one hour by each station. The available GPS observations are analyzed in order to find an optimal parameter estimation for the ZTD data in 15 minutes steps. In addition, the IWV is obtained from the ZTD and surface observations (Dick et al., 2001). Another result of the GPS analysis are slant path delays, obtained using a least square adjustment. The least squares solution considers all individual slants path delays of a given station and makes use of the full temporal resolution of 2.5 minutes. During 2006 the GFZ started to analyze slant path delays operationally for about 130 GPS stations leading to about 1 million slant path delays per day. Data used in this thesis is based on the output of an operational version of EPOS during COPS in summer 2007.

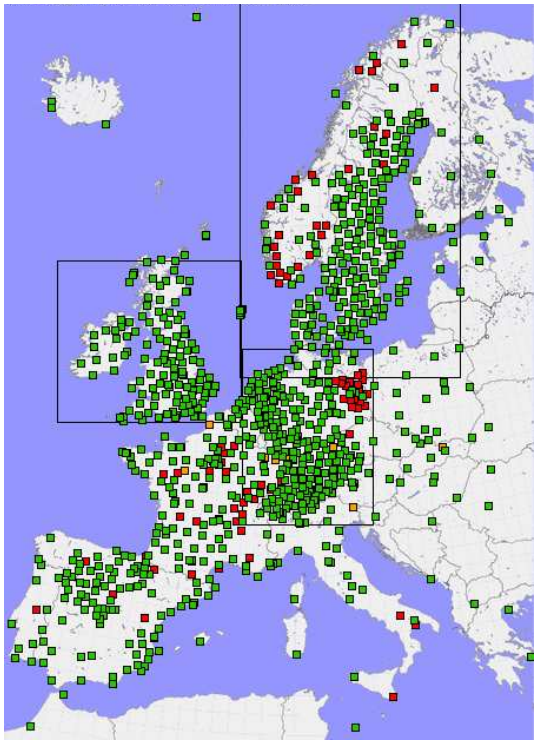


Figure 3: The GPS network status in Europe on July 14, 2009 (URL http://egvap.dmi.dk/). The geographical locations of the ground-based receivers are indicated by green and red dots.

4 Observation modeling

4.1 Propagation of radio signals in the Earth's atmosphere

In this section the basic equations governing the ray trajectory are derived from electromagnetic wave propagation in an inhomogeneous medium. The wavelengths of GPS signals (L_1 and L_2) are small compared to the characteristic scale lengths of atmospheric inhomogeneities which allows for the usage of geometrical optics.

4.1.1 The eikonal equation

An electromagnetic wave can be described by a complex valued function E varying over space and time. To simplify the description the wave function is assumed to be of the form

$$E(\mathbf{r}, t) = U(\mathbf{r}) \exp(-i\omega t) \tag{24}$$

where U denotes the amplitude, ω denotes the angular frequency, $\mathbf{r}$ denotes location in space and t denotes time. In this description, the effects of diffraction and polarization are neglected. From Maxwell's equations it follows that the spatial component of the wave is given by a Helmholtz equation:

$$\triangle U(\mathbf{r}) + k^2 n^2(\mathbf{r}) U(\mathbf{r}) = 0 \tag{25}$$

where n is the refractive index of the medium, $k = 2\pi/\lambda$ is the free-space wave number and λ is the wave length. We seek for a solution in the following form

$$U(\mathbf{r}) = A(\mathbf{r}) \exp(ik\Phi(\mathbf{r})) \tag{26}$$

where A is amplitude and Φ is phase. Substituting Equation (26) in the Helmholtz equation (25) and retaining only the dominant terms, i.e. the ones multiplied by k^2, yields an equation for the phase

$$(\nabla\Phi)^2 = n^2 \tag{27}$$

This equation is also known as the eikonal equation (Kravtsov and Orlov, 1990).

4.1.2 The ray trajectory equation

The surfaces of constant phase, the phase fronts, define the shape of the radiation field. The eikonal equation therefore corresponds to a geometric description of the propagation of the wave. A curve $\mathbf{r} = \mathbf{r}(\varsigma)$ is normal to the phase front if

$$\frac{d\mathbf{r}}{d\varsigma} = \nabla\Phi \tag{28}$$

where ς is the ray trajectory parameter. Differentiating this equation with respect to ς and using the eikonal equation yields

$$\frac{d^2\mathbf{r}}{d\varsigma^2} = n\nabla n \tag{29}$$

The ray trajectory parameter ς can be readily determined from

$$\left|\frac{d\mathbf{r}}{ds}\frac{ds}{d\varsigma}\right| = \frac{ds}{d\varsigma} = |\nabla\Phi| = n$$

to be $d\varsigma = ds/n$. Hence, the ray trajectory equation becomes

$$\frac{d}{ds}\left(n\frac{d\mathbf{r}}{ds}\right) = \nabla n \tag{30}$$

4.1.3 The Euler Lagrange equation

An interesting representation of the ray trajectory equation can be derived as follows: For a curve parameterized with u, i.e. $\mathbf{r} = \mathbf{r}(u)$, the ray trajectory equation reads as

$$\frac{d}{du}\left(n\frac{du}{ds}\frac{d\mathbf{r}}{du}\right) = \frac{ds}{du}\nabla n \tag{31}$$

which can be recast to yield

$$\frac{d}{du}\frac{\partial f}{\partial r_i'} - \frac{\partial f}{\partial r_i} = 0 \tag{32}$$

for $i = 1, ..., 3$ where

$$f = n \left| \frac{d\mathbf{r}}{du} \right| \tag{33}$$

The set of Equations (32) are known as the Euler Lagrange equations.

4.1.4 Fermat's principle

Since the Euler Lagrange equation appear in calculus of variation, an alternative derivation of the ray trajectory equation is obvious; namely the derivation of the ray trajectory equation from Fermat's principle: The path taken by a ray between two points A and B is the path that can be traversed in the least time. The optical path length L is defined as

$$L = \int_A^B n \, ds = \int_a^b n \cdot \left| \frac{d\mathbf{r}}{du} \right| du \tag{34}$$

and from the fundamental lemma of calculus of variation it follows that the ray trajectory minimizes the optical path length if and only if the integrand verifies the Euler-Lagrange equations.

Another representation of the optical path length can be derived using the relation between the tangent vector of the ray trajectory and the gradient of the phase. Multiplying Equation (28) with the tangent unit vector and taking into account that $d\varsigma = ds/n$ leads to

$$n = \nabla\Phi \cdot \frac{d\mathbf{r}}{ds} \tag{35}$$

which allows for the conclusion that

$$L = \Phi_b - \Phi_a \tag{36}$$

Thus, the optical path length is entirely determined by the phase difference. This representation of the optical path length makes the appearance of the slant path delay in the phase pseudorange observation equation (17) more obvious.

4.1.5 The slant path delay in the context of geometrical optics

In terms of the optical path length L and the geometric path length G the slant path delay reads as

$$S = L - G \tag{37}$$

According to Fermat's principle, the signal path from the satellite to the receiver is such that the optical path length, and therefore the slant path delay, takes a minimum value. In order to simulate the slant path delay, at least two strategies exist: (i) solve the eikonal equation for a given field of the refractivity index or (ii) solve the ray trajectory equation (the characteristic equation of the eikonal equation) and integrate the refractivity index along the ray trajectory. In this thesis the latter task is performed. Since a mesoscale model does not provide the dependency of the refractivity index on the location in analytical form, an interpolation routine which allows to compute the refractivity index at an arbitrary location is needed. The gradient of the refractivity index can be computed by differentiation of the interpolation routine.

4.2 Geometrical description of the problem

Figure 4 provides a sketch of the geometry in a two dimensional plane for simulating slant path delays. The satellite S transmits a signal to the ground based receiver R. Due to the bending effect of the atmosphere on the signal path, the apparent elevation angle (arriving elevation angle) of the signal at the receiver is different from the geometric elevation angle at the receiver. The signal path, indicated by the dotted line, obeys the ray trajectory equation. The geometric path corresponds to the straight line connecting the receiver and the satellite. The bending angle, defined as the angle between the tangent vector at the receiver and the tangent vector at the satellite, is introduced to measure the discrepancy between the signal path and the geometric path. The orthonormal system located at Earth's center will be referred as the global frame of reference. The orthonormal system located at the receiver will be referred as the local frame of reference.

From Figure 4 the close relation to GPS radio occultation (RO) is evident. A GPS radio occulation occurs when a transmitting GPS satellite, setting or rising behind the Earth's limb, is viewed by a Low Earth Orbit (LEO) satellite. The relative motion between the GPS and LEO satellite, provides a scanning of the atmosphere. The ray passing through the atmosphere is refracted and delayed due to the gradient in refractivity. For details on remote sensing of the Earth's atmosphere using GPS RO the reader is referred to Kursinski et al. (1997).

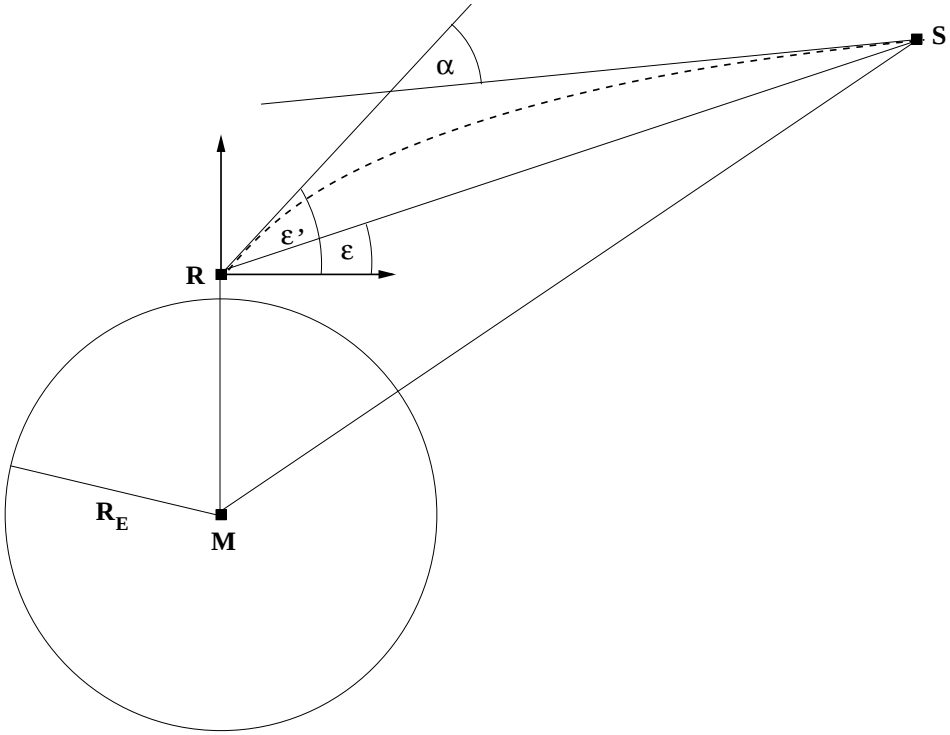


Figure 4: Sketch of the geometry for simulating slant path delays. The location of the receiver is marked with R, the location of the satellite is marked with S and Earth's center is marked with M. R_E denotes the radius of Earth. ε denotes the geometric angle at the receiver, ε' denotes the apparent elevation angle at the receiver and α denotes the bending angle. The signal path from the satellite S to the receiver R is indicated by the dotted line.

4.3 The governing system of differential equations in 3D

The method most suited to the problem consists in applying Fermat's principle, that is, in looking for the curve with cartesian coordinates $(x, y(x), z(x))$ between the receiver with coordinates $(a, y(a), z(a))$ and the satellite with coordinates $(b, y(b), z(b))$ which minimizes the optical path length:

$$L = \int_a^b n(x, y(x), z(x)) \sqrt{1 + y'(x)^2 + z'(x)^2}\, dx \tag{38}$$

This condition is satisfied if and only if the integrand $f = n(x, y, z) \sqrt{1 + y'^2 + z'^2}$ verifies the Euler-Lagrange equation

$$\begin{aligned} \frac{d}{dx}\frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} &= 0 \\ \frac{d}{dx}\frac{\partial f}{\partial z'} - \frac{\partial f}{\partial z} &= 0 \end{aligned}$$

Inserting f and expanding the partial derivative with respect to y' and z' leads to

$$\begin{aligned} \frac{d}{dx}\frac{n\, y'}{\sqrt{1 + y'^2 + z'^2}} &= n_y \sqrt{1 + y'^2 + z'^2} \\ \frac{d}{dx}\frac{n\, z'}{\sqrt{1 + y'^2 + z'^2}} &= n_z \sqrt{1 + y'^2 + z'^2} \end{aligned}$$

where n_y and n_z denote the partial derivative of n with respect to y and z respectively. Evaluating the total derivative with respect to x leads to

$$\begin{aligned} \frac{n\, y''}{\sqrt{1 + y'^2 + z'^2}} - \frac{n\, y'(y'\, y'' + z'\, z'')}{(1 + y'^2 + z'^2)^{3/2}} + \frac{n'\, y'}{\sqrt{1 + y'^2 + z'^2}} &= n_y \sqrt{1 + y'^2 + z'^2} \\ \frac{n\, z''}{\sqrt{1 + y'^2 + z'^2}} - \frac{n\, z'(y'\, y'' + z'\, z'')}{(1 + y'^2 + z'^2)^{3/2}} + \frac{n'\, z'}{\sqrt{1 + y'^2 + z'^2}} &= n_z \sqrt{1 + y'^2 + z'^2} \end{aligned}$$

where n' denotes the total derivative of n with respect to x. Rearranging the system of equations leads to

$$\begin{aligned} n\, y''(1 + z'^2) - n\, z''\, y'\, z' &= n_y\, (1 + y'^2 + z'^2)^2 - n'\, y'(1 + y'^2 + z'^2) \\ n\, z''(1 + y'^2) - n\, y''\, z'\, y' &= n_z\, (1 + y'^2 + z'^2)^2 - n'\, z'(1 + y'^2 + z'^2) \end{aligned}$$

which is a system of two linear equations for the two unknowns y'' and z''. Solving for y'' and z'' and keeping in mind that $n' = n_x + n_y y' + n_z z'$ leads to

$$\begin{aligned} n\,y'' &= (1+y'^2)(n_y\,(1+z'^2) - n_x\,y' - n_z\,y'\,z') + y'\,z'\,(n_z\,(1+y'^2) - n_x\,z' - n_y\,y'\,z') \\ n\,z'' &= (1+z'^2)(n_z\,(1+y'^2) - n_x\,z' - n_y\,y'\,z') + y'\,z'\,(n_y\,(1+z'^2) - n_x\,y' - n_z\,y'\,z') \end{aligned}$$

which can be further simplified to yield

$$\begin{aligned} y'' &= \left(\frac{n_y}{n} - \frac{n_x}{n}y'\right)\,(1+y'^2+z'^2) \\ z'' &= \left(\frac{n_z}{n} - \frac{n_x}{n}z'\right)\,(1+y'^2+z'^2) \end{aligned} \tag{39}$$

This coupled system of ordinary differential equations (ODE) is the basis of the numerical algorithm. With explicit boundary conditions according to

$$\begin{aligned} y_a &= y(a) \\ z_a &= z(a) \\ y_b &= y(b) \\ z_b &= z(b) \end{aligned} \tag{40}$$

it represents a nonlinear two point boundary value problem (BVP). The system of ODEs is exact; that is, no approximations are involved in their derivation from Fermat's principle.

4.3.1 The method of solution

The nonlinear two point BVP can be solved numerically using either a shooting method, a finite difference scheme or a collocation method, see e.g. Ascher et al. (1995). The first method is referred to as a local method, the two latter methods are referred to as global methods.

The shooting method consists of formulating the BVP as an initial value problem (IVP) with a priori unknown initial values. The unknown initial values together with the solution over the integration interval are determined iteratively. This approach requires the solution of the IVP for the ODEs for each initial value. The success of the shooting method depends on a number of factors, the most important of which is the stability of the IVP that must be solved at each iteration. An ODE is stable if a small change to the ODE and/or the initial or boundary conditions leads to a small change in the solution. In fact, it is the case that for many stable BVPs the corresponding IVPs (beginning from either endpoint and integrating towards the other endpoint) are insufficiently stable for shooting to succeed. In particular this can be the case on very long integration intervals. In addition, the shooting method is computational expensive. The inherent problems of local methods can be overcome using global methods.

Global methods for solving BVPs fall into two related categories. The first global approach is a finite difference scheme where a sequence of nodes is defined and the derivatives in the ODEs are approximated by finite differences at the sequence of nodes. The resulting difference equations plus the boundary conditions give a set of non-linear algebraic equations for the solution at the sequence of nodes. The second global approach is to approximate the solution defined in terms of a basis of splines defined piecewise at the sequence of nodes and to collocate this approximate solution. In collocation the approximate solution is substituted in the system of ODEs. The ODE system is required to be satisfied exactly at each collocation point. The number of collocation points plus the number of boundary conditions must equal the number of unknown coefficients in the approximate solution. The spline basis for collocation leads to a non-linear algebraic system of equations which must be solved iteratively. At each iteration a structured linear system of equations must be solved. A similarly structured system arises from the finite difference method.

For the BVP discussed in this thesis the method of choice is a global method. Since the end points of the integration interval, i.e. the location of the receiver and the satellite, are automatically part of the over all solution, and the algorithm does not rely on any other source of information, this approach will be referred to as the direct numerical simulation. A finite difference scheme, to be prescribed hereinafter, is developed.

4.3.2 The governing system of algebraic equations

Let x_i for $i = 0, ..., n+1$ denote a non-uniform sequence of nodes at which the solution y_i and z_i is sought. The derivatives of y and z with respect to x at the sequence of nodes x_j are approximated as

$$\begin{aligned} y'(x_j) &= \sum_{k=j-1}^{j+1} L'_{j,k}(x_j) \cdot y_k \\ z'(x_j) &= \sum_{k=j-1}^{j+1} L'_{j,k}(x_j) \cdot z_k \\ y''(x_j) &= \sum_{k=j-1}^{j+1} L''_{j,k}(x_j) \cdot y_k \\ z''(x_j) &= \sum_{k=j-1}^{j+1} L''_{j,k}(x_j) \cdot z_k \end{aligned} \tag{41}$$

where $L_{j,k}(x)$ denote Lagrange basis polynomials according to

$$L_{j,k}(x) = \prod_{\substack{m=j-1 \\ m \neq k}}^{j+1} \frac{(x - x_m)}{(x_k - x_m)} \tag{42}$$

for $j = 1, ..., n$. Inserting the finite differences into the ODEs leads to a system of $2n$ non-linear algebraic equations for the $2n$ unknowns y_j and z_j

$$\begin{aligned} \sum_{k=j-1}^{j+1} L''_{j,k}(x_j) y_k &- \left[\frac{n_y(x_j, y_j, z_j)}{n(x_j, y_j, z_j)} - \frac{n_x(x_j, y_j, z_j)}{n(x_j, y_j, z_j)} \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) y_k \right) \right] \\ &\cdot \left[1 + \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) y_k \right)^2 + \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) z_k \right)^2 \right] = 0 \\ \sum_{k=j-1}^{j+1} L''_{j,k}(x_j) z_k &- \left[\frac{n_z(x_j, y_j, z_j)}{n(x_j, y_j, z_j)} - \frac{n_x(x_j, y_j, z_j)}{n(x_j, y_j, z_j)} \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) z_k \right) \right] \\ &\cdot \left[1 + \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) y_k \right)^2 + \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) z_k \right)^2 \right] = 0 \end{aligned} \tag{43}$$

for $j = 1, ..., n$ with

$$\begin{aligned} y_0 &= y_a \\ z_0 &= z_a \\ y_{n+1} &= y_b \\ z_{n+1} &= z_b \end{aligned} \tag{44}$$

Note that a uniform sequence of nodes with step size Δx implies

$$\begin{aligned} y'(x_j) &= 0.5 \cdot (y_{j+1} - y_{j-1})/\Delta x \\ z'(x_j) &= 0.5 \cdot (z_{j+1} - z_{j-1})/\Delta x \\ y''(x_j) &= (y_{j+1} - 2y_j + y_{j-1})/\Delta x^2 \\ z''(x_j) &= (z_{j+1} - 2z_j + z_{j-1})/\Delta x^2 \end{aligned}$$

for $j = 1, ..., n$, i.e. the first and second derivatives are discretised by second order accurate central differences. The Ansatz represents a second order accurate non-uniform finite difference scheme.

4.3.3 The numerical solution of the system of algebraic equations

The system of Equations (43) is written in terms of the residuals R_j^y and R_j^z as

$$\begin{aligned} R_j^y(y_{j-1}, z_{j-1}, y_j, z_j, y_{j+1}, z_{j+1}) &= 0 \\ R_j^z(y_{j-1}, z_{j-1}, y_j, z_j, y_{j+1}, z_{j+1}) &= 0 \end{aligned} \tag{45}$$

for $j = 1, ..., n$. In short term notation the system of Equations (45) reads as

$$F(Z) = 0 \tag{46}$$

where $Z = (y_1, z_1, ..., y_n, z_n)$ denotes the solution vector. Newton's iteration is applied to solve the non-linear system of equations. Let Z_r denote the solution vector at the iteration step r. The solution vector Z_{r+1} at the iteration step $r+1$ is obtained by solving the following system of linear equations

$$\mathbf{J}(Z_r)(Z_{r+1} - Z_r) = -F(Z_r) \tag{47}$$

where $\mathbf{J}$ denotes the Jacobian

$$\mathbf{J} = \frac{\partial F}{\partial Z} \tag{48}$$

and explicitly reads as

$$\mathbf{J} = \begin{pmatrix}
\frac{\partial R_1^y}{\partial y_1} & \frac{\partial R_1^y}{\partial z_1} & \frac{\partial R_1^y}{\partial y_2} & \frac{\partial R_1^y}{\partial z_2} & 0 & 0 & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial R_1^z}{\partial y_1} & \frac{\partial R_1^z}{\partial z_1} & \frac{\partial R_1^z}{\partial y_2} & \frac{\partial R_1^z}{\partial z_2} & 0 & 0 & \cdots & \cdots & \cdots & \cdots \\
\frac{\partial R_2^y}{\partial y_1} & \frac{\partial R_2^y}{\partial z_1} & \frac{\partial R_2^y}{\partial y_2} & \frac{\partial R_2^y}{\partial z_2} & \frac{\partial R_2^y}{\partial y_3} & \frac{\partial R_2^y}{\partial z_3} & 0 & 0 & \cdots & \cdots \\
\frac{\partial R_2^z}{\partial y_1} & \frac{\partial R_2^z}{\partial z_1} & \frac{\partial R_2^z}{\partial y_2} & \frac{\partial R_2^z}{\partial z_2} & \frac{\partial R_2^z}{\partial y_3} & \frac{\partial R_2^z}{\partial z_3} & 0 & 0 & \cdots & \cdots \\
0 & 0 & \frac{\partial R_3^y}{\partial y_2} & \frac{\partial R_3^y}{\partial z_2} & \frac{\partial R_3^y}{\partial y_3} & \frac{\partial R_3^y}{\partial z_3} & \frac{\partial R_3^y}{\partial y_4} & \frac{\partial R_3^y}{\partial z_4} & 0 & 0 \\
0 & 0 & \frac{\partial R_3^z}{\partial y_2} & \frac{\partial R_3^z}{\partial z_2} & \frac{\partial R_3^z}{\partial y_3} & \frac{\partial R_3^z}{\partial z_3} & \frac{\partial R_3^z}{\partial y_4} & \frac{\partial R_3^z}{\partial z_4} & 0 & 0 \\
\vdots & & & & & & \ddots & & & \vdots \\
0 & 0 & \cdots & \cdots & 0 & 0 & \frac{\partial R_N^y}{\partial y_{N-1}} & \frac{\partial R_N^y}{\partial z_{N-1}} & \frac{\partial R_N^y}{\partial y_N} & \frac{\partial R_N^y}{\partial z_N} \\
0 & 0 & \cdots & \cdots & 0 & 0 & \frac{\partial R_N^z}{\partial y_{N-1}} & \frac{\partial R_N^z}{\partial z_{N-1}} & \frac{\partial R_N^z}{\partial y_N} & \frac{\partial R_N^z}{\partial z_N}
\end{pmatrix}$$

The non-zero entries of the Jacobian are provided in Appendix A. Since the Jacobian is a band matrix (bandwidth seven), the linear system of equation can be efficiently solved by LU-Decomposition. There is no need to store the entire Jacobian. The Jacobian can be stored in condensed form represented in a $l + u + 1$ times n matrix for its $l = 3$ lower and $u = 3$ upper co-diagonals. The Newton iteration is stopped if a predefined number of iterations is performed.

The ray trajectory equation and the derived system of algebraic equations is form invariant, i.e. the formulation is independent of the coordinate system. It is convenient

to solve the governing system of algebraic equations in the local frame of reference. The first guess vector Z_0 in the Newton iteration is chosen to be equal to the coordinates of points along the straight line connecting the first and the last node. In fact, this choice is an approximate zero, i.e. it provides fast and safe convergence in Newtons iteration. It is advantageous to rotate the local frame of reference such that the x-axis points from the receiver to the satellite in question (the y- and z-axis remain perpendicular to the x-axis). In the rotated local frame of reference the first guess vector in the Newton iteration is the null vector, i.e. $Z_0 = 0$, which in turn does not only represent an approximate zero but also greatly simplifies the computation in the first Newton iteration.

4.3.4 The computation of the refractivity and its derivatives

In MM5, the pressure, the temperature and the water vapor mixing ratio, are stored at gridpoints, specified by longitude, latitude and height (Grell et al., 1995). In the very first step, the gridded dry, wet and total refractivity are computed at each model grid point. To compute the refractivity N at an arbitrary point in the local frame of reference a coordinate transformation and an interpolation is needed. The Earth is assumed to be a perfect sphere with radius R_E being equal to the local curvature radius of the reference ellipsoid at the center point of the mesoscale model domain. Similarly, the gravitational acceleration at mean sea level g_m is computed at the center point of the mesoscale model domain. The variation of the gravitational acceleration g with respect to height h is accounted for through

$$g(h) = \frac{g_m R_E^2}{(R_E + h)^2} \tag{49}$$

The algorithm to compute N at a point with cartesian coordinates (x, y, z) in the local frame of reference consists of four steps:

1) The cartesian coordinates (x, y, z) in the local frame of reference are transformed to the cartesian coordinates (x_g, y_g, z_g) in the global frame of reference:

$$\begin{aligned} x_g &= a_{11} \cdot x + a_{12} \cdot y + a_{13} \cdot z + x_l \\ y_g &= a_{21} \cdot x + a_{22} \cdot y + a_{23} \cdot z + y_l \\ z_g &= a_{31} \cdot x + a_{32} \cdot y + a_{33} \cdot z + z_l \end{aligned}$$

where a_{jk} for $j = 1, ...3$ and $k = 1, ..., 3$ denote the cartesian coordinates of the unit basis vectors of the local frame of reference in the global frame of reference and (x_l, y_l, z_l) denote

the cartesian coordinates of the origin of the local frame of reference in the global frame of reference.

2) The cartesian coordinates (x_g, y_g, z_g) in the global frame of reference are transformed to spherical coordinates (λ, ϕ, h):

$$\begin{aligned} \lambda &= \arctan(y_g/x_g) \\ \phi &= \arctan(z_g/\sqrt{x_g^2 + y_g^2}) \\ h &= \sqrt{x_g^2 + y_g^2 + z_g^2} - R_E \end{aligned}$$

3) The longitude λ and the latitude ϕ are converted to the cross point coordinates $\hat{x}$ and $\hat{y}$ in the uniform model grid. For details on this coordinate conversion the reader is referred to Grell et al. (1995) Appendix 3.

4) The refractivity N at model grid point coordinates $(\hat{x}, \hat{y}, h)$ is computed according to the following algorithm:

a) Floor towards the largest integers I and J less then or equal to $\hat{x}$ and $\hat{y}$, i.e. $I = \lfloor \hat{x} \rfloor$ and $J = \lfloor \hat{y} \rfloor$, compute the increments X and Y according to

$$\begin{aligned} X &= \hat{x} - \hat{X}_{IJ} \\ Y &= \hat{y} - \hat{Y}_{IJ} \end{aligned}$$

and the bilinear interpolation coefficients X_1, X_2, X_3, X_4:

$$\begin{aligned} X_1 &= (1 - X) \cdot (1 - Y) \\ X_2 &= (1 - X) \cdot (0 + Y) \\ X_3 &= (0 + X) \cdot (1 - Y) \\ X_4 &= (0 + X) \cdot (0 + Y) \end{aligned}$$

The two dimensional arrays $\hat{X}_{IJ}$ and $\hat{Y}_{IJ}$ for $I = 1, ..., M_I$ and $J = 1, ..., M_J$ denote the grid point coordinates of the uniform model grid. The integers M_I and M_J denote the horizontal dimensions of the model grid.

b) Determine at the neighboring grid points the adjacent grid point indices with respect to height A, B, C, D by binary search, and compute the interpolation coefficients L_1, L_2, L_3, L_4:

$$
\begin{aligned}
L_1 &= (h - H_{I+0J+0A})/(H_{I+0J+0A} - H_{I+0J+0A-1}) \\
L_2 &= (h - H_{I+0J+1B})/(H_{I+0J+1B} - H_{I+0J+1B-1}) \\
L_3 &= (h - H_{I+1J+0C})/(H_{I+1J+0C} - H_{I+1J+0C-1}) \\
L_4 &= (h - H_{I+1J+1D})/(H_{I+1J+1D} - H_{I+1J+1D-1})
\end{aligned}
$$

c) Perform logarithmic interpolation to compute the refractivity values N_1, N_2, N_3, N_4 at appropriate height:

$$
\begin{aligned}
N_1 &= N_{I+0J+0A} \cdot (N_{I+0J+0A}/N_{I+0J+0A-1})^{L_1} \\
N_2 &= N_{I+0J+1B} \cdot (N_{I+0J+1B}/N_{I+0J+1B-1})^{L_2} \\
N_3 &= N_{I+1J+0C} \cdot (N_{I+1J+0C}/N_{I+1J+0C-1})^{L_3} \\
N_4 &= N_{I+1J+1D} \cdot (N_{I+1J+1D}/N_{I+1J+1D-1})^{L_4}
\end{aligned}
$$

d) Compute N according to

$$
N = \sum_{l=1}^{4} X_l \cdot N_l
$$

The three dimensional arrays H_{IJK} and N_{IJK} for $I = 1, ..., M_I$, $J = 1, ..., M_J$ and $K = 1, ..., M_K$ store the gridded model height and the gridded model refractivity respectively. The integer M_k denotes the vertical dimension of the model grid. Since the MM5 is a limited area model, caution is needed at lateral model boundaries and above the finite model top. At lateral model boundaries the integers I and J are limited in step a) according to:

$$
\begin{aligned}
I &= \min(\max(I, 1), M_I - 2) \\
J &= \min(\max(J, 1), M_J - 2)
\end{aligned}
$$

Above the finite model top hydrostatic equilibrium is enforced and the variation of gravity with respect to height is accounted for. The wet refractivity above the finite model top is neglected. This approximation is justified as long as the model top has sufficient altitude. With the aid of the perfect gas law and the hydrostatic equation the refractivity above the model top reads as

$$N = -\frac{k_1 R_d}{g(h)} \cdot P_h$$

where $R_d = 287.05$ Jkg^{-1}K^{-1} is the gas constant of dry air and P_h denotes the partial derivative of pressure with respect to height. The pressure P is computed in a similar manner as the refractivity N through

$$P = \sum_{l=1}^{4} X_l \cdot P_l$$

where

$$\begin{aligned} P_1 &= P_{I+0J+0A} \cdot (P_{I+0J+0A}/P_{I+0J+0A-1})^{L_1} \\ P_2 &= P_{I+0J+1B} \cdot (P_{I+0J+1B}/P_{I+0J+1B-1})^{L_2} \\ P_3 &= P_{I+1J+0C} \cdot (P_{I+1J+0C}/P_{I+1J+0C-1})^{L_3} \\ P_4 &= P_{I+1J+1D} \cdot (P_{I+1J+1D}/P_{I+1J+1D-1})^{L_4} \end{aligned}$$

denote pressure values at appropriate height. The three dimensional array P_{IJK} for $I = 1, ..., M_I$, $J = 1, ..., M_J$ and $K = 1, ..., M_K$ stores the gridded model pressure. Differentiation with respect to height yields to

$$P_h = \sum_{l=1}^{4} X_l \cdot P_{l,h}$$

where

$$\begin{aligned} P_{1,h} &= P_1 \cdot \log(P_{I+0J+0A}/P_{I+0J+0A-1})/(H_{I+0J+0A} - H_{I+0J+0A-1}) \\ P_{2,h} &= P_2 \cdot \log(P_{I+0J+1B}/P_{I+0J+1B-1})/(H_{I+0J+1B} - H_{I+0J+1B-1}) \\ P_{3,h} &= P_3 \cdot \log(P_{I+1J+0C}/P_{I+1J+0C-1})/(H_{I+1J+0C} - H_{I+1J+0C-1}) \\ P_{4,h} &= P_4 \cdot \log(P_{I+1J+1D}/P_{I+1J+1D-1})/(H_{I+1J+1D} - H_{I+1J+1D-1}) \end{aligned}$$

denote the derivatives of pressure values with respect to height at appropriate height. Note that the derivative of the pressure with respect to height is extrapolated from the two upper most model levels.

As a smooth alternative to bilinear interpolation, bicubic interpolation is implemented. Bicubic interpolation can be accomplished using either Lagrange polynomials, cubic splines or cubic convolution algorithm. A cubic convolution algorithm (Keys, 1981) is applied since bicubic spline interpolation requires the solution of a linear system of equations at each point where the refractivity is to be computed. The cubic convolution algorithm is provided in Appendix B.

The computation of the partial derivative of N with respect to x, y or z is performed by rigorous application of the chain rule of differential calculus. The relevant computations in the last step are provided in detail. The computations in the former steps are carried out in a similar manner. Hereinafter the subscript ξ denotes the partial derivative with respect to x, y or z:

Step a) reads as:

$$\begin{aligned}
X_\xi &= \hat{x}_\xi \\
Y_\xi &= \hat{y}_\xi \\
X_{1,\xi} &= (0 - X_\xi) \cdot (1 - Y) + (1 - X) \cdot (0 - Y_\xi) \\
X_{2,\xi} &= (0 - X_\xi) \cdot (0 + Y) + (1 - X) \cdot (0 + Y_\xi) \\
X_{3,\xi} &= (0 + X_\xi) \cdot (1 - Y) + (0 + X) \cdot (0 - Y_\xi) \\
X_{4,\xi} &= (0 + X_\xi) \cdot (0 + Y) + (0 + X) \cdot (0 + Y_\xi)
\end{aligned}$$

Step b) reads as:

$$\begin{aligned}
L_{1,\xi} &= h_\xi / (H_{I+0J+0A} - H_{I+0J+0A-1}) \\
L_{2,\xi} &= h_\xi / (H_{I+0J+1B} - H_{I+0J+1B-1}) \\
L_{3,\xi} &= h_\xi / (H_{I+1J+0C} - H_{I+1J+0C-1}) \\
L_{4,\xi} &= h_\xi / (H_{I+1J+1D} - H_{I+1J+1D-1})
\end{aligned}$$

Step c) reads as:

$$\begin{aligned}
N_{1,\xi} &= N_1 \cdot \log(N_{I+0J+0A} / N_{I+0J+0A-1}) \cdot L_{1,\xi} \\
N_{2,\xi} &= N_2 \cdot \log(N_{I+0J+1B} / N_{I+0J+1B-1}) \cdot L_{2,\xi} \\
N_{3,\xi} &= N_3 \cdot \log(N_{I+1J+0C} / N_{I+1J+0C-1}) \cdot L_{3,\xi} \\
N_{4,\xi} &= N_4 \cdot \log(N_{I+1J+1D} / N_{I+1J+1D-1}) \cdot L_{4,\xi}
\end{aligned}$$

Step d) reads as:

$$N_\xi = \sum_{l=1}^{4}(X_{l,\xi} \cdot N_l + X_l \cdot N_{l,\xi})$$

The partial derivatives $\hat{x}_\xi, \hat{y}_\xi$ and h_ξ are provided by differentiation of step 3. Higher order and mixed partial derivatives are computed in the same manner. For completeness, the final step of the computation is provided. Hereinafter the subscript η denotes the partial derivative with respect to x, y or z.

Step a) reads as:

$$\begin{aligned}
X_{\xi\eta} &= \hat{x}_{\xi\eta} \\
Y_{\xi\eta} &= \hat{y}_{\xi\eta} \\
X_{1,\xi\eta} &= (0 - X_{\xi\eta}) \cdot (1 - Y) + (1 - X) \cdot (0 - Y_{\xi\eta}) \\
&+ (0 - X_\xi) \cdot (0 - Y_\eta) + (0 - X_\eta) \cdot (0 - Y_\xi) \\
X_{2,\xi\eta} &= (0 - X_{\xi\eta}) \cdot (0 + Y) + (1 - X) \cdot (0 + Y_{\xi\eta}) \\
&+ (0 - X_\xi) \cdot (0 + Y_\eta) + (0 - X_\eta) \cdot (0 + Y_\xi) \\
X_{3,\xi\eta} &= (0 + X_{\xi\eta}) \cdot (1 - Y) + (0 + X) \cdot (0 - Y_{\xi\eta}) \\
&+ (0 + X_\xi) \cdot (0 - Y_\eta) + (0 + X_\eta) \cdot (0 - Y_\xi) \\
X_{4,\xi\eta} &= (0 + X_{\xi\eta}) \cdot (0 + Y) + (0 + X) \cdot (0 + Y_{\xi\eta}) \\
&+ (0 + X_\xi) \cdot (0 + Y_\eta) + (0 + X_\eta) \cdot (0 + Y_\xi)
\end{aligned}$$

Step b) reads as:

$$\begin{aligned}
L_{1,\xi\eta} &= h_{\xi\eta}/(H_{I+0J+0A} - H_{I+0J+0A-1}) \\
L_{2,\xi\eta} &= h_{\xi\eta}/(H_{I+0J+1B} - H_{I+0J+1B-1}) \\
L_{3,\xi\eta} &= h_{\xi\eta}/(H_{I+1J+0C} - H_{I+1J+0C-1}) \\
L_{4,\xi\eta} &= h_{\xi\eta}/(H_{I+1J+1D} - H_{I+1J+1D-1})
\end{aligned}$$

Step c) reads as:

$$
\begin{aligned}
N_{1,\xi\eta} &= (N_{1,\eta} \cdot L_{1,\xi} + N_1 \cdot L_{1,\xi\eta}) \cdot \log(N_{I+0J+0A}/N_{I+0J+0A-1}) \\
N_{2,\xi\eta} &= (N_{2,\eta} \cdot L_{2,\xi} + N_2 \cdot L_{2,\xi\eta}) \cdot \log(N_{I+0J+1B}/N_{I+0J+1B-1}) \\
N_{3,\xi\eta} &= (N_{3,\eta} \cdot L_{3,\xi} + N_3 \cdot L_{3,\xi\eta}) \cdot \log(N_{I+1J+0C}/N_{I+1J+0C-1}) \\
N_{4,\xi\eta} &= (N_{4,\eta} \cdot L_{4,\xi} + N_4 \cdot L_{4,\xi\eta}) \cdot \log(N_{I+1J+1D}/N_{I+1J+1D-1})
\end{aligned}
$$

Step d) reads as:

$$
N_{\xi\eta} = \sum_{l=1}^{4} (X_{l,\xi\eta} \cdot N_l + X_{l,\xi} \cdot N_{l,\eta} + X_{l,\eta} \cdot N_{l,\xi} + X_l \cdot N_{l,\xi\eta})
$$

Again, the partial derivatives $\hat{x}_{\xi\eta}, \hat{y}_{\xi\eta}$ and $h_{\xi\eta}$ are provided by differentiation of step 3.

4.3.5 The sequence of nodes

So far the sequence of nodes for which the solution is sought was not specified. Roughly speaking, the sequence of nodes can be chosen arbitrary and a successive refinement of nodes can be performed. For example, one may chose a uniform sequence of nodes and double the number of nodes until the solution at the nodes which coincide reaches a predefined accuracy. Once the solution is sought on the uniform sequence of nodes, the slant path delay can be evaluated by a standard numerical routine, e.g. the trapezoid or the Simpson's rule. Then the number of nodes is doubled, the computation of the ray trajectory is repeated, and another estimate of the slant path delay is computed. This procedure is repeated until the relative difference of the current and previous computed slant path delay is smaller than some prescribed error tolerance. In principle, (provided that the algorithm converges) this procedure allows to compute the slant path delay with a predefined error tolerance. However, a uniform sequence of nodes turns out to be computationally expensive. It appears to be convenient not to use a uniform sequence of nodes but a non-uniform sequence of nodes that automatically minimizes the numerical error associated to the computation of the slant path delay. A convenient (ad hoc) choice of nodes are Legendre-Gauss-Lobatto nodes (Abramowitz and Stegun, 1972) since (i) the nodes are densified close to the endpoints of the interval where it is expected to be important and (ii) the slant path delay can be readily evaluated with aid of the Gauss-Lobatto quadrature rule, which, compared to other numerical integration formulas, provides an outstanding accuracy. The Legendre-Gauss-Lobatto nodes x_i and weights w_i for $i = 0, ..., n+1$ are provided in Appendix C. The number of Legendre-Gauss-Lobatto nodes is increased until a predefined error tolerance is reached.

4.4 The computation of the slant path delay

In the rotated local frame of reference the slant path delay reads as

$$\begin{aligned} S &= 10^{-6} \cdot \int_a^b N_d(x, y(x), z(x)) \sqrt{1 + y'(x)^2 + z'(x)^2}\, dx \\ &+ 10^{-6} \cdot \int_a^b N_w(x, y(x), z(x)) \sqrt{1 + y'(x)^2 + z'(x)^2}\, dx \\ &+ \int_a^b (\sqrt{1 + y'(x)^2 + z'(x)^2} - 1)\, dx \end{aligned} \tag{50}$$

Two computational models for the slant path delay are implemented.

4.4.1 The rigorous computational model for the slant path delay

Once the coordinates of the ray trajectory (x_i, y_i, z_i) for $i = 0, ..., n+1$ are determined, the slant path delay can be readily numerically evaluated with aid of the Gauss-Lobatto quadrature rule

$$\begin{aligned} S &= 10^{-6} \cdot \sum_{i=0}^{n+1} N_d(x_i, y_i, z_i) \sqrt{1 + y'(x_i)^2 + z'(x_i)^2}\, w_i \\ &+ 10^{-6} \cdot \sum_{i=0}^{n+1} N_w(x_i, y_i, z_i) \sqrt{1 + y'(x_i)^2 + z'(x_i)^2}\, w_i \\ &+ \sum_{i=0}^{n+1} (\sqrt{1 + y'(x_i)^2 + z'(x_i)^2} - 1)\, w_i \end{aligned} \tag{51}$$

The derivatives $y'(x_j)$ and $z'(x_j)$ for $j = 1, ..., n$ are provided by Equation (41). The derivatives at the first and last node are approximated by second order accurate one-sided finite differences

$$
\begin{aligned}
y'(x_0) &= \sum_{k=0}^{2} L'_{1,k}(a) \cdot y_k \\
z'(x_0) &= \sum_{k=0}^{2} L'_{1,k}(a) \cdot z_k \\
y'(x_{n+1}) &= \sum_{k=n-1}^{n+1} L'_{n,k}(b) \cdot y_k \\
z'(x_{n+1}) &= \sum_{k=n-1}^{n+1} L'_{n,k}(b) \cdot z_k
\end{aligned} \tag{52}
$$

With the numerical integration along the ray trajectory, the numerical algorithm to simulate the slant path delay is completed. Equation (51) provides the rigorous computational model for the slant path delay.

4.4.2 The fast computational model for the slant path delay

A fast computational model for the slant path delay can be derived, if the signal path is approximated by the geometric path and the slant path delay below the model top and above the model top is calculated separately (Zus et al., 2008).

The contribution to the slant path delay below the model top S_b is computed with the aid of the Gauss-Lobatto quadrature rule

$$
S_b = 10^{-6} \cdot \sum_{i=0}^{l+1} N(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)\, \tilde{w}_i \tag{53}
$$

where $(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)$ for $i = 0, ..., l+1$ denote the Legendre-Gauss-Lobatto nodes along the geometric path and $\tilde{w}_i$ for $i = 0, ..., l+1$ denote corresponding weights.

The contribution to the slant path delay above the model top S_t is computed by transformation of the integration variable. The derivative of the geometric path s_0 with respect to height h can be computed with the law of cosine

$$
s'_0(h) = \left(1 - \frac{(h_t + R_E)^2}{(h + R_E)^2}\, sin(\psi_t)^2\right)^{-0.5} \tag{54}
$$

where h_t denotes the height at the model top point and ψ_t denotes the angle between the tangent vector of the geometric path and the radial vector at the model top point. The contribution to the slant path delay above the model top S_t is now determined with the hydrostatic equation and the extended first mean value theorem of integral calculation

$$S_t = 10^{-6} \cdot k_1 R_d \int_{\infty}^{h_t} \frac{s_0'(h)}{g(h)} \cdot P'(h) dh = 10^{-6} \cdot k_1 \cdot R_d \frac{s_0'(\xi)}{g(\xi)} \cdot P_t \tag{55}$$

where P' denotes the derivative of pressure with respect to height and P_t denotes the pressure computed at the model top point. The height ξ, for which Equation (55) holds, is now approximated by the height ξ_s, which results from the assumption of the Standard Atmosphere (1976) above the model top

$$S_t = 10^{-6} \cdot k_1 \cdot R_d \frac{s_0'(\xi_s)}{g(\xi_s)} \cdot P_t \tag{56}$$

For elevation angles close to the zenith, Equation (56) is equivalent and turns out to be as accurate as the well known formulation of Saastamoinen (1972).

The slant path delay reads as

$$S \;=\; 10^{-6} \cdot \sum_{i=0}^{l+1} N(\tilde{x}_i, \tilde{y}_i, \tilde{z}_i)\, \tilde{w}_i + 10^{-6} \cdot k_1 \cdot R_d \frac{s_0'(\xi_s)}{g(\xi_s)} \cdot P(\tilde{x}_{l+1}, \tilde{y}_{l+1}, \tilde{z}_{l+1}) \tag{57}$$

Since the slant path delay in this approximation does not require the solution of the ray trajectory equation, i.e. the numerical solution of a non-linear system of equations, the computation of the slant path delay is very efficient. However, this approximation of the slant path delay is inaccurate for low elevation angles.

4.5 The governing system of differential equations in 2D

In the two-dimensional case the single ODE derived from the Euler-Lagrange equation reads as

$$z'' \;=\; \left(\frac{n_z}{n} - \frac{n_x}{n} z'\right)(1 + z'^2) \tag{58}$$

With explicit boundary conditions according to

$$\begin{aligned} z_a &= z(a) \\ z_b &= z(b) \end{aligned} \tag{59}$$

it represents a non-linear two point BVP.

4.5.1 Numerical solution of the boundary value problem (BVP) in 2D

The method of solution in the two-dimensional case is the same as in the three-dimensional case: The non-linear two point BVP is transformed to a system of non-linear algebraic equations which are solved by Newton iteration. Since the Jacobian is a tridiagonal matrix (a consequence of using low order Lagrange basis polynomials) the system of equations can be very efficiently solved by LU-Decomposition. There is no need to store the entire Jacobian. Only non-zero entries must be stored, i.e. the lower, the main and the upper diagonal.

The two-dimensional BVP can be solved in a three-dimensional refractivity field. For example, the two-dimensional BVP can be solved in the plane defined by the center of the Earth, the receiver and the transmitter. In the rotated local frame of reference this corresponds to the x-z plane. In fact, it is convenient to seek for a solution in the x-z plane since the refractivity is more or less stratified with respect to height. It is evident that in terms of computational efficiency solving the two-dimensional BVP is superior compared to solving the three-dimensional BVP (the number of equations is reduced by a factor of two).

4.5.2 Numerical solution of the initial value problem (IVP) in 2D

Alternatively the single ODE can be solved as an IVP. The single second order ODE is transformed to a system of two first order ODEs. With $\tau = z$ and $\upsilon = z'$ the system of first order ODEs explicitly reads as

$$\begin{aligned} \tau' &= f_1(x, \tau, \upsilon) = \upsilon \\ \upsilon' &= f_2(x, \tau, \upsilon) = \left(\frac{n_\tau}{n} - \frac{n_x}{n}\upsilon\right)(1 + \upsilon^2) \end{aligned} \tag{60}$$

With initial conditions according to

$$\begin{aligned} \tau_a &= \tau(a) \\ \upsilon_a &= \upsilon(a) \end{aligned} \tag{61}$$

the IVP can be readily solved by a standard numerical routine, i.e. a Runge-Kutta formula. In contrast to the numerical solution of the BVP, the numerical solution of the IVP requires the tangent of the ray trajectory at the initial point. Since the tangent of the ray trajectory at the initial point is not known, the numerical solution of the IVP can not be regarded as an efficient alternative to the numerical solution of the BVP. Though a shooting method allows to determine iteratively the tangent of the ray trajectory at the initial point such that the final ray traced point coincides with the second boundary condition, the solution of the IVP is computational expensive. The solution of the BVP always guarantees that the first and last node, i.e. the location of the receiver and the transmitter, are part of the overall solution. On the other hand, provided that the tangent of the ray trajectory at the initial point is a priori known, the numerical solution of the IVP can be applied to study whether it is possible to solve the BVP by the proposed numerical algorithm with high accuracy.

4.6 Snell's law for a spherically layered atmosphere

For a spherically layered atmosphere a remarkably simple analytic relation can be derived. Since the atmosphere is assumed to be spherically layered, it is sufficient to consider polar coordinates. Let (ϕ, r) denote the polar coordinates of a point in the global frame of reference. For a curve with polar coordinates $(\phi, r(\phi))$ between two points with polar coordinates $(\alpha, r(\alpha))$ and $(\beta, r(\beta))$ the optical path length L to be minimized reads as

$$L = \int_{\alpha}^{\beta} n(\phi, r(\phi)) \sqrt{r(\phi)^2 + r'(\phi)^2}\, d\phi \tag{62}$$

According to the fundamental lemma of calculus of variation, the integral takes an extreme value if and only if the integrand $f = n(\phi, r)\sqrt{r^2 + r'^2}$ verifies the Euler-Lagrange equation

$$\frac{d}{d\phi}\frac{\partial f}{\partial r'} - \frac{\partial f}{\partial r} = 0$$

A well known identity of calculus of variation, namely Beltrami's identity, states that the Euler Lagrange equation is equivalent to

$$\frac{d}{d\phi}\left(f - r'\frac{\partial f}{\partial r'}\right) = \frac{\partial f}{\partial \phi}$$

For a spherically layered atmosphere the refractivity index depends on the radial distance only, i.e. $n = n(r)$, and therefore

$$\frac{d}{d\phi}\left(f - r'\frac{\partial f}{\partial r'}\right) = 0$$

which immediately gives

$$f - r'\frac{\partial f}{\partial r'} = p$$

where p denotes the so called impact parameter, a constant along the ray trajectory. Inserting f in the above expression and evaluating the partial derivative with respect to r' leads to

$$\frac{n\,r^2}{\sqrt{r^2 + r'^2}} = p$$

Keeping in mind that the angle ψ between the tangent vector and the radial vector at an arbitrary point on a curve in polar coordinates is given by $\psi = \arctan(r/r')$ results in

$$n\,r\,sin(\psi) = p \tag{63}$$

which is Snell's law for a spherically layered atmosphere. Alternatively, Snell's law can be derived from Equation (58), i.e. the underlying equation of the numerical algorithm to solve the BVP, under the assumption of a spherically layered atmosphere.

One may think that the direct (piecewise) application of Snell's law is an alternative to the numerical solution of the BVP. This is not the case since (i) the tangent vector of the ray trajectory at any point is a priori not known and (ii) the refractivity field is generally not spherically layered. However, this remarkable simple relation can be used to check for correctness and to study the accuracy of the numerical solution obtained by solving the BVP: Given a spherically layered atmosphere the impact parameter computed along any simulated ray trajectory must be an invariant.

4.7 Potential applications of the ray tracing algorithm

The proposed ray-tracing algorithm allows to simulate the propagation of electromagnetic waves in an inhomogeneous medium (in the context of geometrical optics) given the location of the receiver and the transmitter. It is foreseeable that the ray-tracing algorithm, either in the proposed general form or with some simplifications, will have a number of alternative applications. For example, an interesting application is the simulation of the signal path of radar signals. In such applications the location of the transmitter, the apparent elevation/azimuthal angle and the time delay between the emitted and received signal is known. Thus, one is concerned with solving an IVP where the integration interval (the location of the backscattered signal) is a priori not known. Two potential applications, which are of particular interest for remote sensing of the atmosphere using ground-based and spaceborne GPS technology, are shortly described in the following.

4.7.1 GPS radio occultation

An immediate by-product of the ray tracing algorithm are bending angles. The bending angle α, defined as the angle between the tangent vector at the receiver $\mathbf{u}_a$ and the tangent vector at the satellite $\mathbf{u}_b$, can be computed according to

$$\alpha = \arccos\left(\frac{\mathbf{u}_a \cdot \mathbf{u}_b^T}{|\mathbf{u}_a| \cdot |\mathbf{u}_b|}\right) \tag{64}$$

In the local frame of reference $\mathbf{u}_a = (1, y'(a), z'(a))$, $\mathbf{u}_b = (1, y'(b), z'(b))$ and the first derivatives at the first and the last node can be computed according to Equation (52). Obviously, the proposed ray-tracing algorithm allows the direct simulation of bending angles. Besides the ability to simulate slant path delays, the ray-tracing algorithm presented in this thesis is believed to be an appropriate tool for simulating GPS RO observations, e.g. the bending angle or the Doppler frequency shift. Formally a LEO satellite (receiver) can be treated just as a ground based receiver and vice versa.

4.7.2 Direct mapping

In GPS observation processing mapping functions are used to map the a priori tropospheric delay from the zenith direction to lower elevation angles under the assumption of spherical symmetry. If T_z denotes the tropospheric delay estimate in the zenith direction and m denotes the mapping function the tropospheric delay estimate T at an elevation angle ε is determined by

$$T(\varepsilon) = m(\varepsilon) \cdot T_z \tag{65}$$

Marini (1972) showed that the elevation dependency of the mapping function can be approximated by a continued fraction expansion according to

$$m(\varepsilon) = \cfrac{1}{\sin(\varepsilon) + \cfrac{c_1}{\sin(\varepsilon) + \cfrac{c_2}{\sin(\varepsilon) + ...}}} \tag{66}$$

where the parameters c_i are determined by curve fitting.

To date various parameterized mapping functions using a continued fraction expansion similar to equation (66) exist (Davis et al., 1985; Niell, 1996; Boehm et al., 2006). In atmospheric science there is an increasing interest in near-horizon GPS observations because they contain atmospheric profile information (Sokolovskiy et al., 2001; Lowry et al., 2002). To process such near-horizon GPS observations, very accurate mapping functions are required. At low elevation angles parameterized mapping functions, in particular the parameterized dry mapping function, can introduce significant errors. Therefore Rocken et al. (2001) introduced a concept known as direct mapping. Direct mapping differs from parameterized mapping in that it requires an atmospheric profile of temperature, pressure and humidity for a specific location and time. Such a profile can be obtained from radio soundings or from NWP analysis fields. Direct mapping functions are determined by evaluating ray integrals under the assumption of spherical symmetry for a given apparent elevation angle. The geometric elevation angle is then determined from the apparent elevation angle. However, the mapping function must provide a delay correction for a given geometric elevation angle. This task can be accomplished using the ray-tracing algorithm presented in this thesis.

5 The simulation of slant path delays

5.1 The MM5 configuration and the GPS network

The MM5 is a limited-area, non-hydrostatic, terrain-following sigma-coordinate model designed to simulate or predict mesoscale and regional-scale atmospheric circulation. It is the latest in a series developed from a mesoscale model used by Anthes at Penn State in the early 70's that was later documented by Anthes and Warner (1978). Since that time, it has undergone many changes designed to broaden its usage (Grell et al., 1995).

The MM5 version used in this thesis is based on version 3 (MM5v3). The MM5 allows the user to specify the model grid configuration and the selection of different parameterization schemes for various physical processes. Table 1 summarizes the MM5 grid configuration and the basic state variables. Table 2 summarizes the MM5 forecast configuration and the corresponding namelist parameter. The selected parameterization schemes are believed to be the most comprehensive schemes available in the MM5v3. The non-hydrostatic MM5 version is used. For details on the dynamical core and the parameterization schemes the reader is referred to Grell et al. (1995). The horizontal diffusion scheme was modified by the author and is described in detail in an upcoming section. The modified horizontal diffusion scheme can be activated by setting the namelist parameter ITPDIF=3.

MM5 grid	Selected option
Projection type	Lambert conformal
Horizontal resolution	18 km
Center longitude	5.7E
Center latitude	48.9N
Mesh size	64×70×36
Ref. model top pressure	100 hPa
Ref. sea level pressure	1000 hPa
Ref. sea level temperature	275 K

Table 1: Table summarizing the MM5 grid configuration and basic state variables.

Since MM5 is a limited area research model, it requires initial conditions as well as lateral boundary conditions. Gridded data from the European Centre for Medium-Range Weather Forecasts (ECMWF) operational analysis is interpolated to the MM5 grid configuration and the standard MM5 initialization is performed. Hydrometeors, i.e. cloud water and cloud ice, are reset in the initial state. Ground based GPS data are not assimilated at the ECMWF.

The MM5 domain and an overview of the geographical locations of the receivers used in this thesis is given in Figure 5. The receivers are assembled into two networks, hereinafter

MM5 physics	Selected option	Namelist parameter
Coriolis force	Full 3D Coriolis force	ICOR3D=1
Upper boundary condition	Radiative	IFUPR=1
Lateral boundary condition	Relaxation	IBOUDY=3
Cloud microphysics	Reisner	IMPHYS=7
Planetary boundary layer	MRF	IBLTYP=5
Ground temperature	Simulated	ITGFLG=1
Heat/Moist fluxes	Simulated	ISFFLX=1
Soil model	5-layer soil model	ISOIL=1
Radiation	RRTM LW/Dudhia SW	IFRAD=4
Deep convection	Grell	ICUPA=3
Horizontal diffusion	Modified	ITPDIF=3

Table 2: Table summarizing the MM5 forecast configuration and the corresponding namelist parameter.

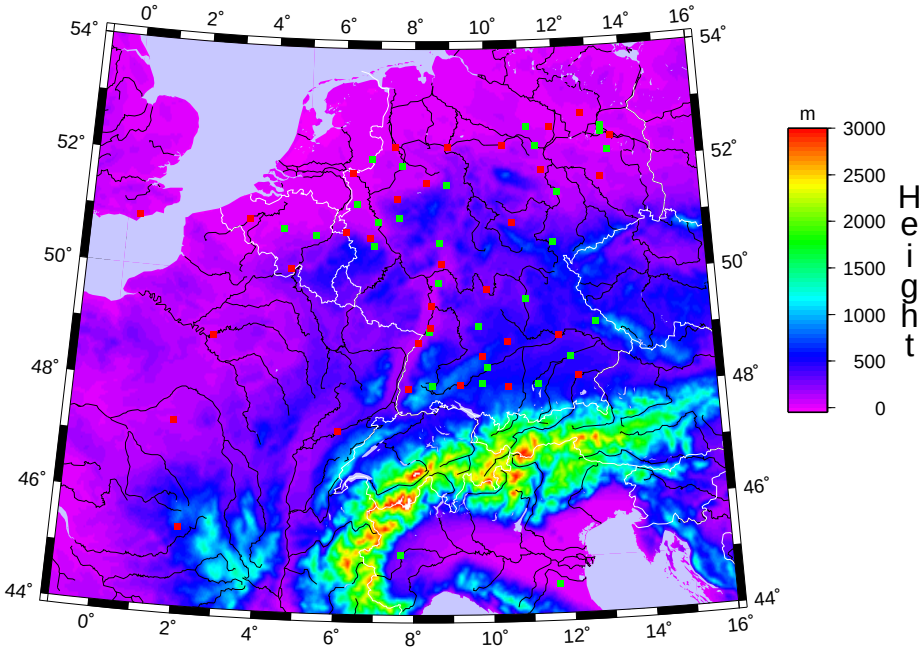


Figure 5: The MM5 domain and the geographical locations of the ground-based receivers (indicated by green and red dots).

referred to as network red and network green. The location of the receivers are indicated by red and green dots. The membership of a specific receiver to one of the networks is more or less randomly chosen. Each receiver is simultaneously tracking several satellites with different azimuth and elevation angels. Notably, the number of receivers used in this study is less then the number of receivers which are actually available. The reason for the selection of a subset of receivers is that the model topography does not match the actual topography in the entire model domain. Though interpolation/extrapolation of model grid point variables to the receiver locations is performed, receivers are not believed to be represented by the model if located far below or above the model topography. Therefore, only receivers are considered which are located at an adequate distance to the model topography: The magnitude of the difference between a bilinear interpolation of adjacent model terrain heights and the receiver height does not exceed a threshold of 50 m. In addition, receivers located close to lateral boundaries are not used. Note that the coordinates of the receivers are provided with respect to the IGb2000 reference frame (www.ngs.noaa.gov/CORS). The ellipsoidal heights of the receivers are converted to the geoidal heights during the preprocessing. For details the reader is referred to Lemoine et al. (1998).

In order to measure the discrepancy between two sets of data a set of statistical quantities is introduced in Appendix D.

5.2 Intercomparison of the BVP solution and the IVP solution

The numerical algorithm to solve the BVP in 2D has been implemented in MATLAB to allow the computation of the ray trajectory and the simulation of the slant path delay. The IVP is solved using the ode45 algorithm available in the MATLAB ODE suite (Shampine and Reichelt, 1997). The ode45 is based on an explicit fourth-order Runge-Kutta formula with adaptive stepsize control. Once the ray trajectory is determined the trapezoidal rule is applied to compute the slant path delay. The solution obtained by the ode45 algorithm deals as the reference solution.

The dry refractivity N_d and the wet refractivity N_w are assumed to be functions of the height h only

$$\begin{aligned} N_d &= D_0 \cdot \exp(-h/H_d) \\ N_w &= W_0 \cdot \exp(-h/H_w) \end{aligned} \tag{67}$$

where $D_0 = 260$ and $W_0 = 120$ denote the dry and wet refractivity at sea level respectively, and $H_d = 8000$ m and $H_w = 2700$ m denote the scale height of the dry and wet refractivity

(Melbourne et al., 1994). The individual refractivity profiles are shown in Figure 6. The Earth's radius is chosen to be $R_E = 6369$ km in this simulation. The partial derivatives of the refractivity with respect to the cartesian coordinates in the local frame of reference are computed by rigorous application of the chain rule of differential calculus.

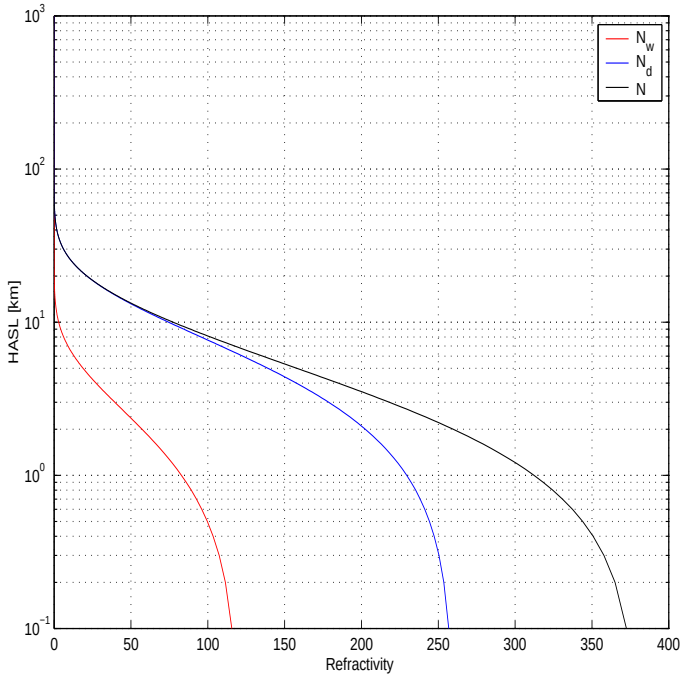


Figure 6: The wet/dry and total refractivity profile according to equation (67).

In the first step the ray trajectory is computed by solving the IVP. A hypothetical ground based receiver is located 100 m above sea level. The explicit numerical integration is carried out to about 20200 km above sea level in order to mimic the approximate orbit of a hypothetical satellite. The tangent of the ray trajectory at the initial point is calculated from a given apparent elevation angle at the receiver. The apparent elevation angle is varied from 5 to 90 degree. As a matter of the bending effect, the apparent elevation angle is different from the geometric elevation angle. Once the final ray traced point is computed, the geometric elevation angle can be determined. For each apparent elevation angle the first and the final ray trace point are fixed. These two points determine the boundary values for solving the BVP in the second step. The number of Legendre-Gauss-Lobatto nodes is chosen to be equal to the number of nodes obtained by the solution of the IVP. The first guess vector in the Newton iteration is chosen to be equal to the coordinates along the geometric path. Over the entire elevation range a single Newton iteration is performed.

Figure 7 shows the altitude of points along the geometric path versus the difference in

altitude of points along the signal path and the geometric path. Each line represents the differences in altitude for a specific elevation angle. According to Figure 7 the agreement of the ray trajectory computed by solving the IVP and the BVP is excellent over the entire elevation range. In fact, this is not suprising since in both cases the underlying ODE is the same. However, the numerical algorithm to solve the ODE is very different. It is remarkable that a single Newton iteration is sufficient for elevation angles down to 5°. This can be explained by the fact that the nonlinear system of equations is well-conditioned and the first guess vector in the Newton iteration is an approximate zero.

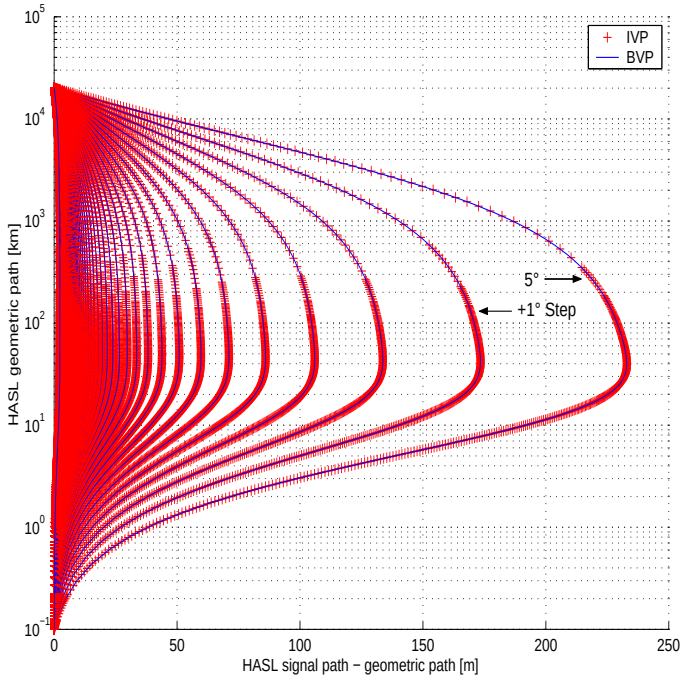


Figure 7: Altitude of points along the geometric path versus the difference in altitude of points along the signal path and the geometric path. The red crosses correspond to the solution of the IVP, the blue solid line corresponds to the solution of the BVP. Each line corresponds to a specific apparent elevation angle (ranging from 5 to 90 degree).

The agreement of the simulated slant path delay computed from the solution of the IVP and the BVP is on a sub-millimeter level over the entire elevation range (not shown). The simulation reveals that the number of nodes increases from about 350 close to the zenith to about 550 for low elevation angles. In the following, the main features of the slant path delay are analyzed. Since the numerical solution of the IVP and the BVP is practically identical, the discussion is restricted to the numerical solution of the BVP.

According to Figure 7, the bending of the ray trajectory is increasing with decreasing elevation angles. The apparent elevation angle is larger then the geometric elevation

angle. Therefore the signal from the satellite to the receiver traverses higher altitudes, corresponding to lower refractivities, when ray-tracing is applied. At the same time the bending effect increases the curvature of the ray trajectory and thus increases the path length. The former effect decreases the slant path delay, where as the latter effect increases the slant path delay. However, these two complementary effects do not compensate, since, according to Fermat's principle, the ray traced solution is the one which minimizes the optical path length and therefore the slant path delay.

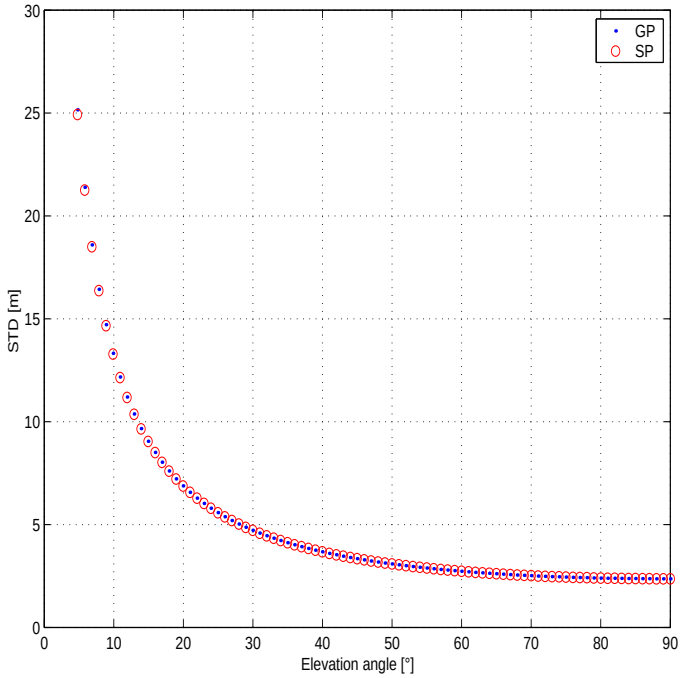


Figure 8: The slant path delay versus the elevation angle. The blue dots correspond to the solution obtained by numerical integration along the geometric path, the red circles correspond to the solution obtained by numerical integration along the signal path.

Figure 8 shows the slant path delay versus the geometric elevation angle. The blue dots correspond to the solution obtained by numerical integration along the geometric path, the red circles correspond to the solution obtained by numerical integration along the signal path. For both options the slant path delay varies from about 2.5 to 25 m for elevation angles from 90 to 5 degree. The elevation dependency of the slant path delay almost obeys a cosecant law. For low elevation angles, the curvature of the Earth and to lesser extent the bending effect lead to significant deviations from a simple cosecant law. In fact, an accurate description of the elevation dependency of the slant path delay is given through the continued fraction expansion provided by Equation (66) for an appropriate set of parameters. According to Figure 8, the slant path delay computed along the signal path is slightly reduced when compared to the slant path delay computed along the geometric

path. The difference between the slant path delay computed along the geometric path and the slant path delay computed along the signal path versus the elevation angle is shown in Figure 9. The systematic differences in the slant path delay at elevation angles of 30, 20, 10 and 5 degree are about 1, 4, 30 and 230 mm respectively. Though the differences are small, they are not negligible. The computation of the slant path delay along the geometric path is of limited value for elevation angles below 30°, if a reasonable accuracy constraint of 1 mm is set.

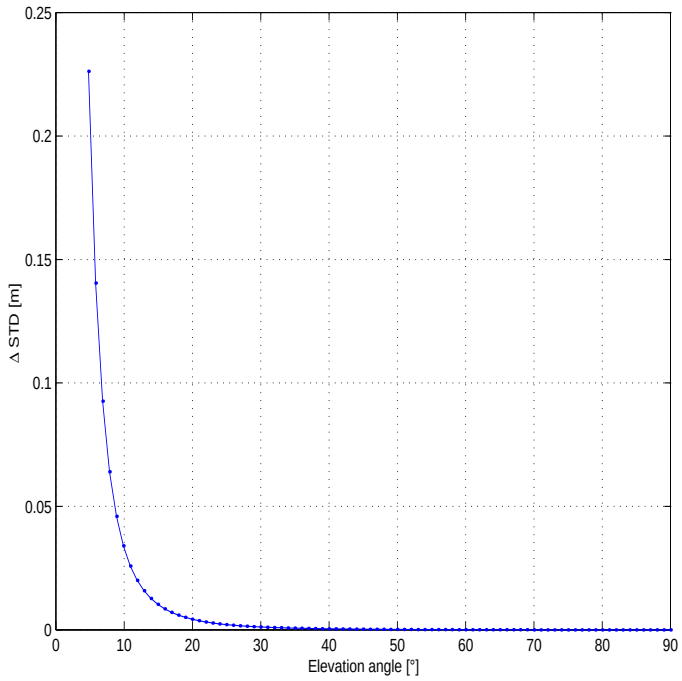


Figure 9: The difference in the slant path delay versus the elevation angle.

Figure 10 shows the slant geometric delay versus the geometric elevation angle. The differences at elevation angles of 30, 20, 10 and 5 degree are about 1, 3, 25 and 220 mm respectively. It is concluded that at elevation angles below 30° degree the slant geometric delay can not be neglected.

Finally, Figure 11 shows the bending angle versus the geometric elevation angle. As expected, the bending angle rapidly increases with decreasing elevation angles.

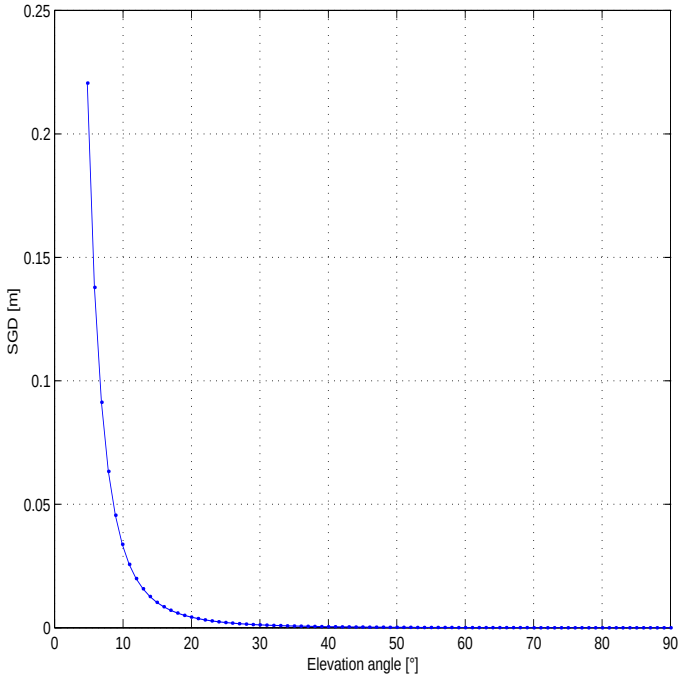


Figure 10: The slant geometric delay versus the elevation angle.

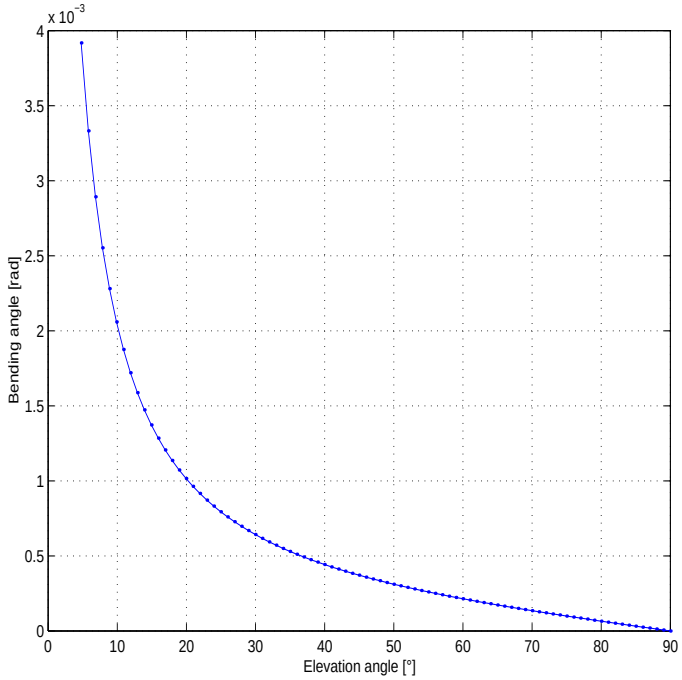


Figure 11: The bending angle versus the elevation angle.

5.3 Intercomparison of the BVP solution and an analytic solution

The numerical algorithm to solve the BVP in 2D as well as in 3D was implemented in FORTRAN to allow the computation of the ray trajectory and the simulation of the slant path delay in the MM5 environment. In order to check the implementation for correctness, Snell's law for a spherically layered atmosphere, is used.

The refractivity at each grid point of the MM5 domain is computed according to the Standard Atmosphere (1976). The MM5 grid configuration is given in Table 1. An hypothetical ground based receiver is located in the center of the MM5 domain at an altitude of 300 m. Hypothetical satellites are distributed in circular orbits at an altitude of 20200 km. The elevation angle at the receiver varies from 1 to 90 degree. The azimuthal angle at the receiver varies from 0 to 270 degree with a stepsize of 45°. Over the entire elevation and azimuthal range 550 nodes are used. The first guess vector in the Newton iteration is chosen to be equal to the geometric path. Two Newton iterations are performed. For a given elevation and azimuthal angle at the receiver the ray trajectory and the impact parameter at each point along the ray trajectory are computed. The BVP is solved in 2D and 3D.

The results of this simulation are summarized as follows: At first, no differences in the simulated ray trajectory obtained by solving the BVP in 2D and 3D are found. This is expected, since the atmosphere is spherically layered. Any difference in the simulated ray trajectory would indicate a problem in either one or both implementations. Second, no azimuthal dependency on the quality of the solution is found. That is, the numerical algorithm works equally well over the entire azimuthal range. Third, along each simulated ray trajectory the numerical solution of the BVP almost exactly obeys Snell's law. For example, at an elevation angle of 85° the impact parameter evaluated at the receiver is of the order of 10^5 m and varies along the ray trajectory on a centimeter level. Likewise, at an elevation angle of 5° the impact parameter evaluated at the receiver is of the order of 10^6 m and varies along the ray trajectory on a meter level. This is the case if either one or two Newton iterations are performed. If the ray trajectory is approximated by the geometric path, i.e. no Newton iteration is performed, the variation of the impact parameter at an elevation angle of 85° is of the order of 10^2 m and the variation in the impact parameter at an elevation angle of 5° is of the order of 10^3 m. In other words, the impact parameter computed along a straight line does not obey Snell's law.

Table 3 summarizes the impact parameters computed at the receiver and the satellite for a variety of elevation angles. The variation of the impact parameter is on a meter level for elevation angles above 3° if one/two Newton iterations is/are performed. No clear tendency in the variation of the impact parameter with respect to the elevation angle is visible, indicating that the variation of the impact parameter is due to numerical noise.

For elevation angles below 5°, the simulation with two Newton iterations performs best, i.e. the agreement in the impact parameter remains on a meter level. A single Newton iteration is sufficient as long as the elevation angle is not below 5°.

Iteration	0		1		2	
ε[deg]	p_a[m]	p_b[m]	p_a[m]	p_b[m]	p_a[m]	p_b[m]
1	6369980	6368285	6369228	6369192	6369214	6369213
3	6362219	6360526	6360940	6360930	6360932	6360933
5	6346707	6345018	6345234	6345229	6345232	6345229
7	6323462	6321779	6321904	6321905	6321903	6321905
9	6292513	6290838	6290918	6290921	6290917	6290921
11	6253898	6252233	6252298	6252291	6252297	6252291
21	5947795	5946211	5946231	5946229	5946231	5946229
41	4808217	4806937	4806950	4806942	4806950	4806942
61	3088698	3087876	3087884	3087877	3087884	3087877
81	996636	996371	996371	996371	996371	996371
85	555264	555117	555117	555117	555117	555117
89	111188	111158	111158	111158	111158	111158

Table 3: The impact parameter at the receiver p_a and the satellite p_b for a variety of elevation angles ε at the receiver and number of Newton iterations.

5.4 Accuracy and speed

The previous simulation confirms that both implementations in the MM5 environment, i.e. the numerical algorithm to solve the BVP in 2D and 3D, work as intended, and are sufficiently accurate under standard atmospheric conditions. However, the standard refractivity field differs considerably to the refractivity field simulated by a mesoscale model. It is natural to assume that for a given refractivity field the accuracy of simulated slant path delays depends on the number of nodes, the interpolation scheme involved in the computation of the refractivity and the number of Newton iterations.

In order to estimate the accuracy and the speed of the numerical algorithm, an experiment was conducted for a particular day and hour. A short-range forecast from 0-6UTC is performed for the 7 of August, 2007. The MM5 configuration is given in Table 1 and 2. This particular day is chosen since strong gradients in the refractivity with respect to longitude, latitude and height are present in the model atmosphere. Figure 12 shows the IWV field and Figure 13 shows the vertical cross sections of the water-vapor mixing ratio at constant latitude/longitude valid at 3UTC. Hypothetical ground-based receivers are located at model terrain grid points and hypothetical satellites are evenly distributed

in circular orbits at an altitude of 20200 km. The elevation angle varies from 5 to 90 degree and the azimuthal angle varies from 0 to 270 degree with a step size of 45° at each receiver. The number of simulated slant path delays is 54180.

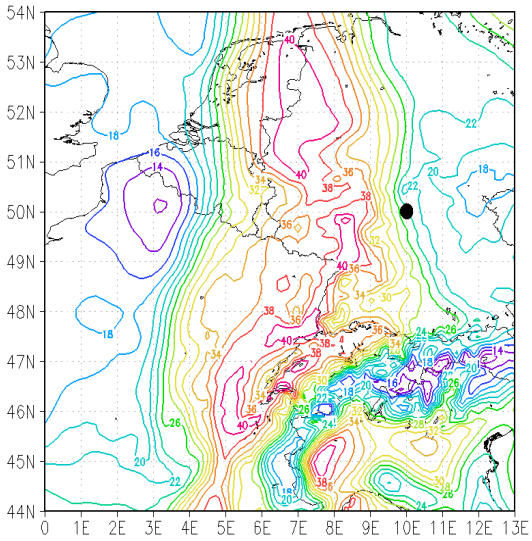


Figure 12: The IWV field [mm] at 3UTC on August 7, 2007. The black dot indicates the location of a hypothetical receiver (10E 50N).

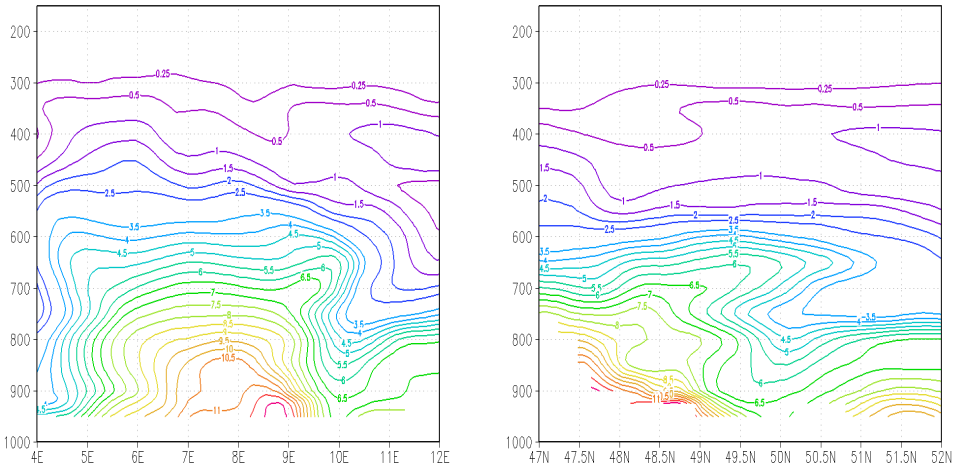


Figure 13: Vertical cross section of the water vapor mixing ratio [g/kg] at latitude 10E (left panel) and longitude 50N (right panel) at 3UTC on August 7, 2007.

Table 4 summarizes the configuration for three selected simulations A, B and C. The simulations differ by the type of the BVP being solved, i.e. the two dimensional BVP or the three dimensional BVP, the number of Netwon iterations, the number of nodes and the interpolation scheme. The speed is measured in terms of the number of simulated slant path delays per second. The simulations are performed with a single CPU. The simulation A, which solves the BVP in the three dimensional case using 2000 nodes, 2 Newton iterations and bi-cubic interpolation deals as the reference simulation. Increasing the number of nodes to 3000 and the number of Newton iterations to 3 does not alter simulated slant path delays significantly, i.e. the magnitude of maximal differences in the computed slant path delays are < 1 mm.

Configuration	BVP type	Iteration	Nodes	Interpolation	No. delays/second
A	3D	2	2000	bi-cubic	1.5
B	2D	1	1500	bi-cubic	10.0
C	2D	1	1500	bi-linear	45.0

Table 4: Table summarizing the configuration for three selected simulations A, B and C. The configurations differ by the number of Newton iterations, the number of nodes and the interpolation method. The computational speed is defined as the number of simulated slant path delays per second using a single CPU.

The differences of simulated slant path delays versus the elevation angle are shown in Figure 14. The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD versus the elevation angle. The magnitudes of differences increase with decreasing elevation angles. The magnitude of maximal differences between the simulation B and A are on a millimeter level over the entire elevation range. The magnitude of maximal differences between the simulation C and A reach at low elevation angles the centimeter level. Apparently, the differences are due to the different interpolation schemes. No systematic difference is obvious. The RMS and the magnitude of the MD are on a sub-millimeter level close to the zenith. Both configurations B and C are concluded to be sufficiently accurate.

The experiment reveals that the accuracy of simulated slant path delays depends primarily on the number of nodes, the interpolation scheme for the refractivity and the number of Newton iterations. Slant path delays simulated in the refractivity field of a mesoscale model do not deviate significantly if the BVP is solved in the two-dimensional or three-dimensional case. This can be explained by the fact that the refractivity is more or less stratified with respect to height. At an arbitrary location the refractivity gradient in the x-z plane is much larger then the refractivity gradient in the x-y plane. Though it is observed that the projection of the ray trajectory in the x-y plane is not a straight line, the variations in the x-y plane are too small to be significant in the computation of slant path delay. A number of additional simulations in a variety of refractivity fields were

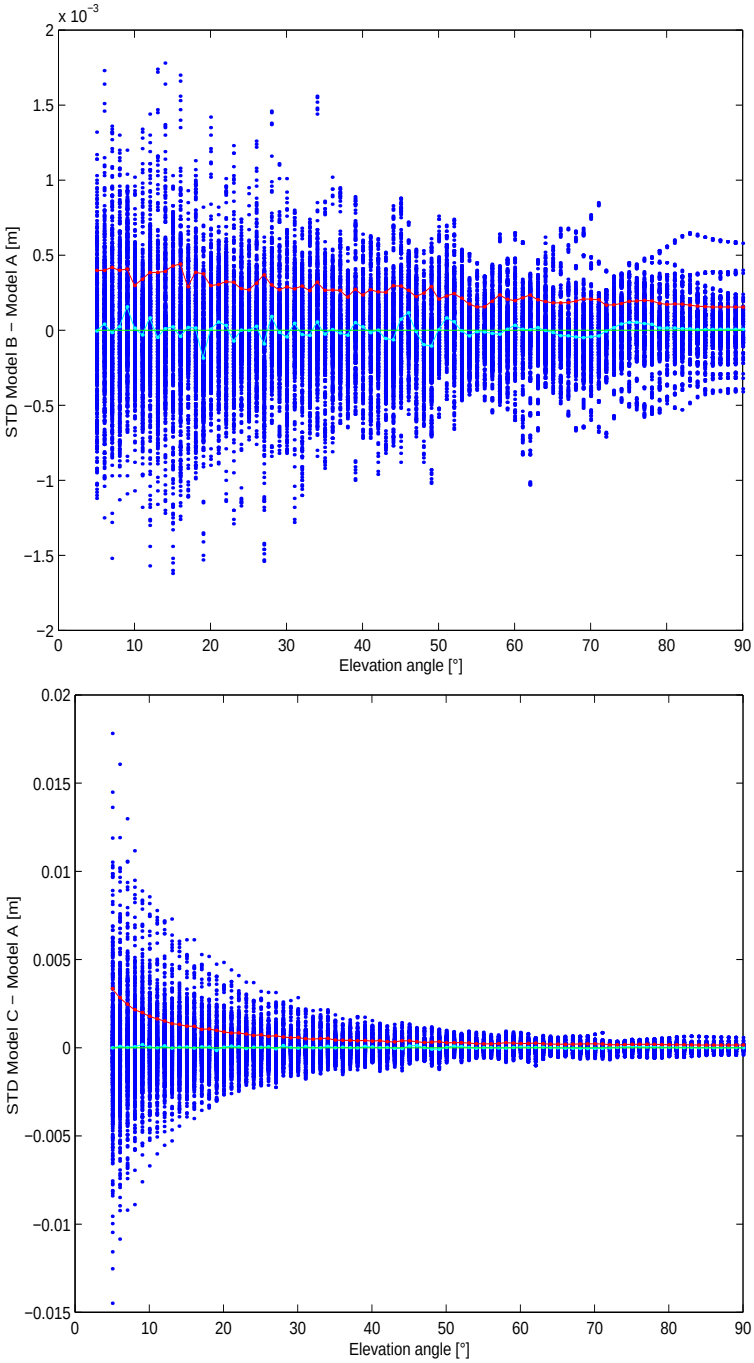


Figure 14: Differences of simulated slant path delays versus the elevation angle for the simulation B and A (upper panel) and the simulation C and A (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD versus the elevation angle.

performed and supported this finding. Nevertheless, it is not precluded that under certain atmospheric conditions, e.g. if the ray trajectory is simulated in the vicinity of a deep low pressure system, this conclusion holds. A similar conclusion was drawn by Hobiger et al. (2008) who analyzed the differences of simulated slant path delays in the vicinity of a typhoon. In the remaining simulations performed in this thesis the configuration B is used, unless otherwise explicitly noted.

5.5 Asymmetry

In this section the potential asymmetric information content of slant path delays is estimated. Simulated slant path delays are analyzed for a hypothetical ground based receiver located in the vicinity of a strong IWV gradient. The particular day and hour in this experiment is the same as in the previous simulation. The receiver is located at 10E 50N at an altitude of 280 m. The elevation angle varies from 5 to 50 degree and the azimuthal angle varies from 0 to 355 degree at the receiver. Figure 12 shows the IWV field in the vicinity of the receiver. A moisture front ranging from south to north located at about 10E separates moist from dry air. The vertical cross sections of the water vapor mixing ratio indicate that the water vapor mixing ratio increases form east to west and north to south in the vicinity of the receiver. How does the asymmetry present in the model space convert to the asymmetry in observation space? The relative asymmetry in observation space A can be defined as

$$A(\varepsilon, \phi) = \frac{S(\varepsilon, \phi) - S(\varepsilon, \phi + \pi)}{S(\varepsilon, \phi)} \tag{68}$$

where the slant path delay S is computed for the elevation angle ε and the azimuthal angle ϕ.

Figure 15 shows the relative asymmetry of the slant path delay in a sky plot. The plot is oriented such that the northern (eastern) horizon is on top (right) and the zenith is in the middle of the panel. The magnitude of the relative asymmetry of the slant path delay increases with decreasing elevation angles and takes a maximum value at an elevation angle of about 5° and an azimuthal angle of about 90°. Even for low elevation angles the magnitude of the relative asymmetry of the slant path delay is small. This can be explained by the fact that the relative asymmetry in the dry slant delay is fairly small (not shown). However, the gradient in the wet refractivity with respect to longitude and latitude is relatively large. Indeed, the main contribution to the asymmetry of the slant path delay stems from the asymmetry of the slant wet delay. Figure 16 shows the relative asymmetry of the slant wet delay in a sky plot. Virtually, the relative asymmetry in Figure 15 and 16 is the same. However, the magnitude of the relative asymmetry of the

slant wet delay reaches values of up to 30 percent at low elevation angles. The sign of the relative asymmetry of the slant wet delay changes from negative at an azimuthal angle of about 90° to positive at an azimuthal angle of about 180°. This finding agrees very well with the water vapor distribution in the vicinity of the receiver as shown in Figure 13. Noteably, the magnitude of the relative asymmetry of the slant wet delay at an elevation angle of 5° is about 10 times larger then at an elevation angle of 30°.

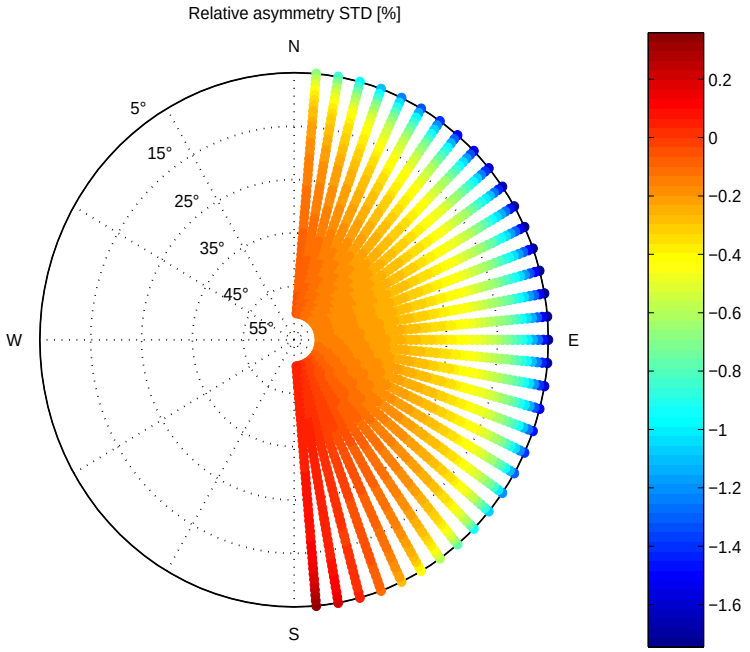


Figure 15: Sky plot of the relative asymmetry of the slant path delay in percent at the hypothetical receiver located at 10E 50N. The elevation angle varies from 5° (boundary of the panel) to 50° (middle of the panel).

The simulation suggests, that in particular slant path delays at low elevation angles can contain valueable asymmetric information. However, in the case of realistic satellite orbits it is not possible to recover asymmetry in observation space with such great detail. In the current satellite constellation maximal 10 satellites are simultaneously in view at a time. Elevation and azimuthal angels at a receiver are unevenly distributed and there are segments in the sky plot, depending on the geographical location of the receiver, where satellits are never in view of the receiver. For a comprehensive analysis on the information content provided by slant path delays from a network of receiver stations in Germany the reader is referred to (Bender et al., 2009).

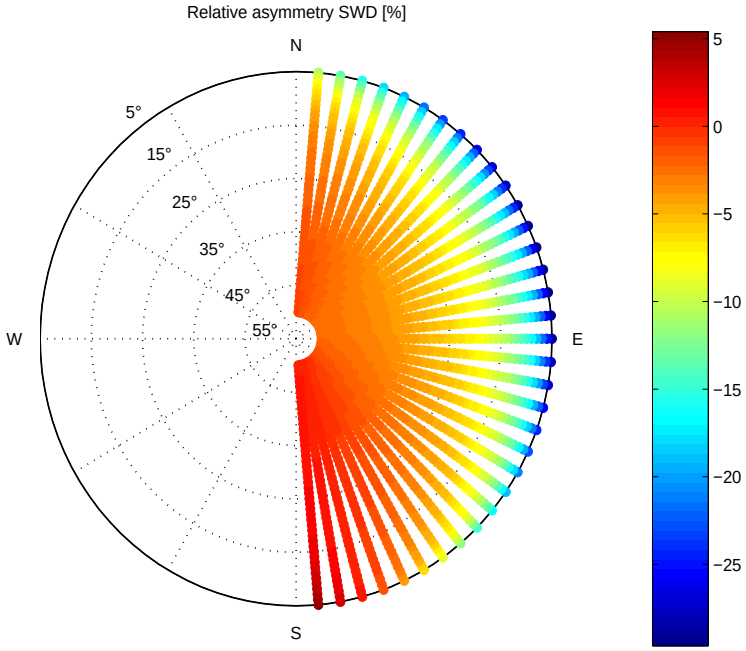


Figure 16: Sky plot of the relative asymmetry of the slant wet delay in percent at the hypothetical receiver located at 10E 50N. The elevation angle varies from 5° (boundary of the panel) to 50° (middle of the panel).

5.6 Statistical intercomparison of simulated and observed data: Part 1

In this section a statistical intercomparison of simulated and observed data is provided. The time period under investigation is 18-31 of August 2007. In order to minimize errors associated to the forecast model, short-range forecasts from 0-6UTC are performed for each day. The MM5 configuration is given in Tables 1 and 2. The forecast model output frequency is 15 minutes. Slant path delay observations collected at all receiver locations indicated in Figure 5 enter the intercomparison. The full time resolution of 2.5 minutes is used. In case the time of the observation does not match the forecast model output time, the nearest available model output time is used. With this setup, 732668 slant path delay observations enter the intercomparison. Hereinafter, the model prediction, i.e. the simulated slant path delays at appropriate time, will be referred to as the background. Thus, the background is obtained by integrating the initial state in time using the forecast model and mapping the forecast model prediction to observation space using the forward model. Background minus Observation (BmO) differences are therefore a composite of forward model errors, forecast model errors, errors in the initial state and observation errors. In this description, remaining error sources, i.e. representative errors in space and time, are not considered. In the following, the BmO differences are computed and compared for two different backgrounds L and R, which are generated using two different forward models. The forward models differ by the number of Newton iterations which are performed in the signal path determination. The background L is generated using no Newton iteration and the background R is generated using one Newton iteration. Since the same observations, the same forecast model and the same initial state is used in both cases, the BmO differences that are revealed are due to differences in the forward model.

Figure 17 shows the BmO differences versus the elevation angle for the background L and R. The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD versus the elevation angle. The RMS and the MD are computed from the BmO differences with a sliding average of 4°. The magnitudes of BmO differences for both options are small close to the zenith and become progressively larger for lower elevation angles. The relative level of agreement of simulated and observed slant path delays, i.e. the normalized deviation, is independent of the elevation angle. For elevation angles above 30° the BmO differences for both options are practically identical. For elevation angles below 30° however the BmO differences for both options start do deviate from each other considerably. Whereas the MD of the BmO differences for the background R remains close to zero, the MD of the BmO differences for the background L rapidly increases with decreasing elevation angles. This is expected: Since the forward model used to determine the background L does not account for the bending effect of the signal path, simulated slant path delays at low elevation angles are overestimated relative to the observations.

According to Figure 17 the BmO differences are on a centimeter level over the entire elevation range. Since slant path delays range from about 2.5 m close to the zenith to about 25 m at an elevation angle of 5° degree, the correlation of simulated and observed slant path delays is high. This is not too surprising, since the main contribution to the slant path delay stems from the slant dry delay S_d, which is believed to be well represented in both, the model prediction and the observation. From its definition the BmO statistics does not allow for a strict separation of BmO differences of individual contributions, e.g. BmO differences of the slant wet delay S_w. However, assuming that the main contribution to the BmO differences stem from the BmO differences of the S_w, allows one to estimate the MD and the RMS of the S_w. In addition, the conversion factor Π allows one to estimate the MD and the RMS of the slant water vapor SWV. Close to the zenith the MD and the RMS of the S_w are -3 mm and 8 mm. Therefore, the MD and the RMS of the IWV are about -0.5 mm and 1.3 mm. That is, the estimated agreement of simulated and observed IWV is good. Noteably, the RMS of IWV, estimated from the BmO statistics of the slant path delays, is consistent with statistical intercomparisons of the IWV estimates processed at the GFZ Potsdam and the Lokal Model (LM) of the German Weather Service (DWD) (Dick et al., 2001).

From the RMS of the slant wet delay the variation of the water vapor mixing ratio in the planetary boundary layer (PBL) can be estimated. In the zenith case the variation of the slant wet delay S_w with respect to the wet refractivity N_w reads as

$$\delta S_w \;=\; 10^{-6} \cdot \sum_{i=0}^{n+1} \delta N_w(x_i, y_i, z_i)\, w_i$$

The variation of the wet refractivity N_w with respect to the water vapor mixing ratio q is given by

$$\delta N_w \;=\; \frac{0.622 \cdot N_w}{q \cdot (q + 0.622)} \delta q$$

Assuming that the water vapor mixing ratio is 6.5 gkg^{-1} throughout the PBL of depth 2500 m and zero above and that the temperature and the pressure are given by the Standard Atmosphere (1976), leads to a slant wet delay of about 110 mm. Therefore, the variation of the slant wet delay of ± 8 mm corresponds to a variation of the water vapor mixing ratio throughout the PBL of ± 0.5 gkg^{-1}.

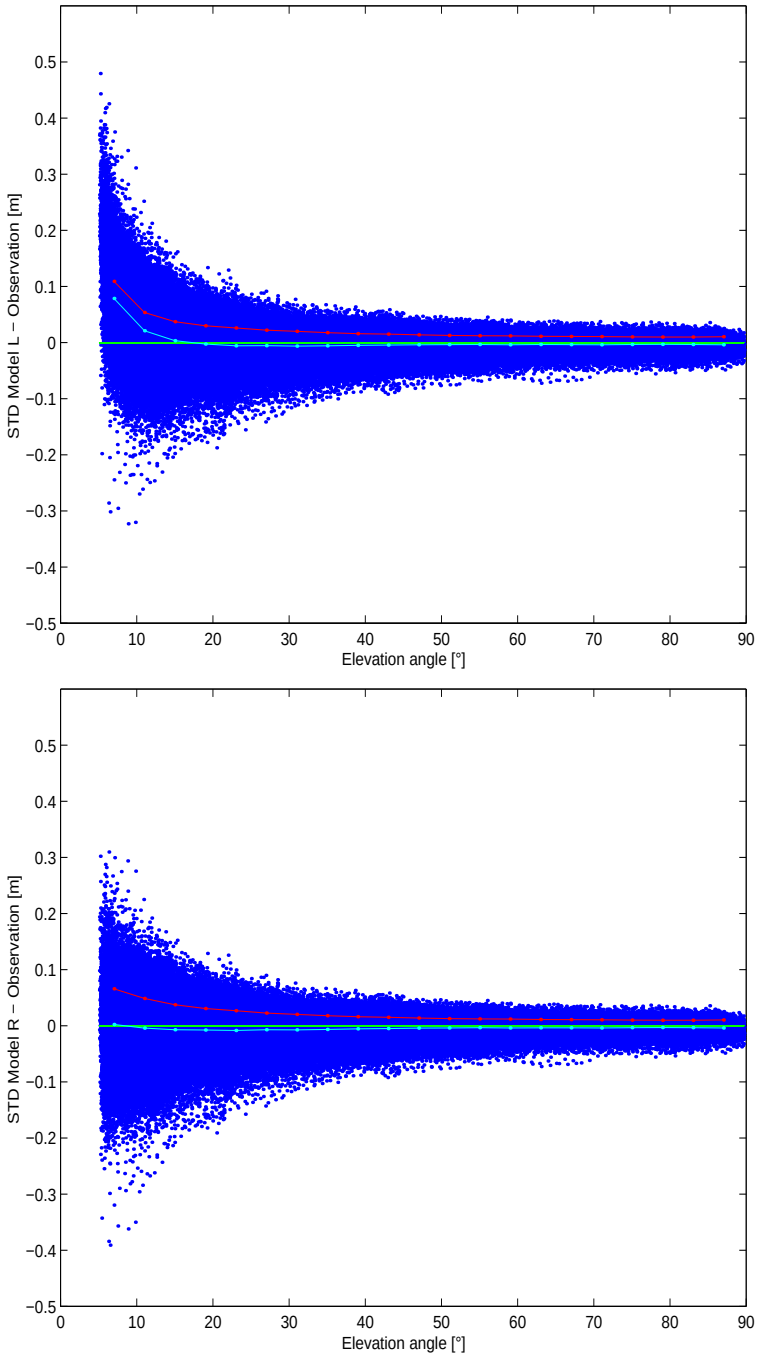


Figure 17: BmO differences versus the elevation angle for the background L (upper panel) and the background R (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD versus the elevation angle.

6 Sensitivity study

The MM5 is an open source research model. It differs to operational NWP models in that it allows the user to specify the model grid configuration and to select one of different parameterization schemes. As an open source research model, it offers the unique possibility to study the sensitivity of physical processes with respect to individual model components. In addition, the continuously operating GPS receiver network in Germany provides data, which can be used for validation purposes. In this section the sensitivity of simulated slant path delays with respect to two forecast model components, namely the horizontal diffusion scheme and the parameterization of deep convection, is investigated.

6.1 Horizontal diffusion

6.1.1 The current horizontal diffusion scheme

In the MM5, two types of horizontal diffusion are used to control non-linear instability: second-order diffusion is used for the row and column of the grid points next to the lateral boundaries, while the more scale-selective fourth-order diffusion is used in the interior of the domain (Grell et al., 1995). In analytical form, the fourth-order diffusion F reads:

$$F = -K_h \left(\frac{\partial^4 \alpha}{\partial x^4} + \frac{\partial^4 \alpha}{\partial y^4} \right) \tag{69}$$

where α is any prognostic variable and K_h denotes the horizontal diffusion coefficient. The horizontal diffusion is applied on constant sigma levels and not on constant height levels. Therefore serious errors over complex terrain, particularly for atmospheric variables having a strong vertical gradient, like water vapor and temperature, are introduced (Li and Atkinson, 1999).

In the MM5 version used in this thesis, two options for calculating the horizontal diffusion of temperature are available. In the first option, actual temperature is used for computing the diffusion on constant sigma levels. In the second option, the perturbation temperature, i.e. the difference between the actual temperature and a given reference temperature, is used for computing the diffusion on constant sigma levels. The first option tries to equalize the temperature differences on the terrain following model surfaces, and therefore tends to cool valleys and to heat mountains. The second option tries to establish a vertical temperature gradient within valleys leading to warmer valleys as compared to the first option. The second option effectively suppresses the erroneous temperature tendency introduced by the first option. However, both options have deficiencies (Zängl,

2002): Above mountainous terrain, a strange and noisy pattern appears in the temperature that is reminiscent of breaking gravity waves. This is explained by the fact that diffusion on constant sigma levels induces horizontal temperature gradients over mountainous topography. As a result, pressure gradients occur and gravity waves are excited. In both options, diffusion is computed along sigma levels for water vapor. Therefore, in both options, a spurious moisture tendency exists.

To prevent spurious temperature or moisture tendencies, diffusion should be computed truly horizontally as suggested by Zängl (2002). At high elevations, this can be done by simply interpolating between vertical coordinate surfaces. The vertical interpolation is performed linearly with height for the temperature and an exponential interpolation is used for the water-vapor mixing ratio. At lower elevations, truly horizontal computation may be impossible without intersecting the ground. Solutions are either to switch back to diffusion along sigma surfaces, using one-sided truly horizontal diffusion or to apply the full coordinate transformation to the horizontal diffusion operator. From an analytical point of view, the latter method is superior to vertical interpolation but it is computationally expensive. Since the first options retains some error in complex terrain, the second option, a modified horizontal diffusion scheme similar to the one proposed by Zängl (2002) was implemented.

6.1.2 The modified horizontal diffusion scheme

The focus is on the approximation of the fourth partial derivative of the water vapor mixing ratio. The approximation of the fourth partial derivative of the temperature is treated in a similar manner except that the vertical interpolation is linear. Let I, J and K denote the grid point indices of a point in the model grid with coordinates y, x and z at which the fourth partial derivative of water vapor with respect to y is to be computed. The second-order accurate discretization of the fourth partial derivative on sigma levels is altered to be:

$$\Omega_f \;=\; \left(Q_A - 4 \cdot Q_B + 6 \cdot Q_{IJK} - 4 \cdot Q_C + Q_D\right) / \Delta y^4$$

where Δy denotes the grid increment,

$$\begin{aligned} Q_A &= Q_{I+2JA} \cdot \left(Q_{I+2JA} / Q_{I+2JA-1}\right)^{L_A} \\ Q_B &= Q_{I+1JB} \cdot \left(Q_{I+1JB} / Q_{I+1JB-1}\right)^{L_B} \\ Q_C &= Q_{I-1JC} \cdot \left(Q_{I-1JC} / Q_{I-1JC-1}\right)^{L_C} \\ Q_D &= Q_{I-2JD} \cdot \left(Q_{I-2JD} / Q_{I-2JD-1}\right)^{L_D} \end{aligned}$$

and Q denotes the gridded water-vapor mixing ratio. The vertical interpolation coefficients are computed as

$$\begin{aligned} L_A &= (H_{IJK} - H_{I+2JA})/(H_{I+2JA} - H_{I+2JA-1}) \\ L_B &= (H_{IJK} - H_{I+1JB})/(H_{I+1JB} - H_{I+1JB-1}) \\ L_C &= (H_{IJK} - H_{I-1JC})/(H_{I-1JC} - H_{I-1JC-1}) \\ L_D &= (H_{IJK} - H_{I-2JD})/(H_{I-2JD} - H_{I-2JD-1}) \end{aligned}$$

and the adjacent grid point indices with respect to height A, B, C and D are determined by binary search. Obviously, the expression for sigma diffusion is recovered in flat terrain. The discretization scheme Ω_f is applied at all model grid points where the vertical extrapolation does not intersect the model topography. At model grid points where the vertical extrapolation intersects the model topography, i.e. the discretization scheme Ω_f is not applicable, the altered second order accurate discretization of the second partial derivative on sigma levels

$$\Omega_s = (Q_B - 2 \cdot Q_{IJK} + Q_C)/\Delta y^2$$

is applied. At model grid points where the discretization Ω_s is not applicable one of the following discretization schemes is used

$$\begin{aligned} \Omega_{f-} &= (Q_D - 4 \cdot Q_C + 3 \cdot Q_{IJK})/\Delta y^4 \\ \Omega_{f+} &= (Q_A - 4 \cdot Q_B + 3 \cdot Q_{IJK})/\Delta y^4 \\ \Omega_{s-} &= (Q_B - Q_{IJK})/\Delta y^2 \\ \Omega_{s+} &= (Q_C - Q_{IJK})/\Delta y^2 \end{aligned}$$

where Ω_{f-}/Ω_{f+} and Ω_{s-}/Ω_{s+} can be derived from Ω_f and Ω_s assuming a transparent boundary. At model grid points where none of the discretization schemes introduced above is applicable, i.e. at a model grid point located in a narrow valley, the second order accurate discretization of the second partial derivative of the water vapor mixing ratio/temperature perturbation on sigma levels is used. In a narrow valley, the discretization scheme thus retains some error. Note that the right hand side of the horizontal diffusion equation needs a negative/positive sign for the fourth-/second-order discretization to damp noise. The same procedure is applied in the x direction. Once the partial derivative with respect to y and x is computed, the update of the water vapor tendency due to horizontal diffusion is straightforward.

Hereinafter, the modified horizontal diffusion scheme will be referred to as HD scheme, sigma diffusion applied to the water vapor mixing ratio and the temperature will be referred to as SD scheme and sigma diffusion applied to the water vapor mixing ratio and the temperature perturbation will be referred to as PD scheme.

6.2 Sensitivity test 1: Horizontal diffusion

The results of three model simulations are presented for the 22 of August, 2007. Short-range forecasts from 0-6UTC are performed, initialized from the same initial state valid at 0UTC. The forecast model configuration is given in Table 1. Instead of the Grell cumulus scheme the Kuo cumulus scheme is used in all three simulations. The three simulations differ by the horizontal diffusion scheme only. The simulation A, B and C use the SD, the PD and the HD scheme respectively. To show the erroneous moisture tendency introduced by the SD and the PD scheme, the focus is on complex terrain (south Germany).

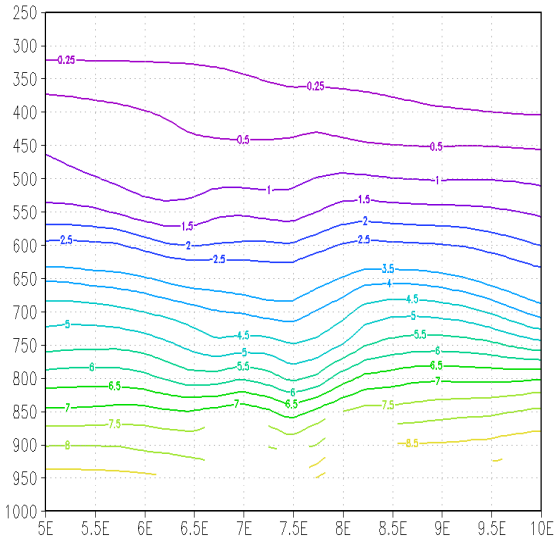


Figure 18: Vertical cross section of the water vapor mixing ratio [g/kg] at constant latitude 48N valid at 0UTC on August 22, 2007.

Figure 18 shows the vertical cross section at constant latitude 48N of the water vapor mixing ratio of the initial state. The white shaded area close to the ground at 7E and 8E represent the Vogues mountains to the west and the Blackforest to the east of the Rhine valley, which is located at about 7.5E. The water vapor mixing ratio in the initial state is more or less stratified with respect to height.

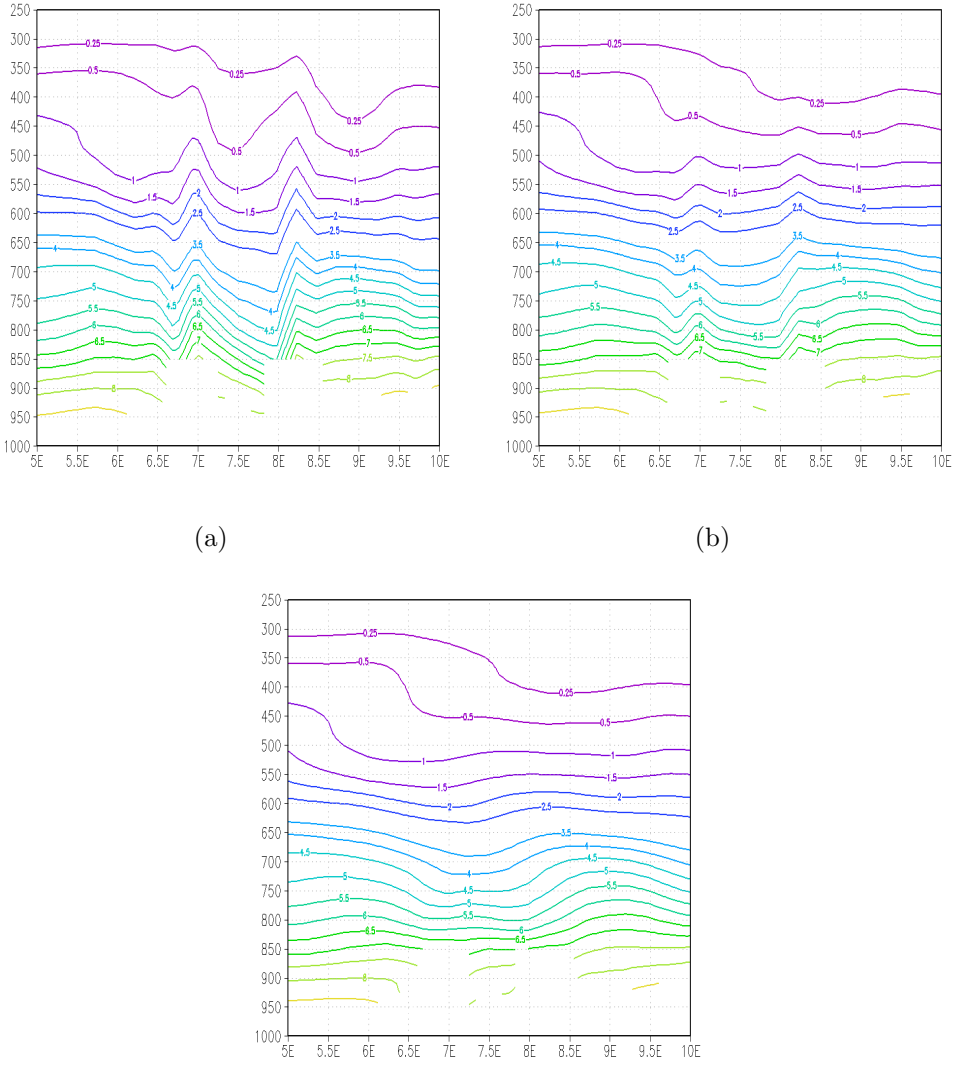
Figure 19 shows the same vertical cross section valid at 2UTC for the three simulations. The simulation A and B show a peak at 7E and 8E, i.e. above the mountains crests, and a trough at 7.5E, i.e. the valley, in the water vapor mixing ratio. The peaks and the trough in the water vapor mixing ratio in the simulation A and B appear instantly after model initialization and vary only slightly over the entire model forecast range (not shown). The peaks and the trough in the water vapor mixing ratio are absent in the simulation C.

The peak and trough in the water vapor mixing ratio appear in simulation A and B due to the erroneous moisture tendency introduced by the SD and the PD schemes. The erroneous moisture tendency is most pronounced using the SD scheme. A spurious moisture tendency however is also present using the PD scheme. The difference is that the SD scheme applies diffusion on sigma levels for both the water vapor mixing ratio and the temperature whereas the PD scheme applies diffusion on sigma levels for the water vapor mixing ratio only. Thus, there is not only a moistening but also warming effect on mountain crests when the SD scheme is used. The moistening and the warming on the mountain crests triggers deep convection, i.e. on top of the Vogues mountains and the Blackforest convective rainfall is enhanced (see next section). The HD scheme behaves more or less neutral to the stratification of the water vapor mixing ratio with respect to height. In fact, under atmospheric conditions where the temperature and the water vapor mixing ratio are uniformly stratified with respect to height the HD scheme will behave neutral, whereas this will not be the case for the SD or the PD scheme.

It is interesting to note that the structure in the water vapor mixing ratio, visible in the simulations A and B, is completely absent in the initial state. A very interesting detail, with simple explanation: The initial state is obtained from an analysis generated at the ECMWF using the Integrated Forecast System (IFS). In the IFS, fourth order diffusion is applied on hybrid vertical surfaces to moisture and on pressure surfaces to temperature. The hybrid coordinate system of the IFS differs to the sigma coordinate system of the MM5 in that the vertical surfaces are not steeply sloped in regions where the terrain is steeply sloped. The application of fourth order diffusion on hybrid vertical surfaces to moisture is not too crucial. The horizontal diffusion scheme for temperature used in the IFS is very similar to horizontal diffusion scheme for temperature used in the HD scheme. In this sense, the simulation C can be roughly regarded as the continuation of the IFS integration.

6.3 Sensitivity test 2: Deep convection

Another issue in the application of the MM5 is related to the convection parameterization scheme, which at best can be regarded as a crude representation of deep convection. The MM5 offers the user to select one of various more or less sophisticated convection parameterization schemes. A general recommendation does not exist. However, some comments



(a) (b)

(c)

Figure 19: Vertical cross section of the water vapor mixing ratio [g/kg] at constant latitude 48N valid at 2UTC on August 22, 2007 for different model forecasts. a) SD scheme b) PD scheme c) MD scheme.

on the performance of individual convection parameterization schemes are provided. For example, it is known that the Kuo cumulus scheme tends to produce much convective rainfall and less resolved-scale precipitation. In particular the Kuo cumulus scheme is mostly applicable to larger grid sizes > 30 km whereas the Grell cumulus scheme is especially usefull for smaller grid sizes $10 - 30$ km and tends to allow a balance between resolved-scale rainfall and convective rainfall (Grell et al., 1995).

To illustrate the differences between at least two convection parameterization schemes an additional simulation for the 22 of August, 2007 from 0-6UTC were performed. The simulation D uses the Grell cumulus scheme in conjunction with the HD scheme. To outline the differences between the Grell cumulus scheme and the Kuo cumuls scheme, it is sufficient to inspect by eye the 15 minutes accumulated precipitation pattern. For a qualitative intercomparison with observed data, Figure 20 shows the measured reflectivity of the DWD radar composite at 2 UTC. Figure 20 indicates two strong precipitation events, i.e. a rainband covering large parts of north-western Germany and an isolated strong precipitation event located over eastern Germany. A time series of the measured reflectivity of the radar composite shows that the rainband is close to stationary and the isolated strong precipitation event vanishes after a few hours. Southern Germany is hardly effected by any precipitation.

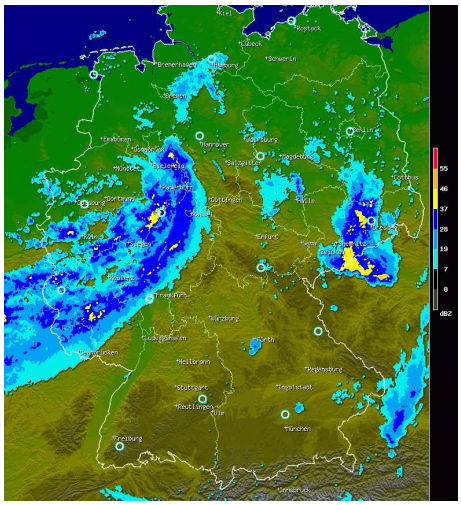


Figure 20: Composite reflectivity of the DWD radar network at 2 UTC on August 22, 2007.

Figure 21 shows the 15 minutes accumulated precipitation for the model simulations A, B, C and D valid at 2 UTC. Obviously the rainband covering large parts of north-western Germany visible in Figure 20 is captured qualitatively by all model simulations. The isolated strong precipitation event in east Germany is captured only by the model simulations A, B and C, i.e. the model simulations using the Kuo cumulus scheme. The

model simulation D, i.e. the model simulation using the Grell cumulus scheme does not capture this precipitation event. The most striking differences between the model simulations are however found in southern Germany. In the simulation A, B and C convection is initiated instantly after model initialization and the 15 minutes accumulated precipitation pattern does not change significantly in time (not shown). The ability of the Kuo cumulus scheme to predict the isolated strong precipitation event in east Germany has to be regarded with caution. The Kuo cumulus scheme predicts precipitation almost in the entire model domain.

From Figure 21 the following conclusion can be drawn: Precipitation is enhanced if (i) the Kuo cumulus scheme is used instead of the Grell cumulus scheme and (ii) the horizontal diffusion is computed on sigma levels. The enhancement of precipitation when the Kuo cumulus scheme is used is consistent with the general comment provided by Grell et al. (1995). The enhancement of precipitation when horizontal diffusion is computed on sigma levels can be explained by the fact that in complex terrain sigma diffusion transports moist (and warm) air masses into the upper atmosphere which in turn triggers and/or enhances convection. This finding is consistent with the conclusion drawn by Zängl (2004). To summarize, a precipitation forecast in mountainous terrain using the Kuo cumulus scheme in conjunction with the SD or PD scheme is doubly penalized. The simulation using the Grell cumulus scheme in conjunction with the HD scheme provides the most accurate precipitation forecast, though the precipitation event in eastern Germany is not well captured.

6.4 Statistical intercomparison of simulated and observed data: Part 2

In this section, a statistical intercomparison of simulated and observed data is provided. The forecast model output frequency is 15 minutes. Data collected at all receiver locations indicated in Figure 5 enter the intercomparison. The full time resolution of 2.5 minutes is used. In case the time of the observation does not match the forecast model output time, the nearest available model output time is used. The BmO differences are computed and compared for the model simulations A, B, C and D. Since the same forward model, the same observations, and the same initial state are used in all cases, the revealed BmO differences are concluded to be due to differences in the forecast model.

Figure 22 and 23 show the BmO differences versus the elevation angle for the background A, B, C and D respectively. The red line indicates the RMS, the cyan line indicates the MD and the green line indicates zero MD versus the elevation angle. The significant MD of the BmO differences for the background A and B can be readily related to the erroneous behavior of the Kuo cumulus scheme in conjunction with the SD and PD schemes. Since in the model simulation A and B the IWV content (or equivalently the PW content)

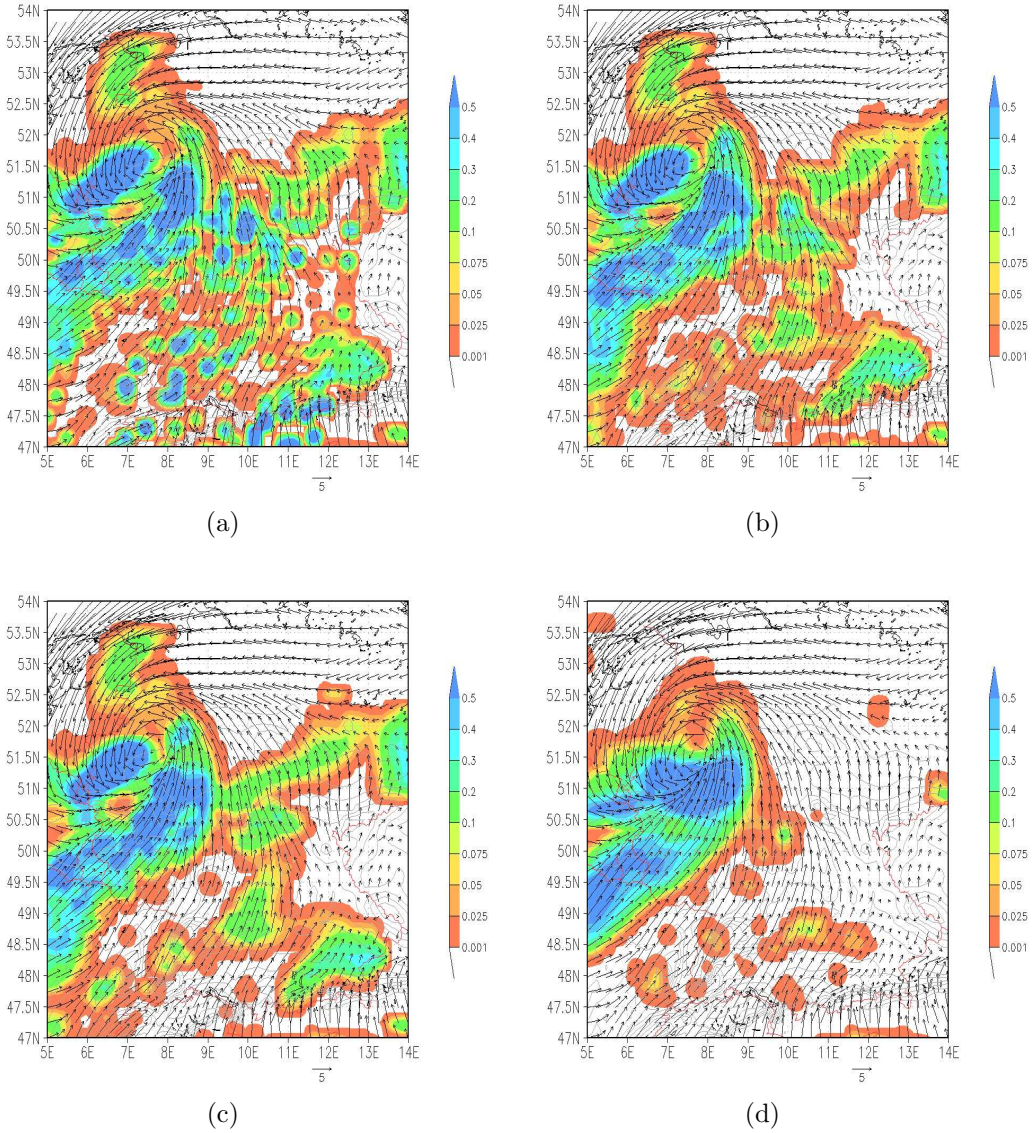


Figure 21: 15 min accumulated precipitation [mm] for different model forecasts valid at 2 UTC on August 22, 2007. The arrows indicate 10 m horizontal wind field. a) Kuo scheme in conjunction with the SD scheme b) Kuo scheme in conjunction with the PD scheme c) Kuo scheme in conjunction with the HD scheme d) Grell scheme in conjunction with the HD scheme.

present in the initial state is immediately converted to precipitation, the simulated slant path delays are underestimated relative to the observed slant path delays. The difference between the magnitude of the MD of the BmO differences for the background C and D can be explained by the tendency of the Kuo cumulus scheme to over predict convective rainfall irrespectively of the used horizontal diffusion scheme.

Figure 13 and 14 reveal that the RMS and the magnitude of the MD become successively smaller over the entire elevation range from background A to D. Table 5 provides the RMS and the MD of the BmO differences for the background A, B, C and D for elevation angles of 7 and 87 degree. The BmO statistics confirm the findings from the previous section and thus demonstrate the ability to use the GPS data for verification purposes.

Statistics	ε [deg]	Model D	Model C	Model B	Model A
RMS [mm]	87	9	11	13	17
RMS [mm]	7	60	68	78	97
MD [mm]	87	-6	-8	-10	-14
MD [mm]	7	-9	-25	-36	-61

Table 5: BmO statistics of the four different model simulations.

It is important to note that the particular day chosen for this experiment is not an exceptional day. The discussed forecast model deficiencies are systematic; that is, the main findings of this single case are valid for an arbitrary chosen day at an arbitrary chosen time. This particular day was chosen in order to highlight the erroneous behavior of forecast model components incorporated in the MM5.

From the sensitivity experiments the following conclusion can be drawn: A spurious tendency present in the model minus observation statistics of the slant path delay data is strongly related to the misplacement and/or the enhancement/suppression of precipitation. Slant path delay observations are potentially useful for verification purposes. The analysis of the model minus observation statistics of the slant path delay data can be used to improve forecast model components.

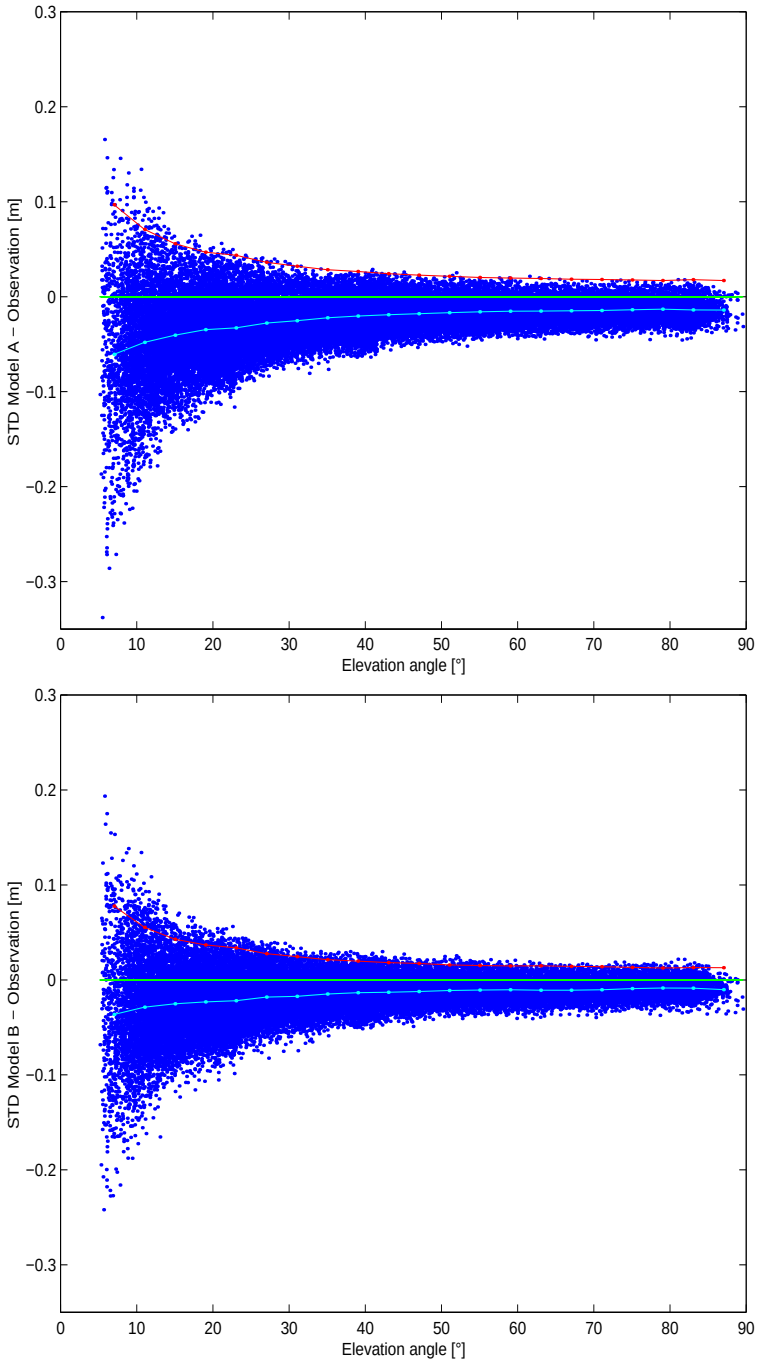


Figure 22: BmO differences (0-6UTC on August 22, 2007) versus the elevation angle for the background A (upper panel) and the background B (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD versus the elevation angle.

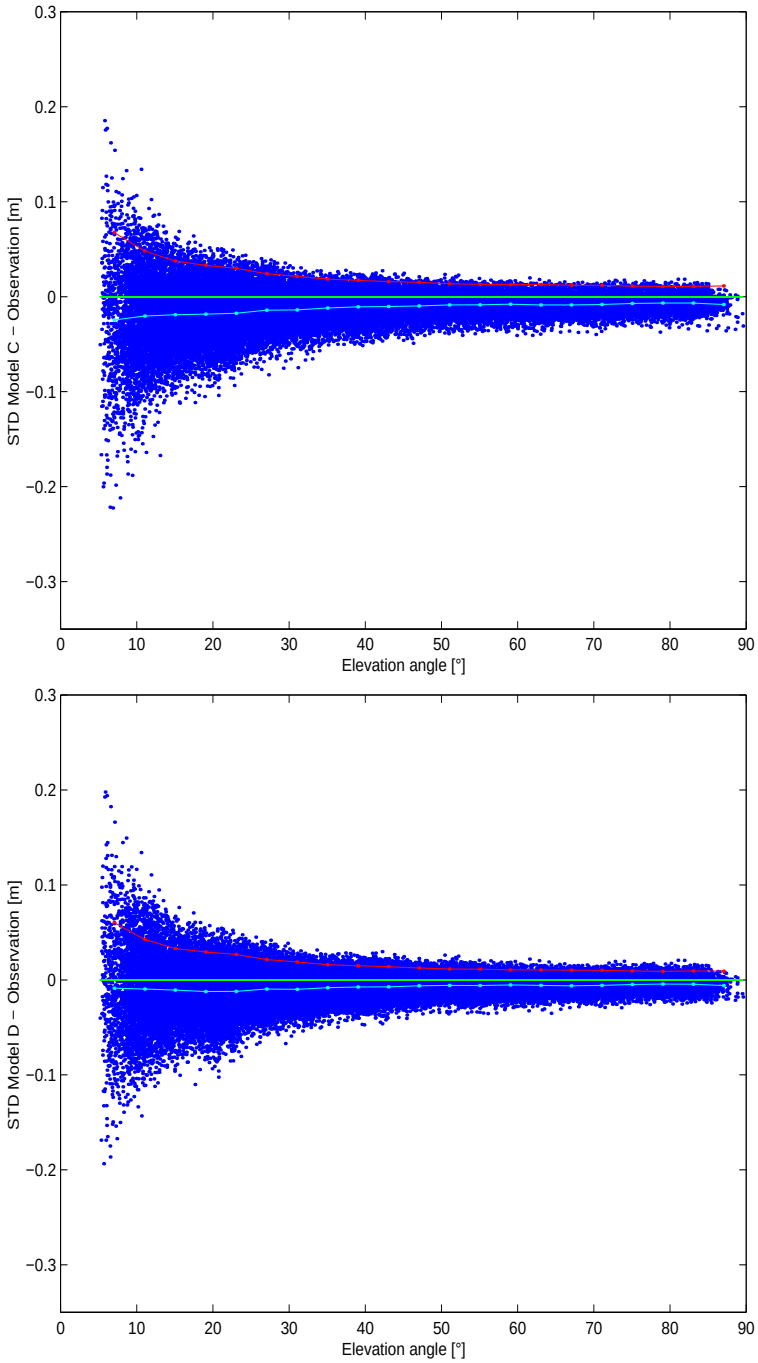


Figure 23: BmO differences (0-6UTC on August 22, 2007) versus the elevation angle for the background C (upper panel) and the background D (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD versus the elevation angle.

7 Tangent linear (TLM) and adjoint (ADJ) code construction

Tangent linear and adjoint models are developed for inverse modelling of physical systems. Among the applications of tangent linear and adjoint models in atmospheric science are model tuning, sensitivity analysis, the determination of singular vectors and variational assimilation. In model tuning data are used to optimize the model equations themselves. Optimization is performed analogously to variational data assimilation, but the control variables are parameters in the underlying equations. In sensitivity analysis, the tangent linear model can be used to analyze the impact of small disturbances. The determination of singular vectors allows to determine which initial perturbations amplify most rapidly. In variational assimilation the control variables are initial conditions. The adjoint model is used to determine the gradient of the cost function. From the diverse applications of tangent linear and adjoint models it follows that data with good spatial coverage, such as data provided from a continuously operating network of GPS receivers, can be useful in many different ways. For example, in the previous section it was demonstrated that simulated slant path delays are very sensitive to particular forecast model components. It is therefore evident that slant path delays can be useful in model tuning.

7.1 Theory

In the following, a brief description of the basic principles of tangent linear and adjoint code construction are given. The focus is on the forward model (a forecast model can be treated in exactly the same manner). For details on tangent linear and adjoint code construction the reader is referred to Giering and Kaminski (1998).

Incorporation of observations in a variational assimilation system requires the development of three operators (i) The forward operator, which maps from model to observation space (ii) the tangent linear operator, which is defined as the gradient of the forward operator with respect to model variables and (iii) the adjoint operator, which is defined as the transpose of the tangent linear operator. In the following, the non-linear operator H represents all required operations to compute the slant path delay S from the gridded temperature, water vapor mixing ratio and pressure fields of the model; that is

$$S = H[\mathbf{x}] \tag{70}$$

where $\mathbf{x}$ denotes the model state. The tangent linear operator $\mathbf{H}$ is defined as

$$\mathbf{H} = \frac{\partial H}{\partial \mathbf{x}} \tag{71}$$

and the adjoint operator $\mathbf{H}^*$ is defined as

$$\mathbf{H}^* = \mathbf{H}^T \tag{72}$$

A close inspection of the forward model H, i.e. the computation of the slant path delay, reveals that the determination of $\mathbf{H}$ and $\mathbf{H}^*$ is by no means trivial. To allow for a strict (exact) and step-by-step derivation, it is convenient to view the computation of the slant path delay as a composition of operators. The computation of the slant path delay consists of two fundamental numerical procedures, the computation of the ray trajectory and the integration along the ray trajectory. Therefore H can be viewed as a composition of two operators H_s and H_r, that is

$$H = H_s\, H_r[\mathbf{x}] \tag{73}$$

Likewise, each operator can be seen as the composition of individual operators acting successively on intermediate results. If m and n denote the numbers of individual operators representing H_s and H_r respectively, the non-linear operator reads as

$$H = H_{sm} ... H_{s1}\, H_{rn} ... \, H_{r1}[\mathbf{x}] \tag{74}$$

The tangent linear operator is obtained by rigorous application of the chain rule of differential calculus

$$\mathbf{H} = \frac{\partial H_{sm}}{\partial H_{sm-1}} \cdot ... \cdot \frac{\partial H_{s1}}{\partial H_{rn}} \cdot \frac{\partial H_{rn}}{\partial H_{rn-1}} \cdot ... \cdot \frac{\partial H_{r1}}{\partial \mathbf{x}} \tag{75}$$

and the adjoint operator is obtained by the transpose of the tangent linear operator

$$\mathbf{H}^* = \left(\frac{\partial H_{r1}}{\partial \mathbf{x}}\right)^T \cdot ... \cdot \left(\frac{\partial H_{rn}}{\partial H_{rn-1}}\right)^T \cdot \left(\frac{\partial H_{s1}}{\partial H_{rn}}\right)^T \cdot ... \cdot \left(\frac{\partial H_{sm}}{\partial H_{sm-1}}\right)^T \tag{76}$$

Notice that in the tangent linear and adjoint model development the above matrices are not explicitly determined. In fact, one is only concerned with obtaining the result of a matrix/vector product $\mathbf{H}\,\delta\mathbf{u}$ or $\mathbf{H}^*\,\delta\mathbf{v}$ where $\delta\mathbf{u}$ and $\delta\mathbf{v}$ denote vectors. Hereinafter the entries of the vectors $\delta\mathbf{u}$ and $\delta\mathbf{v}$ will be referred as the tangent-linear and the adjoint variables respectively. The matrix/vector products are referred as the tangent linear model (TLM) code and the adjoint (ADJ) code respectively. The forward model is referred as non-linear model (NLM) code. In the remainder of this section, δr denotes the tangent linear and $\delta\hat{r}$ denotes the adjoint variable of an arbitrary variable r.

7.2 Practice

The TLM code is developed directly from the NLM code. The ADJ code is developed from the TLM code. As outlined before, the NLM code is an algorithm that can be viewed as a composition of differentiable operators or functions, each representing a statement in the numerical code. The order of evaluation of the individual functions is imposed by the algorithm. Differentiation of the composition can be done by rigorous application of the chain rule of differential calculus on a coding level. The numerical code can be interpreted as a sequence of assignments. An assignment can be considered as an operator acting on the vector of control variables. In general, not all control variables are involved in an assignment. Hence, for the representation of the assignment it is sufficient to use a restricted operator acting only on the subset of involved control variables. The restricted vector of control variables consists of the left-hand side variable and all control variables on the right-hand side. To construct the adjoint statement, the Jacobian of the operator is determined. This is equivalent to constructing the tangent-linear assignment. The partial derivatives are the entries of the Jacobian. The adjoint matrix is the transposed Jacobian. From this matrix the adjoint assignments are formulated. For illustration, consider the following assignment performing one arbitrary step of an numerical algorithm. Suppose d is obtained by applying a nonlinear function on the control variables a, b and c.

$$d \;=\; f(a,b,c)$$

As it stands, this statements can be seen as a function f with input variables and output variables. The TLM code is obtained by differentiation of f

$$\delta d \;=\; f_a \cdot \delta a + f_b \cdot \delta b + f_c \cdot \delta c$$

Here the subscripts denote the partial derivative with respect to a, b and c respectively. To derive the ADJ code, we need to express the TLM code in matrix form. In order to

define this expression, we need to consider all the input and output variables. It is likely that the variables a, b, c and d will be used in other parts of the code. As a consequence, these variables should be considered as input and output. In other words, the complete expression of the TLM code is:

$$\begin{aligned} \delta a &= \delta a \\ \delta b &= \delta b \\ \delta c &= \delta c \\ \delta d &= f_a \cdot \delta a + f_b \cdot \delta b + f_c \cdot \delta c \end{aligned}$$

In matrix notation we have

$$\begin{pmatrix} \delta a \\ \delta b \\ \delta c \\ \delta d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ f_a & f_b & f_c & 0 \end{pmatrix} \cdot \begin{pmatrix} \delta a \\ \delta b \\ \delta c \\ \delta d \end{pmatrix}$$

The ADJ code is obtained after transposition of the matrix

$$\begin{pmatrix} \delta \hat{a} \\ \delta \hat{b} \\ \delta \hat{c} \\ \delta \hat{d} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & f_a \\ 0 & 1 & 0 & f_b \\ 0 & 0 & 1 & f_c \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \delta \hat{a} \\ \delta \hat{b} \\ \delta \hat{c} \\ \delta \hat{d} \end{pmatrix}$$

or simply

$$\begin{aligned} \delta \hat{a} &= \delta \hat{a} + f_a \cdot \delta \hat{d} \\ \delta \hat{b} &= \delta \hat{b} + f_b \cdot \delta \hat{d} \\ \delta \hat{c} &= \delta \hat{c} + f_c \cdot \delta \hat{d} \\ \delta \hat{d} &= 0 \end{aligned}$$

Note that, the statements must be read from top to bottom and right to left. Adjoint statements accumulate information, global adjoint variables need to be reset at the beginning of the adjoint program, local adjoint variables need to be reset at the beginning of the adjoint routines. The trajectory variables that enter in non-linear forward statements, here a,b and c, need to be recomputed and passed to the adjoint statement. In the following two sections TLM and ADJ code construction is illustrated for the forward model and a specific component of the forecast model.

7.3 The forward model

7.3.1 The TLM and ADJ code

The TLM and ADJ code construction is illustrated for the fast computational model for the slant path delay. Hereinafter, D and W denote local variables and V denotes a dummy variable. In the very first step the refractivity is computed at each grid point according to:

$$\begin{aligned} D_{IJK} &= k_1 \cdot P_{IJK}/T_{IJK} \\ W_{IJK} &= k_2 \cdot P_{IJK} \cdot Q_{IJK}/(0.622 + Q_{IJK})/T_{IJK}^2 \\ N_{IJK} &= D_{IJK} + W_{IJK} \end{aligned}$$

for $I = 1, .., M_I$, $J = 1, ..., M_J$ and $K = 1, ..., M_k$. Hereinafter, the index i indicates the location along the geometric path. The first statement in the NLM code computes the refractivity along the geometric path:

$$\begin{aligned} N_1(i) &= N_{I(i)+0J(i)+0A(i)} \cdot (N_{I(i)+0J(i)+0A(i)}/N_{I(i)+0J(i)+0A(i)-1})^{L_1(i)} \\ N_2(i) &= N_{I(i)+0J(i)+1B(i)} \cdot (N_{I(i)+0J(i)+1B(i)}/N_{I(i)+0J(i)+1B(i)-1})^{L_2(i)} \\ N_3(i) &= N_{I(i)+1J(i)+0C(i)} \cdot (N_{I(i)+1J(i)+0C(i)}/N_{I(i)+1J(i)+0C(i)-1})^{L_3(i)} \\ N_4(i) &= N_{I(i)+1J(i)+1D(i)} \cdot (N_{I(i)+1J(i)+1D(i)}/N_{I(i)+1J(i)+1D(i)-1})^{L_4(i)} \end{aligned}$$

for $i = 0, ..., l + 1$ followed by

$$N(i) = \sum_{k=1}^{4} X_k(i) \cdot N_k(i)$$

for $i = 0, ..., l + 1$.

The second statement in the NLM code computes the pressure at the model top point:

$$\begin{aligned} P_1 &= P_{I(l+1)+0J(l+1)+0A(l+1)} \cdot (P_{I(l+1)+0J(l+1)+0A(l+1)}/P_{I(l+1)+0J(l+1)+0A(l+1)-1})^{L_1(l+1)} \\ P_2 &= P_{I(l+1)+0J(l+1)+1B(l+1)} \cdot (P_{I(l+1)+0J(l+1)+1B(l+1)}/P_{I(l+1)+0J(l+1)+1B(l+1)-1})^{L_2(l+1)} \\ P_3 &= P_{I(l+1)+1J(l+1)+0C(l+1)} \cdot (P_{I(l+1)+1J(l+1)+0C(l+1)}/p_{I(l+1)+1J(l+1)+0C(l+1)-1})^{L_3(l+1)} \\ P_4 &= P_{I(l+1)+1J(l+1)+1D(l+1)} \cdot (P_{I(l+1)+1J(l+1)+1D(l+1)}/P_{I(l+1)+1J(l+1)+1D(l+1)-1})^{L_4(l+1)} \end{aligned}$$

followed by

$$P = \sum_{k=1}^{4} X_k(l+1) \cdot P_k$$

The third statement in the NLM code computes the slant path delay:

$$S = 10^{-6} \cdot \sum_{i=0}^{l+1} N(i) \cdot \tilde{w}_i + 10^{-6} \cdot k_1 \cdot R_d \cdot s_0'(\xi_s)/g(\xi_s) \cdot P$$

Due to the simplicity of the NLM code the construction of the TLM and ADJ code is straight forward. The TLM code is omitted for sake of brevity. The ADJ code is provided in detail.

The ADJ code is the transpose statement of the TLM code. This implies that the ADJ code operates in reverse order. The trajectory variables that enter in NLM code need to be recomputed and passed to the ADJ code. The first statement in the ADJ code reads as:

$$\begin{aligned}
\delta\hat{P} &= \delta\tilde{P} + 10^{-6} \cdot k_1 \cdot R_d \cdot s_0'(\xi_s)/g(\xi_s) \cdot \delta\hat{S} \\
\delta\hat{N}(l+1) &= \delta\hat{N}(l+1) + 10^{-6} \cdot \tilde{w}_{l+1} \cdot \delta\hat{S} \\
\delta\hat{N}(l+0) &= \delta\hat{N}(l+0) + 10^{-6} \cdot \tilde{w}_{l+0} \cdot \delta\hat{S} \\
\delta\hat{N}(l-1) &= \delta\hat{N}(l-1) + 10^{-6} \cdot \tilde{w}_{l-1} \cdot \delta\hat{S} \\
&\cdots \\
\delta\hat{N}(1) &= \delta\hat{N}(1) + 10^{-6} \cdot \tilde{w}_1 \cdot \delta\hat{S} \\
\delta\hat{N}(0) &= \delta\hat{N}(0) + 10^{-6} \cdot \tilde{w}_0 \cdot \delta\hat{S} \\
\delta\hat{S} &= 0
\end{aligned}$$

The second statement in the ADJ code reads as:

$$\begin{aligned}
\delta\hat{P}_4 &= \delta\hat{P}_4 + X_4(l+1) \cdot \delta\hat{P} \\
\delta\hat{P}_3 &= \delta\hat{P}_3 + X_3(l+1) \cdot \delta\hat{P} \\
\delta\hat{P}_2 &= \delta\hat{P}_2 + X_2(l+1) \cdot \delta\hat{P} \\
\delta\hat{P}_1 &= \delta\hat{P}_1 + X_1(l+1) \cdot \delta\hat{P} \\
\delta\hat{P} &= 0
\end{aligned}$$

followed by the assignments

$$\begin{aligned}
V_8 &= (L_4(l+1)+1)\cdot P_4/P_{I(l+1)+1J(l+1)+1D(l+1)} \\
V_7 &= (L_3(l+1)+1)\cdot P_3/P_{I(l+1)+1J(l+1)+0C(l+1)} \\
V_6 &= (L_2(l+1)+1)\cdot P_2/P_{I(l+1)+0J(l+1)+1B(l+1)} \\
V_5 &= (L_1(l+1)+1)\cdot P_1/P_{I(l+1)+0J(l+1)+0A(l+1)} \\
V_4 &= L_4(l+1)\cdot P_4/P_{I(l+1)+1J(l+1)+1D(l+1)-1} \\
V_3 &= L_3(l+1)\cdot P_3/P_{I(l+1)+1J(l+1)+0C(l+1)-1} \\
V_2 &= L_2(l+1)\cdot P_2/P_{I(l+1)+0J(l+1)+1B(l+1)-1} \\
V_1 &= L_1(l+1)\cdot P_1/P_{I(l+1)+0J(l+1)+0A(l+1)-1}
\end{aligned}$$

and the update of adjoint variables

$$\begin{aligned}
\delta\hat{P}_{I(l+1)+1J(l+1)+1D(l+1)} &= \delta\hat{P}_{I(l+1)+1J(l+1)+1D(l+1)} + V_8\cdot\delta\hat{P}_4 \\
\delta\hat{P}_{I(l+1)+1J(l+1)+0C(l+1)} &= \delta\hat{P}_{I(l+1)+1J(l+1)+0C(l+1)} + V_7\cdot\delta\hat{P}_3 \\
\delta\hat{P}_{I(l+1)+0J(l+1)+1B(l+1)} &= \delta\hat{P}_{I(l+1)+0J(l+1)+1B(l+1)} + V_6\cdot\delta\hat{P}_2 \\
\delta\hat{P}_{I(l+1)+0J(l+1)+0A(l+1)} &= \delta\hat{P}_{I(l+1)+0J(l+1)+0A(l+1)} + V_5\cdot\delta\hat{P}_1 \\
\delta\hat{P}_{I(l+1)+1J(l+1)+1D(l+1)-1} &= \delta\hat{P}_{I(l+1)+1J(l+1)+1D(l+1)-1} - V_4\cdot\delta\hat{P}_4 \\
\delta\hat{P}_{I(l+1)+1J(l+1)+0C(l+1)-1} &= \delta\hat{P}_{I(l+1)+1J(l+1)+0C(l+1)-1} - V_3\cdot\delta\hat{P}_3 \\
\delta\hat{P}_{I(l+1)+0J(l+1)+1B(l+1)-1} &= \delta\hat{P}_{I(l+1)+0J(l+1)+1B(l+1)-1} - V_2\cdot\delta\hat{P}_2 \\
\delta\hat{P}_{I(l+1)+0J(l+1)+0A(l+1)-1} &= \delta\hat{P}_{I(l+1)+0J(l+1)+0A(l+1)-1} - V_1\cdot\delta\hat{P}_1 \\
\delta\hat{P}_4 &= 0 \\
\delta\hat{P}_3 &= 0 \\
\delta\hat{P}_2 &= 0 \\
\delta\hat{P}_1 &= 0
\end{aligned}$$

The third statement in the ADJ code reads as:

$$\begin{aligned}
\delta\hat{N}_4(i) &= \delta\hat{N}_4(i) + X_4(i)\cdot\delta\hat{N}(i) \\
\delta\hat{N}_3(i) &= \delta\hat{N}_3(i) + X_3(i)\cdot\delta\hat{N}(i) \\
\delta\hat{N}_2(i) &= \delta\hat{N}_2(i) + X_2(i)\cdot\delta\hat{N}(i) \\
\delta\hat{N}_1(i) &= \delta\hat{N}_1(i) + X_1(i)\cdot\delta\hat{N}(i) \\
\delta\hat{N}(i) &= 0
\end{aligned}$$

for $i = l+1, ..., 0$, followed by the assignments

$$\begin{aligned}
V_8 &= (L_4(i)+1) \cdot N_4(i)/N_{I(i)+1J(i)+1D(i)} \\
V_7 &= (L_3(i)+1) \cdot N_3(i)/N_{I(i)+1J(i)+0C(i)} \\
V_6 &= (L_2(i)+1) \cdot N_2(i)/N_{I(i)+0J(i)+1B(i)} \\
V_5 &= (L_1(i)+1) \cdot N_1(i)/N_{I(i)+0J(i)+0A(i)} \\
V_4 &= L_4(i) \cdot N_4(i)/N_{I(i)+1J(i)+1D(i)-1} \\
V_3 &= L_3(i) \cdot N_3(i)/N_{I(i)+1J(i)+0C(i)-1} \\
V_2 &= L_2(i) \cdot N_2(i)/N_{I(i)+0J(i)+1B(i)-1} \\
V_1 &= L_1(i) \cdot N_1(i)/N_{I(i)+0J(i)+0A(i)-1}
\end{aligned}$$

and the update of adjoint variables

$$\begin{aligned}
\delta\hat{N}_{I(i)+1J(i)+1D(i)} &= \delta\hat{N}_{I(i)+1J(i)+1D(i)} + V_8 \cdot \delta\hat{N}_4(i) \\
\delta\hat{N}_{I(i)+1J(i)+0C(i)} &= \delta\hat{N}_{I(i)+1J(i)+0C(i)} + V_7 \cdot \delta\hat{N}_3(i) \\
\delta\hat{N}_{I(i)+0J(i)+1B(i)} &= \delta\hat{N}_{I(i)+0J(i)+1B(i)} + V_6 \cdot \delta\hat{N}_2(i) \\
\delta\hat{N}_{I(i)+0J(i)+0A(i)} &= \delta\hat{N}_{I(i)+0J(i)+0A(i)} + V_5 \cdot \delta\hat{N}_1(i) \\
\delta\hat{N}_{I(i)+1J(i)+1D(i)-1} &= \delta\hat{N}_{I(i)+1J(i)+1D(i)-1} - V_4 \cdot \delta\hat{N}_4(i) \\
\delta\hat{N}_{I(i)+1J(i)+0C(i)-1} &= \delta\hat{N}_{I(i)+1J(i)+0C(i)-1} - V_3 \cdot \delta\hat{N}_3(i) \\
\delta\hat{N}_{I(i)+0J(i)+1B(i)-1} &= \delta\hat{N}_{I(i)+0J(i)+1B(i)-1} - V_2 \cdot \delta\hat{N}_2(i) \\
\delta\hat{N}_{I(i)+0J(i)+0A(i)-1} &= \delta\hat{N}_{I(i)+0J(i)+0A(i)-1} - V_1 \cdot \delta\hat{N}_1(i) \\
\delta\hat{N}_4(i) &= 0 \\
\delta\hat{N}_3(i) &= 0 \\
\delta\hat{N}_2(i) &= 0 \\
\delta\hat{N}_1(i) &= 0
\end{aligned}$$

for $i = l+1, ..., 0$.

The very last step of the ADJ code reads as:

$$\begin{aligned}
\delta\hat{D}_{IJK} &= \delta\hat{D}_{IJK} + \delta\hat{N}_{IJK} \\
\delta\hat{W}_{IJK} &= \delta\hat{W}_{IJK} + \delta\hat{N}_{IJK} \\
\delta\hat{N}_{IJK} &= 0
\end{aligned}$$

for $I = M_I, .., 1$, $J = M_J, ..., 1$ and $K = M_K, ..., 1$, followed by

$$\begin{aligned}
\delta\hat{P}_{IJK} &= \delta\hat{P}_{IJK} + W_{IJK}/P_{IJK} \cdot \delta\hat{W}_{IJK} \\
\delta\hat{Q}_{IJK} &= \delta\hat{Q}_{IJK} + W_{IJK}/Q_{IJK} \cdot \delta\hat{W}_{IJK} \\
\delta\hat{T}_{IJK} &= \delta\hat{T}_{IJK} - W_{IJK}/T_{IJK} \cdot 2 \cdot \delta\hat{W}_{IJK} \\
\delta\hat{Q}_{IJK} &= \delta\hat{Q}_{IJK} - W_{IJK}/(0.622 + Q_{IJK}) \cdot \delta\hat{W}_{IJK} \\
\delta\hat{P}_{IJK} &= \delta\hat{P}_{IJK} + D_{IJK}/P_{IJK} \cdot \delta\hat{D}_{IJK} \\
\delta\hat{T}_{IJK} &= \delta\hat{T}_{IJK} - D_{IJK}/T_{IJK} \cdot \delta\hat{D}_{IJK} \\
\delta\hat{W}_{IJK} &= 0 \\
\delta\hat{D}_{IJK} &= 0
\end{aligned}$$

for $I = M_I, .., 1$, $J = M_J, ..., 1$ and $K = M_K, ..., 1$. To understand the effectiveness of the ADJ code in a variational assimilation framework, let us introduce the cost-function C, depending on the model state $\mathbf{x}$, measuring the difference between a single slant path delay observation y and the model prediction $H[\mathbf{x}]$ in a least square sense

$$C[\mathbf{x}] = \frac{1}{2}\,(y - H[\mathbf{x}])^T \frac{1}{\sigma_r^2}(y - H[\mathbf{x}]) \tag{77}$$

where σ_r denotes the observation error. The gradient of the cost function C with respect to the $\mathbf{x}$ reads as

$$\nabla C[\mathbf{x}] = -\mathbf{H}^T(\mathbf{x})\,\frac{(y - H[\mathbf{x}])}{\sigma_r^2} \tag{78}$$

If one performs the assignment

$$\delta\hat{S} = -\frac{(y - H[\mathbf{x}])}{\sigma_r^2} \tag{79}$$

at the beginning of the ADJ code, it is not difficult to see that the gridded adjoint variables $\delta\hat{Q}_{IJK}, \delta\hat{T}_{IJK}$ and $\delta\hat{P}_{IJK}$ for $I = 1, ..., M_I$, $J = 1, ..., M_J$ and $K = 1, ..., M_K$ at the end of the ADJ code store the gradient of the cost function C with respect to the gridded control variables Q_{IJK}, T_{IJK} and P_{IJK} for $I = 1, ..., M_I$, $J = 1, ..., M_J$ and $K = 1, ..., M_K$. The computed gradient is exact. It is interesting to observe how information is accumulated in the ADJ code: a single scalar, i.e. the normalized departure, is input, and three gridded (though sparse) arrays are output.

7.3.2 Impact of a single slant path delay in variational assimilation

Since the slant path delay is an integrated measurement, it is interesting to estimate the impact in variational assimilation. Linearizing the cost function J in the vicinity of the background state $\mathbf{x}_b$ leads to

$$\begin{aligned} J[\mathbf{x}] &= \frac{1}{2}\left(\mathbf{x}-\mathbf{x}_b\right)^T \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}_b\right) + \\ &+ \frac{1}{2}\left(y - H[\mathbf{x}_b] - \mathbf{H}[\mathbf{x}_b]\left(\mathbf{x}-\mathbf{x}_b\right)\right)^T \mathbf{R}^{-1}\left(y - H[\mathbf{x}_b] - \mathbf{H}[\mathbf{x}_b]\left(\mathbf{x}-\mathbf{x}_b\right)\right) \end{aligned} \quad (80)$$

The gradient of the cost function J with respect to the model state $\mathbf{x}$ reads as

$$\nabla J[\mathbf{x}] = \mathbf{B}^{-1}\left(\mathbf{x}-\mathbf{x}_b\right) - \mathbf{H}[\mathbf{x}_b]^T \mathbf{R}^{-1}\left(y - H[\mathbf{x}_b] - \mathbf{H}[\mathbf{x}_b]\left(\mathbf{x}-\mathbf{x}_b\right)\right)$$

The gradient of the cost function vanishes at the analysis $\mathbf{x}_a$, i.e. $\nabla J[\mathbf{x}_a] = 0$, and thus the analysis increment reads as

$$\mathbf{x}_a - \mathbf{x}_b = \left(\mathbf{B}^{-1} + \mathbf{H}[\mathbf{x}_b]^T \mathbf{R}^{-1} \mathbf{H}[\mathbf{x}_b]\right)^{-1} \mathbf{H}[\mathbf{x}_b]^T \mathbf{R}^{-1}\left(y - H[\mathbf{x}_b]\right)$$

A variant of the Sherman-Morrison-Woodbary formula states that (Bouttier and Courtier, 1999)

$$\left(\mathbf{B}^{-1} + \mathbf{H}[\mathbf{x}_b]^T \mathbf{R}^{-1} \mathbf{H}[\mathbf{x}_b]\right)^{-1} \mathbf{H}[\mathbf{x}_b]^T \mathbf{R}^{-1} = \mathbf{B}\mathbf{H}[\mathbf{x}_b]^T \left(\mathbf{H}[\mathbf{x}_b]\mathbf{B}\mathbf{H}[\mathbf{x}_b]^T + \mathbf{R}\right)^{-1}$$

Hence, the analysis increment reads as

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{B}\mathbf{H}[\mathbf{x}_b]^T \left(\mathbf{H}[\mathbf{x}_b]\mathbf{B}\mathbf{H}[\mathbf{x}_b]^T + \mathbf{R}\right)^{-1}\left(y - H[\mathbf{x}_b]\right)$$

For a single slant path delay the observation error covariance matrix $\mathbf{R}$ and its model equivalent $\mathbf{H}[\mathbf{x}_b]\mathbf{B}\mathbf{H}[\mathbf{x}_b]^T$ are scalar quantities. Introducing $\sigma^2 = \mathbf{R} + \mathbf{H}[\mathbf{x}_b]\mathbf{B}\mathbf{H}[\mathbf{x}_b]^T$ allows one to write the analysis increment as

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{B}\mathbf{H}[\mathbf{x}_b]^T \frac{(y - H[\mathbf{x}_b])}{\sigma^2} \tag{81}$$

This expression reveals that the background error covariance matrix $\mathbf{B}$ determines to a large extent which variable at which location is affected by the observation minus model departure. In addition, the correlations in $\mathbf{B}$ perform the spatial information spreading and smoothing.

Noteably, the algebra above holds in the case the observation time does not match the initial time. The difference to the previous example is that the mapping from model to observation space includes the forecast model that allows a comparison between the model state and the observation at the appropriate time. However, the underlying approximation, i.e. the linearization of the cost function, must be viewed with caution (the forward model and in particular the successive application of the forecast model on the initial state are highly non-linear). For a single slant path delay observation at the time step i the analysis increment at initial time can be written by rigorous application of the chain rule of differential calculus as

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{B}\mathbf{M}_1[x_b]^T ... \mathbf{M}_i[x_{bi-1}]^T \mathbf{H}[\mathbf{x}_{bi}]^T \frac{(y_i - H\, M_i ... M_1[\mathbf{x}_b])}{\sigma_i^2} \tag{82}$$

where $\sigma_i^2 = \mathbf{R}_i + \mathbf{H}[\mathbf{x}_{bi}]\mathbf{M}_i[\mathbf{x}_{bi-1}]...\mathbf{M}_1[\mathbf{x}_b]\mathbf{B}\mathbf{M}_1[x_b]^T ... \mathbf{M}_i[x_{bi-1}]^T\mathbf{H}[\mathbf{x}_{bi}]^T$. This expression reveals that the background error covariance matrix $\mathbf{B}$ is evolved in time by the adjoint of the tangent linear operator of the forecast model $\mathbf{M}^T$.

7.4 The modified horizontal diffusion scheme in the forecast model

7.4.1 The TLM and ADJ code

In this section, the focus is on the discretization scheme Ω_f of the water vapor mixing ratio at a grid point with grid point coordinates I, J and K. The alternative discretization schemes are treated in a similar manner. The control variables are Q and K_h. The horizontal diffusion coefficient K_h appears to be a control variable since it is a function of the local deformation of the wind field (Grell et al., 1995). The first statement in the TLM code reads as:

$$
\begin{aligned}
\delta Q_A &= (L_A+1)\cdot Q_A/Q_{I+2JA}\cdot\delta Q_{I+2JA} - L_A\cdot Q_A/Q_{I+2JA-1}\cdot\delta Q_{I+2JA-1}\\
\delta Q_B &= (L_B+1)\cdot Q_B/Q_{I+1JB}\cdot\delta Q_{I+1JB} - L_B\cdot Q_B/Q_{I+1JB-1}\cdot\delta Q_{I+1JB-1}\\
\delta Q_C &= (L_C+1)\cdot Q_C/Q_{I-1JC}\cdot\delta Q_{I-1JC} - L_C\cdot Q_C/Q_{I-1JC-1}\cdot\delta Q_{I-1JC-1}\\
\delta Q_D &= (L_D+1)\cdot Q_D/Q_{I-2JD}\cdot\delta Q_{I-2JD} - L_D\cdot Q_D/Q_{I-2JD-1}\cdot\delta Q_{I-2JD-1}
\end{aligned}
$$

The second statement in the TLM code reads as:

$$
\delta\Omega_f = \left(\delta Q_A - 4\cdot\delta Q_B + 6\cdot\delta Q_{IJK} - 4\cdot\delta Q_C + \delta Q_D\right)/\Delta y^4
$$

and the TLM code of the third statement reads as:

$$
\delta F = -K_h\cdot\delta\Omega_f - \Omega_f\cdot\delta K_h
$$

The first statement of the ADJ code reads as:

$$
\begin{aligned}
\delta\hat{\Omega}_f &= \delta\hat{\Omega}_f - K_h\cdot\delta\hat{F}\\
\delta\hat{K}_h &= \delta\hat{K}_h - \Omega_f\cdot\delta\hat{F}\\
\delta\hat{F} &= 0
\end{aligned}
$$

The second statement of the ADJ code reads as:

$$
\begin{aligned}
\delta\hat{Q}_A &= \delta\hat{Q}_A + 1\cdot\delta\hat{\Omega}_f/\Delta y^4\\
\delta\hat{Q}_B &= \delta\hat{Q}_B - 4\cdot\delta\hat{\Omega}_f/\Delta y^4\\
\delta\hat{Q}_C &= \delta\hat{Q}_C - 4\cdot\delta\hat{\Omega}_f/\Delta y^4\\
\delta\hat{Q}_D &= \delta\hat{Q}_D + 1\cdot\delta\hat{\Omega}_f/\Delta y^4\\
\delta\hat{Q}_{IJK} &= \delta\hat{Q}_{IJK} + 6\cdot\delta\hat{\Omega}_f\Delta y^4\\
\delta\hat{\Omega}_f &= 0
\end{aligned}
$$

and the third statement of the ADJ code reads as:

$$
\begin{aligned}
\delta\hat{Q}_{I+2JA} &= \delta\hat{Q}_{I+2JA} + (L_A + 1) \cdot Q_A/Q_{I+2JA} \cdot \delta\hat{Q}_A \\
\delta\hat{Q}_{I+1JB} &= \delta\hat{Q}_{I+1JB} + (L_B + 1) \cdot Q_B/Q_{I+1JB} \cdot \delta\hat{Q}_B \\
\delta\hat{Q}_{I-1JC} &= \delta\hat{Q}_{I-1JC} + (L_C + 1) \cdot Q_C/Q_{I-1JC} \cdot \delta\hat{Q}_C \\
\delta\hat{Q}_{I-2JD} &= \delta\hat{Q}_{I-2JD} + (L_D + 1) \cdot Q_D/Q_{I-2JD} \cdot \delta\hat{Q}_D \\
\delta\hat{Q}_{I+2JA-1} &= \delta\hat{Q}_{I+2JA-1} - L_A \cdot Q_A/Q_{I+2JA-1} \cdot \delta\hat{Q}_A \\
\delta\hat{Q}_{I+1JB-1} &= \delta\hat{Q}_{I+1JB-1} - L_B \cdot Q_B/Q_{I+1JB-1} \cdot \delta\hat{Q}_B \\
\delta\hat{Q}_{I-1JC-1} &= \delta\hat{Q}_{I-1JC-1} - L_C \cdot Q_C/Q_{I-1JC-1} \cdot \delta\hat{Q}_C \\
\delta\hat{Q}_{I-2JD-1} &= \delta\hat{Q}_{I-2JD-1} - L_D \cdot Q_D/Q_{I-2JD-1} \cdot \delta\hat{Q}_D \\
\delta\hat{Q}_A &= 0 \\
\delta\hat{Q}_B &= 0 \\
\delta\hat{Q}_C &= 0 \\
\delta\hat{Q}_D &= 0
\end{aligned}
$$

7.4.2 Impact of a forecast model component in variational assimilation

Similar to the forward model, the forcast model can be viewed as a composition of forecast model components, e.g. horizontal diffusion, vertical diffusion, advection etc. If m denotes the numbers of individual forecast model components, a single forward model step from time step $i-1$ to i can be written as

$$\mathbf{x}_i = M_{im}...M_{i1}[\mathbf{x}_{i-1}] \tag{83}$$

According to equation (82) the analysis is affected by a forecast model component twice: (i) through the observation minus model departure and (ii) through the background error covariance which is evolved in time. It is obvious that in case of a large error in a forecast model component, variational assimilation leads to an erroneous analysis. The error in a forcast model component is accumulated into the analysis. Minimizing the discrepancy between the observation and the model prediction through the 4DVAR does not imply that the analysis is optimal. On the contrary, if a model component incorporated in the 4DVAR is far from being perfect, minimizing the discrepancy between the observation and the model prediction leads to an erroneous analysis. In other words, variational assimilation works if and only if the forecast model is sufficiently accurate (Bouttier and Courtier, 1999).

7.5 Check for correctness

The TLM and ADJ code construction of the forward model, including the simulation of the ray trajectory, is relatively laborious but there is no intrinsic difficulty. Some caution is needed in the LU-decomposition, since it consists of a forward and backward substitution.

TLM and ADJ routines are tested for correctness using the standard comparison of the TLM and finite difference-derived NLM gradients to check correctness of the TLM

$$\lim_{\delta u \to 0} H(\mathbf{u} + \delta\mathbf{u}) - H(\mathbf{u}) \quad = \quad \mathbf{H}\,\delta\mathbf{u} \tag{84}$$

and the definition of the adjoint

$$(\mathbf{H}\,\delta\mathbf{u}, \delta\mathbf{v}) \quad = \quad (\delta\mathbf{u}, \mathbf{H}^*\,\delta\mathbf{v}) \tag{85}$$

to check for consistency of the TLM and the ADJ (Zou et al., 1997). The check for correctness is applied to individual forward model components, i.e. the interpolation routines, the computation of the ray trajectory, the numerical integration along the ray trajectory, as well as to the entire forward model.

The check for consistency of the TLM and the ADJ using FORTRAN double precision for a slant path delay with an elevation angle of 5° and an azimuth angle of 0° (computed at a receiver located at the center point of the MM5 domain) results in

$$(\mathbf{H}_r\,\delta\mathbf{u}, \delta\mathbf{v})/(\delta\mathbf{u}, \mathbf{H}_r^*\,\delta\mathbf{v}) = 1.00000000000000$$

for the ray-tracing and

$$(\mathbf{H}\,\delta\mathbf{u}, \delta\mathbf{v})/(\delta\mathbf{u}, \mathbf{H}^*\,\delta\mathbf{v}) = 1.00000000000000$$

for the entire forward model.

8 Four-dimensional variational assimilation

8.1 The MM5 4DVAR system

A description of the MM5 4DVAR system used in this thesis can be found in Ruggiero et al. (2002). Though the 4DVAR system is based on MM5v3, the tangent linear and adjoint model only contains the same set of physics that are in the serial 4DVAR based on MM5v1. This includes the Kuo cumulus scheme, a moist stable precipitation scheme, a simple radiative cooling scheme and a bulk PBL parameterization. To allow the use of a greater number of vertical levels than is practical with the single-level bulk PBL scheme, the user is given the option to run the MRF (Medium Range Forcast) PBL scheme of Hong and Pan (1996).

The analysis vector of the MM5 4DVAR system consists of the following control variables: the horizontal wind components, the temperature, the water-vapor mixing ratio, the pressure perturbation and the vertical velocity. Cloud water, cloud ice, surface variables, and the variables at lateral boundaries are not included. The MM5 4DVAR only constructs diagonal background error covariance matrices. This approximation has proven to work well for assimilation studies conducted with the system (Wulfmeyer et al., 2006; Grzeschik et al., 2008). This can be explained by the ability of 4DVAR to self-generate physically consistent structure functions during model integration. For each control variable, the diagonal elements of the background error covariance matrix are specified by constructing the differences between a short-range model forecast of 15 minutes and the initial values at each grid point. At each vertical level, the maximum value of the difference is found and assigned to all grid points on that level. This creates a vertical profile of background errors valid at all geographical locations of the model. The forecast errors are then squared and assigned to the diagonal elements of the background error covariance matrix $\mathbf{B}$.

8.2 The implementation in the MM5 4DVAR

The forward operator and the adjoint operator implemented in the MM5 4DVAR MPI environment are based on the fast computational model for the slant path delay. It can be used as long as the elevation angle is not below 30°. Therefore, slant path delay observations with elevation angles below 30° are not assimilated. As a consequence, the potential asymmetric information content present in the GPS data can not be completely utilized. The implementation of the fast computational model for the slant path delay was motivated by the requirements to run the MM5 4DVAR in an operational mode for a period of six months during COPS (Wulfmeyer et al., 2008) and the demonstration project Probabilistic Hydrological and Atmospheric Simulation of flood Events in the alpine region (D-PHASE) (Rotach et al., 2009). The results of this assimilation experiment are not

discussed in this thesis. The implementation of the rigorous computational model for the slant path delay in the MM5 4DVAR MPI environment and simulations over a period of at least one month together with a statistical evaluation of the results goes far beyond the scope of this thesis. Instead, the author attaches great importance to the model physics/numerics incorporated in the MM5 4DVAR. To date, two serious deficiencies in the MM5 4DVAR were identified, which became obvious during COPS and D-PHASE, namely the deep convection scheme and the implementation of the horizontal diffusion. Noteably, these deficiencies could be identified by screening the model minus observation departure of the GPS data.

To bypass the tendency introduced by the Kuo cumulus scheme, the Grell cumulus scheme and its adjoint were implemented into the the MM5 4DVAR. The forward and adjoint model of the Grell cumulus scheme already existed in an advanced version of the serial 4DVAR based on MM5v2 (Zou et al., 1997), so that only minor changes were necessary to include it in the parallel 4DVAR based on MM5v3. In addition, the modified horizontal diffusion scheme and its adjoint were implemented.

The MM5 grid configuration and the basic state variables in the following assimilation experiments remain unchanged, see Table 1. Table 6 summarizes the MM5 forecast configuration and the configuration in the MM5 4DVAR.

MM5 physics	Forecast MM5v3.7	Assimilaton MM5v3.4
Coriolis force	Full 3D Coriolis force	Full 3D Coriolis force
Upper boundary condition	Radiative	Rigid top
Lateral boundary condition	Relaxation	Relaxation
Cloud microphysics	Reisner 2	not simulated
Planet boundary layer	MRF	MRF
Ground temperature	Simulated	fixed
Heat/Moist fluxes	Simulated	not simulated
Soil model	5-layer soil model	not simulated
Radiation	RRTM LW/Dudhia SW	simple radiative cooling
Convection	Kain Fritsch	Grell*
Horizontal diffusion	Zängl	Modified*

Table 6: Table summarizing the parameterization schemes used in the forecast and the assimilation. The Grell cumulus scheme and the modified horizontal diffusion scheme were implemented by the author as an alternative to the Kuo cumulus scheme and the sigma diffusion scheme.

To speed up computation, mesoscale models are implemented for parallel computers such that each processor handles only a geographical sub-area. The parallelization was implemented such that the same source architecture for the serial and parallel versions of the code was maintained (Ruggiero et al., 2006). In principle, the parallelization involves

spatial domain decomposition where the horizontal analysis domain is partitioned into smaller subdomains or patches. A separate processor is dedicated to the computation for a specific subdomain. During the execution of the 4DVAR code, each patch requires only the boundary information from its neighboring patches. Communication between various patches is accomplished with the use of the standard MPI environment. The signal from the receiver to the satellite can intersect two or more subdomains. Some additional logic is required in the routines when compared to the single processor version: In the forward integration, the contributions to the slant path delay determined by the processors in question are gathered to compute the total slant path delay and the corresponding normalized departure. In the backward integration, the update of the adjoint variables is restricted to the processors in question. The forward model and the adjoint model were implemented into the MPI environment of the 4DVAR system, such that speedy assimilation of GPS data became routinely possible. With a horizontal domain size of 64 times 70 grid points and 30 processors the speed up factor is about 10 compared to a single processor configuration.

8.3 The observation screening in the 4DVAR

The BmO differences used for validation purpose include all GPS data. In the 4DVAR some medium filtering of the GPS data based on an estimate of the relative departure of the slant wet delay is applied in order to reject questionable data and to avoid problems in the minimization due to large differences between the model prediction and the observations. Assuming that the slant dry delay is well represented in both, the model prediction and the observation, allows to estimate the relative departure of the slant wet delay R through

$$R = \frac{S - S_o}{S_w} \tag{86}$$

where S and S_o denote the simulated and the observed slant path delay respectively, and S_w denotes the simulated slant wet delay. In the first forward integration of the 4DVAR an observation is rejected if $|R| > 0.2$. As can be seen in the following simulations, this threshold is located at the tail of the Gaussian distribution of the relative departure of the slant wet delay and thus removes potential outliers in the assimilation. These data is also not considered in the following forward and backward integrations and is therefore permanently black listed during minimization.

8.4 The forward model error and the observation error

The forward model error is estimated as follows: The number of Legendre-Gauss-Lobatto nodes in the numerical integration of the contribution to the slant path delay below the

model top is fixed and chosen to be 42, which leads to sub-millimeter accuracy over the elevation range from 30 to 90 degree. Note, that the number of nodes is considerably smaller than the number of nodes used in the rigorous computational model for the slant path delay. This is due to the fact that the integration interval is much smaller. In the zenith case the error associated to the mean gravity computed above the model top can be readily neglected, see e.g. Elgered et al. (1991). Since the numerical integration below the model top has sub-millimeter accuracy, the overall error of the forward model in the zenith case is estimated to be 1 mm. In the slant case, the numerical integration below the model top and the assumptions made above the model top (spherical symmetry of the atmosphere) suggest that the error of the forward model is roughly proportional to the cosecant of the elevation angle at the receiver.

Several studies were undertaken to determine the error of GPS data by means of intercomparison with independent observations like water vapor radiometers (Bender et al., 2008) However, it remains difficult to estimate the error of the GPS data. An initial comparison of GPS stations in Germany with the local 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. Noteably, the determination of both the IWV and the ZTD data from the raw measurement contains numerous assumptions. Further studies using different methods are required. In the absence of more accurate estimates and on the basis of the error of ZTD data provided by the GFZ Potsdam, the overall error of a slant path delay is estimated to be of the order of a few millimeters (about 1 to 6 mm) times the cosecant of the elevation angle at the receiver. An elevation dependent observation error σ_r according to

$$\sigma_r(\varepsilon) = \frac{c}{sin(\varepsilon)} \tag{87}$$

is used where $c = 3$ mm. Due to insufficient knowledge of error correlation in the PPP algorithm, the correlation of the GPS data is disregarded. Therefore, the observation error covariance matrix $\mathbf{R}$ is strictly diagonal.

8.5 Assimilation experiment 1

The chosen day for this assimilation experiment was 14 of August, 2007. The network red provides GPS data for the assimilation. On the contrary, the network green is the observing network, meaning that GPS data provided by this network was not assimilated. The assimilation window was 3 hours, ranging from 0-3UTC. The assimilation frequency was 30 minutes. For validation purposes, GPS data were evaluated every 15 minutes. The forecast range is 0-6UTC. Two simulations are performed, one with assimilation of the GPS data, the other one without. Since it is of interest whether independent upper

air measurements support the results gained by the GPS data assimilation, radiosonde measurements are used for intercomparison. The model equivalents of the radiosonde observations at a specific location are computed by the same interpolation routine that is also used in the forward model of the slant path delay.

To emphasize the erroneous behavior of the former physics incorporated in the MM5 4DVAR, two model simulations starting from the same initial state valid at 0UTC were performed. The first simulation uses the Kuo cumulus scheme in conjunction with the SD scheme and the second simulation uses the Grell cumulus scheme in conjunction with the HD scheme. The frequency histogram of the relative departure of the slant wet delay is shown in Figure 24 and confirms the finding from the previous section. Again, the systematic model error introduced by the Kuo cumulus scheme in conjunction with the SD scheme is obvious: The distribution of the relative departure of the slant wet delay is far from being Gaussian. The IWV content present in the initial state is immediately converted to precipitation such that simulated slant path delays are underestimated relative to observations. On contrary, if the Grell cumulus scheme is used in conjunction with the HD scheme the distribution of the relative departure of the slant wet delay is almost Gaussian. This simple example already indicates that the physics incorporated in 4DVAR plays a fundamental role. The Grell cumulus scheme in conjunction with the HD scheme is preferable for the use in 4DVAR (there is no indication that this is not the case), however, it is not clear whether this configuration is accurate enough for meaningful assimilation of observations linked to water vapor.

The following questions regarding 4DVAR are addressed:

1. Does the 4DVAR algorithm work properly?
2. Is there a detectable and significant impact of the GPS data in the assimilation window as well as after the assimilation window?
3. Is there any indication that the humidity content is better represented with assimilation than without assimilation?
4. Is it possible to easily validate the results with independent upper air measurements?
5. Is there an impact in precipitation?

In connection question 1, it is noted that the 4DVAR works properly from a technical point of view. The incorporated TLM and ADJ routines were tested for correctness using the standart comparison of the TLM and finite difference-derived NLM gradients to check correctness of the TLM and the definition of the adjoint to check for consistency of the TLM and the ADJ. For details on the nature of those tests the reader is referred to

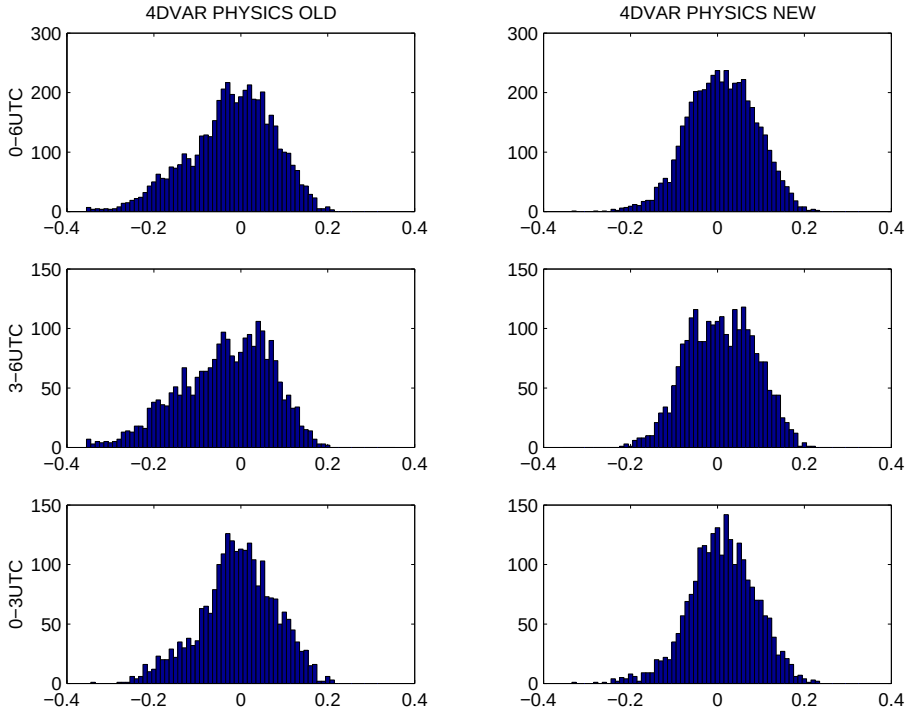


Figure 24: Network red (0-6UTC on August 14, 2007): Frequency histogram of the relative departure of the slant wet delay. The different panels show the different time intervals of the model integration. The overall statistics is shown in the upper most panel. The left panel shows the performance of the old physics incorporated in the MM5 4DVAR, the right panel shows the performance of the new physics incorporated in the MM5 4DVAR.

Ruggiero et al. (2002). The cost function and the norm of the gradient of the cost function versus the iteration number are shown in Figure 25. The decrease in both quantities confirm that the minimization in the 4DVAR works as intended. A sharp decrease of the cost function and the gradient is visible in the first few iterations, which in fact is typical for the quasi-Newton method.

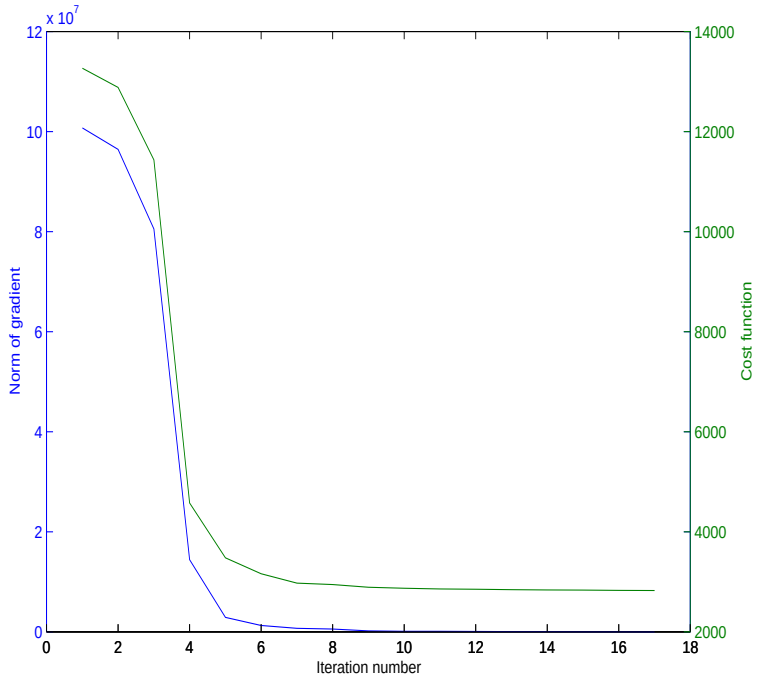


Figure 25: The cost function and the norm of the gradient versus the iteration number.

In order to address question 2, the BmO differences for the red network are computed for the control experiment and the assimilation experiment. Figure 26 shows the BmO differences versus the time for the red network. The small bias showing up in the control experiment is reduced to zero in the assimilation experiment. The RMS is significantly reduced in the assimilation window as well as after the assimilation window. However, the impact seems to fade away quickly. This is due to the fact that air masses are advected above receivers with the ambient wind direction from the west. Figure 27 shows the BmO differences versus the elevation angle for the red network. The RMS is reduced over the entire elevation range in the assimilation experiment. For elevation angles from 90 to 25 degree the small bias showing up in the control experiment is reduced to zero in the assimilation experiment. At lower elevation angles the reduction of the bias is smaller. This can be explained by the fact that the forward model used in this experiment is strictly not applicable at elevation angles as low as 20°.

The frequency histogram of the relative departure of the slant wet delay for the red

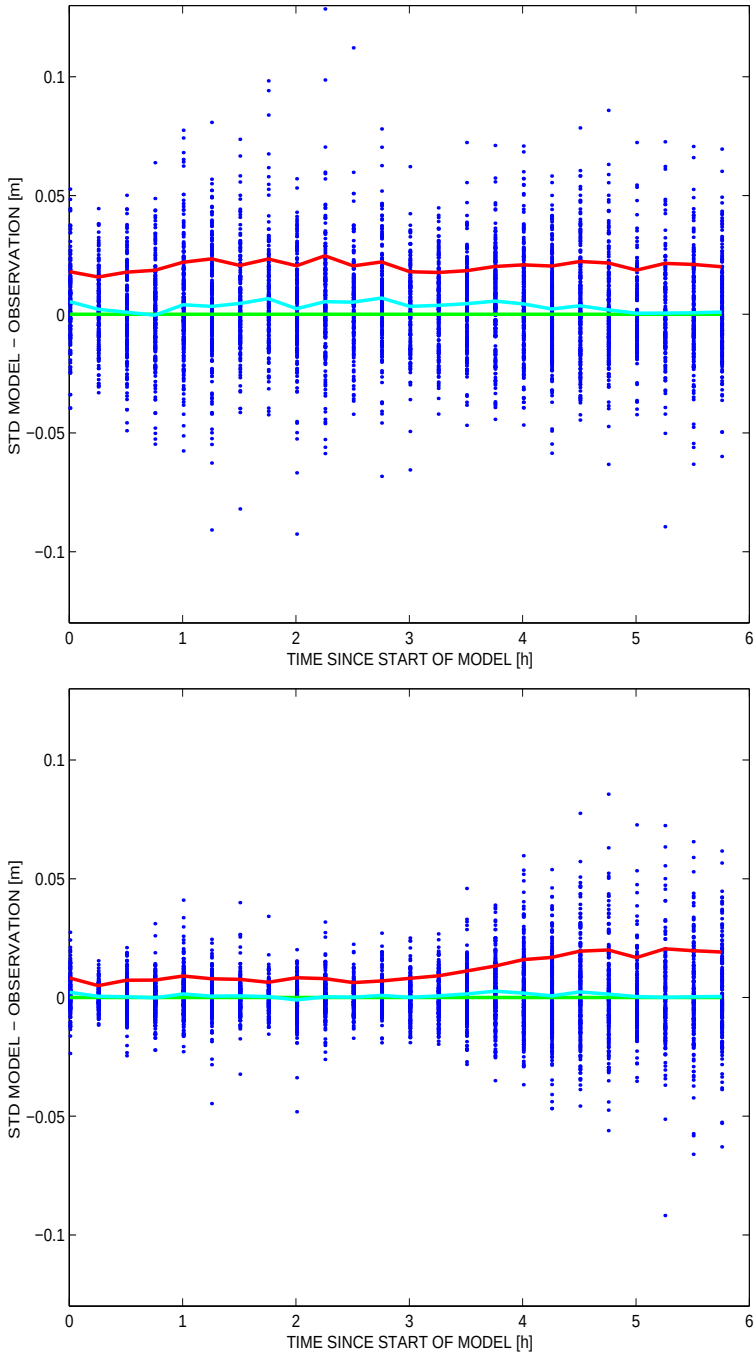


Figure 26: Network red (0-6UTC on August 14, 2007): BmO differences versus the time. Control experiment (upper panel), assimilation experiment (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD.

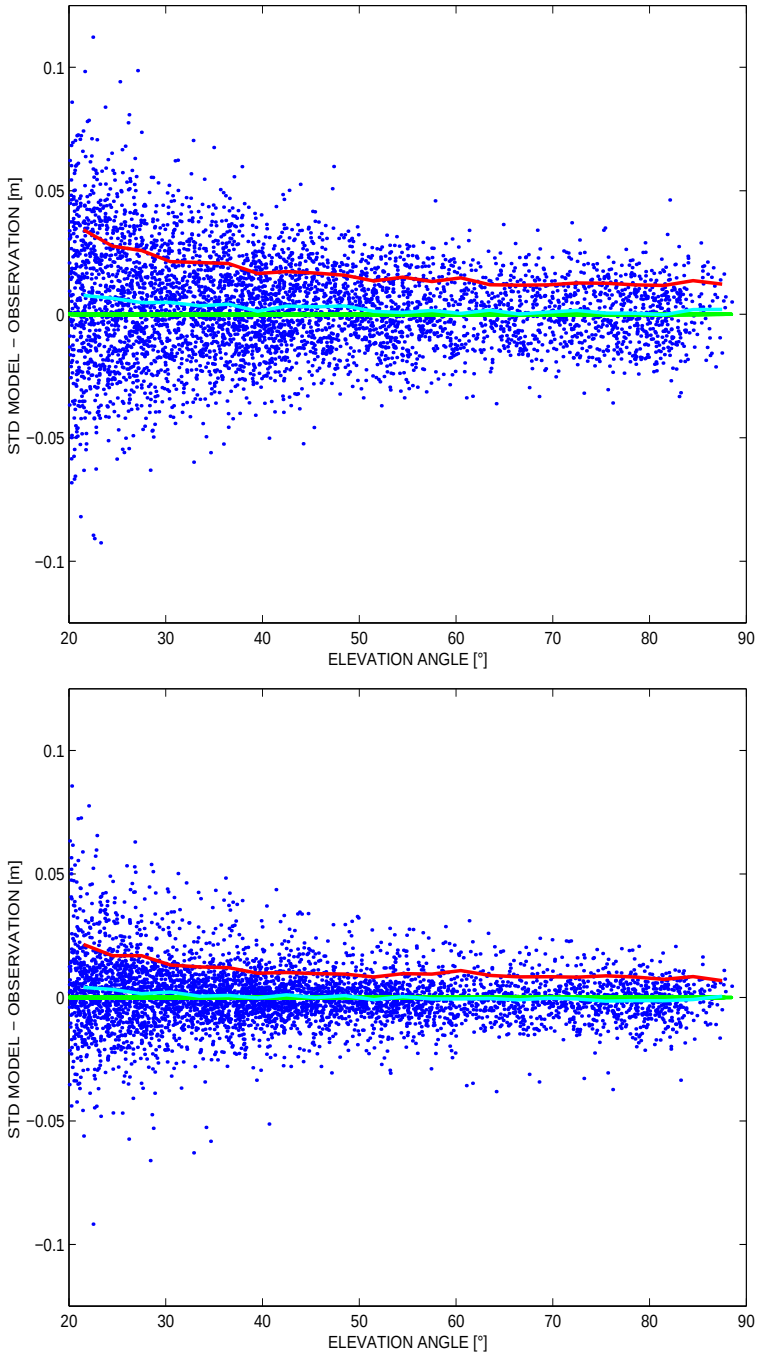


Figure 27: Network red (0-6UTC on August 14, 2007): BmO differences versus the elevation angle. Control experiment (upper panel), assimilation experiment (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD.

network is shown in Figure 28 and confirms that the assimilation has a detectable and significant impact in the assimilation window as well as after the assimilation window in observation space.

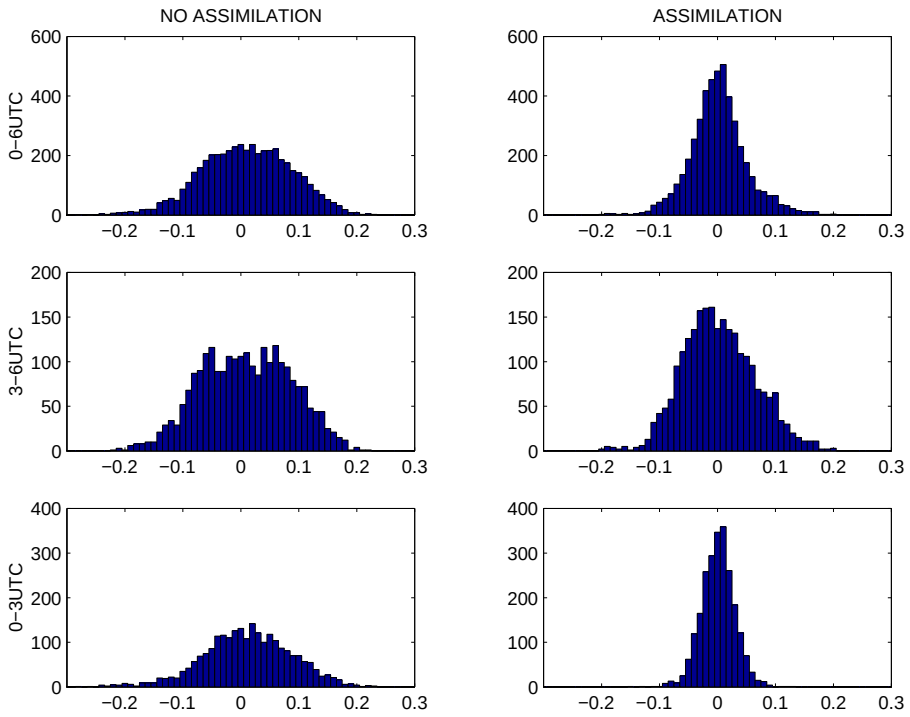


Figure 28: Network red (0-6UTC on August 14, 2007): Frequency histogram of the relative departure of the slant wet delay. The different panels show the different time intervals of the model integration. The overall statistics is shown in the upper most panel.

In model space the difference between the assimilation experiment and the control experiment in the water vapor field reaches ± 1 gkg^{-1}. The differences in the temperature field reach about ± 0.5 K and in the wind field ± 0.5 ms^{-1}. Figure 29 and Figure 30 show the difference in 850 hPa between the assimilation experiment and the control experiment in the water vapor field, the temperature field and the wind field. Air masses are advected above receivers with the ambient wind direction from the west. It is concluded that the impact of the GPS data in the assimilation window as well as after the assimilation window is reasonable and significant.

In order to address question 3, the BmO differences for the green network are computed for the control experiment and the assimilation experiment. Figure 31 shows the BmO differences versus the time for the green network. The RMS is slightly reduced in the assimilation window as well as after the assimilation window. Though the reduction of the RMS for the green network is smaller than the reduction of the RMS for the red

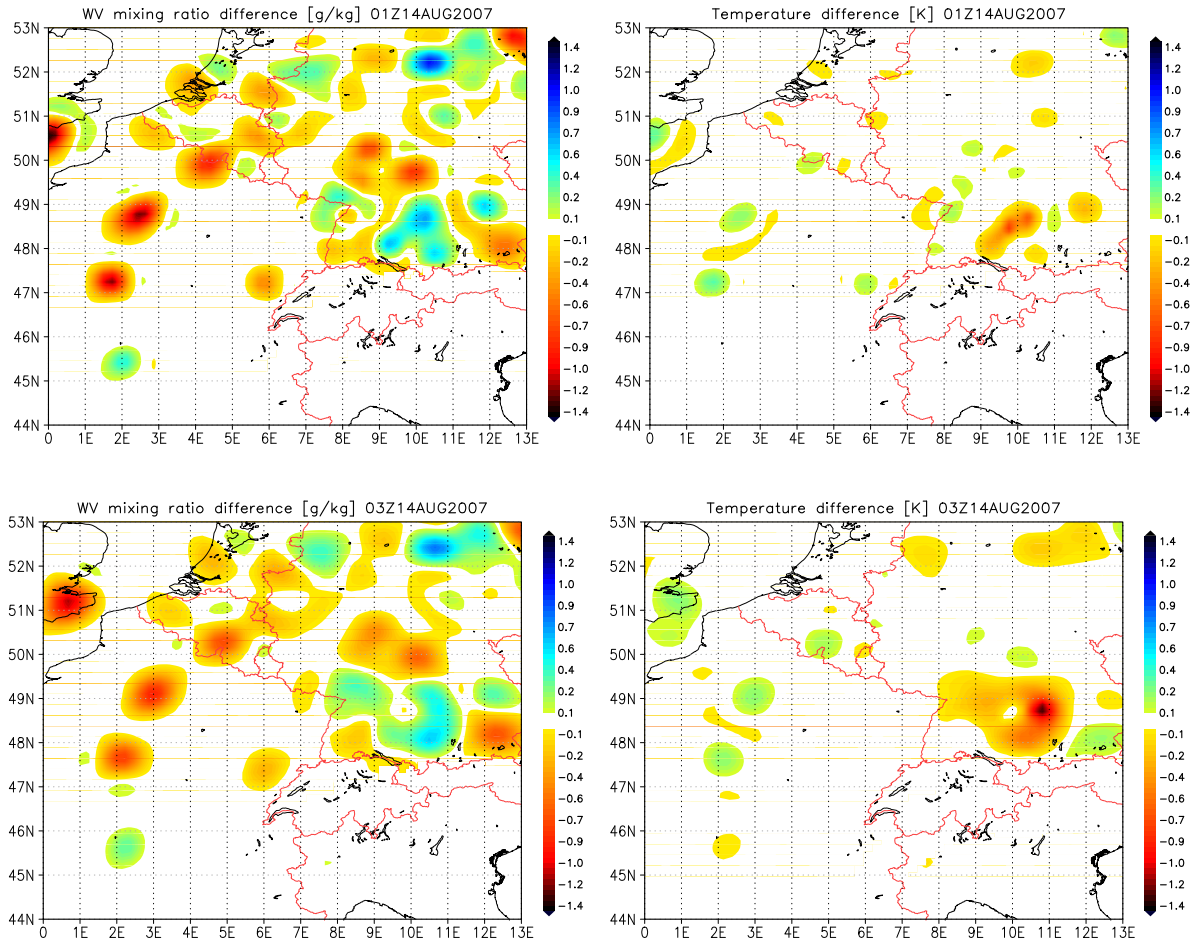


Figure 29: Difference between assimilation and no assimilation in water vapor mixing ratio [g/kg] (left) and temperature [K] (right) in 850 hPa at 1UTC (upper panel) and 3UTC (lower panel) on August 14, 2007.

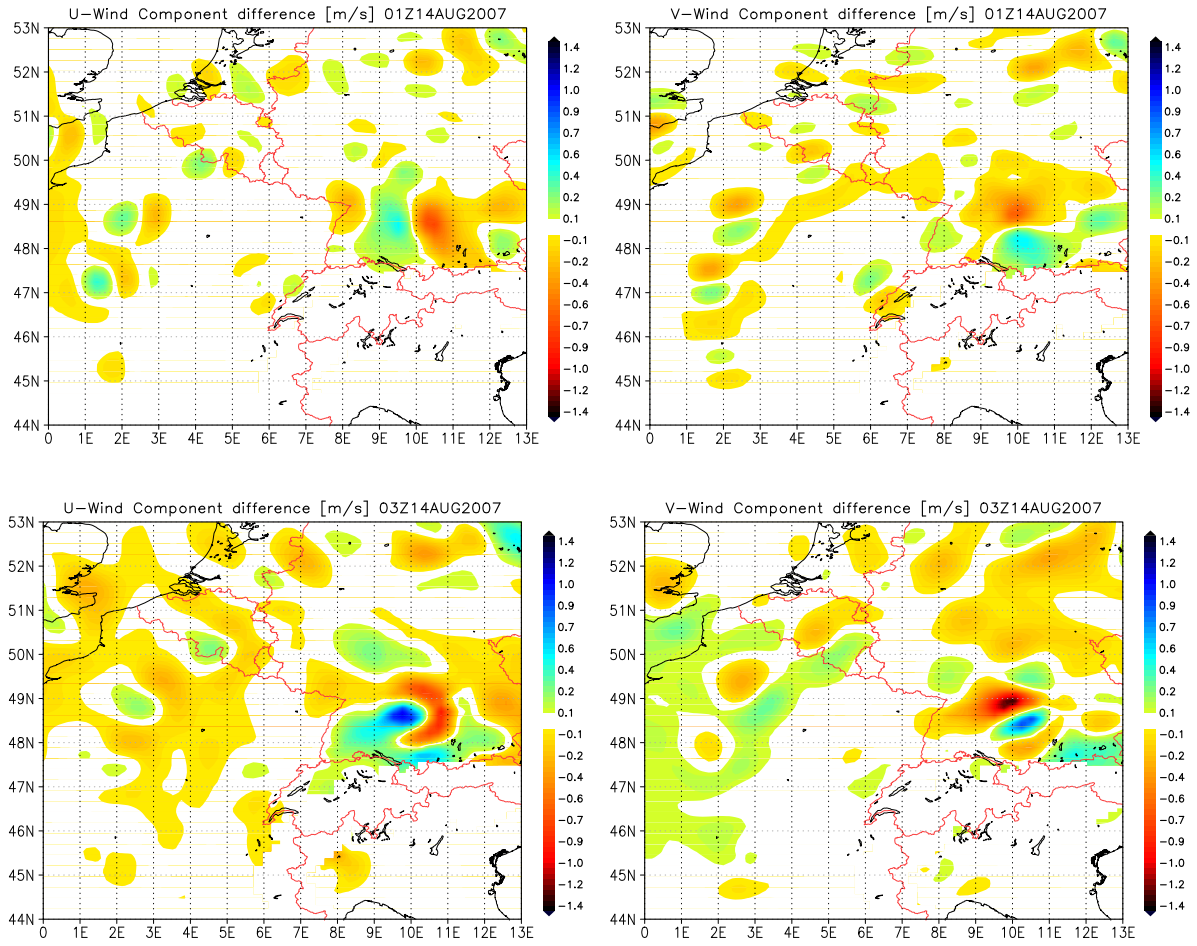


Figure 30: Difference between assimilation and no assimilation in u-wind component [m/s] (left) and v-wind component [m/s] (right) in 850 hPa at 1UTC (upper panel) and 3UTC (lower panel) on August 14, 2007.

network, the reduction is more continuous in time. Figure 32 shows the BmO differences versus the elevation angle for the green network and indicates that the RMS is reduced over the entire elevation range in the assimilation experiment.

The frequency histogram of the relative departure of the slant wet delay for the green network is shown in Figure 33 and confirms that the assimilation has a detectable impact in the assimilation window as well as after the assimilation window in observation space. The statistics of the green network support the results gained by the assimilation of the GPS data of the red network. Figure 33 indicates that the humidity content is better represented in the assimilation experiment than in the control experiment.

In connection with question 4, an intercomparison with independent measurements is performed. All radiosondes from central Europe from 0-6UTC which are available from the ECMWF Meteorological Archival and Retrieval System (MARS) are evaluated. Numerous radiosondes are far away from the impact region as some of the receivers in the green network. For those radiosondes no difference between the assimilation experiment and the control experiment is found. Most of the radiosondes are launched at 0UTC and some at 6UTC. Regarding those profiles where differences are seen, it is observed that the structure, which is present in the radiosonde measurement but not in the control experiment, is partly generated by the assimilation of the GPS data. However, there are specific locations where the water vapor content seems to be better represented in the control experiment when compared to the assimilation experiment.

Figure 34 and Figure 35 show four radiosonde measurements and their model equivalents. The first radiosonde measurement is of particular interest since it is close to Stuttgart, Germany, a convective active region on this particular day. At this radiosonde location, the assimilation enhanced the water-vapor content in the lower troposphere. The simulation with assimilation is in better agreement with the radiosonde measurement compared to the simulation without assimilation except at one height level in the lower troposphere. At this height level another interesting feature is observed: A small kink in the water-vapor profile is introduced by the assimilation, which is also present in the radiosonde measurement, but not in the simulation without assimilation. This structure is attributed to the the ability of the 4DVAR to self-generate physically consistent structure functions during model integration, provided that the model physics is accurate enough and the background errors are reasonable. Similar effects can be observed for the other radiosonde measurements. In particular, the intercomparison with the second radiosonde measurement indicates the remarkable ability of the 4DVAR to generate a complex structure even though the GPS data are integrated measurements. However, it would be beneficial to assimilate GPS data together with conventional and non-conventional data, since otherwise the ability to generate meaningful structure functions relies entirely on model physics and the background error covariance matrix, which, in particular in the MM5 4DVAR must be regarded as a crude estimate. In any case, due to the sparse

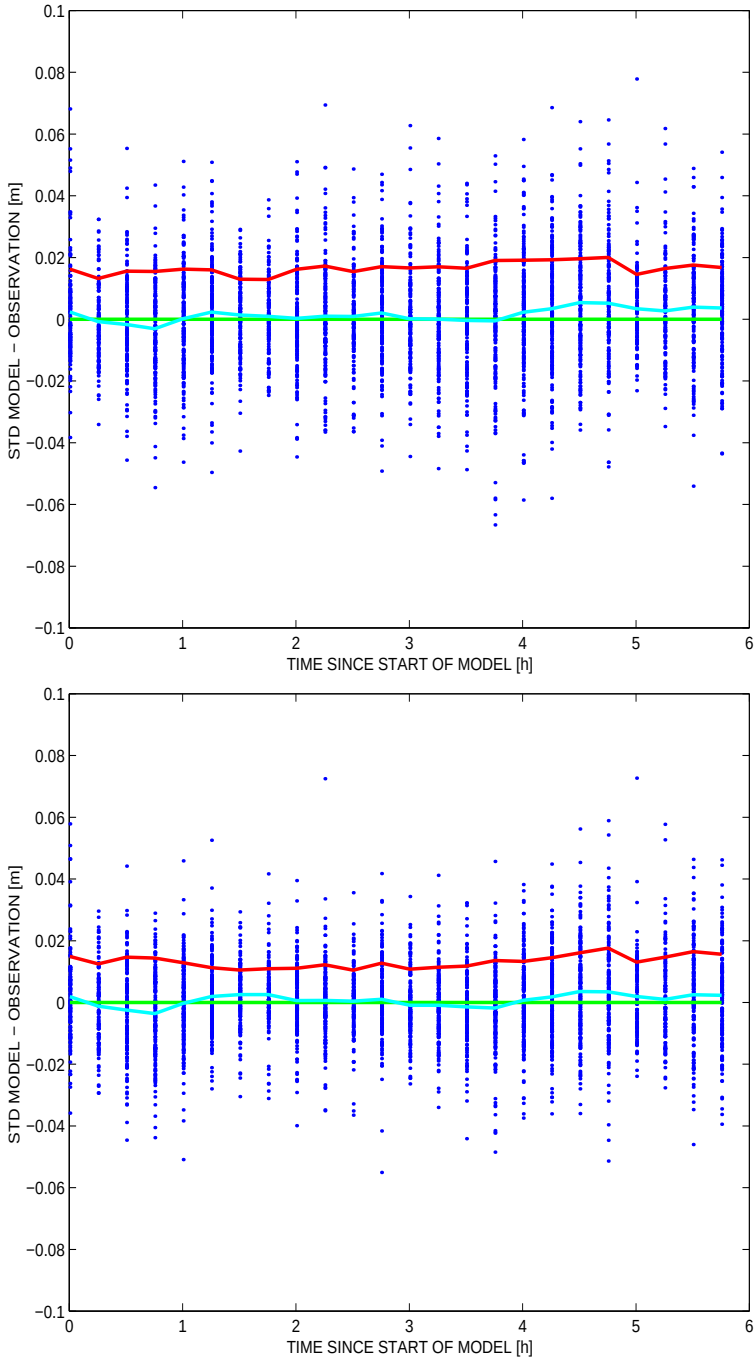


Figure 31: Network green (0-6UTC on August 14, 2007): BmO differences versus the time. Control experiment (upper panel), assimilation experiment (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD.

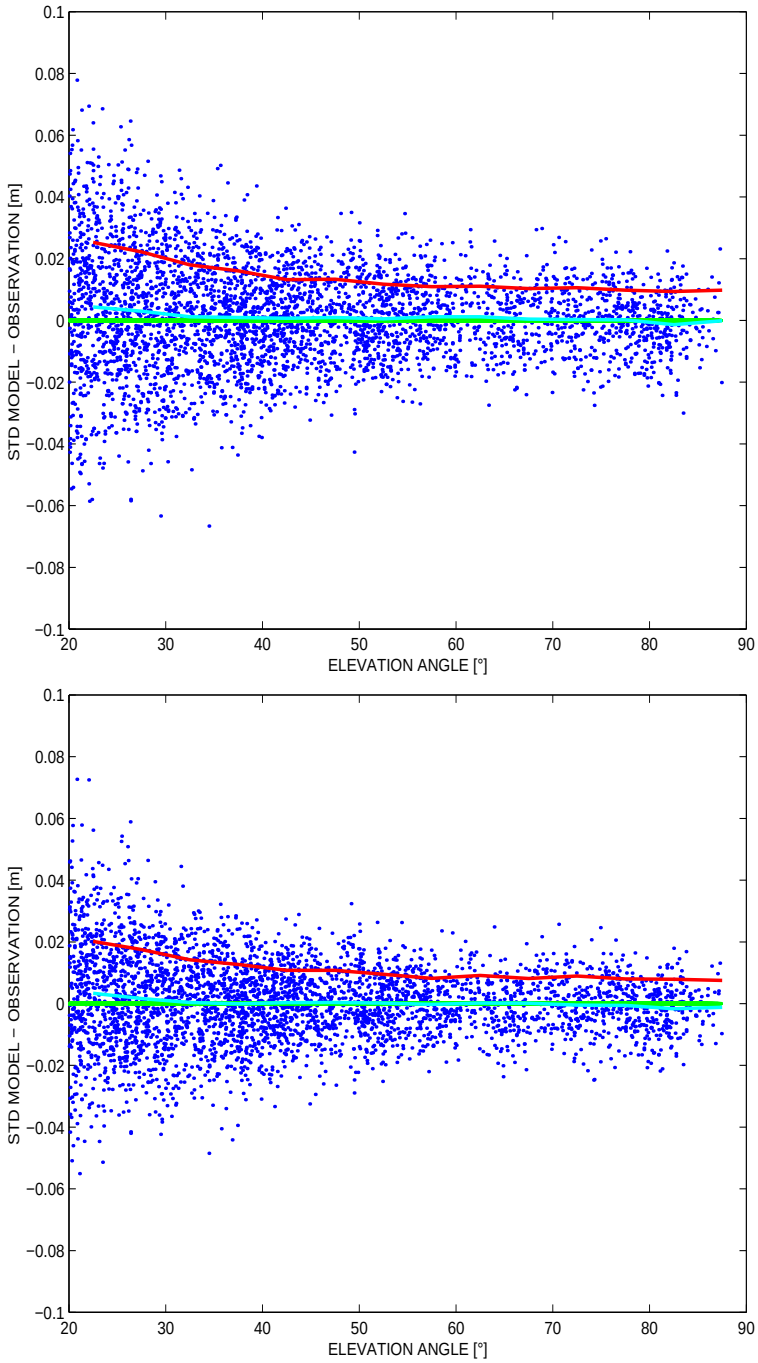


Figure 32: Network green (0-6UTC on August 14, 2007): BmO differences versus the elevation angle (bottom). Control experiment (upper panel), assimilation experiment (lower panel). The red line indicates the RMS, the cyan line indicates the MD and the green line indicates the zero MD.

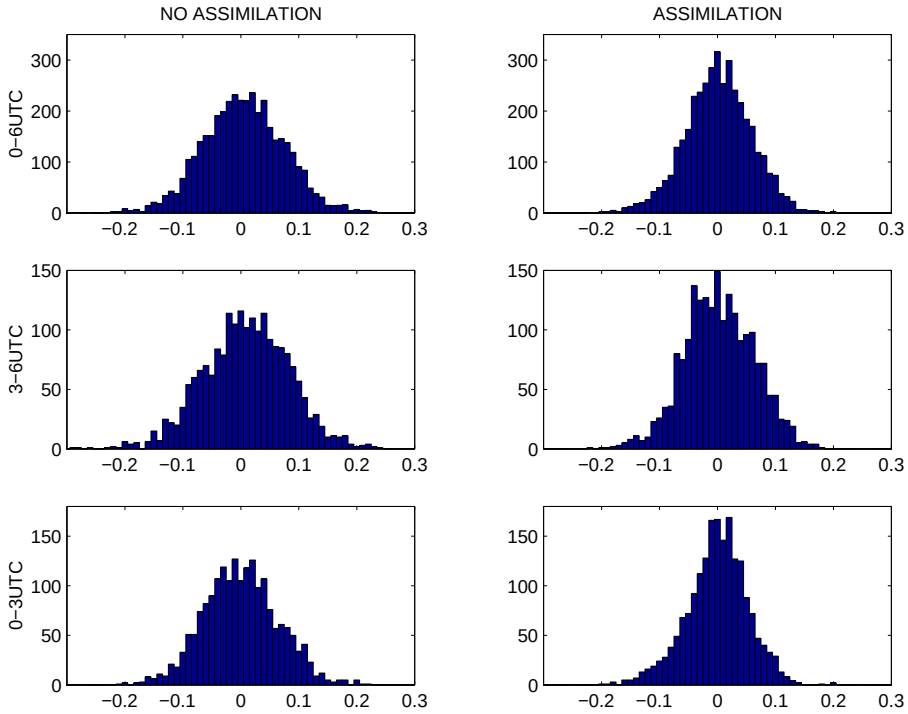


Figure 33: Network green (0-6UTC on August 14, 2007): Frequency histogram of the relative departure of the slant wet delay. The different panels show the different time intervals of the model integration. The overall statistics is shown in the upper most panel.

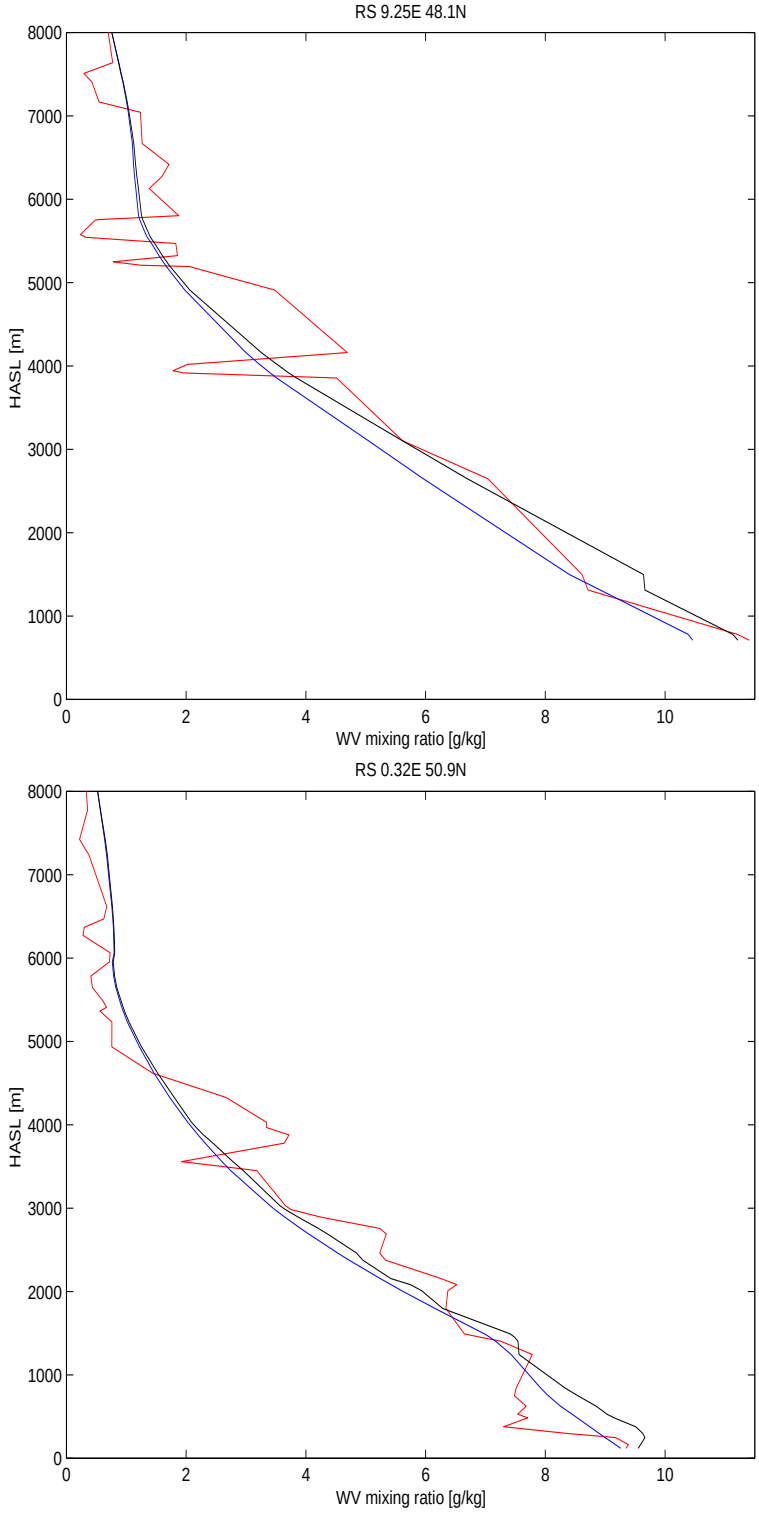


Figure 34: WV mixing ratio profiles at specific radiosonde locations (0-6UTC on August 14, 2007). The red line indicates the radiosonde measurement, the blue line indicates the model prediction without assimilation and the black line indicates the model prediction with assimilation.

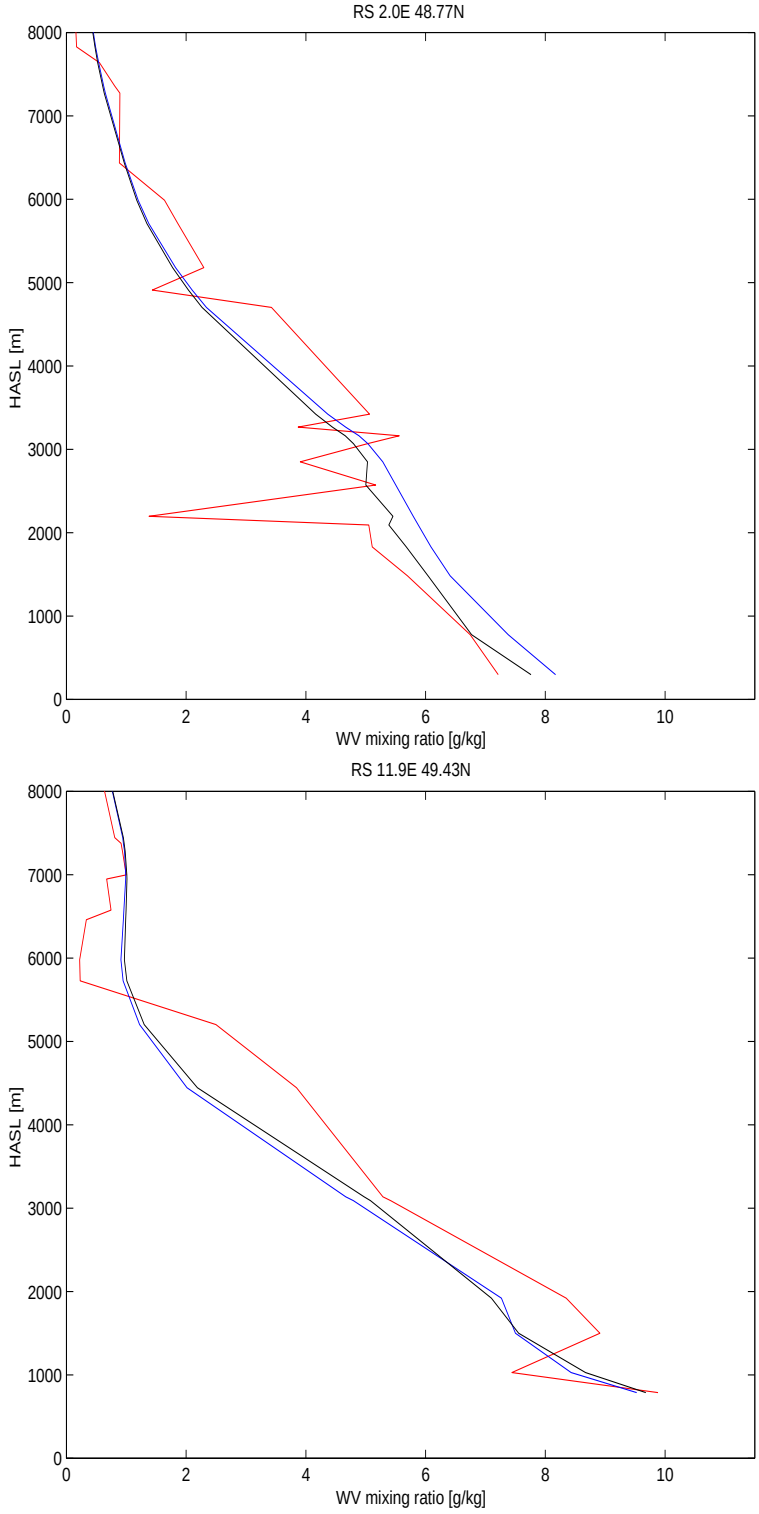


Figure 35: WV mixing ratio profiles at specific radiosonde locations (0-6UTC on August 14, 2007). The red line indicates the radiosonde measurement, the blue line indicates the model prediction without assimilation and the black line indicates the model prediction with assimilation.

character of water-vapor observations, it is difficult to draw a final conclusion regarding the validation with independent upper air field measurements, at least for a short-range forecast.

In connection with question 5, a qualitative comparison with the DWD radar composite is performed to gain insight into the precipitation predicted by the simulations with and without assimilation. Figure 36 shows the 15 minutes accumulated precipitation qualitatively compared to the radar composite of the DWD. The simulation with assimilation predicts the initiation of convection very close to the region where it was observed. Similar to the radar image, the convective system moves with the ambient flow from southwest to northeast before it vanishes in the simulation. A close inspection in the area of interest indicates that the local enhancement of water vapor in the lower troposphere of the order of 1 gkg^{-1} due to the assimilation of the GPS data, also indicated by the water vapor profile of the nearby radiosonde location Stuttgart Figure 34, is the primary source that permits to simulate the observed convective system.

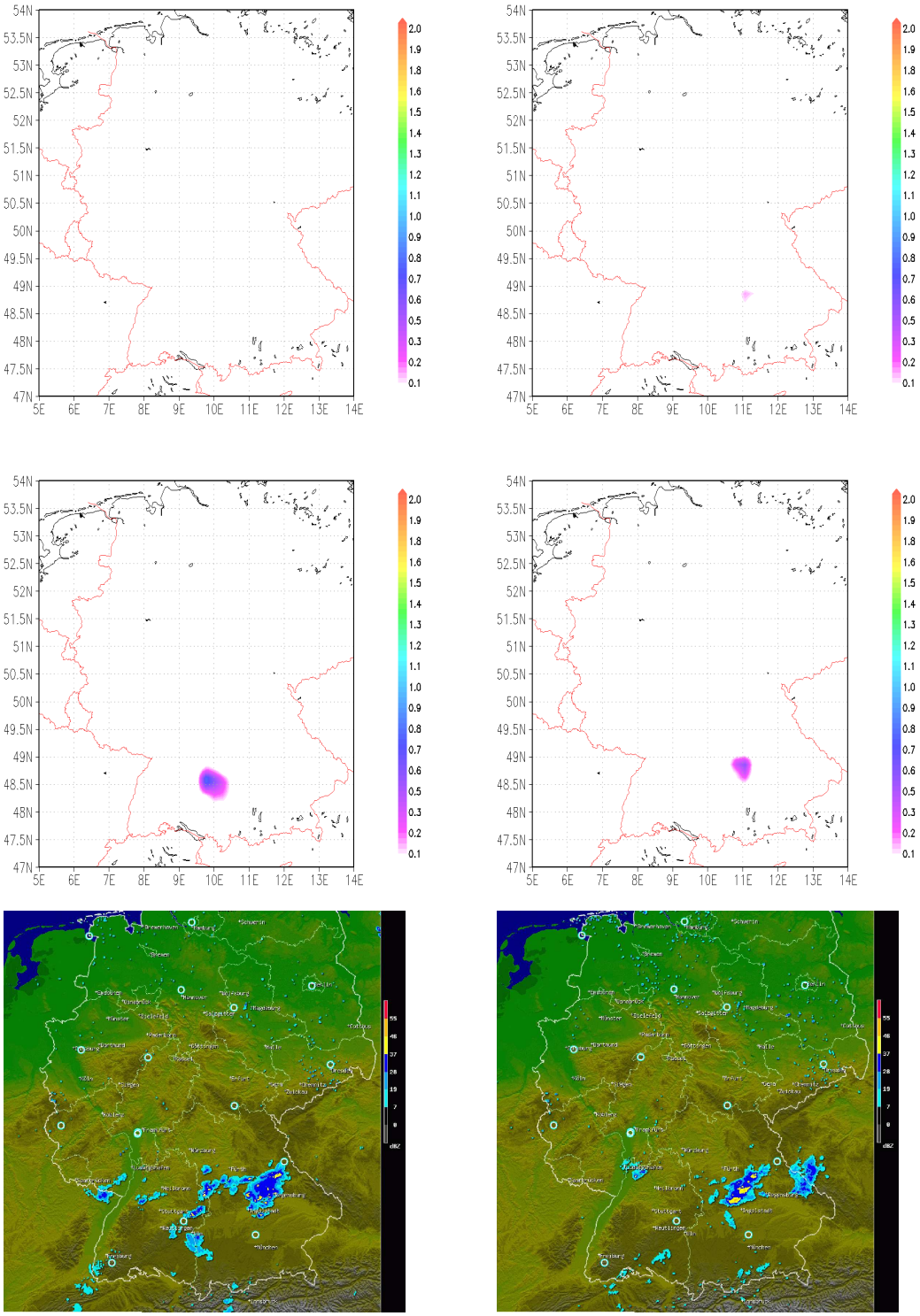


Figure 36: 15 min accumulated precipitation [mm] with assimilation and without assimilation compared qualitatively with the reflectivity of the DWD radar composite at 01:30UTC (left panel) and at 04:00UTC (right panel) on August 14, 2007. No assimilation (upper panel), assimilation (middle panel) and radar image (lower panel).

8.6 Assimilation experiment 2

To estimate the impact of the assimilation of the GPS data on a statistically significant basis, an assimilation experiment was conducted for a period of two weeks. The time period under investigation was 1-14 of August 2007. For each day two simulations were performed, one with assimilation of the GPS data, the other one without. The network red provides GPS data for the assimilation, The network green was the observing network, meaning that GPS data provided by this network were not assimilated. The assimilation window was 3 hours, ranging from 0-3UTC with an assimilation frequency of 30 minutes. The forecast range was 0-24UTC and 15 minute data were evaluated.

Figure 37 and Figure 38 shows the ND distribution of the slant path delay for the red and green network for different time intervals. At first, we take a look at the ND distribution of the red network. Concerning the first three hours of the simulation, the expected behavior is observed. The scatter of the ND distribution is strongly reduced in the first few hours of model integration. The impact of the assimilation fades away after a few hours of model integration. One may think that the reduction of the scatter of the ND distribution in the assimilation window is a trivial result. This is not the case: The simple physics used in the MM5 4DVAR is different from the full and more comprehensive physics used in the forecast. No artefacts (significant biases) are observed since the important components, i.e. the cumulus scheme and the horizontal diffusion scheme, have a similar performance in both, the assimilation and the forecast configuration.

It is interesting to observe that the ND distribution of the control experiment is not completely Gaussian in the first hours of the model integration (0-3UTC), whereas it is close to Gaussian when the forecast model integration continuous in time (3-18UTC). The Gaussian distribution in the first few hours of model integration is slightly skewed. This feature can be problematic regarding variational assimilation, since the underlying hypothesis is that the observation and background statistics are Gaussian. The deviation from a Gaussian distribution is not attributed to deficiencies in the forward/forecast model or errors in the observations. It is rather associated to the model spin up, an inherent problem of limited area mesoscale models which are not in a cycling NWP mode. No spin-up runs to remove artefacts resulting from the resolution jump and the accompanied interpolation errors from the ECMWF analysis to the MM5 model grid configuration were performed. The resolution jump from the ECMWF analysis to the MM5 model grid configuration is small (25 to 18 km) and therefore this contribution to the spin-up is limited. However, the spin-up effect due to the different dynamics of the IFS and the MM5 and the spin-up effect caused by hydrometeors which are all reseted in the MM5 initial state, is not precluded.

The reduction of the scattering of the ND distribution of the green network is much smaller than the reduction of the scattering of the ND distribution for the red network. Again,

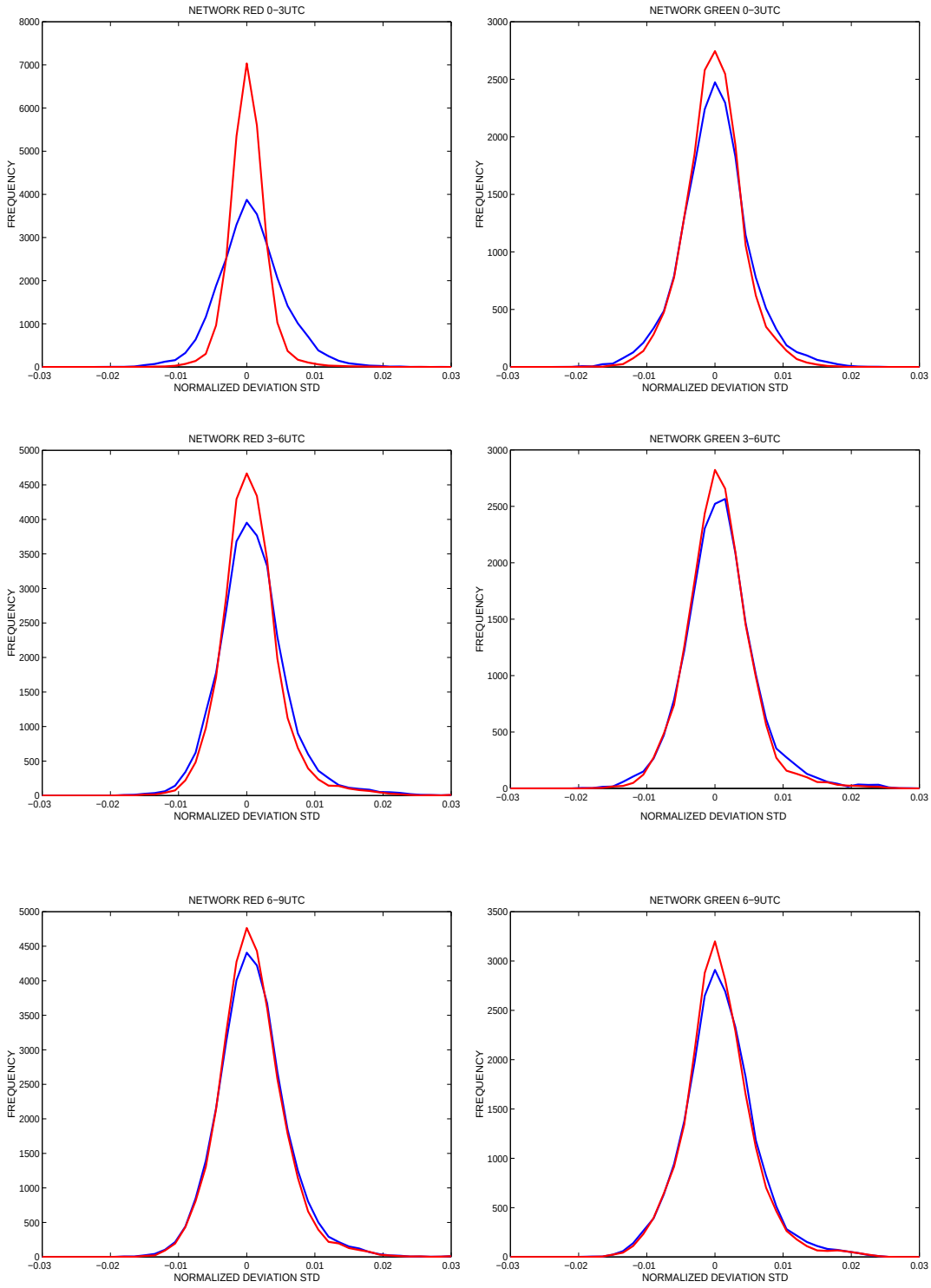


Figure 37: Frequency histogram of the normalized deviation of the slant path delay for the red and the green network. The red line corresponds to the assimilation experiment and the blue line corresponds to the control experiment. The different panels show the different time intervals of the model integration (1-14 of August 2007).

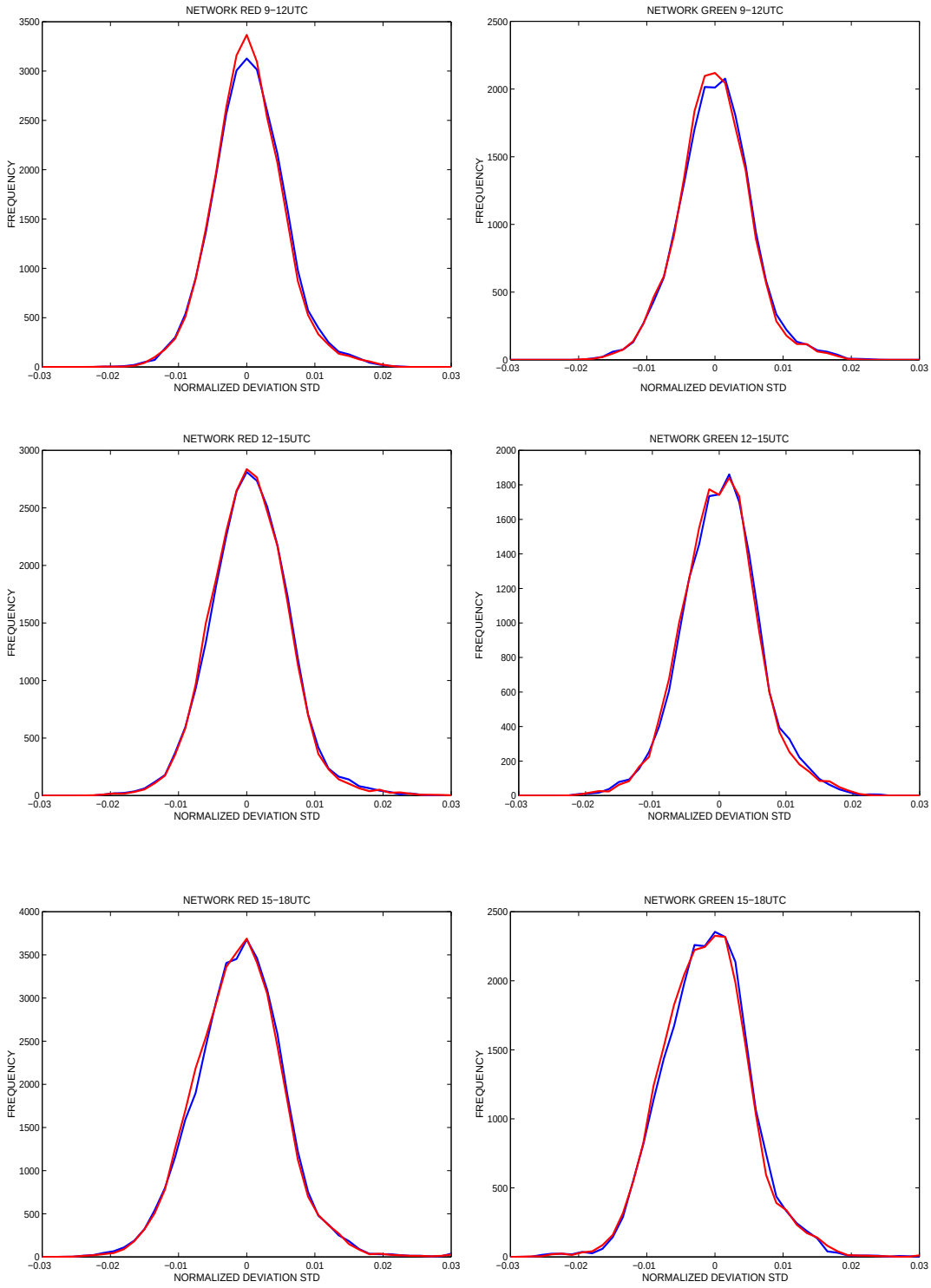


Figure 38: Frequency histogram of the normalized deviation of the slant path delay for the red and the green network. The red line corresponds to the assimilation experiment and the blue line corresponds to the control experiment. The different panels show the different time intervals of the model integration (1-14 of August 2007).

it is observed that the reduction is much more continuous in time. The ND distribution of both networks is virtually the same for the time interval 6-9UTC. As a whole the green network supports the results gained by the assimilation of the GPS data of the red network, and thus indicates, on a statistical significant basis, a slight improvement in the humidity content with the assimilation of the GPS data. Inspection for both, the control experiment and the assimilation experiment, for the red and the green network reveals that at about 12-15UTC the mean of the ND distribution starts shifting considerably. The reason for this tendency is not yet understood. Further studies using different model configurations are required to understand this tendency.

8.7 Assimilation experiment 3

To investigate the impact of the assimilation of the GPS data on the precipitation forecast, an assimilation experiment was conducted for a period of one month. The time period under investigation was 1-31 of August 2007. For each day two simulations were performed, one with assimilation of the GPS data, the other one without. In this experiment, both networks red and green provided GPS data for the assimilation. The assimilation window was 3 hours, ranging from 0-3UTC, and the forecast range was 0-24UTC. The assimilation frequency was 30 minutes.

The hourly measured precipitation data is taken from stations of the RR network of the DWD. The precipitation data is converted to the surface grid points of the mesoscale model. For a specific grid point, the precipitation data is averaged for stations which have a distance $\leq 0.2^\circ$ to that grid point. The gridded precipitation data set is used for intercomparison with the predicted precipitation of the model if at least two stations contribute to the average.

Figure 39 shows the accumulated precipitation field for August 2007 for the observation, the control experiment and the assimilation experiment. The precipitation patterns are similar for all cases. However, the precipitation maximum visible in the observation located at about 8E 51.5N is absent in both the control and the assimilation experiment. On contrary, the control experiment and the assimilation experiment show two precipitation maxima located at about 8E 50.5N and 9E 50N respectively. Apart from the misplacement of the maxima in the accumulated precipitation the agreement between the observation and the two model simulations is good.

Figure 40 shows the difference in the accumulated precipitation between the control experiment and the observation and the assimilation experiment and the observation. Inspection by eye reveals that the differences are very similar for both model simulations. However, there are areas where the differences are more pronounced. The large differences occurring through the misplacement of the precipitation maxima in the simulations are

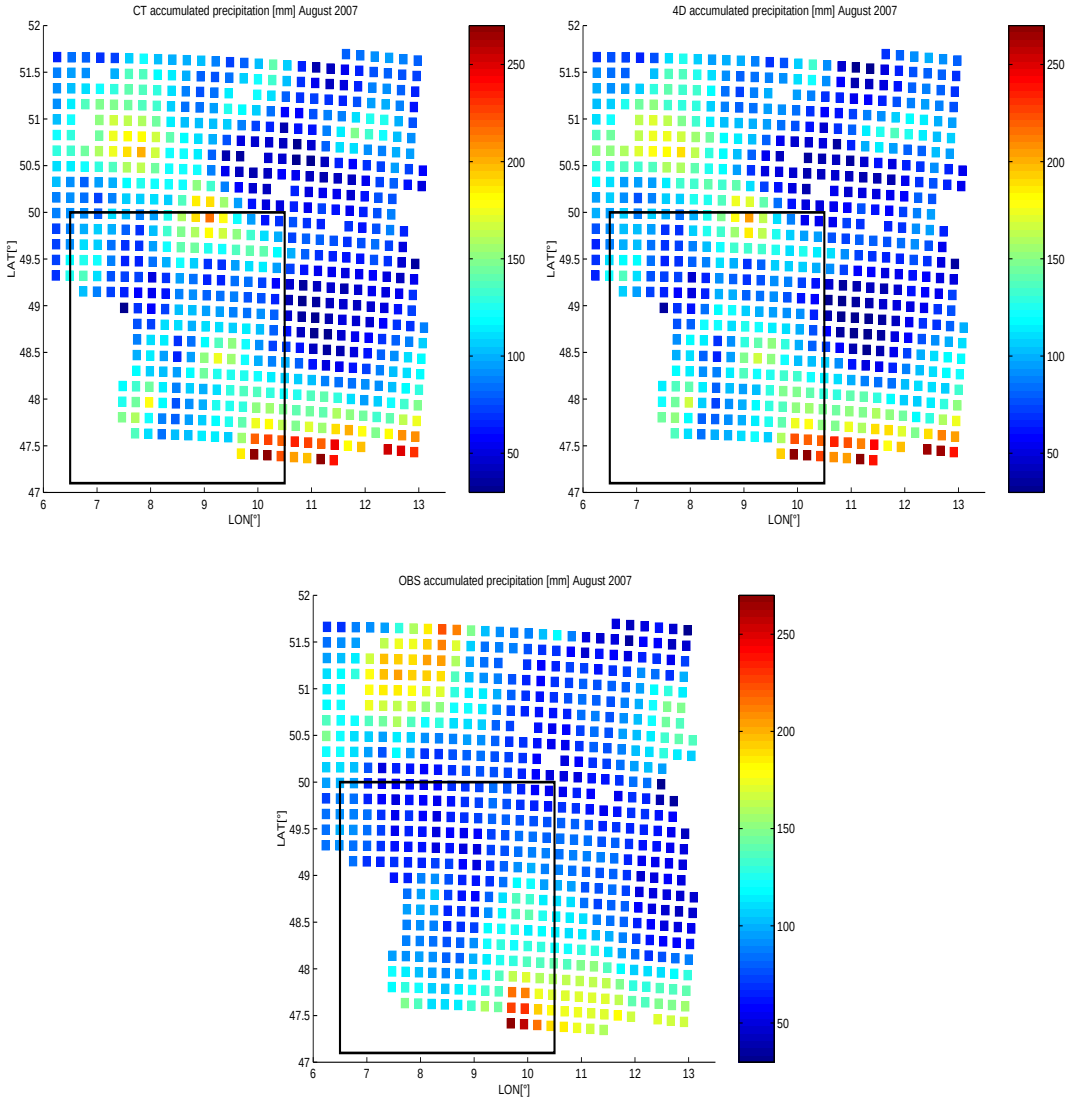


Figure 39: Accumulated precipitation [mm] (1-31 of August 2007) for the control experiment, the the assimilation experiment and the observation. The black rectangle indicates the COPS area.

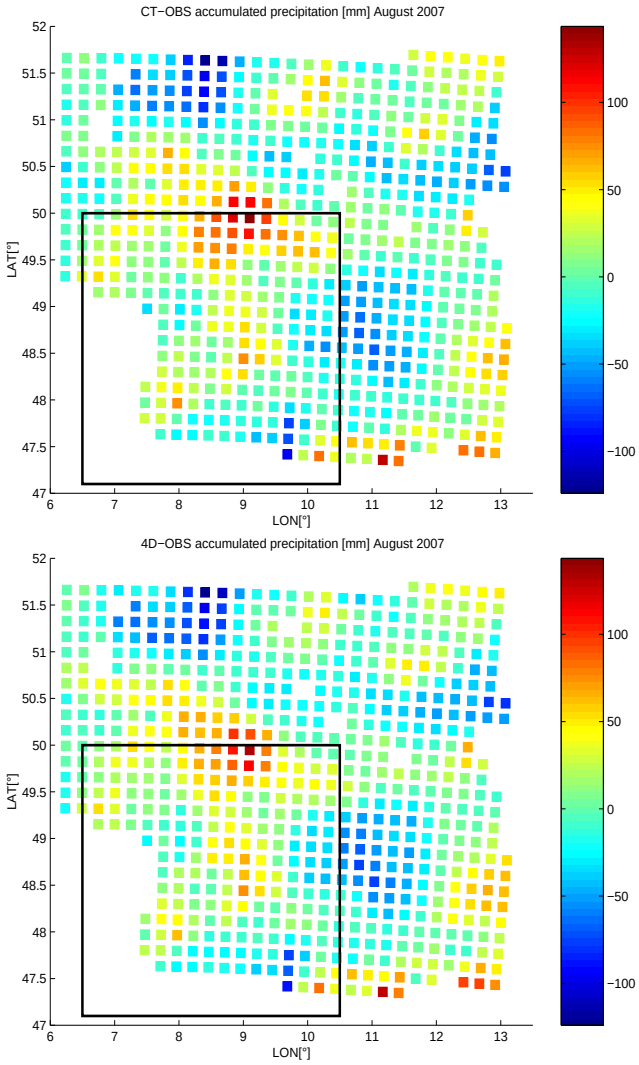


Figure 40: Difference in the accumulated precipitation [mm] (1-31 of August 2007) between the control experiment and the observation (upper panel) and between the assimilation experiment and the observation (lower panel). The black rectangle indicates the COPS area.

slightly reduced in the assimilation experiment. In addition, large differences visible at isolated locations are slightly reduced in the assimilation experiment. This might be an indication that outliers in the precipitation forecast can be reduced by the assimilation of the GPS data.

Table 7 summarized the RMS, the MD, the MAD and the PMCC for the control experiment and the assimilation experiment. The statistical quantities show that the assimilation of the GPS data has a weak but positive impact. In addition, Figure 41 shows the POD, the FAR, the TSS and the ETS, for the control experiment and the assimilation experiment versus various precipitation thresholds. Essentially all skill scores for all precipitation thresholds confirm that the assimilation of the GPS data has a weak positive impact.

Statistics	RMS [mm]	MD [mm]	MAD [mm]	PMCC
CT	36.65	3.39	28.24	0.62
4D	35.89	3.05	27.64	0.64

Table 7: Statistics of the control and the assimilation experiment.

To get some qualitative insight into the temporal evolution of the accumulated precipitation in the COPS area Figure 42 shows the averaged daily cycle of precipitation for the observation, the control experiment and the assimilation experiment. Two features are obvious. At first, both simulations underestimate precipitation in the first few hours of the model integration (0-6UTC). Second, both simulations overestimate precipitation during day time (6-24UTC). The underestimation of precipitation can be readily associated with the model spin-up. The overestimation of precipitation might be to some extend a consequence of the underestimation in the first few hours of the model integration. It is interesting to observe that qualitatively the evolution of the accumulated precipitation of the assimilation experiment is closer to the the observation. The 4DVAR seems to be able to reduce the spin-up phase and to improve the evolution of the accumulated precipitation during day time. Again, it can not be precluded that the ability of the 4DVAR to reduce the model spin-up is the reason for the improved evolution during day time. However, the differences between the assimilation experiment and the control experiment are rather small. The averaged daily cycle indicates a weak but positive impact in the temporal evolution of precipitation.

As a whole, the impact of the assimilation of the GPS data must be regarded as slightly positive to neutral. It is concluded, that in the current configuration the impact on a precipitation forecast is rather weak. This finding supports results obtained from comparable studies utilizing the impact of the GPS data on precipitation (Vedel and Huang, 2004; Poli et al., 2007), and in fact, this result is not too surprising. The assimilation experiment 2 already indicated that the impact of the assimilation in the humidity field

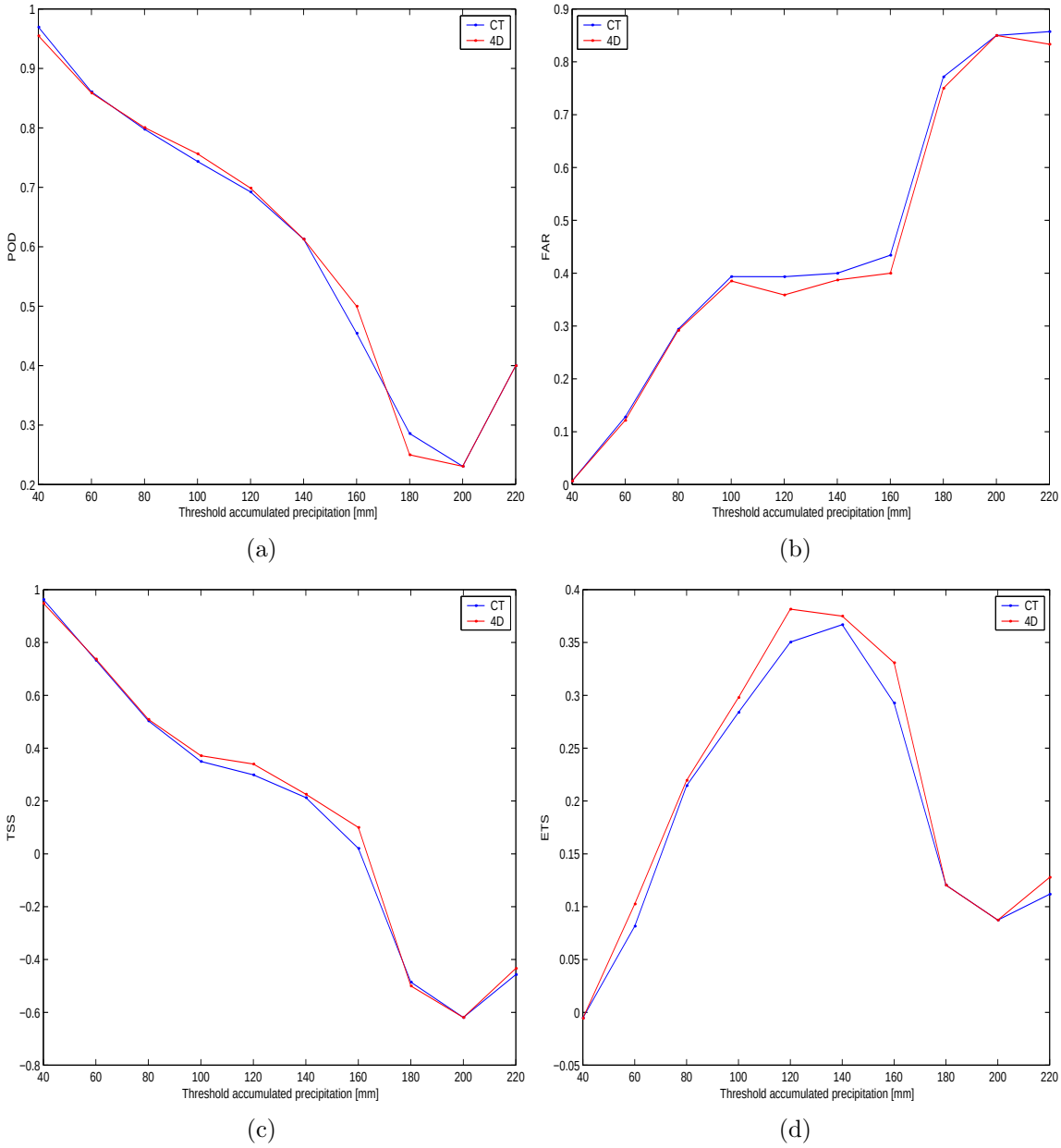


Figure 41: Skill scores for the control experiment and the assimilation experiment versus the precipitation threshold [mm] (1-31 of August 2007). a) POD b) FAR c) TSS and d) ETS.

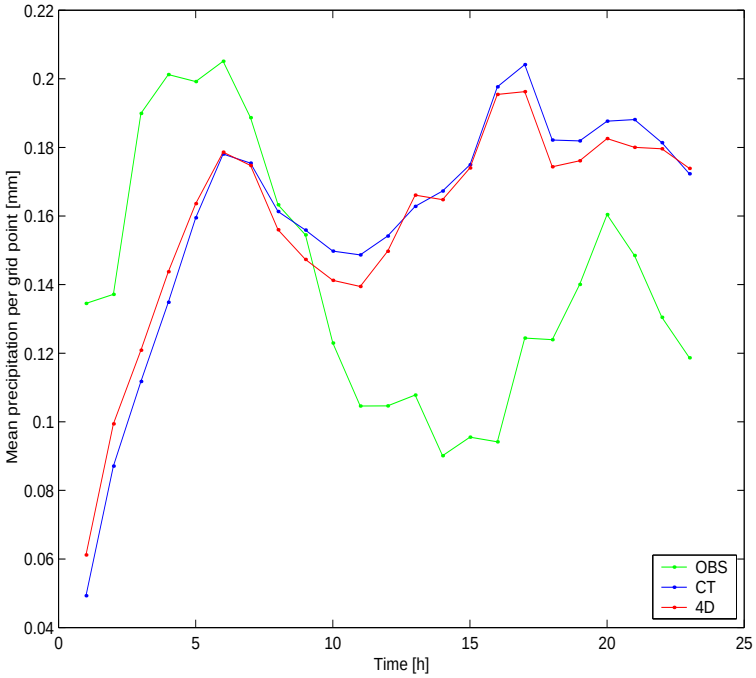


Figure 42: The averaged daily cycle of precipitation [mm] in the COPS area (1-31 of August 2007) for the observation, the control experiment and the assimilation experiment.

is weak and lasts for no more than 12 hours. Since precipitation is strongly linked to the humidity field, and there is no reason to assume that a precipitation forecast can be improved without an improved humidity field, the results from the assimilation experiment 3 supports the results from the assimilation experiment 2 and vice versa.

9 Summary and outlook

Ray-tracing is necessary to simulate the propagation of radio signals through a mesoscale model. In this thesis, a rigorous ray-tracing algorithm, based on a global approach, was developed which allows the direct numerical simulation of GPS slant path delays in the MM5. The ray-tracing algorithm is based on a finite difference scheme, involves the exact location of the receiver and the satellite automatically in the overall solution, and does not rely on any other information or empirical model. The structured non-linear system of equations, arising due to the finite difference scheme, can be efficiently solved by Newton's iteration.

Extensive tests were performed to validate the ray-tracing algorithm. The main features of ray-traced slant path delays were utilized. For elevation angles as low as 5° at the ground-based receiver a single Newton iteration turns out to be sufficient. The developed ray-tracing algorithm proves to be a robust and accurate numerical tool, applicable over the entire elevation range. An immediate by-product of the ray-tracing algorithm is the bending angle and/or the Doppler frequency shift. Thus, besides the ability to simulate slant path delays, the ray-tracing algorithm developed in this thesis is believed to be an appropriate tool for simulating measurements derived from GPS RO. As an alternative to the rigorous computational model for the slant path delay, a fast computational model for the slant path delay was developed.

The ray-tracing algorithm was used to estimate the potential azimuthally asymmetric information content. The simulation indicates that in particular slant path delays for elevation angles $< 20°$ can contain valueable asymmetric information. It was the first time that slant path delay data, collected by a large network of continuously operating ground-based receivers, are compared with their mesoscale model equivalents over the entire elevation range using a highly sophisticated observation model. The statistical intercomparison of simulated and observed slant path delays indicates a good agreement over the entire elevation range. The estimated RMS of the IWV is about 1.3 mm. The sensitivity of simulated slant path delays with respect to forward model components was investigated. It was found that for elevation angles $< 30°$ the ray bending effect due to refractivity gradients cannot be neglected.

The sensitivity of simulated slant path delays with respect to two forecast model components, namely the horizontal diffusion scheme and the parameterization of deep convection, was investigated. Regarding the use of ground-based GPS data for model evaluation, the following conclusion can be drawn from the sensitivity experiments: A spurious tendency present in the model minus observation statistics of the slant path delay data is strongly related to the misplacement and/or the enhancement/suppression of precipitation. It is concluded that slant path delay observations are potentially useful for verification purposes. Noteably, GPS slant path delay data are not merely valueable for

verification purposes. In this thesis, the analysis of the model minus observation statistics of the GPS slant path delay data was actually used to improve a forecast model component. The sensitivity studies were restricted to two forecast model components. In future, sensitivity studies shall be extended to remaining parts of the forecast model, in particular the vertical diffusion scheme and parameterization schemes associated with the soil surface exchange.

A description of basic principles of tangent linear and adjoint code construction was given to provide insight into an interesting coding technique applicable over a wide range of applications in scientific computing. Tangent linear and adjoint code construction was illustrated for the fast computational model for slant path delay and a forecast model component. The provided tangent linear and adjoint code can be directly used as a recipe for implementation in a different mesoscale model, e.g. the Weather Research and Forecasting Model (WRF). The impact of a single slant path delay in variational assimilation was estimated analytically. Difficulties arising due to imperfect model components incorporated in variational assimilation were discussed.

A set of modifications to the existing MM5 4DVAR system was introduced. A modified horizontal diffusion scheme was developed, which eliminates the erroneous temperature and moisture tendencies in complex terrain. Furthermore, the Grell cumulus scheme and its adjoint were implemented in the MM5 4DVAR system in order to eliminate the erroneous over-prediction of precipitation introduced by the Kuo cumulus scheme. Although the author does not preclude any other remaining deficiencies in the MM5 4DVAR, he believes that these were the most significant systematic model errors.

In this thesis, tools for routine data assimilation of slant path delay observations for the MM5 4DVAR were developed, which allow the assimilation of slant path delay data in the MPI environment. The implemented observation model and its adjoint rely on the fast computational model for the slant path delay. As a consequence, the potential asymmetric information content present in the GPS data can not be completely utilized. In future, the bending of the ray trajectory shall be accounted for in 4DVAR in order to take full advantage of the GPS data.

A single case study was discussed in detail. The results from this single case study can be summarized as follows: (1) The 4DVAR of GPS data is efficient in reducing the model minus observation departure. (2) The impact is area wide and reaches ± 1 gkg^{-1} in the water vapor mixing ratio, ± 0.5 K in the temperature and ± 0.5 ms^{-1} in the wind field. (3) GPS data of an independent network of receivers confirm the results gained by 4DVAR. (4) In a convectively active region, radiosonde measurements confirm the results gained by 4DVAR. The vertical structure in the water vapor field, present in radiosonde measurements, is partly generated by the assimilation of the GPS data. The intercomparison indicates the remarkable ability of the 4DVAR to generate a complex structure in the water vapor field even though the GPS data are integrated measurements.

(5) A qualitative comparison with radar data shows that the 4DVAR of GPS data improves the simulation of an observed convective system.

An assimilation experiment over a period of two weeks indicates that the impact in observation space of an independent network of receivers was positive. The experiment revealed, that, in the current configuration, the impact in the humidity content lasts for about 12 h. Finally, an assimilation experiment over a period of one month was performed. Both, subjective verification (inspection of the precipitation pattern by eye) and objective verification revealed that the impact of the assimilation on the precipitation forecast was positive.

Though the results of the assimilation experiments presented in this thesis are promising, there is room left for improvement. In particular, a Rapid Update Cycle (RUC) would be beneficial to circumvent the model spin up, an inherent problem of limited area mesoscale models which are not in a cycling NWP mode. In addition, GPS data shall be assimilated together with conventional data and non-conventional data, e.g. surface data and radar radial wind data, in order to complement one another and to reduce the degrees of freedom in the assimilation. The tuning of background/observation errors as well as the development of a more sophisticated data selection algorithm is also a subject to future research.

A The Jacobian

The non-zero entries of the Jacobian **J** are

$$
\begin{aligned}
J_{2j-1,2j-3} &= \frac{\partial R_j^y}{\partial y_{j-1}} = L''_{j,j-1}(x_j) - L'_{j,j-1}(x_j) \cdot \beta_j^y \\
J_{2j+0,2j-2} &= \frac{\partial R_j^z}{\partial z_{j-1}} = L''_{j,j-1}(x_j) - L'_{j,j-1}(x_j) \cdot \beta_j^z \\
J_{2j-1,2j-1} &= \frac{\partial R_j^y}{\partial y_j} = L''_{j,j}(x_j) - L'_{j,j}(x_j) \cdot \beta_j^y - \gamma_j^y \\
J_{2j+0,2j+0} &= \frac{\partial R_j^z}{\partial z_j} = L''_{j,j}(x_j) - L'_{j,j}(x_j) \cdot \beta_j^z - \gamma_j^z \\
J_{2j-1,2j+1} &= \frac{\partial R_j^y}{\partial y_{j+1}} = L''_{j,j+1}(x_j) - L'_{j,j+1}(x_j) \cdot \beta_j^y \\
J_{2j+0,2j+2} &= \frac{\partial R_j^z}{\partial z_{j+1}} = L''_{j,j+1}(x_j) - L'_{j,j+1}(x_j) \cdot \beta_j^z \\
J_{2j-1,2j+2} &= \frac{\partial R_j^y}{\partial z_{j+1}} = -L'_{j,j+1}(x_j) \cdot \alpha_j^y \\
J_{2j+0,2j+1} &= \frac{\partial R_j^z}{\partial y_{j+1}} = -L'_{j,j+1}(x_j) \cdot \alpha_j^z \\
J_{2j-1,2j+0} &= \frac{\partial R_j^y}{\partial z_j} = -L'_{j,j}(x_j) \cdot \alpha_j^y - \delta_j^y \\
J_{2j+0,2j-1} &= \frac{\partial R_j^z}{\partial y_j} = -L'_{j,j}(x_j) \cdot \alpha_j^z - \delta_j^z \\
J_{2j-1,2j-2} &= \frac{\partial R_j^y}{\partial z_{j-1}} = -L'_{j,j-1}(x_j) \cdot \alpha_j^y \\
J_{2j+0,2j-3} &= \frac{\partial R_j^z}{\partial y_{j-1}} = -L'_{j,j-1}(x_j) \cdot \alpha_j^z
\end{aligned}
$$

for $j = 2, ..., n-1$ and

$$
\begin{aligned}
J_{1,1} &= \frac{\partial R_1^y}{\partial y_1} = L''_{1,1}(x_1) - L'_{1,1}(x_1) \cdot \beta_1^y - \gamma_1^y \\
J_{2,2} &= \frac{\partial R_1^z}{\partial z_1} = L''_{1,1}(x_1) - L'_{1,1}(x_1) \cdot \beta_1^z - \gamma_1^z \\
J_{1,3} &= \frac{\partial R_1^y}{\partial y_2} = L''_{1,2}(x_1) - L'_{1,2}(x_1) \cdot \beta_1^y \\
J_{2,4} &= \frac{\partial R_1^z}{\partial z_2} = L''_{1,2}(x_1) - L'_{1,2}(x_1) \cdot \beta_1^z \\
J_{1,4} &= \frac{\partial R_1^y}{\partial z_2} = -L'_{1,2}(x_1) \cdot \alpha_1^y \\
J_{2,3} &= \frac{\partial R_1^z}{\partial y_2} = -L'_{1,2}(x_1) \cdot \alpha_1^z \\
J_{1,2} &= \frac{\partial R_1^y}{\partial z_1} = -L'_{1,1}(x_1) \cdot \alpha_1^y - \delta_1^y \\
J_{2,1} &= \frac{\partial R_1^z}{\partial y_1} = -L'_{1,1}(x_1) \cdot \alpha_1^z - \delta_1^z
\end{aligned}
$$

and

$$
\begin{aligned}
J_{2n-1,2n-3} &= \frac{\partial R_n^y}{\partial y_{n-1}} = L''_{n,n-1}(x_n) - L'_{n,n-1}(x_n) \cdot \beta_n^y \\
J_{2n+0,2n-2} &= \frac{\partial R_n^z}{\partial z_{n-1}} = L''_{n,n-1}(x_n) - L'_{n,n-1}(x_n) \cdot \beta_n^z \\
J_{2n-1,2n-1} &= \frac{\partial R_n^y}{\partial y_n} = L''_{n,n}(x_n) - L'_{n,n}(x_n) \cdot \beta_n^y - \gamma_n^y \\
J_{2n+0,2n+0} &= \frac{\partial R_n^z}{\partial z_n} = L''_{n,n}(x_n) - L'_{n,n}(x_n) \cdot \beta_n^z - \gamma_n^z \\
J_{2n-1,2n+0} &= \frac{\partial R_n^y}{\partial z_n} = -L'_{n,n}(x_n) \cdot \alpha_n^y - \delta_n^y \\
J_{2n+0,2n-1} &= \frac{\partial R_n^z}{\partial y_n} = -L'_{n,n}(x_n) \cdot \alpha_n^z - \delta_n^z \\
J_{2n-1,2n-2} &= \frac{\partial R_n^y}{\partial z_{n-1}} = -L'_{n,n-1}(x_n) \cdot \alpha_n^y \\
J_{2n+0,2n-3} &= \frac{\partial R_n^z}{\partial y_{n-1}} = -L'_{n,n-1}(x_n) \cdot \alpha_n^z
\end{aligned}
$$

with coefficients according to

$$
\begin{aligned}
\alpha_j^y &= 2\left[\frac{n_y(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)}{n(x_j,y_j,z_j)}\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)y_k\right)\right]\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)z_k\right)\\
\alpha_j^z &= 2\left[\frac{n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)}{n(x_j,y_j,z_j)}\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)z_k\right)\right]\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)y_k\right)\\
\beta_j^y &= 2\left[\frac{n_y(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)}{n(x_j,y_j,z_j)}\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)y_k\right)\right]\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)y_k\right)\\
&\quad -\frac{n_x(x_j,y_j,z_j)}{n(x_j,y_j,z_j)}\cdot\sigma_j\\
\beta_j^z &= 2\left[\frac{n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)}{n(x_j,y_j,z_j)}\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)z_k\right)\right]\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)z_k\right)\\
&\quad -\frac{n_x(x_j,y_j,z_j)}{n(x_j,y_j,z_j)}\cdot\sigma_j\\
\gamma_j^y &= \left[\frac{n_{yy}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_y(x_j,y_j,z_j)n_y(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j -\\
&\quad \left[\frac{n_{xy}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)n_y(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)y_k\right)\\
\gamma_j^z &= \left[\frac{n_{zz}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_z(x_j,y_j,z_j)n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j -\\
&\quad \left[\frac{n_{xz}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)z_k\right)\\
\delta_j^y &= \left[\frac{n_{yz}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_y(x_j,y_j,z_j)n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j -\\
&\quad \left[\frac{n_{xz}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)y_k\right)\\
\delta_j^z &= \left[\frac{n_{yz}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_y(x_j,y_j,z_j)n_z(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j -\\
&\quad \left[\frac{n_{xy}(x_j,y_j,z_j)}{n(x_j,y_j,z_j)} - \frac{n_x(x_j,y_j,z_j)n_y(x_j,y_j,z_j)}{n(x_j,y_j,z_j)n(x_j,y_j,z_j)}\right]\cdot\sigma_j\cdot\left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j)z_k\right)
\end{aligned}
$$

and

$$\sigma_j = \left[1 + \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) y_k\right)^2 + \left(\sum_{k=j-1}^{j+1} L'_{j,k}(x_j) z_k\right)^2\right]$$

for $j = 1, ..., n$.

B The bicubic interpolation

In matrix notation, the general rule for an increment $q \in [0, 1]$ for one dimension reads as (Keys, 1981)

$$\Gamma(q) = 0.5 \cdot \begin{pmatrix} 1 & q & q^2 & q^3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 2 & -5 & 4 & -1 \\ -1 & 3 & -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

where a_k for $k = 0, ..., 3$ denote the function values at a uniform sequence of nodes and Γ denotes the interpolant. For two dimensions, the general rule is applied in I direction with the increment X according to

$$b_k = \Gamma(X; a_0^k, a_1^k, a_2^k, a_3^k)$$

for $k = 0, ..., 3$ and then again in J direction with the increment Y according to

$$N = \Gamma(Y; b_0, b_1, b_2, b_3)$$

Since bicubic interpolation involves 16 grid points, the integers I and J are limited by the grid point indices next to the lateral boundaries according to:

$$\begin{aligned} I &= \min(\max(I, 2), M_I - 3) \\ J &= \min(\max(J, 2), M_J - 3) \end{aligned}$$

C The Legendre-Gauss-Lobatto nodes

The Legendre-Gauss-Lobatto nodes are defined on the interval $[-1, 1]$. The endpoints of the interval $[-1, 1]$ are included in a total of m abscissas, giving $m - 2$ free abscissas. Abscissas are symmetrical about the origin. The free abscissas ζ_i for $i = 2, ..., m - 1$ are the roots of the polynomial $P'_{m-1}(\zeta)$ where $P_m(\zeta)$ is a Legendre polynomial

$$P_m(\zeta) = \frac{1}{2^m m!} \frac{d^m}{d\zeta^m} (\zeta^2 - 1)^m \tag{88}$$

No explicit formula of the nodes is known. However, they can be computed numerically. With the weights of the free abscissas

$$\kappa_i = \frac{2}{m(m-1)[P_{m-1}(\zeta_i)]^2} \tag{89}$$

for $i = 2, ..., m - 1$ and of the endpoints

$$\kappa_1 = \kappa_m = \frac{2}{m(m-1)} \tag{90}$$

the quadrature rule reads as (Abramowitz and Stegun, 1972)

$$\int_{-1}^{1} g(\zeta) d\zeta = \kappa_1 g(-1) + \kappa_m g(1) + \sum_{i=2}^{m-1} \kappa_i g(\zeta_i) \tag{91}$$

where $g(\zeta)$ denotes a function defind on the interval $[-1, 1]$. Before applying the quadrature rule the integral over $[a, b]$ must be changed into an integral over $[-1, 1]$. This change of interval is done in the following way:

$$\int_a^b g(\zeta) d\zeta = \frac{b-a}{2} \int_{-1}^{1} g\left(\frac{b-a}{2}\zeta + \frac{a+b}{2}\right) d\zeta \tag{92}$$

After applying the quadrature rule, the following approximation is obtained:

$$\int_a^b g(\zeta) d\zeta = \frac{b-a}{2} \sum_{i=1}^{m} \kappa_i g\left(\frac{b-a}{2}\zeta_i + \frac{a+b}{2}\right) \tag{93}$$

Therefore, the non-uniform sequence of nodes for which the solution in the ray tracing is sought is given by

$$x_i = \frac{b-a}{2}\zeta_i + \frac{a+b}{2} \tag{94}$$

for $i = 1, ..., m$. The weights are given by

$$w_i = \frac{b-a}{2}\kappa_i \tag{95}$$

for $i = 1, ..., m$.

D Statistical quantities

In order to measure the discrepancy between two sets of data a_i and b_i for $i = 1, ..., n$ in a statistical sense, the following quantities are introduced:

The Mean Deviation MD is defined as

$$\text{MD} = \frac{1}{n}\sum_{i=1}^{n}(a_i - b_i) \tag{96}$$

The Normalized Deviation ND is defined as

$$\text{ND} = \frac{2}{n}\sum_{i=1}^{n}\frac{a_i - b_i}{a_i + b_i} \tag{97}$$

The Mean Absolute Deviation MAD is defined as

$$\text{MAD} = \frac{1}{n}\sum_{i=1}^{n}|a_i - b_i| \tag{98}$$

The Root Mean Square RMS is defined as

$$\text{RMS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - b_i)^2}$$

The correlation coefficient, also known as the Product-Moment Coefficient of Correlation PMCC, is defined as

$$\text{PMCC} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{a_i - \bar{a}}{s_a}\right)\left(\frac{b_i - \bar{b}}{s_b}\right)$$

where

$$\begin{aligned} \overline{a} &= \frac{1}{n}\sum_{i=1}^{n} a_i \\ \overline{b} &= \frac{1}{n}\sum_{i=1}^{n} b_i \end{aligned}$$

denote the mean and

$$\begin{aligned} s_a &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(a_i - \overline{a})^2} \\ s_b &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(b_i - \overline{b})^2} \end{aligned}$$

denote the standard deviation.

Categorical forecasts are verified by applying 2×2 contingency tables

	Observed	
Forecast	Yes	No
Yes	a	b
No	c	d

where a denotes the number of hits, b denotes the number of false alarms, c denotes the number of misses and d denotes the number of correct negatives. A number of categorical statistics can be computed from the elements in the contingency table to describe particular aspects of the forecast performance:

The Probability of Detection POD is defined as

$$\text{POD} = \frac{a}{a+c} \tag{99}$$

The False Alarm Rate FAR is defined as

$$\text{FAR} = \frac{b}{b+a} \tag{100}$$

The True Skill Score TSS is defined as

$$\text{TSS} = \frac{a}{a+c} - \frac{b}{b+a} \tag{101}$$

The Equitable Threat Score ETS is defined as

$$\text{ETS} = \frac{a-e}{a+b+c-e} \tag{102}$$

where

$$e = \frac{(a+b)(a+c)}{a+b+c+d}$$

E Acronyms

3DVAR	Three-dimensional variational assimilation
4DVAR	Four-dimensional variational assimilation
ADJ	Adjoint model
BFGS	Broyden-Fletcher-Goldfarb-Shanno algorithm
BmO	Background minus Observation
BVP	Boundary Value Problem
COPS	Convective and Orographically-induced Precipitation Study
CPU	Central Processing Unit
DFG	Deutsche Forschungsgemeinschaft
DKRZ	Deutsches Klimarechenzentrum
D-PHASE	Demonstration of Probabilistic Hydrological and Atmospheric Simulation of flood Events in the alpine region
DWD	German Weather Service
ECMWF	European Centre for Medium Range Weather Forecasts
EUMETNET	The Network of European Meteorological Services
EPOS	Earth Parameters and Orbit determination System
E-GVAP	EUMETNET GPS water vapour programme
ETS	Equitable Threat Score
FAR	False Alarm Rate
FTP	File Transfer Protocol
GFZ	German Research Center for Geosciences
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HASL	Height Above Sea Level
IFS	Integrated Forecast System
IGS	International GNSS Service
IPM	Institut für Physik und Meteorologie
IVP	Initial Value Problem
IWV	Integrated Water Vapor
L-BFGS	Limited Memory BFGS
LEO	Low Earth Orbit
LM	Local Model
LW	Long Wave
MD	Mean Deviation
MAD	Mean Absolute Deviation
MARS	Meteorological Archival and Retrieval System
MATLAB	Matrix Laboratory

MM5	Mesoscale Model 5
MM5v1	MM5 version 1
MM5v2	MM5 version 2
MM5v3	MM5 version 3
MPI	Message Passing Interface
MRF	Medium Range Forecast
ND	Normalized Deviation
NLM	Non-Linear Model
NWP	Numerical Weather Prediction
ODE	Ordinary Differential Equation
PBL	Planetary Boundary Layer
PMCC	Product-Moment Coefficient of Correlation
POD	Probability of Detection
PPP	Precise Point Positioning
PW	Precipitable Water
QPF	Quantitative Precipitation Forecasting
RO	Radio Occulation
RMS	Root Mean Square
RRTM	Rapid Radiative Transfer Model
RUC	Rapid Update Cycle
SW	Short Wave
SWV	Slant Water Vapor
TLM	Tangent-Linear Model
TSS	True Skill Score
UK	United Kingdom
UTC	Universal Time Coordinated
WRF	Weather Research and Forecasting Model
WV	Water Vapor
WWRP	World Weather Research Program
ZTD	Zenith Total Delay

References

Abramowitz, M. and I. A. Stegun, 1972: *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, 9th printing.* Dover, New York.

Alber, C., R. Ware, C. Rocken and J. Braun, 2000: Obtaining single path phase delays from GPS double differences. *Geophys. Res. Lett*, **27**, 2661–2664.

Anthes, R. A. and T. T. Warner, 1978: Development of hydrodynamic models suitable for air pollution and other mesometeorological studies. *Mon. Wea. Rev.*, **106**, 1045–1078.

Ascher, U. M., R. M. Mattheij and R. D. Russell, 1995: *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations.* SIAM Classics in Applied Mathematics 13, SIAM, Philadelphia.

Bender, M., G. Dick, J. Wickert, M. Ramatschi, G. Gendt, M. Ge, M. Rothacher, A. Raabe and G. Tetzlaff, 2009: Estimates of the information provided by GPS slant data observed in Germany regarding tomographic applications. *J. Geophys. Res.*, **114**, D06303.

Bender, M., G. Dick, J. Wickert, T. Schmidt, S. Song, G. Gendt, M. Ge and M. Rothacher, 2008: Validation of GPS slant delays using water vapour radiometers and weather models. *Meteorologische Zeitschrift*, **17**, 807–812.

Bevis, M., S. Businger, T. Herring, C. Rocken, R. Anthes and R. Ware, 1992: GPS meteorology: Remote sensing of atmospheric water vapor using the Global Positioning System. *J. Geophys. Res.*, **97**, 15787–15801.

Bevis, M., S. Businger, T. Herring, C. Rocken, R. Anthes and R. Ware, 1994: GPS meteorology: Mapping zenith wet delays onto precipitable water. *J. Appl. Meteorol.*, **33**, 379–1386.

Boehm, J., A. Niell, P. Tregoning and H. Schuh, 2006: Global Mapping Function (GMF): A new empirical mapping function based on numerical weather model data. *Geophys. Res. Lett*, **33**, L07304,doi:10.1029/2005GL025546.

Bouttier, F. and P. Courtier, 1999: Data assimilation concepts and methods. Technical report, Meteorological Training Course Lecture Series, ECMWF, European Center for Medium Range Weather Forecasts, Reading, UK.

Braun, J., C. Rocken and R. Ware, 2001: Validation of single slant water vapor measurements with GPS. *Radio Science*, **36**, 459–472.

Crook, N., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785.

Davis, J. L., T. A. Herring, I. I. Shapiro, A. E. Rogers and G. Elgered, 1985: Geodesy by radio interferometry: Effects of the atmospheric modelling errors on estimation of baseline length. *Radio Sci.*, **20**, 1593–1607.

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

Elgered, G., J. Davis, T. Herring and I. Shapiro, 1991: Geodesy by radio interferometry: Water vapor radiometry for estimation of the wet delay. *J. Geophys. Res.*, **96**, 6541–6555.

Eresmaa, R., S. Healy, H. Järvinen and K. Salonen, 2008: Implementation of ray-tracing operator for ground-based GPS slant delay observation modelling. *J. Geophys. Res.*, **113**, D11114 DOI:10.1029/2007JD009256.

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

Gendt, G., G. Dick and W. Söhne, 1999a: GFZ Analysis Center of IGS-Annual Report 1998. Technical report, edited by K. Gowey, R. Neilan and A. Moore, IGS Central Bureau, Pasadena, California, 1999.

Gendt, G., P. Fang and J. F. Zumberge, 1999b: Moving IGS products towards real-time. Technical report, edited by K. Gowey, R. Neilan and A. Moore, IGS Central Bureau Jet Propulsion Laboratory, Pasadena, California, 1999.

Giering, R. and T. Kaminski, 1998: Recipes for adjoint code construction. *ACM Transactions on Mathematical Software*, **24**, 4, 437–474.

Grell, G. A., J. Dudhia and R. Stauffer, 1995: A description of the fith-generation Penn State/NCAR mesoscale model (MM5). *Technical Note TN-398+STR.*

Grzeschik, M., H. S. Bauer, V. Wulfmeyer, D. Engelbart, U. Wandinger, I. Mattis, D. Althausen, R. Engelmann, M. Tesche and A. Riede, 2008: Four-dimensional Variational Data Analysis of Water Vapor Raman Lidar Data and Their Impact on Mesoscale Forecasts. *J. Atmos. Oceanic Technol.*, **25**, 1437–1453.

Ha, S., Y.-H. Kuo, Y.-R. Guo and G.-H. Lim, 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**, 11, 2635–2655.

Hobiger, T., R. Ichikawa, Y. Koyama and T. Kondo, 2008: Fast and accurate ray-tracing algorithms for real-time space geodetic applications using numerical weather models. *J. Geophys. Res.*, **113**, D20302 DOI:10.1029/2008JD010503.

Hofmann-Wellenhof, B., H. Lichtenegger and J. Collins, 1992: *GPS: Theory and Practice*. Springer-Verlag.

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

Järvinen, H., R. Eresmaa, H. Vedel, K. Salonen, S. Niemela and J. deVries, 2007: A variational data assimilation system for ground-based GPS slant delays. *Quart. J. Roy. Meteorol. Soc.*, **133**, 969–980.

Keys, R., 1981: Cubic convolution interpolation for digital processing. *Acoustics, Speech and Signal Processing*, **29**, 1153–1160.

Kravtsov, Y. A. and Y. I. Orlov, 1990: *Geometrical Optics of Inhomogeneous Media*. Springer-Verlag.

Kursinski, E. R., G. A. Hajj, W. I. Bertiger, S. S. Leroy, T. K. Meehan, L. J. Romans, J. T. Schofield, D. J. McCleese, W. G. Melbourne, C. L. Thornton, T. P. Yunck, J. R. Eyre and R. N. Nagatani, 1996: Initial results of radio occultation observations of earth's atmosphere using the Global Positioning System. *Science*, **271**, 1107–1110.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res.*, **102**, 23429–23465.

Lemoine, F. G., S. C. Kenyon, J. K. Factor, R. G. Trimmer, N. K. Pavlis, D. S. Chinn, C. M. Cox, S. M. Klosko, S. B. Luthcke, M. H. Torrence, Y. M. Wang, R. G. Williamson, E. C. Pavlis, R. H. Rapp and T. R. Olson, 1998: The Development of the Joint NASA GSFC and NIMA Geopotential Model EGM96. Technical report, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, 1998.

Li, J. G. and B. W. Atkinson, 1999: Transition regimes in valley airflows. *Bound.-Layer Meteor.*, **91**, 385–411.

Liu, D. C. and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, **45**, 503–528.

Lowry, A. R., C. Rocken, S. V. Sokolovskiy and K. D. Anderson, 2002: Vertical profiling of atmospheric refractivity from ground-based GPS. *Radio Sci.*, **37**. DOI:10.1029/2000RS002565.

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

Marini, J. W., 1972: Correction of satellite tracking data for an arbitrary tropospheric profile. *Radio Sci.*, **7**, 223–231.

Melbourne, W. G., E. Davis, C. Duncan, G. A. Hajj, K. Hardy, E. Kursinski, T. Meechan and L. Young, 1994: The application of spaceborn GPS to atmospheric limb sounding and global change monitoring. Technical report, Publication 94-18, Jet Propulsion Laboratory, Pasadena, California, 1994.

Niell, A. E., 1996: Global mapping functions for the atmosphere delay at radio wavelengths. *J. Geophys. Res.*, **101**, 3227–3246.

Persson, A. and F. Grazzini, 2005: User Guide to ECMWF forecast products. Technical report, European Centre for Medium Range Weather forecasting, Reading, UK.

Poli, P., P. Moll, F. Rabier, G. Desroziers, B. Chapnik, L. Berre, S. B. Healy, E. Andersson and F.-Z. E. Guelai, 2007: Forecast impact studies of zenith total delay data from European near real-time GPS stations in Meteo France 4DVAR. *J. Geophys. Res.*, **112**, 1–16.

Rocken, C., S. V. Sokolovskiy, J. M. Johnson and D. Hunt, 2001: Improved Mapping of Tropospheric Delays. *J. Atmos. Oceanic Technol.*, **18**, 1205–1213.

Rotach, M., M. Arpagaus, M. Dorninger, C. Hegg, J. Frick, A. Montani, R. Ranzi, F. Bouttier, A. Buzzi, G. Frustaci, K. Mylne, E. Richard, A. Rossa, C. Schär, M. Staudinger, H. Volkert, V. Wulfmeyer, H.-S. Bauer, F. Ament, M. Denhard, F. Fundel, U. Germann and M. Stoll, 2009: MAP D-PHASE: Real-time Demonstration of Weather Forecast Quality in the Alpine Region. *Bull. Amer. Meteorol. Soc.* Submitted for publication.

Ruggiero, F. H., J. G. Michalakes, T. Nehrkorn, G. D. Modica and X. Zou, 2006: Development and Tests of a New Distributed-Memory MM5 Adjoint. *J. Atmos. Oceanic Technol.*, **23**, 424–436.

Ruggiero, F. H., G. D. Modica, T. Nehrkorn, M. Cerniglia, J. G. Michalakes and X. Zou, 2002: Beta test code documentation for CWO-5. Technical report, Air Force Research Laboratory.

Saastamoinen, J., 1972: *The use of artificial satellites for geodesy, Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites*, volume 15 of *Geophys. Monogr. Ser.*, chapter Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites, pp. 247–251. American Geophysical Union.

Shampine, L. F. and N. W. Reichelt, 1997: The MATLAB ODE suite. *SIAM Journal of Scientific Computing*, **18**, 1–22.

Sokolovskiy, S. V., C. Rocken and A. Lowry, 2001: Use of GPS for estimation of bending angles of radio waves at low elevations. *Radio Sci.*, **36**, 473–482.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller, 2007: Climate change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Technical report, Cambridge University Press, Cambridge, UK.

Standard Atmosphere, U. S., 1976: U.S. Standard Atmosphere 1976. *U.S. Government Printing Office, Washington, D.C.*

Vedel, H. and X.-Y. Huang, 2004: Impact of Ground Based GPS Data on Numerical Weather Prediction. *J. Meteorol. Soc. Japan*, **82**, 1B, 459–472.

Ware, R., C. Alber, C. Rocken and F. Solheim, 1997: Sensing integrated water vapor along GPS ray paths. *Geophys. Res. Lett.*, **24**, 417–420.

Weckwerth, T., 2000: The effect of small-scale moisture variability on thunderstorm initiation. *Mon. Wea. Rev.*, **128**, 4017–4030.

Wulfmeyer, V., H.-S. Bauer, M. Grzeschik, A. Behrendt, F. Vandenberghe, E. Browell, S. Ismail and R. Ferrare, 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., A. Behrendt, H. S. Bauer, C. Kottmeier, U. Corsmeier, A. Blyth, G. Craig, U. Schumann, M. Hagen, S. Crewell, P. D. Girolamo, C. Flamant, M. Miller, A. Montani, S. Mobbs, E. Richard, M. Rotach, M. Arpagaus, H. Russchenberg, P. Schlüssel, M. König, V. Gärtner, R. Steinacker, M. Dorninger, D. Turner, T. Weckwerth, A. Hense and C. Simmer, 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. Meteorol. Soc.* DOI:10.1175/2008BAMS2367.1.

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.

Zängl, G., 2004: The sensitivity of simulated orographic precipitation to model components other than cloud microphysics. *Quart. J. Roy. Meteorol. Soc.*, **130**, 1857–1875.

Zou, X., F. Vandenberghe, M. Pondeca and Y.-H. Kuo, 1997: Introduction to adjoint techniques and the MM5 adjoint modeling system. Technical report, NCAR Tech. Note NCAR/TN-435-STR, NCAR, P.O. Box 3000, Boulder, CO, 80307.

Zumberge, J. F., M. B. Heflin, D. C. Jefferson, M. M. Watkins and F. H. Webb, 1997: Precise point positioning for the efficient and robust analysis of GPS data from large networks. *J. Geophys. Res.*, **102**, 5005–5018.

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

Acknowledgements

I would like to express my profound respect and gratitude to my principal supervisor Prof. Dr. Volker Wulfmeyer, Institute of Physics and Meteorology, University of Hohenheim, for the help, support and patience.

I take this opportunity to thank my co-advisor Dr. Hans-Stefan Bauer, Institute of Physics and Meteorology, University of Hohenheim, for continuous and extensive support throughout the last years.

I thank Dr. Galina Dick, GFZ Potsdam, for not only providing the GPS data but also for the continious support. I am thankful to Dr. Michael Bender, GFZ Potsdam, for the highly valuable tips regarding the GPS data and the fruitfull discussions on the numerical simulation of phase delays.

I am thankful to the Deutsches Klimarechenzentrum GmbH (DKRZ) in Hamburg and the ECMWF for the computing facilities and the Deutsche Forschungsgemeinschaft (DFG) for the funding in the framework of the priority program SPP1176 Quantitative Precipitation Forecast.

I am thankful to my colleagues Matthias Grzeschik and Kirsten Warach-Sagi for the discussions about assimilation, Thomas Schwitalla for discussions about the MM5, and Marcus Radlach and Sandip Pal for the discussions about meteorology.

I would like to thank my colleges Marcus Radlach, Sandip Pal, Maxim Schiller and Matthias Grzeschik for the nice evenings with Indian kitchen.

Finally, I would like to thank Tina, Mina, Eric, my sister, my brothers and my parents for continuous support.

Erklärung:

Ich erkläre, daß ich die Arbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe und daß alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, durch die Angabe von Quellen als Entlehnungen kenntlich gemacht worden sind.

Florian Zus
Stuttgart, den 23. Juli 2009