The use of Global Navigation Satellite Systems (GNSS) signals in Numerical Weather Prediction (NWP)

Doctoral Thesis

by

Lídia Cucurull Molina

Advisor: Dr. Antonio Rius (CSIC)

co-Advisor: Dr. Angel Redaño (UB)

Departament d'Astronomia i Meteorologia
Facultat de Física
Universitat de Barcelona
Insitut d'Estudis Espacials de Catalunya, IEEC-CSIC

Barcelona, May 2001

To my mother

To Marcel

Acknowledgements

This work is the result of many years of working and has never been possible without the help and collaboration of many people.

First of all, I would like to acknowledge Antoni Rius for guiding me along this GPS world and for his perseverance in approaching me to the computer technology. He gave me the opportunity to do research and to be involved in an international environment. The door of his office has been open all the time for discussion on many topics and for learning a piece of meteorology both together.

I am also grateful to my co-advisor at the Universitat de Barcelona (UB, Spain) Angel Redaño for his help on the meteorological part of this research work.

I also thank François Vandenberghe (NCAR, U.S.), Jordi Vilà (Wageningen University, The Netherlands), Beatriz Navascués [Instituto Nacional de Meterología (INM), Spain], Pedro Elósegui (Harvard-Smithsonian Center for Astrophysics, U.S.), and Ingrid Bárcena [Centre de Supercomputació de Catalunya (CESCA), Spain] for helping me on several aspects of this research study and for their patience in answering me all my questions and doubts.

The IEEC Earth Sciences Department for being my friends during all this time and for preprocessing the GPS data: Alex, Ana, David, Dirk, Estel, Giulio, Josep Maria, Pepa, and of course Josep, who spent a lot of time in solving my problems with the computers. The administrative team Anna, Pilar, Mireia, and Eva for making life easier.

Thanks also to Manuel de Pondeca (Florida State University, U.S.) for kindly providing me the MM5 4D-VAR code, and to The Met. Office (UK) for the 1D-VAR algorithm. I must also acknowledge José Antonio García-Moya and Bill Kuo for hosting me at the INM and NCAR, respectively. Josep Vidal (UB) for helping me with the comparison with the MASS model, Joan Bech (Servei de Meteorologia de Catalunya, Spain) for giving a lot of meteorological data used in this thesis, and Wei Wang (NCAR) for kindly answering all my questions on the MM5 model.

The GPS data used for the analyses was gathered from publicly available servers,

maintained by organizations like IGS, CODE, JPL and the meteorological analyses and topography data sets were obtained from ECMWF and USGS, respectively. I thank NCAR and Caltech/JPL for providing the MM5 model and GIPSY/OASIS–II package, respectively. The Insituto Geográfico Nacional (IGN, Spain) loaned and operated the GPS receivers used in Section (4.1). The GPS data analyzed in Sections (4.2) and (6.2) were provided by the Institut Cartogràfric de Catalunya (ICC, Spain). I am also grateful to CESCA for the computer facilities.

Thanks a lot to my family specially to my mother Isabel for encouraging me since the very beginning and to my dearest Marcel for his support and comprehension during my long working hours.

This research work was partially supported by EC Grant WAVEFRONT PL-952007, MAGIC PL-972065, CLIMAP ENV4-CT97-0387, and the Spanish Climate CICYT grant CLI95-1781.

Contents

Chapter 1

Motivation

The aim of this work is to use precise and continuous measurements of GPS-derived observations together with Numerical Weather Prediction (NWP) models. Along this thesis it has been attempted to present what GPS provides to the meteorological community and how these data can be used to improve the prediction of moisture and precipitation fields.

The water vapor distribution and content are critical variables for the description of the state and evolution of many physical processes in the Earth's atmosphere. Although water vapor constitutes only a small fraction of all the atmospheric gases (fractional volume 0-4 percent), its importance is far greater than this measure would indicate. For example, water vapor plays an important role in atmospheric processes that range, in spatial scales, from micro to global meteorology (Rowland and Isaksen 1988). In addition, water vapor is a greenhouse gas and long-term variations in its total global content could potentially be used as an indicator of global climate change (Yuan et al. 1993). However, the distribution of water vapor is a highly variable function of both time and space and correlates poorly with surface humidity measurements. Lack of precise and continuous water vapor data is one of the major error sources in short-term forecasts of precipitation (Kuo et al. 1993, 1996). Although ground-based techniques such as radiosondes or water vapor radiometers (WVRs) are sensitive to the water vapor content present in the atmosphere, they can be expensive to operate and they provide either poor temporal resolution, poor spatial coverage, or both. Radiosondes are launched typically only once every twelve hours and they are sparse over wide areas in the globe; in contrast to space-based WVR, ground-based WVR have good temporal resolution but poor spatial coverage. New observational techniques that are sensitive to the spatial and temporal distribution

of the water vapor content in the atmosphere have made now possible the retrieval of precise and continuous estimates of water vapor with high spatial density. This is the case of the Global Positioning System (GPS). Even though GPS was originally designed for military navigation and positioning, the applications of this technique already abound in areas such as geodesy, volcanology, oceanography, or glaciology to cite a few. See, e.g., Herring (1999), and references therein, for a review of some applications for geodesy. For an overview of previous studies on the applications of the GPS technique carried out in our group, we refer to Rius et al. (1997); Cucurull et al. (1998); Ruffini et al. (1998), (1999b); Flores (1999), Behrend et al. (2000); Flores et al. (2000a), (2001).

The structure of the present thesis is as follows:

- Chapter (2) briefly reviews some general aspects of the NWP systems and how these models assimilate the observational data sets.
- Chapter (3) summarizes the time delay effects on GPS electromagnetic waves traveling through the atmosphere (excluding the ionosphere), and how the atmospheric water vapor content can be determined through the modeling of these effects.
- Chapter (4) contains the comparison of the GPS-derived observations with the values obtained from three different NWP models and for different meteorological situations.
- Chapter (5) presents a simple approach of the variational assimilation of GPS data and surface pressure measurements, in which a 1-dim algorithm performs a local analysis of one atmospheric column.
- Chapter (6) describes the global 4-dim variational assimilation of GPS-derived observations into a NWP model. Two different cases are analyzed depending on the type of variables and the number of observations to be assimilated.
- Chapter (7) concludes the thesis summarizing the main results of this study and finally discussing the future lines of research.

Chapter 2

Introduction to NWP models

2.1 Basics of meteorological models

We assume in this work that the state of the atmosphere can be described by $\mathbf{x}(\mathbf{r}, t)$ for a given position $\mathbf{r}$ and time t,

$$\mathbf{x}(\mathbf{r}, t) = \mathbf{x}[\mathbf{v}(\mathbf{r}, t), T(\mathbf{r}, t), p(\mathbf{r}, t), \rho(\mathbf{r}, t)] \tag{2.1}$$

where $\mathbf{v}$ is the wind, T the temperature, p the pressure, and ρ the density. These fields are not completely independent since they have to verify some constraints. The pressure, temperature, and density are related by the equation of state for perfect gases,

$$\rho = \frac{M_r p}{RT} \tag{2.2}$$

where R is the gas constant per mole and M_r is the molecular weight of the perfect gas.

The evolution of the components of the state vector $\mathbf{x}(\mathbf{r}, t)$ are related by the fundamental laws of fluids mechanics and thermodynamics expressed in terms of partial differential equations. These governing laws are the conservation of mass (continuity equation), the conservation of momentum, and the conservation of energy (first law of thermodynamics). We next briefly describe these conservation laws. See, e.g., Holton (1992) for a detailed description of these equations.

- **conservation of mass:**

  The conservation of mass (or continuity equation) can be written as follows:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{v}) \tag{2.3}$$

where $-\nabla \cdot (\rho \mathbf{v})$ is the mass inflow per unit volume. This equation states that the rate of increase in mass per unit volume is equal to the rate of change of density. In the expression above, we have made use of the relationship $\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla$

- **conservation of momentum:**

  The momentum equations come from the Newton's second law, where the motion in an inertial reference frame is now referenced in a frame rotating with the Earth. If we assume that the only forces acting on the atmosphere are the pressure gradient force, gravitation, and friction, these equations can be written as follows:

$$\frac{D\mathbf{v}}{Dt} = -2\mathbf{\Omega} \wedge \mathbf{v} - \frac{1}{\rho}\nabla p + \mathbf{g} + \mathbf{F}_r \tag{2.4}$$

  where $\mathbf{\Omega}$ is the angular velocity of the Earth, and $\mathbf{F}_r$ is the frictional force.

- **conservation of energy:**

  The first law of thermodynamics relates the rate of addition of heat energy Q, the internal energy E and the work W done by the system on its surroundings when expanding for a system which is in thermodynamic equilibrium. The mass of fluid of a control volume in the atmosphere is not in thermal equilibrium. However, if we consider that the internal energy of the fluid is composed of the internal kinetic energy of the particles and the kinetic energy due to the motion of the fluid, the first law of thermodynamics still applies and can be written as follows:

$$\frac{DE}{Dt} = \frac{DQ}{Dt} + \frac{DW}{Dt} \tag{2.5}$$

  Equation (2.5) states that the rate of change of thermodynamical energy is equal to the rate of heating plus the rate at which work is done on the fluid parcel by external forces. Considering the external forces that do work, expanding and working with equation (2.5), and substrating the terms involved in the mechanical energy equation, we obtain the usual form of the thermal energy equation:

$$Q = c_v \frac{DT}{Dt} + p\frac{D\alpha}{Dt} \tag{2.6}$$

where c_v is the specific heat at constant volume, and α is the specific volume ($\alpha = 1/\rho$).

Since vertical motion in the atmosphere is generally very small, in some cases it is a useful approximation to assume that there is a balance between the pressure gradient and gravity forces given by

$$\Delta p = -\rho g \Delta z \tag{2.7}$$

where g is the acceleration of gravity and z is the height variable. (Here, ρ is the density of dry air). Note that Δp is negative since pressure decreases with altitude. The hydrostatic equation (2.7) provides a good approximation for the vertical dependence of the pressure variable. Assuming a constant value for g and making use of the equation of state for an ideal gas (2.2), the integration of (2.7) gives

$$p(z) = p(0)\exp(-z/h) \tag{2.8}$$

where the scale height h is given by $h = RT/M_r g$, and is the increase in altitude necessary to reduce the pressure by a factor e. The value of pressure variable at $z = 0$ is given by $p(0)$.

The conservation equations together with the equation of state for perfect gases and the hydrostatic approximation are known collectively as the primitive equations. A NWP model consists of these equations together with initial conditions (i.e. the model state at the initial time), and boundary condition (i.e. the condition that the model needs to satisfy at the boundary of the domain) if the model domain does not cover all the Earth. This model is supposed to reflect the evolution of the state of the atmosphere and has usually the form of a differential equation. We can write such a model F in a general continuous form

$$\frac{\partial \mathbf{x}(\mathbf{r}, t)}{\partial t} = F[\mathbf{x}(\mathbf{r}, t)] \tag{2.9}$$

$$\mathbf{x}(\mathbf{r}, t)|_{t=t_0} = \mathbf{x}_0(\mathbf{r}) \tag{2.10}$$

$$\mathbf{x}(\mathbf{r}, t)|_{\Gamma} = \mathbf{y}(t) \tag{2.11}$$

where $\mathbf{x}_0(\mathbf{r})$ and $\mathbf{y}(t)$ represent the initial condition at time t_0, and the lateral boundary condition at the boundary of the domain Γ, respectively.

However, the equations that drive the evolution of the state of the atmosphere and which are described in (2.9), (2.10), and (2.11) in a generic form, are non-linear and can only be solved by using numerical techniques. Furthermore, the non-linearity of the equations means that any small error in the initial conditions are likely to grow making the solution useless in a deterministic sense after a number of days. Therefore, these continuous equations are solved by discretization in space (providing a grid size) and time (providing a time step).

Given an adequate description of the initial conditions, NWP models can satisfactorily predict the atmospheric state for periods ahead of hours up to a few days. For longer periods deficiencies in the models become apparent. The most fundamental of these is the parameterization of sub grid scale motions due to the discretization in space of the model. Even when the grid size is reduced, there will be smaller scale motion to be considered which, because of the lack of a comprehensive theory describing the interactions between different scales which always occur in turbulent motions, can only be dealt with in a crude empirical manner.

If the horizontal grid size of the model is larger than around 10 km, we can use the hydrostatic approximation (2.7). With this approximation, the vertical acceleration is neglected when comparing with gravitational acceleration. The models that use this approximation are known as hydrostatic models (e.g. the version of HIRLAM and MASS models which are described in this thesis). However, when the representation of a smaller-scale phenomena is required, finer grid size is needed and the hydrostatic approximation can not be used. This is for instance the case of the MM5 model, which is a non-hydrostatic model. In general, the higher the resolution, the more accurate the model. This is because the orographic influences are particularly strong in some areas and non-hydrostatic models tend to resolve the orographic effects reasonable well provided they work at sufficient resolution. Since the non-hydrostatic models allow more localized studies, they also can solve the deep convection effect better than hydrostatic models (Dudhia et al. 1993). The impact of the grid resolution on the comparison between the MM5 model and the GPS-derived observations is analyzed in Section (4.2).

There are different methods which are used in the solution of the differential equations of motions of the atmosphere. We should distinguish between two broad categories: finite differences (e.g. MM5 and MASS models), and spectral (e.g. HIRLAM

model).

- *finite differences*: values and derivatives of the variables are represented at discrete points on a grid. Usually the grid is regular in the horizontal coordinate and most models have a non-uniform grid in the vertical direction to have a finer resolution in the boundary layer. See, e.g., Arakawa (1966), Lilly (1964), Sadourny (1975), and Arakawa (1984) for a detailed description of this technique.

- *spectral method*: the horizontal representation is in terms of truncated series of spherical harmonic functions, whose variation is described by sines and cosines in the east/west direction and by associated Legendre functions in the north/south direction. See, e.g., Machenhauer (1979), and Jarrau and Simmons (1984) for a description of this method.

Two different types of models are in use for NWP depending on their coverage area: global and regional models. Global models cover the whole Earth and are generally used for medium range forecasts (more than two days), while regional models run for limited areas and are used for shorter range predictions (between 1 and 3 days). More recently the resolution of regional models has been increased to just a few kilometres in order to resolve better mesoscale phenomena. Along this thesis, we are going to deal with regional models only.

An important characteristic of NWP is the nesting capability, i.e. the ability to perform simulations over areas enclosed in a larger simulation domain which forecast output is used as initial and boundary conditions for the smaller domain. This interaction between different scales can be 1-way nesting (e.g. like in HIRLAM and MASS models) or 2-way nesting (like in MM5). In both cases, the nest's input comes from the coarser simulation, while there is a feedback to the coarser mesh over the nest interior for the 2-way interaction case.

The useful forecast of a NWP model depends not only on the resolution and accuracy with which the dynamical processes are represented; it critically depends also on the initial conditions used to integrate the model. The available observations should be used to initialize the model and to dynamically modify its evolution in order to make good predictions. In the following section, we introduce the concept of data assimilation and summarize the different methods used to assimilate observational data in order to improve the forecasts of a meteorological model.

2.2 Data assimilation techniques

An analysis is the production of an accurate image of the true state of the atmosphere at a given time. Here, the atmosphere is assumed divided in N voxels. The best possible state $\mathbf{x^t}$ that can be represented by a model at time t can be described, as was seen in the previous section, by a N-dimensional discrete vector:

$$\mathbf{x^t}_t = [\mathbf{v}_1, T_1, p_1, \rho_1, ...\mathbf{v}_n, T_n, p_n, \rho_n...] \; ; \; n = 1, ..., N \tag{2.12}$$

where $\mathbf{x^t}(\mathbf{r}, t)$ has been renamed as $\mathbf{x^t}_t$ for clarity. Since the NWP models have lower resolution than reality, this state will never be completely realistic (representativeness error). Therefore, there will exist some discrepancies between the observations (even if they do not have any instrumental error) and their equivalent modeled values in the analysis.

The information that can be used to produce the analysis is a collection of observed values provided by observations of the true state. If the model state is overdetermined by the observations, then the analysis reduces to an interpolation problem. In most cases the data is sparse and the analysis problem is under-determined. In current operational meteorological models, the dimension of the model state $\mathbf{x}$ is of order of 10^6, and the dimension of the observation vector is of order of 10^5 per analysis.

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. Physical constraints on the analysis problem can also help. The background information can be climatology or a trivial state, it can also be generated from the output of a previous analysis or the evolution predicted by a forecast model. In a well-behaved system, one expects that the information accumulates in time into the model state, and propagates to all variables of the model.

Data assimilation is the general process by which observational meteorological data are merged together with a dynamical model of the atmospheric flow in order to obtain as accurate and as consistent as possible a description of the atmosphere. Depending on the problem under consideration, the final product can be the description of the atmosphere at a given time, the evolution of the atmospheric state over a given period of time or the estimates of averaged climatic quantities. See, e.g., Bouttier and Courtier (1998) for a review on this subject.

An ideal assimilation should be able to process all the available information:

- observations (surface stations, radiosondes and pilot balloons, aircraft reports,

vertical temperature soundings, cloud-drift winds, ...) and their covariances.

- physical laws governing the temporal evolution of the atmosphere (i.e. the equations of the model.)

- a priori information (background state and, in case of regional models, boundary condition) and its covariance.

The variances and correlations of the observational error are needed for each kind of data to be assimilated. In current meteorology, a diagonal matrix (no correlation) generally works well.

The variances and correlations of the background error are also needed to spread and smooth information from observations location to surrounding grid points (a diagonal matrix in the grid point space does not work). The (linearized) balance properties of the atmosphere (geostrophy, hydrostatic approximation, etc) are also included in the background correlations. Thus, these correlation terms provide information about all the variables that are balanced with one model variable.

A crude estimate of the background error variances can be obtained from climatological variances of the meteorological fields. A better average estimate is achieved from the differences between the forecast and a verifying analysis. The estimation of the error covariances is more difficult because they are not observed directly. In general, the only way to estimate background covariances is to assume that they are stationary over a period of time and uniform over a domain. The background error covariances should vanish for very large separations.

In practice, two empirical methods are used for the background covariance matrix:

1) built from temporal series of the differences between a verifying analysis and forecasted values, and

2) estimated from temporal series of the the differences between two model predictions (f_1 and f_2) at two different times [called NCEP (National Centers for Environmental Prediction) method, because it was introduced by the U.S. weather service]. Usually, f_1 and f_2 are a 24 h and 48 h forecast runs.

There are two basic approaches to data assimilation: *continuous assimilation*, that introduces observations and analyze them, using the first guess at the appropriate time, and continues the forecast from the updated atmospheric state; and *intermittent assimilation*, that treats observations simultaneously, usually every 6 hours. We next briefly describe the main data assimilation techniques.

Cressman analysis and related methods:

One may like to design the analysis procedure as an algorithm in which the model state is set equal to the observed values in the vicinity of available observations, and to an "arbitrary" state otherwise. This formed the basis of the *Cressman analysis scheme* which is still widely used for simple assimilation systems.

The model state is assumed represented as grid-point values. If we denote by $\mathbf{x_b}$ a previous estimate of the model state (*background*) provided by climatology or a previous forecast, and by $\mathbf{y}_i$ a set of $i = 1...n$ observations of the state, the Cressman analysis provides the model state $\mathbf{x_a}$ defined at each grid-point j according to the following *update* equation:

$$\mathbf{x_a}(j) = \mathbf{x_b}(j) + \sum_{i=1}^{n} W(i,j)[\mathbf{y}(i) - \mathbf{x_b}(j)] \tag{2.13}$$

$$W(i,j) = \frac{w(i,j)}{\sum_{i=1}^{n} w(i,j)} \tag{2.14}$$

The weight function is given by

$$w(i,j) = \frac{R_j^2 - d_{i,j}^2}{R_j^2 + d_{i,j}^2} \quad ; d(i,j) < R_j \text{ and } \mathbf{y}(i) \text{ is observed} \tag{2.15}$$

$$w(i,j) = 0 \quad ; d(i,j) \geq R_j \text{ or } \mathbf{y}(i) \text{ is not observed} \tag{2.16}$$

where $d_{i,j}$ is a measure of the distance between points i and j. The weight function $w(i,j)$ equals one if the grid-point i is colocated with observation j. It is a decreasing function of distance which is zero if $d_{i,j} \geq R_j$, where R_j is a user-defined constant beyond which the observations have no weight. There are many variants of Cressman method. One is the *successive correction method* or *observation nudging* where the weights can be less than one for i=j, which means that a weighted average between the background and the observation is performed. This method is not satisfactory in practice for the following reasons:

- if we have a preliminary estimate of the analysis with a good quality, we don't want to replace it by values provided from poor quality observations,
- in a region with dense observations (denser than the model grid-points), the Cressman method does not define the analysis in an intrinsic way,

- when going away from an observation, it is not clear how to relax the analysis toward the "arbitrary" state, and
- the Creesman method does not guarantee some basic known properties of the true system, like smoothness of the fields, or relationship between the variables.

The method has been quite successful especially when enough data are available to define the atmosphere and is very cheap in computation and easily programmed. It is however very empirical in its nature and proper used of the known statistical properties of the atmosphere cannot be made.

Statistical analysis:

We denote by $\mathbf{x_a}$ the analysis model state, $\mathbf{y}$ the vector of observations, H the observation operator, B the covariance matrix of the background errors $(\mathbf{x_b} - \mathbf{x^t})$, R the covariance matrix of the observation errors $(\mathbf{y} - \mathrm{H}[\mathbf{x^t}])$, and A the covariance matrix of the analysis errors $(\mathbf{x_a} - \mathbf{x^t})$. The estimate $\mathbf{x_a}$ can be obtained as a solution to the variational problem (or "least-squares") (Lorenc 1986):

$$J(\mathbf{x}) = (\mathbf{x} - \mathbf{x_b})^T B^{-1} (\mathbf{x} - \mathbf{x_b}) + (\mathbf{y} - \mathrm{H}[\mathbf{x}])^T R^{-1} (\mathbf{y} - \mathrm{H}[\mathbf{x}]) \tag{2.17}$$

$$= J_b(\mathbf{x}) + J_o(\mathbf{x}) \tag{2.18}$$

where J is called the *cost function* of the analysis, J_b is the *background term*, and J_o is the *observation term*. The estimate $\mathbf{x_a}$ is the value which minimizes the cost function J,

$$\frac{\partial J}{\partial \mathbf{x}} = 2B^{-1}(\mathbf{x} - \mathbf{x_b}) + 2\mathbf{H}^T R^{-1}(\mathbf{y} - \mathrm{H}[\mathbf{x}]) = 0 \tag{2.19}$$

The analysis $\mathbf{x_a}$ is optimal (i.e. it is closest in a least-squares sense to the true state $\mathbf{x^t}$). If the background and observation error probability density functions (pdfs) are Gaussian, then the optimal state $\mathbf{x_a}$ is also the maximum likelihood estimator of $\mathbf{x^t}$. External weak constraints can be added to the cost function (2.18).

If the operator H can be linearized as $\mathbf{H}$,

$$\mathrm{H}(\mathbf{x} + \delta\mathbf{x}) = \mathrm{H}(\mathbf{x}) + \mathbf{H}(\delta\mathbf{x}) \tag{2.20}$$

the expectation of the background and observation errors is zero, and the observation and background errors are mutually uncorrelated, then the solution of (2.19) is the

optimal least-squares estimator, or the *Best Linear Unbiased Estimator* (*BLUE*), defined by the following equations:

$$\mathbf{x_a} = \mathbf{x_b} + K(\mathbf{y} - \mathrm{H}[\mathbf{x_b}]) \tag{2.21}$$

$$K = B\mathbf{H}^T(\mathbf{H}B\mathbf{H}^T + R)^{-1} \tag{2.22}$$

where the linear operator K is called the *gain*, or *weight matrix*, of the analysis. The analysis error covariance matrix is, for any K:

$$A = (I - K\mathbf{H})B(I - K\mathbf{H})^T + KRK^T \tag{2.23}$$

If K is the optimal least-squares gain, the expression becomes $A = (I - K\mathbf{H})B$. Weakly non-linear observation operators can be used, with a small loss in the optimality of the result.

The linear least-squares solution requires in principle the specification of covariance matrices B and R. Except in analysis problems of very small dimension, it is impossible to compute exactly the least-squares analysis. Some approximations are necessary, like the *Optimal interpolation* (OI) analysis.

The OI is an algebraic simplification of the computation of the weight K in the analysis equations (2.21) and (2.22). The equation for $\mathbf{x_a}$ can be regarded as a list of scalar analysis equations, one per model variable in the vector $\mathbf{x}$. For each model variable the analysis increment is given by the corresponding line of K times the vector of background departures $(\mathbf{y} - \mathrm{H}[\mathbf{x_b}])$. The fundamental hypothesis in OI is that for each model variable, only a few observations are important in determining the analysis increment.

The selection of observations should in principle provide all the observations which would have a significant weight in the optimal analysis. In principle, background error covariances are assumed to be small for large separation, so that only the observations in a limited geometrical domain around the model variable need to be selected. The advantage of OI is its simplicity of implementation and its relatively small cost if the right assumptions can be made on the observation selection. A drawback of this method is that spurious noise is produced in the analysis fields because different sets of observations are used on different parts of the model state. Also, it is impossible to guarantee the coherence between small and large scales of the analysis.

In order to avoid the computation of the gain K, we can look for the analysis as an approximate solution to the minimization problem defined by the cost function (2.18).

This is the principle of the *three-dimensional variational analysis* (3D-VAR). The solution is sought iteratively by performing several evaluations of the cost function and of its gradient (2.19) in order to approach the minimum using a suitable descent algorithm. The approximation lies in the fact that only a small number of iterations are performed. The initial point of the minimization, or *first guess*, is usually taken equal to the background $\mathbf{x_b}$.

The popularity of 3D-VAR stems from its conceptual simplicity and from the ease with which complex observation operators can be used, since only the operators and the adjoints of their tangent linear need to be provided. The method removes the local data selection in the OI algorithm and performs a global analysis of the 3-dim meteorological fields. In OI the weights are obtained using suitable simplifications, while in 3D-VAR the minimization is performed directly, and therefore allows for additional flexibility.

In the case of time evolving systems when one wants to reconstruct the state of the system at the final time of the observing period, one might implement sequential estimation over the observing period $[0,R]$. The cost function defined in (2.18) is extended now over a given time interval with the observations distributed among n epochs in the interval,

$$J(\mathbf{x}) = (\mathbf{x}-\mathbf{x_b})^T B^{-1}(\mathbf{x}-\mathbf{x_b}) + \sum_0^n (\mathbf{y}_t - \mathrm{H}_t[\mathbf{x}_t])^T R_t^{-1}(\mathbf{y}_t - \mathrm{H}_t[\mathbf{x}_t]) \qquad (2.24)$$

where $(\mathbf{x}_0.....\mathbf{x}_n)$ is the sequence of state vectors over the period $[0...n]$, and $(\mathbf{y}_0.....\mathbf{y}_n)$ is the set of observation vectors during the observing period. The covariance matrix of $\mathbf{y}_t$ is defined as R_t, and is the sum of the observational and representativeness error. The background error covariance matrix B is only defined at initial time.

The equation (2.24) is subject to the strong constraint that the sequence of model states $\mathbf{x}_t$ must be a solution of the model equations:

$$\mathbf{x}_t = M_t(\mathbf{x}_0) \qquad (2.25)$$

for all model states $\mathbf{x}_t$, where M_t is the model forecast operator from the initial time to t with covariance matrix Q_t. The minimization of (2.24) is thus a nonlinear constrained optimization problem which is very difficult to solve in the general case. Two hypotheses simplify the general minimization problem: *causality* and the *tangent linear hypothesis*.

- Causality: The forecast model can be expressed as the product of the intermediate forecast steps. If the dates t are sorted with respect to time, with $\mathbf{x}_0$ as initial condition so that M_0 is the identity, and by denoting M_t the forecast step from $t-1$ to t we have $\mathbf{x}_t = M_t(\mathbf{x}_{t-1})$ or

$$\mathbf{x}_t = M_t M_{t-1} ... M_1(\mathbf{x}_0) \tag{2.26}$$

- Tangent linear hypothesis: The cost function can be made quadratic by assuming that the operators H_t and M_t can be linearized,

$$M_t(\mathbf{x}_t + \delta\mathbf{x}_t) = M_t(\mathbf{x}_t) + \mathbf{M}_t(\delta\mathbf{x}_t) \tag{2.27}$$
$$\mathrm{H}_t(\mathbf{x}_t + \delta\mathbf{x}_t) = \mathrm{H}_t(\mathbf{x}_t) + \mathbf{H}_t(\delta\mathbf{x}_t) \tag{2.28}$$

Under these approximations, the estimate of the state vector at time t is given by the following expressions, known as the *Linear Forward Kalman Filter* solution:

$$\mathbf{x}_t = \mathbf{x}_{t,f} + K_t(\mathbf{y}_t - \mathrm{H}_t[\mathbf{x}_{t,f}]) \tag{2.29}$$
$$K_t = P_{t,f}\mathbf{H}_t^T(\mathbf{H}_t P_{t,f}\mathbf{H}_t^T + R_t)^{-1} \tag{2.30}$$
$$P_t = (I - K_t\mathbf{H}_t)P_{t,f} \tag{2.31}$$

where $\mathbf{x}_{t,f}$ is the a priori estimate of the state vector at t with a covariance matrix $P_{t,f}$. The estimate $\mathbf{x}_{t,f}$ is a forecast from the estimate valid at previous time $t-1$.

The forecast step transports in time the model state and the covariance matrix in the following way:

$$\mathbf{x}_{t,f} = M_t(\mathbf{x}_{t-1}) \tag{2.32}$$
$$P_{t,f} = \mathbf{M}_t P_{t-1}\mathbf{M}_t^T + Q_t \tag{2.33}$$

The (linear) Forward Kalman Filter, under the assumption that all observations are processed if they were available at the end of the time period R and that the matrix covariances are dropped ($P_t = P_{t,f} = B$), leads to OI or 3D-VAR, depending on the type of estimation used in the analysis step.

The filter approach works sequentially, i.e. the estimate of the state of the atmosphere at time t is performed with the observations taken before t. However, when

all the observations over the interval [0,R] are available, the only way to estimate the optimal state of the atmosphere at t=0 is to run a *smoothing* or *backward filter*. In this case, the model solution over the assimilation period is globally adjusted to the available observations, with propagation of the information contained in the observations both forward and backward in time. This is the idea which underlies the Kalman smoother and the *four-dimensional variational* (4D-VAR) assimilation technique. The equations are the same, provided the observation operators are generalized to include a forecast model that will allow a comparison between the model state and the observations at the appropriate time. These methods require the assimilation to wait for the observations over the whole time interval to be available before the analysis procedure can begin. The Kalman smoother, under the assumption that the model is perfect ($Q = 0$) leads to 4D-VAR.

With the present computer technology, both the Kalman filter and the Kalman smoother cannot be implemented for NWP problems. The Kalman smoother runs at a much larger cost than the 4D-VAR approach. Chapter (6) analyzes in detail the assimilation of GPS-derived observations into MM5 in a 4D-VAR approach.

4D-VAR is a simple generalization of 3D-VAR for observations that are distributed in time (see, e.g., Talagrand 1991). A schematic diagram of the 3D-VAR and 4D-VAR approaches are shown in Figure (2.1). In the diagram, the optimal state of the atmosphere is found for a period of 6 h. The evaluation of the 4D-VAR cost function and its gradient, requires one direct model integration from dates 0 to n and one suitably modified *adjoint integration* made of transposes of the tangent linear model time stepping operators $\mathbf{M}_t$.

Over a given time interval, under the assumption that H and M can be linearized, and the model is perfect, with the same input data (initial background and its covariance B, distribution of observations and their covariances R_t), the 4D-VAR analysis at the end of the time interval is equal to the Kalman filter analysis at the same time.

The Hessian of the cost function of the variational analysis is equal to twice the inverse of the analysis error covariance matrix. In 4D-VAR, the error covariance matrix P_t is implicitly computed. The flow dependence of the estimation error is therefore accounted for. However, there is no explicit expression for the forecast error, such in the Kalman filtering approach.

In NWP, the adjoint code development requires the development of the tangent linear model. Any code, no matter how complex it is, boils down to a series of transformations of values through the application of basic operations (i.e. sum, multiplication,

division). The tangent linear code is obtained by differentiation of the forward code. To derive the adjoint code, one needs to transpose the tangent linear code. There are automatics adjoint generators (e.g., TAMC, http://www.mit.edu/giering/tamc; ADIFOR, http://www.mcs.anl.gov/autodiff), although most adjoint model developments are carried out by hand. This is because the pre-existing models are usually not 'clean' and content many approximations which makes the adjoint coding not directly applicable to most problems. See Zou et al. (1997) for details on the MM5 adjoint coding.

New observing systems, especially remote sensors such as satellites and radars, have been available during the last years. Data assimilation techniques should be adapted for this large number of new instruments. This is the case of the GPS system which is described in the following chapter.

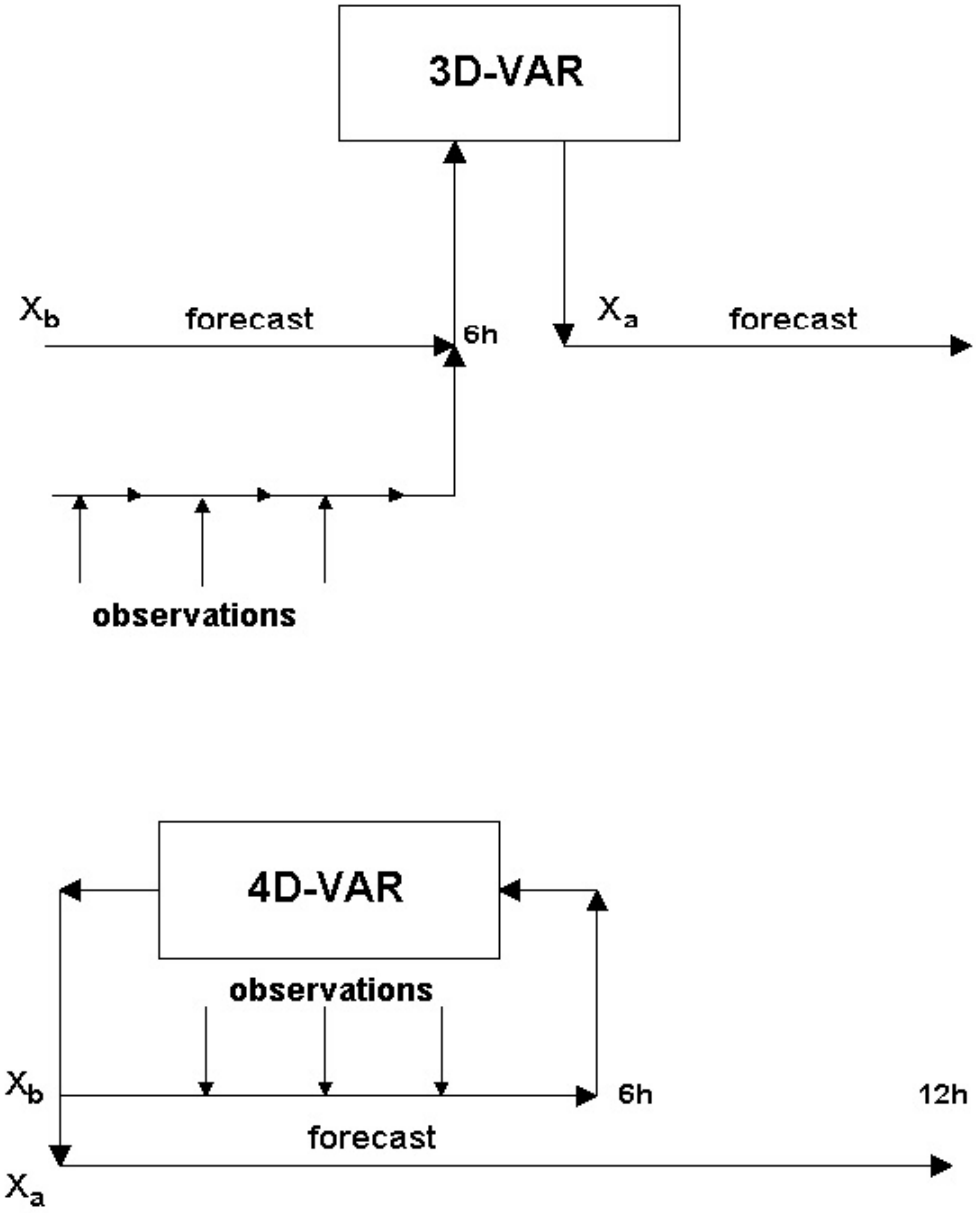


Figure 2.1: Schematic diagrams of the 3D-VAR and 4D-VAR techniques.

Chapter 3

A new observational data source: GPS system

3.1 Time delay effects on GPS signal

The GPS constellation consists of 24 operational satellites (orbiting at an altitude of about 20,000 km) which transmit ranging data at two radio frequencies for positioning (1.2276 and 1.57542 GHz). The use of two frequencies is motivated by the dispersive nature of the effect introduced by the ionosphere. The carrier frequencies are modulated by pseudo-random noise codes and a navigation message, which is designed to inform the user about the health and positioning of the satellite. The carrier frequencies and the modulations are controlled by on-board atomic clocks.

These GPS signals are picked up either by ground-based receivers or by moving platforms (like buoys, planes, low Earth orbit satellites, etc). This thesis has been addressed to the observations gathered at ground receivers forming networks of sizes ranging from 5 to 300 km. A study on the use of GPS signals gathered at smaller networks (less than 5 km) can be found in Flores (1999) and Flores et al. (2001).

The GPS signal, as it propagates through the atmosphere, is affected by the atmosphere in two different ways (Bevis et al. 1992):

1. the signal experiences an extra delay relative to the straight line were the atmosphere replaced by vacuum, and

2. the signal travels in a curved path instead of in a straight line.

The delay in signal arrival time can be stated in terms of an equivalent increase in travel path length. This excess path length is given by

$$\Delta L = \int_L n(s) ds - G \tag{3.1}$$

where $n(s)$ is the refractive index as a function of position s along the curved ray path L, and G is the straight-line geometrical path length trough the atmosphere. The expression above can be written as follows:

$$\Delta L = \int_L [n(s) - 1]\, ds + [S - G] \tag{3.2}$$

where S is the path length along L. The first term in (3.2) is due to the slowing effect and is because the index of refraction of the atmosphere differs from one. The second term is due to bending and vanishes for rays oriented along the zenith direction and in the absence of horizontal gradients.

The extra delay of the GPS signal, known as atmospheric delay, is commonly regarded as a nuisance parameter for geodetic purposes. However, due to the inherent sensitivity of the GPS system, this error source can be converted into the object of study. The use of GPS for the determination of water vapor content in the zenith direction with a precision of a few millimeters is a well established technique (e.g., Bevis et al. 1992; Rocken et al. 1993, 1995; Businger et al. 1996; Duan et al. 1996).

Some of the atmospheric estimates obtained with GPS have been compared with estimates obtained with other techniques which are also sensitive to the amount of water vapor in the atmosphere, such as very-long baseline interferometry (VLBI) and WVR. Carlsson et al. (1996), for example, showed that the root-mean-square (rms) agreement between the water vapor delay estimates from these techniques is better than 10 mm, though the VLBI estimates presented an unresolved systematic bias of about 5 mm compared to the other two methods. The measurements from these three techniques, in turn, have been compared with integrated values of water vapor obtained with radiosondes, which measures the in-situ state of the atmosphere. For example, the agreement found between estimates of precipitable water (PW) obtained using GPS and radiosondes is at the 2 mm rms level (Rocken et al. 1993, 1995; Duan et al. 1996; Tregoning et al. 1998; Emardson et al. 1998).

We next discuss the equivalence between water vapor delay (commonly known as wet delay) and precipitable water.

3.2 Modeling of the neutral atmospheric delay

The neutral atmosphere affects the propagation of electromagnetic signals by retarding and bending them. These effects can be determined if one knows the index of refraction n, or more conveniently, the refractivity N, defined as $10^6(n-1)$. If we neglect the contribution of the liquid water content (Kursinski, 1997), the total refractivity of the neutral atmosphere is given by (Smith and Weintraub 1953; Thompson et al. 1986)

$$N = k_1 \left(\frac{P_d}{T}\right) Z_d^{-1} + k_2 \left(\frac{P_w}{T}\right) Z_w^{-1} + k_3 \left(\frac{P_w}{T^2}\right) Z_w^{-1}, \tag{3.3}$$

where P_d and P_w are the partial pressures (in hPa) of dry air and water vapor, respectively, T is the temperature (in K) of air, $k_1 = (77.604 \pm 0.014)$ K hPa^{-1}, $k_2 = (64.79 \pm 0.08)$ K hPa^{-1}, $k_3 = (3.776 \pm 0.004) \times 10^5$ K^2 hPa^{-1} are the refractivity constants (Thayer 1974), and Z_d^{-1} and Z_w^{-1} are the inverse compressibility factors of the dry gases and water vapor (Owens 1967), respectively. For an ideal gas $Z = 1$, and for the atmosphere differs from unity by a few parts per thousand.

The atmospheric delay is the integral of the refractivity along the ray path. We use mapping functions to relate atmospheric delays as measured in any sky direction to zenith delays. See Niell (1996) for a review and error discussion on the various mapping functions employed in geodetic analysis that do not incorporate azimuthal variation, and Chen and Herring (1997) for mapping functions that do incorporate azimuthal asymmetry.

It is useful to separate the zenith atmospheric delay as the sum of two terms (Davis et al. 1985): the hydrostatic delay and the wet delay. These two delay terms are given by the expressions

$$\Delta L_h^z = \mathrm{ZHD} = 10^{-6} \textstyle\int_0^\infty dz\, N_h = \displaystyle\int_0^\infty dz \left(k_1 \frac{P_d}{T} Z_d^{-1} + (k_2 - k_2') \frac{P_w}{T} Z_w^{-1}\right) \tag{3.4}$$

$$\Delta L_w^z = \mathrm{ZWD} = 10^{-6} \textstyle\int_0^\infty dz\, N_w = \displaystyle\int_0^\infty dz \left(k_2' \frac{P_w}{T} Z_w^{-1} + k_3 \frac{P_w}{T^2} Z_w^{-1}\right) \tag{3.5}$$

where h refers to hydrostatic, w to wet, z is the zenith direction, $k_2' = k_2 - k_1(M_w/M_d) \approx 17 \pm 10$ K hPa^{-1}, and M_w and M_d are the molar weights of water vapor and dry gases, respectively. Hence, ΔL^z=ZTD=ZHD+ZWD. The zenith hydrostatic delay is a result of the induced dipole moment and has a typical value of around 2300 mm at sea level. The zenith wet delay, which is associated with the atmospheric water vapor, has a smaller value (0-300 mm at sea level), and is due to the permanent dipole moment of the molecules of water vapor present in the troposphere. The wet contribution is very difficult to model because it is highly variable in space and time.

The zenith hydrostatic delay can be accurately modeled if measurements of total surface pressure are available (Saastamoinen 1972),

$$\mathrm{ZHD} = (2.2779 \pm 0.0024)\,\frac{P_s}{f(\lambda, H)} \tag{3.6}$$

where ZHD is expressed in mm, P_s is the surface pressure expressed in hPa, and $f(\lambda, H) = 1 - 0.00266\cos(2\lambda) - 0.00028H$ accounts for the variation in gravitational acceleration with latitude λ and the height H of the surface above the ellipsoid (in km). This delay can be predicted to better than a millimeter with surface pressure accuracies of 0.4 hPa (see, e.g., Elgered et al. 1991 for a discussion on the uncertainty associated to the zenith hydrostatic delay).

Once estimated, the zenith wet delays can be accurately converted into PW by using the expression (Bevis et al. 1994)

$$\mathrm{PW} = \mathrm{ZWD} \cdot \Pi(T_m, \rho_v) \tag{3.7}$$

where PW and ZWD are expressed in mm and $\Pi(T_m, \rho_v)$ is given by

$$\Pi = \frac{10^6}{\rho_v R_v [(k_3/T_m) + k_2']} \tag{3.8}$$

In (3.8), R_v is the specific gas constant for water vapor (461.5 J kg^{-1}K^{-1}), T_m is the mean temperature of the atmosphere, defined as $T_m = (\int_0^\infty dz\, P_w/T)/(\int_0^\infty dz\, P_w/T^2)$ (Davis et al. 1985), and ρ_v is the density of liquid water. The mean temperature depends on the vertical profile of P_w and atmospheric temperature T. Fortunately, the mean temperature has been empirically found to be well correlated with surface temperature (Bevis et al. 1992). For example, based on more than 120 000 radiosonde profiles from 38 sites in Europe, Emardson and Derks (1999) determined a value of $\Pi \sim 0.15$ for a time-averaged ground temperature of 275 K with a relative rms error of 1.14%. Because this error is significantly smaller than the error of the estimates of zenith wet delays, an error of 10 mm in zenith wet delay propagates to 1.5 mm in PW after (3.8). It is now possible to recover PW routinely from GPS data with a rms error of less than 2 mm + 1% of the PW and a long-term bias of less than 2 mm (Dixon et al. 1990; Herring et al. 1990; Bevis et al. 1992; Rocken et al. 1997; Ware et al. 1997).

One of the most suitable atmospheric application of GPS is perhaps the assimilation of water vapor content estimates into NWP and climate models. The lack of humidity measurements that could potentially be assimilated into NWP forecast

models is the main reason of its (sometimes) low reliability (Kuo et al. 1993, 1996). The fact that GPS can supply these data in near-real time (Rocken et al. 1997) and at low cost is changing, at the algorithmic level, the way these models are being used to assimilate the GPS estimates (Zou and Kuo, 1996; Kuo et al. 1996). GPS slant delay measurements (Ware et al. 1997), the delay along the lines-of-sight from the receiver to the satellites (as opposed to zenith delays, the delay in the zenith direction) provide information that can be used to extract vertical profiles of the index of refraction of the atmosphere. However, this thesis is focused on zenith atmospheric delays, while slant delays are not treated through the work.

GPS data will possibly be assimilated into operational NWP models in the future in a variational assimilation context. In preparation for these efforts, it is necessary to determine how NWP models will fare in simulating GPS measurements. This is analyzed in the following chapter.

Chapter 4

The use of GPS to validate NWP models: atmospheric delays

The methods used to obtain GPS-derived observations from geodetic techniques and their comparison with NWP models is the subject of the following sections.

In this chapter, we analyze several meteorological situations using three different mesoscale numerical models. The study shows that the observations estimated from the GPS atmospheric delays are an accurate measurement to validate some aspects of NWP models.

Section (4.1) analyzes an experiment characterized by high precipitable water values associated to rainfall events by means of the HIRLAM (HIgh Resolution Limited Area Modeling) model. Particular emphasis is devoted to describe our implementation of a NWP monitoring system with the use of a near real time (NRT) GPS data processing. A mesoscale convective system is treated in Section (4.2) with the MM5 (Mesoscale Modeling System v.5) model. We finally present first results for a selected period of intense precipitation by using the MASS (Mesoscale Atmospheric Simulation System) model in Section (4.3).

4.1 The HIRLAM model: December 1996

In this section, we analyze a meteorological situation using GPS observations and a NWP model in the vicinity of the Madrid Sierra, Spain, between 2 and 15 December 1996. See Elósegui et al. (1998) for a detailed description of this GPS campaign and the GPS data processing.

The modeled calculations of precipitable water are carried out by HIRLAM, the

hydrostatic NWP model operational at the Instituto Nacional de Meteorología (INM, Spain). We have run HIRLAM in two modes: analysis or HIRLAM/A and forecast or HIRLAM/F.

We have concentrated on estimates of PW derived from zenith delay measurements acquired at several ground-based GPS sites. The results from this study will provide an upper bound on the size of the GPS errors that are acceptable so that these GPS-derived PW estimates have a positive impact on numerical weather prediction models once they are correctly assimilated.

Yang et al. (1999) have studied the ability of the HIRLAM model to reproduce the spatial and temporal evolution of PW by comparing the model values with GPS estimates obtained on a spatial scale of 100-1500 km in northern Europe. Our goal here is to perform a comparative study of the spatial and temporal distribution of the water vapor content as obtained with GPS and modeled with HIRLAM on a smaller scale (5-50 km) and high topographic relief. Data from radiosonde launches were also available in the area of study and have been used as an additional check. The geographical region selected for the experiment, the Madrid Sierra, Spain, is characterized by complex topography. The GPS-derived PW data will be used to check the reliability of the model performance in an area with strong topography (Cucurull et al. 1999, 2000).

4.1.1 Experimental setup and meteorological situation

We conducted a GPS experiment during 2-15 December 1996 in the Madrid Sierra of Spain to (1) study the spatial and temporal variations of PW and (2) compare the GPS-derived estimates of PW with PW values simulated with the HIRLAM numerical weather prediction model.

The GPS network consisted of five Trimble 400SSE receivers forming baselines ranging in length from 5 to 50 km. Figure (4.1) shows the geographical location of the GPS sites. We employed Trimble geodetic L1/L2 antennas with ground plane (Elósegui et al. 1998). The geographical region selected for the experiment is characterized by complex topography with maximum altitude difference between GPS sites of about 400 m. Figure (4.2) shows a contour plot of the topography of the region. Meteorological data were collected at one of the sites (Robledo) during the experiment. As part of the meteorological package, Robledo operates a ($\sim$0.3 hPa) barometer. Vertical profiles of temperature, pressure and relative humidity were also available from 12-hourly radiosonde launches at the Barajas airport.

Two frontal systems crossed the GPS sites during the experiment. The first at around 4-6 December and the second at around 12-14 December. We selected the second front to perform a more detailed analysis of GPS-derived PW because this presented the largest PW values. The synoptic regime corresponding to this meteorological situation can be observed in the 36-h sequence of Figure (4.3). Each map shows a temporal snapshot of the analyzed mean sea level pressure and temperature at 850 hPa as determined by the HIRLAM/A low resolution ($0.5° \times 0.5°$) model. A low pressure system seen in the North Atlantic (somewhat to the northwest of the Iberian Peninsula) at 12 UTC 12 December moves east reaching the continent and overpassing the area of interest during 13 December. This low pressure system advected warm, moist air from the southwest as it approached the coast. The front (see Figure 4.4) brought steady precipitation to the area of the experiment, with heaviest rain at around the end of 12 December. Figure (4.3c) shows that the flow is changing from southwesterly to northwesterly in the center of the Iberian Peninsula on 14 December.

4.1.2 GPS data analysis

The GPS observations at each site consisted of data streams, simultaneously received from 6 to 8 satellites, of undifferenced dual-frequency carrier-phase and pseudo-range measurements obtained every 30 sec. We used GPS satellite precise orbits and clocks as well as consistent earth-rotation parameters provided by the International GPS Service (IGS), together with the GIPSY/OASIS-II (v.4) software package (Webb and Zumberge 1993 and references therein) to estimate zenith total delays at the five GPS sites with a precision of about 0.5 cm (Elósegui et al. 1998). This software uses a stochastic filter to provide time-dependent estimates of the atmospheric delays for each site. The dynamics of these delays were modeled as a random-walk stochastic process, with a drift rate of 0.25 cm $h^{-1/2}$. This drift rate was chosen to be consistent with measurements obtained with a colocated WVR (see Elósegui et al. 1998 and Ruffini et al. 1999a for a more thorough discussion.)

To derive precipitable water from the estimates of zenith total delay we first calculated and subtracted out the hydrostatic contribution. We used pressure values at each site to compute zenith hydrostatic delays via (3.6). In the absence of surface pressure measurements at all sites but Robledo, we have used HIRLAM/A to calculate 6-hourly pressure values at all sites. The required pressure measurements between these modeled values we filled up using the pressure data collected at Rob-

ledo corrected for height differences between sites. This is justified because of the small bias (0.4 hPa) between the readings of the barometer and the surface pressure values calculated by the NWP model at the Robledo site over the course of the whole experiment. This small bias will account for 1 mm bias of zenith hydrostatic delay. Verification scores of both HIRLAM analysis and forecasts against observations were produced routinely during the objective verification procedure at the INM. The pressure rms error of HIRLAM/A at mean sea level is 1.5 hPa. We have therefore adopted this value as the expected error of the estimates of surface pressures at the other GPS sites, which is equivalent to 3.4 mm error in terms of zenith hydrostatic delays. The resulting uncertainty of the zenith wet delays estimates is about 6 mm (the errors in the zenith total delay and the hydrostatic delay added in quadrature) or, equivalently, about 1 mm PW.

Precipitable water estimates for the whole campaign every 150 sec are shown in Figure (4.5) for all five GPS sites. The time series for all the stations of the network are quite similar though small differences between stations due to short-term water vapor variations can be observed. It is particularly important to determine how sensitive are the GPS data to specific meteorological situations. For this reason, we have selected the front passage of 12-14 December described above, which will be studied in greater detail in the following section. Table (4.1) shows the average and rms PW values for this front as well as for the first front and the entire campaign. From the table, it is clear the influence of orography on the measurements. This dependence is basically due to a scaling of the amount of integrated water vapor with altitude. In order to carry out a proper intercomparison of the data it would be necessary to correct for this height-dependent scaling factor. However, we have not found any empirical function (e.g. an exponential law) that could model adequately this dependence, perhaps due to the complex topography of the network and/or the highly unstable atmospheric conditions during the experiment. The table reveals a significant increase of PW due the passage of the second front. Also, the rms values for the entire campaign are larger than those of the two fronts because the data span is five times larger and PW deviates more and more from a mean value the longer the time period.

In order to compare PW values derived from GPS data with the other two techniques (radiosonde and HIRLAM) we first transformed the ellipsoidal GPS heights (WGS84 ellipsoid) to sea level heights using the OSU91A geoid model (Rapp et al. 1991). In the area of interest the geoid undulation (i.e. the height of the geoid above

the ellipsoid) amounts to about 50 m. This value we subtracted from the ellipsoidal heights.

4.1.3 Radiosonde data

Radiosonde balloons were launched twice a day (at approximately 00 and 12 UTC) from the Madrid-Barajas airport. The balloons, which are operated by the INM, were equipped with Vaisala RS-80s radiosondes with A-humicap humidity sensor (brand names are mentioned for identification purposes only). We obtained zenith hydrostatic and wet delays by integrating the atmospheric profiles sensed by the sonde along its ascending path using (3.4) and (3.5), respectively.

Since Barajas is close to the IGNE site (horizontal distance of $\sim$12 km, height difference of $\sim$82 m) (see Figure 4.1), we have also compared atmospheric delays derived from GPS at IGNE and radiosonde at Barajas by integrating the measurements of the latter from the height of IGNE upwards. (However, we should note that significant water vapor gradients can occur even over a 12 km distance, specially during the passage of a frontal system.) The balloons reach a maximum height above the geoid of approximately 25$-$30 km. To compare formally GPS and radiosonde delays, one would have to integrate the radiosonde measurements up to the altitudes of the GPS satellites. Since this is not obviously possible, we have used the following procedure to extrapolate the radiosonde data (temperature, dew point temperature and pressure) above the last measurement available and derive radiosonde delays:

- *temperature*: we used the same values as in the upper level of the HIRLAM model (see Section 4.1.4).

- *dew point temperature*: we assigned it low values to get a zero wet-pressure value—i.e., we assumed that there will not be any significant amount of water vapor at, and above, these high altitudes (i.e. 30 km upwards), which is a realistic assumption (Bertin et al. 1996.)

- *pressure*: we used a constant temperature (isothermal) condition from an altitude of about 12 km upwards throughout the stratosphere, which is a quite realistic approximation. The equation of hydrostatic equilibrium $dP/P = -(Mg/RT)dz$ can be readily integrated because T is constant throughout this region. This isothermal condition leads to the equation for pressure P at altitudes z above 12 km (the approximate altitude of the tropopause z_{TP}),

$$P = P_{TP} \cdot \exp\left[\frac{-Mg(z - z_{TP})}{RT_{TP}}\right] \tag{4.1}$$

where P_{TP} is the pressure at the tropopause, M is the molar mass of the air, g is the effective acceleration of gravity in the stratosphere (assumed constant), R is the universal gas constant, and T_{TP} is the temperature at the tropopause (Bertin et al. 1996). Since from an altitude of about 12 km pressure decreases exponentially with altitude, we fitted the radiosonde pressure data obtained from the tropopause upwards (we took z_{TP} as 12 km) with an exponential law, from which we obtained the lapse rate. The values of P_{TP} and T_{TP} in (4.1) are provided by the soundings. We found an average standard deviation of about 0.6 hPa when fitting the pressure with an exponential function. This same law allowed us to extrapolate the pressure values to higher altitudes. Note that as the pressure is almost zero at about 70 km and all the functions involved in the delays calculation have pressure in the numerator, it will be enough to extrapolate the height data until about 70 km. We found that the extrapolated "upper atmosphere", that is, the atmosphere from ∼30 km to ∼70 km, contributes an average of 3% to the zenith total delay. Radiosonde zenith wet delays can be converted to PW using (3.7).

The average PW rms error resulting from considering the contribution of the instrumental errors in the integrating of the radiosonde profiles are about 1 mm. For the Vaisala radiosonde the humidity sensor has a nominal resolution of 1% and a repeatability (i.e., the standard deviation of differences between two successful calibrations) of 2%, the temperature sensor has a resolution of 0.1 K and repeatability of 0.2 K, and the pressure sensor has a resolution of 0.1 hPa and repeatability of 0.5 hPa. (See the Appendix for more details on the propagation of statistical errors associated to radiosonde estimates of atmospheric delays.)

4.1.4 The HIRLAM simulation

The hydrostatic HIRLAM short-range weather forecasting system is a complete analysis and forecast system over a limited area (Källen 1996). A fourth-order implicit horizontal diffusion is used to prevent the enstrophy accumulate at the smaller scales. The physics contains a parameterization for solar and longwave radiations and for simple surface processes. In the cloud parameterization, the large scale convection is a Kuo type with a specific treatment of microphysical processes.

At the INM, the HIRLAM system is run at two different horizontal resolutions, 0.5° latitude by 0.5° longitude [operational low-resolution run (OPR)], and 0.2° latitude by 0.2° longitude [high-resolution run (HIR)], both with the same 31 p-sigma, hybrid levels and vertical resolution. The OPR model domain covers the area between 15.5° and 65.0° north and between -66.5° and 30.0° east. Global forecasts from the European Center for Medium Range Forecasts (ECMWF) are used as boundary conditions to the OPR model. The high resolution model, HIR, 1-way nested into the OPR, has been specially designed to cover the Iberian Peninsula. OPR simultaneous fields provide the lateral boundaries to HIR. The topography of Spain, due to its complexity, is much better represented in the HIR than in the OPR run.

The HIRLAM model has a 6-h data assimilation cycle. It is based on a limited area version of the old ECMWF 3D Optimum Interpolation (see Section 2.2) scheme for the analysis of the upper air fields (Lönnberg and Shaw 1987). It is multivariate in the mass and wind fields and univariate in the relative humidity. A short-range first-guess forecast (6 h) is corrected by observations from a 3-h period spanning the nominal analysis time. Single level data from the surface land stations, ships, buoys as well as from aircrafts are used. Also, winds from geostationary satellites are introduced into the analysis. Multilevel observation reports processed include information from radiosondes and pilot balloons (e.g. TEMP and PILOT). A later step of normal modes initialization is performed after the analysis. OPR and HIR runs have their separated assimilation cycles.

At the INM, 6-h temporal series for the HIRLAM/A model and the 3-h interval for the HIRLAM/F are the only available simulations in an operational way. Accordingly, we have used the 00, 06, 12 and 18 UTC HIR analysis (HIRLAM/A) and the 00 UTC HIR forecasts at regular intervals from 3 to 24 h (HIRLAM/F) to simulate HIRLAM precipitable water. The data for this analysis and forecast come from the HIRLAM archive at the INM. Since the shortest GPS baseline is about 5 km and the HIRLAM higher grid resolution is about 20 km, we have interpolated, both horizontally and vertically, the HIR model variables to the locations of the GPS sites before estimating PW. For this, we have used the same operator used by the HIRLAM model to interpolate the boundary condition fields (Källen 1996). Bilinear interpolation from the four closest grid points values is used in the horizontal. To move the whole model profile from the model topography to the height of each GPS station, the vertical interpolation is carried out by using tension splines but with emphasis on preserving the stability properties inside the planetary boundary layer. To interpolate

vertically the pressure field from the model topography to the height of each GPS station, we have integrated the hydrostatic equation from the HIRLAM surface level to the height of the GPS site using a virtual temperature profile expressed as a linear function of the logarithm of pressure in the vicinity of the GPS level. In case that the GPS station is below the HIRLAM surface level, the virtual temperature is obtained by extrapolation from the three lowest HIRLAM levels, otherwise the temperature profile is obtained by interpolation from the three closest HIRLAM levels to the GPS surface level. This linear virtual temperature profile is obtained by regression from the three selected HIRLAM temperatures.

PW was calculated at every GPS station by integrating the specific humidity in the vertical, PW=$\int dP\,(1/\rho_v)(q/g)$, where q is the specific humidity, P the pressure, g the acceleration of gravity and ρ_v the density of water. The scores of verification against observations of the HIRLAM products at the INM have been used to estimate the error of the calculated PW. The resulting error associated to the PW HIRLAM model varies from 1.4 to 2.1 mm.

In addition, we used a second method to check for the HIRLAM-derived PW values. We used the temperature, pressure and humidity profiles of HIRLAM to generate refractivity profiles. The integration of these profiles along the zenith direction gives us the zenith wet delays after applying (3.5). These simulated zenith delays can be converted to PW using (3.7). The PW obtained by the HIRLAM model (from humidity and pressure profiles) are consistent with the PW values calculated from the integration of the profiles of the vertical refractivity of HIRLAM. We found an average PW bias of 0.5 mm (rms of 0.2 mm) between both techniques. This good agreement confirms the validity of the value of 0.15 used for Π in (3.8) to infer PW from estimates of zenith wet delays with GPS.

4.1.5 Results and discussion

We have used estimates of PW to study the ability of GPS to describe the evolution of a frontal system which crossed our local network. These PW estimates have an accuracy of 1 mm and can provide information on certain features associated to the small spatial and short temporal scales of variation of atmospheric water vapor. Comparable PW accuracies have been obtained in previous studies (e.g., Rocken et al. 1995; Tregoning et al. 1998). The frontal passage selected corresponds to the time period of 12-14 December, which is associated with the largest PW values. Figure (4.6) shows the GPS-derived estimates of PW for two of the sites. We have selected

Robledo and Escorial as the best suited site pair to study the passage of the front because they are the two westermost sites of our network and, in consequence, the first ones to detect any noticeable change in water vapor. (The horizontal distance between Robledo and Escorial is 20 km.) A distinctive element of Figure (4.6) is the larger PW values at Robledo compared to Escorial. This is mostly due, as explained above, to an inverse dependence of the water vapor content with the altitude of the site. Although both series have similar trends, differences between them can be observed. For example, the PW time series present a relative time shift of about 40 min at around 12 December at 14 h, the time series of Robledo leading that of Escorial. This can be interpreted by the air masses with large content of water vapor reaching Robledo earlier than Escorial. The predominant southwesterly winds associated with the front passage (see Figure 4.3) bring the air masses rich in water vapor first to Robledo. This example illustrates that relatively short temporal variations of PW over small spatial scales can be accurately determined using GPS.

We will next compare these PW estimates obtained with GPS data (hereafter PW estimates) with the ones calculated using the HIRLAM/A model (hereafter PW modeled) and the radiosonde data. The intercomparison of PW using three different methods is useful to assess their differences and to validate future parameterizations of NWP models. The ability of HIRLAM to simulate topographically induced effects is limited by its spatial resolution, especially in complex terrain. As it was mentioned in the previous section, the HIRLAM lowest resolution model, OPR, was used to drive the large scale flow in the high-resolution HIRLAM runs. For the comparison between PW estimates and PW modeled we will only use the high-resolution model.

Figure (4.7) shows PW estimates of GPS, HIRLAM/A and radiosonde for the 15-day period at the IGNE site. The GPS estimates are average values over 30 min periods. (This filtering of the high-frequency component of PW has been performed solely for clarity.) All three data sets agree with each other to a few mm level. However, radiosonde estimates appear systematically lower than the PW obtained with the other two methods. For example, the bias between PW estimates from GPS and radiosondes is 1 mm, with radiosondes lower than GPS, and the rms difference is 1.6 mm. This rms difference falls within the expected PW error of the combined GPS and radiosonde errors.

The HIRLAM/A model simulates PW during periods of high amount of precipitable water (between 11-13 December) reasonably well when compared to GPS PW estimates. The two PW maxima, which occurred at 12 UTC on 4 December, and at

00 UTC on 13 December, are associated to the two rainy intervals. The PW modeled values and the PW estimates are comparable during the precipitation periods, that is, at around 12 UTC on December 4, and 00 UTC on December 13. A strong decrease in total atmospheric moisture occurs shortly after the frontal passage (5 and 15 December). The HIRLAM model does not drop the total water vapor amount so sharply. Rather, it seems to slightly overestimate this low level moisture conditions with respect to what it is observed with the GPS data. Thus, HIRLAM seems to perform quite well compared to GPS if the PW does not change too rapidly. Indeed, the bias between PW estimates and PW modeled for the entire experiment is 0.2 mm, with GPS lower than HIRLAM. The rms difference is 2.1 mm and is in agreement with the expected error derived from the standard deviations of GPS and HIRLAM.

The humidity profiles from which the PW modeled are obtained are largely influenced by the relative humidity measured by the radiosondes and assimilated into the HIRLAM model. On the other hand, the relative weight carried by radiosonde measurements on relative humidity at the surface is smaller because the number of sites from which radiosondes are launched in the Iberian Peninsula is, of course, considerably smaller than the volume of surface observations that are assimilated into HIRLAM. Also, the (temporal) sparseness of the radiosonde launches, typically once every 12 h, contributes to sometimes worsen the PW modeled, which are calculated every 6 h. Fortunately, the Barajas radiosonde site is near one of the GPS stations of this study, which gives us some confidence on the PW calculated from HIRLAM/A profiles.

The PW values retrieved from the radiosonde profiles are lower than the PW calculated using the HIRLAM model (average bias of 1.2 mm). However, the rms value of 1.3 mm shows that the precision of the comparison is at the same level as the standard deviation of both techniques, though they are not totally independent. The algorithm used in the OI analysis filters the observation increment (observation departure from first guess) according to the assumed data to first guess error ratio. The comparison of calculated PW by using HIRLAM/A and radiosonde atmospheric profiles have been included to show the analysis filter performance.

Table (4.2) summarizes the site-by-site comparison between PW estimates and PW modeled for the entire experiment. The PW rms value of 2 mm is in agreement with the standard deviation of each technique and indicates the level of precision expected for future studies in which GPS data will be assimilated in NWP models. Figure (4.8) compares PW estimates and PW modeled for all the GPS sites and entire

campaign. The HIRLAM model reproduces reasonably well the PW measurements around values of about 15 mm. However, below this value there is a tendency for the numerical model values of PW to be higher than the PW estimates. In contrast, the PW calculated values are lower than the PW estimates for values higher than 15 mm. A straight-line fit to this data yields a slope of 0.69 ± 0.03, and the χ^2 (per degree of freedom) is 3.1.

We also analyzed the PW values obtained with the HIRLAM/F in order to assess the ability of the NWP model to forecast the precipitable water. Figure (4.9) shows GPS-derived estimates of PW at Robledo during the passage of the second front. The figure also shows the 6-hourly HIRLAM/A PW and the 3-hourly HIRLAM/F values. Unfortunately, it was only possible to obtain values of the PW calculated by HIRLAM/F every 3 h since these were the only available simulations in an operational way at the INM. For the HIRLAM/F results, and comparing with the HIRLAM/A, the rms increases with the forecast range up to 3 mm in 24 h. A similar feature is found for the bias (-1.2 mm). An underestimation of the PW modeled is found when comparing to PW estimates from GPS. As it was expected, the analysis at 00 and 12 UTC, which have made use of radiosonde data, produces a PW value very close to PW estimates by GPS.

4.1.6 Conclusions

We have studied the spatial distribution and the temporal evolution of atmospheric water vapor in terms of precipitable water using GPS. The GPS data used in this experiment were acquired during 2-15 December, 1996, in the Madrid Sierra, Spain. We operated a total of five GPS stations that spanned a maximum horizontal distance of 50 km. The atmospheric flow during the time of the experiment was mainly driven by synoptic scale disturbances. Two frontal systems both associated to large values of PW crossed over the network on 4-6 and 12-14 December, respectively. The HIRLAM numerical weather prediction model simulates the PW for this region and time period. The radiosonde data from a nearby site is used to estimate PW during this time period. We have carried out a comparison of the PW obtained using all three methods.

The PW estimates derived from GPS every 30 min and the PW values obtained from the 6-hourly HIRLAM analysis agree with each other to within 2 mm root-mean-square (rms). This rms value is consistent with the standard deviation of each technique. Hence, to improve the products derived from a numerical weather model

such as HIRLAM by assimilating PW in real time the accuracy of these estimates should be of mm level, provided that the meteorological conditions are similar to those encountered in the Madrid Sierra. GPS-derived and radiosonde-derived PW estimates agree with each other to within 1.6 mm rms. The radiosonde PW estimates are generally lower than those obtained using GPS and HIRLAM.

In addition, we have used HIRLAM in its forecast mode (HIRLAM/F) to predict PW from a given period to a maximum time prediction of 24 h at regular intervals of 3 h. We have found that the bias and the rms between the PW estimates derived from GPS and the HIRLAM/F predictions increase as the extrapolation time becomes larger. For example, the PW bias between both techniques is -1.2 mm and the rms is 3 mm for a 24-h prediction.

The good PW agreement found among all three methods (GPS, HIRLAM and radiosondes) is very encouraging for the possible use of GPS atmospheric products in NWP models in the near future. These are promising results since the disparity in the prognostic skill for precipitation is a consequence of the formation of precipitation on scales essentially smaller than those resolved by present-day global models and the lack of mesoscale data with which to initialize regional fine-mesh models. The availability of such measurements would potentially be useful for studying the distribution of PW on phenomena of smaller spatial scales. The high rate of the GPS data retrievals suggests the assimilation of PW into NWP models in a four-dimensional variational context.

The results found in this section show that GPS measurements can detect small scale fluctuations and therefore can be used to evaluate NWP models with finer resolution. This will be analyzed in Section (4.2).

An unattended system for the comparison of the GPS-derived ZTD estimates with the simulated ZTD values carried out with the HIRLAM NWP model has been established in collaboration with the INM. In the following, we discuss the selected approach as well as an example of the results of the comparison.

4.1.7 Near real time GPS data processing

We have developed a system to retrieve zenith total delay values at several European GPS sites in a NRT mode to improve geophysical parameters estimates (Flores at al. 2000b; Cucurull et al. 2001a) and to develop experience to assist the validation of

some model aspects (Rius et al. 2000).

Figure (4.10) shows the location of the real-time GPS stations. For each station, 1-h interval data is collected from the IGN (ftp://igs.engs.ign.fr) and CDDIS (NASA-GSFC, ftp://cddis.nasa.gov) servers. The data quality check is based on TEQC software from UNAVCO. The GPS data are processed with GIPSY-OASIS II package (Webb and Zumberge 1993). Data files of the last 24 h are merged for each GPS station.

The objective of the process is to provide estimates of the zenith total delay at the GPS stations (and clock corrections for the GPS satellites) within the next hour after data acquisition. The obtained GPS zenith total delays are then compared with the values modeled with the operational forecast of HIRLAM. The comparisons run operationally at the IEEC.

At the INM, 6-h temporal series for the analysis and 1-h interval (since year 2000) for the short range forecast are the available simulations in an operational way. Accordingly, we have used the 00, 06, 12 and 18 UTC HIR and OPR analysis to simulate the precipitable water variable. In the HIRLAM/F case, we have simulated the PW values with the HIR and OPR modes for predictions of 1 h up to 24 h for the analyses verified at 06 and 18 UTC. The forecast range has been extended to 48 h for the 00 and 12 UTC analyses.

Since the HIRLAM higher grid resolution is about 20 km, we have interpolated, both horizontally and vertically, the HIR and OPR model variables to the locations of the GPS sites before estimating PW by using the methodology described in Section (4.1.4).

Forward operators of zenith wet delay and zenith hydrostatic delay have been developed for the HIRLAM model. Simulations of the ZWD, ZHD as well as the ZTD are output from the model for both the HIR and the OPR modes. The OPR run products are interpolated at VILL, MAD2, ZIMM, ONSA, BRUS and WTZR stations whereas the comparison with the HIR run is carried out at VILL, MAD2, ZIMM, and WTZR sites.

Figure (4.11) shows the comparison of the GPS-derived ZTD values (black with error bar in green) with the ones obtained with the HIRLAM model for one selected day (12 UTC 7 December 2000). The modeled ZTD values are plotted for the HIR (in blue) and the OPR (in red) runs. In general, the HIRLAM model simulates ZTD reasonably well when compared to GPS observations. However, short temporal variations of ZTD are not accurately modeled by HIRLAM. One of the stations

(MAD2) presents a bigger bias (and a rms value) when compared with the other GPS sites. This disagreement between the model and the observations at MAD2 station also appears during other days of comparison and reflects the present non-operational situation regarding the limitation in the availability of the real time raw data.

So far, we have used the ZTD times series to monitor the operational forecast of the HIRLAM model. However, the system that we have developed at the IEEC can be extended to other meteorological models.

Table 4.1: Mean and rms PW for the entire 15-day period and during the two front passages at all GPS stations.

		PW Dec. 2-15		**PW Dec. 4-6**		**PW Dec. 12-14**	
station	**height (m)**	**Mean (mm)**	**rms (mm)**	**Mean (mm)**	**rms (mm)**	**Mean (mm)**	**rms (mm)**
Escorial	1026	12.2	3.6	11.3	2.1	17.5	2.0
IGNE	715	13.2	3.9	12.4	2.5	18.7	2.4
Robledo	777	12.9	3.7	11.6	2.3	18.4	2.2
Valdemorillo	794	13.1	3.7	12.2	2.3	18.5	2.2
Villafranca	596	14.7	3.9	13.7	2.4	20.5	2.5

Table 4.2: PW-modeled and PW estimates bias and rms for the entire period of the experiment.

station	**PW bias (mm)**	**PW rms (mm)**
Escorial	-0.7	2.0
IGNE	0.2	2.1
Robledo	-0.4	2.2
Valdemorillo	-0.4	1.9
Villafranca	-0.9	2.0

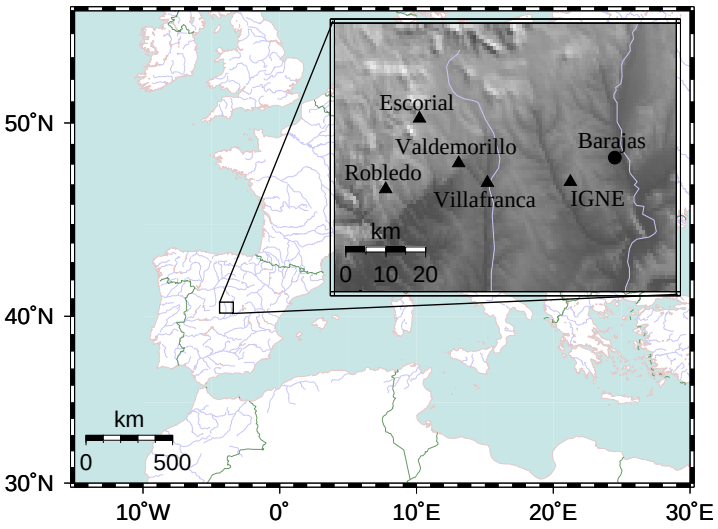


Figure 4.1: Geographical location of the GPS (triangles) and radiosonde (circle) sites involved in the experiment. The geoid altitude of the sites is as follows: Barajas (633 m), Escorial (1026 m), IGNE (715 m), Robledo (777 m), Valdemorillo (794 m) and Villafranca (596 m).

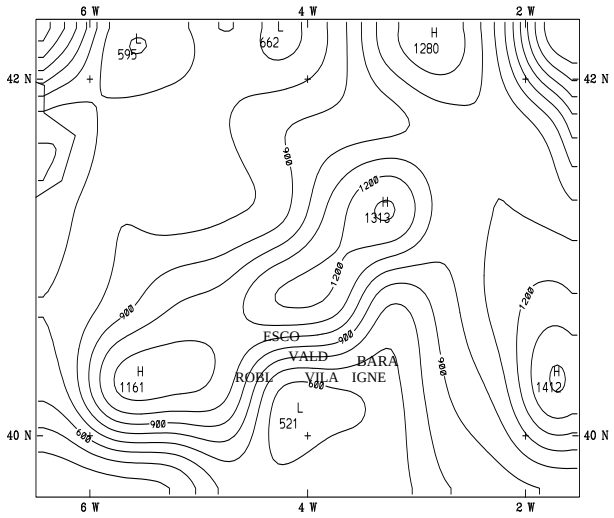
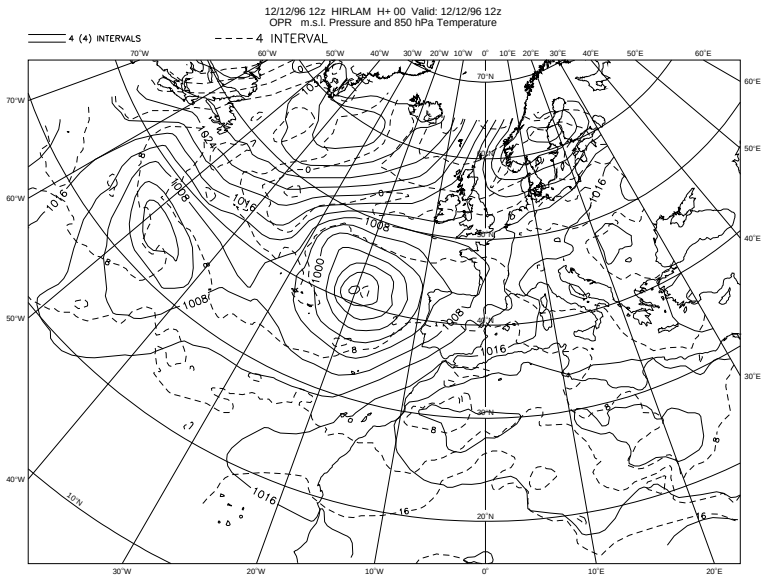


Figure 4.2: Contour plot of the topography of the region and the location of the five GPS stations. Altitude is in meters above sea level. (BARA: Barajas, ESCO: Escorial, ROBL: Robledo, VALD: Valdemorillo, VILA: Villafranca.)

a)



b)

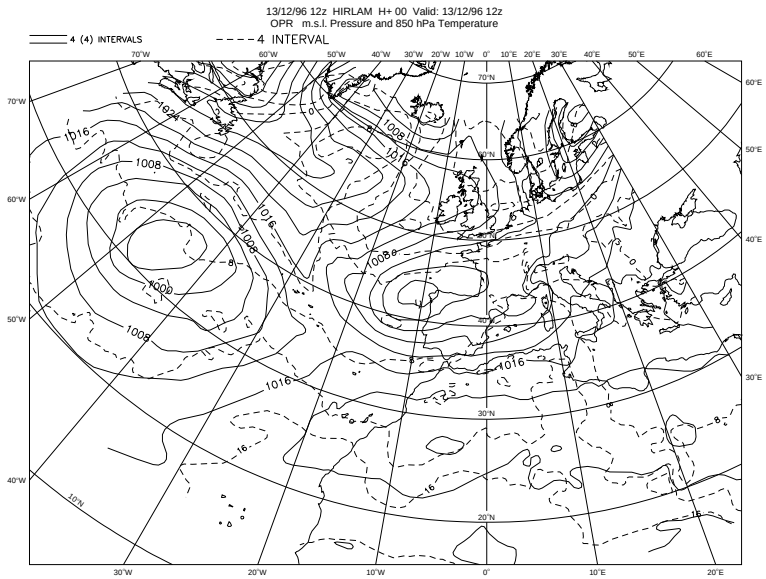


Figure 4.3: Low-resolution, HIRLAM/A maps of low level atmospheric flow for (a) 12 UTC 12 Dec, (b) 12 UTC 13 Dec, and (c) 00 UTC 14 Dec. The contours represent mean sea level pressure (solid) and temperature at 850 hPa (dashed).

c)

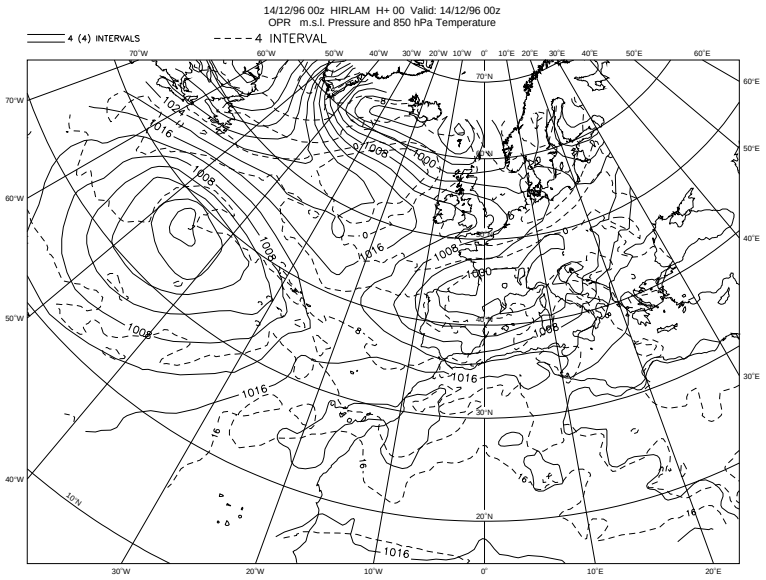


Figure 4.3. (Continued)

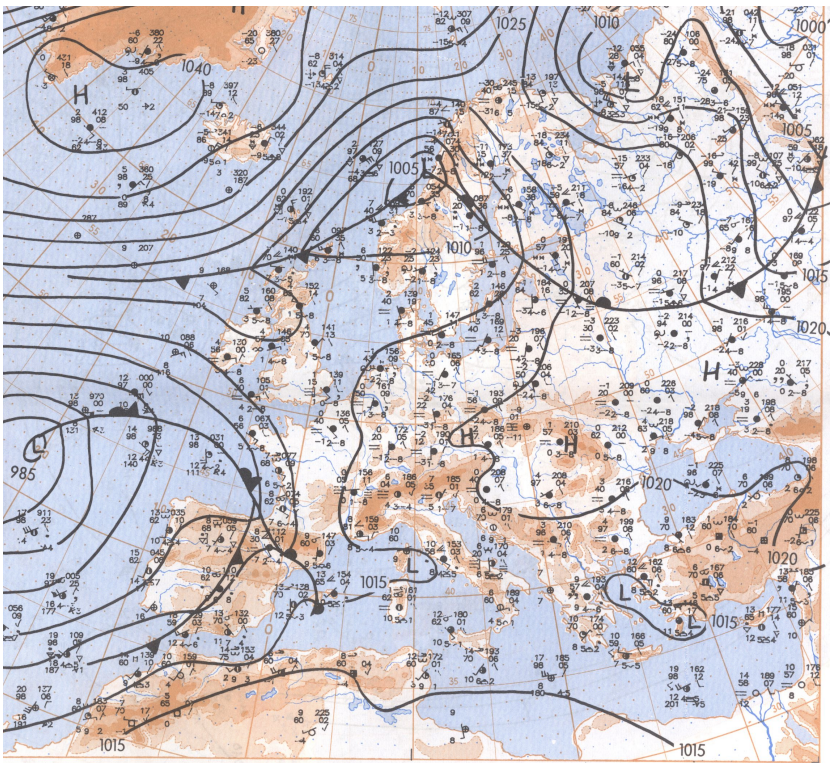


Figure 4.4: Surface analysis [from the Deutscher Wetterdienst (DWD), Germany] of atmospheric flow for 00 UTC 12 December 1996

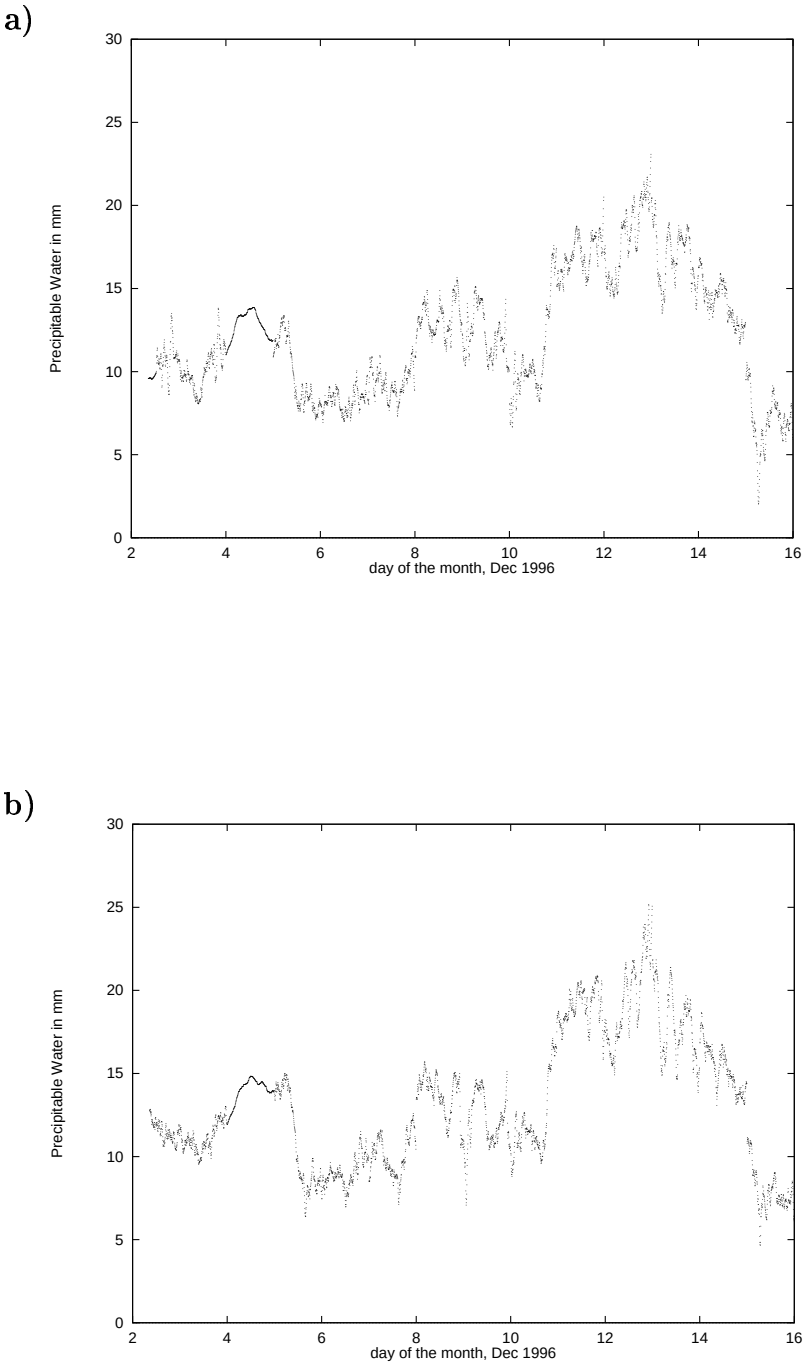
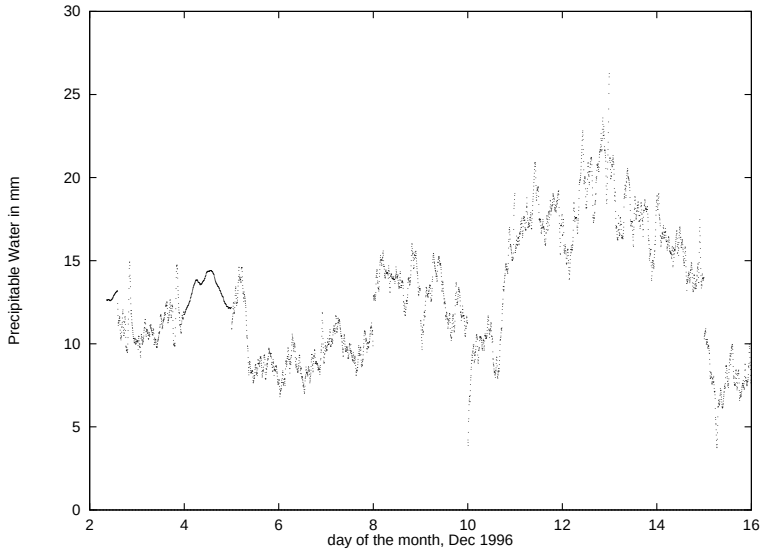


Figure 4.5: GPS-derived precipitable water as a function of time for, from top to bottom: (a) Escorial, (b) IGNE, (c) Robledo, (d) Valdemorillo, and (e) Villafranca.

c)



d)

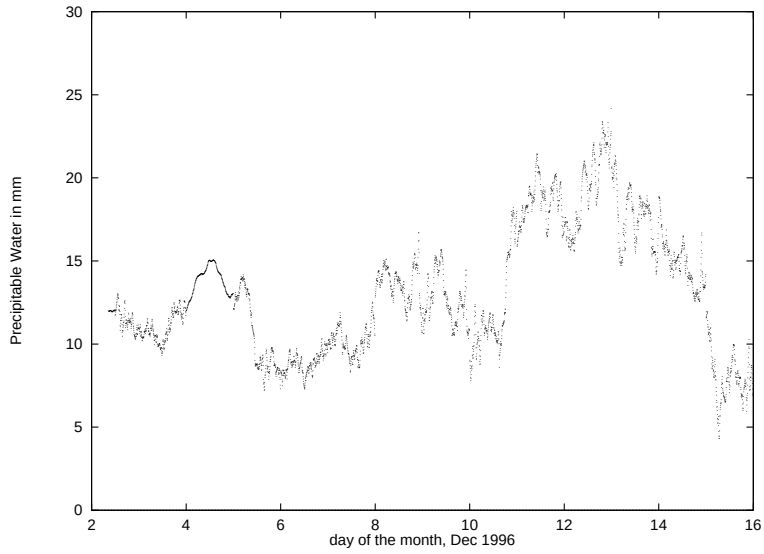


Figure 4.5. (Continued)

e)

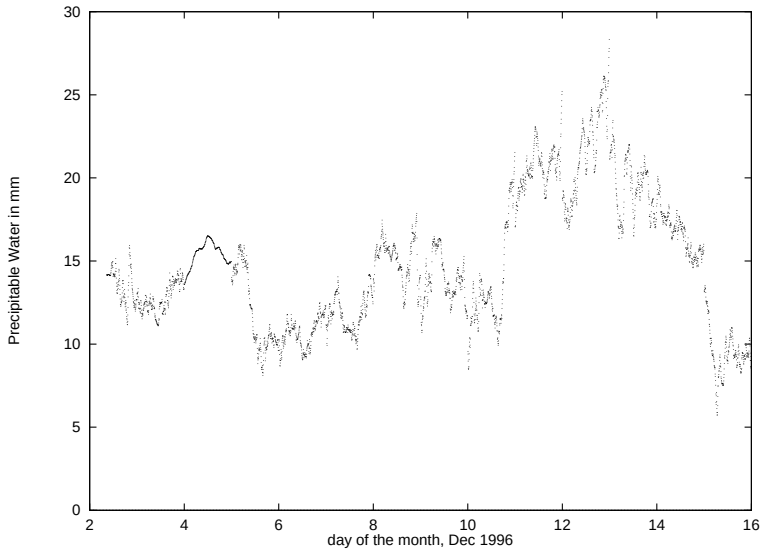


Figure 4.5. (Continued)

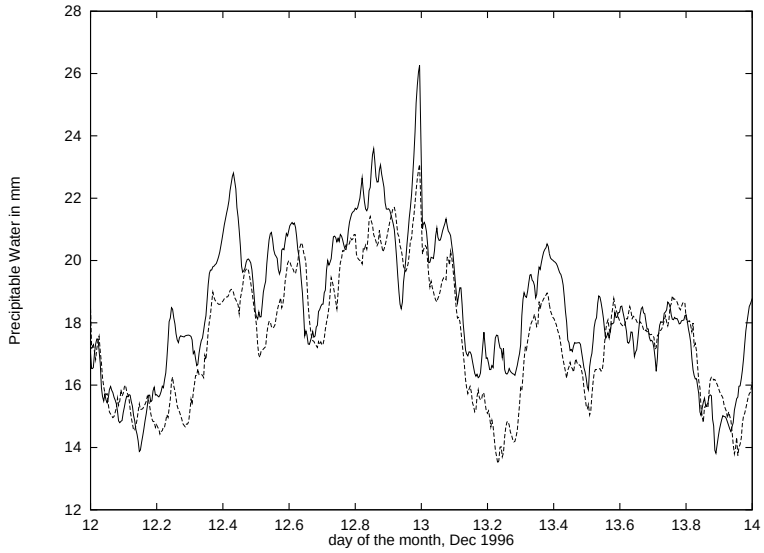


Figure 4.6: The 30-min averaged GPS-derived precipitable water as a function of time at Robledo (continuous line) and Escorial (dashed line) during the passage of a second frontal system in 12-14 Dec.

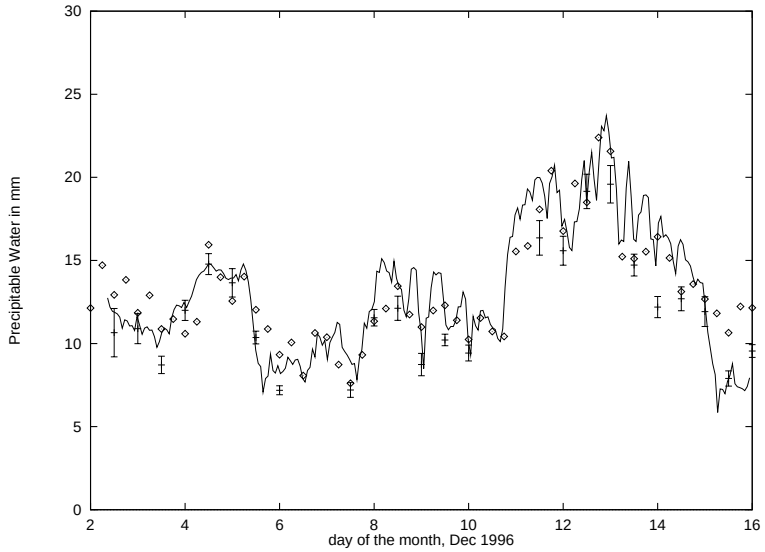


Figure 4.7: Estimates of PW from GPS measurements (continuous line), from HIRLAM calculations (diamond) and from integrated radiosonde profiles (stripe with error bars) vertically integrated at IGNE site, as a function of time. The standard deviations of the PW estimates are about 1 mm (see text) and have not been plotted for clarity. The uncertainties on the radiosonde PW estimates due to the meteorological sensors are shown with their error bars. The HIRLAM PW uncertainties are about 1.7 mm and are not shown.

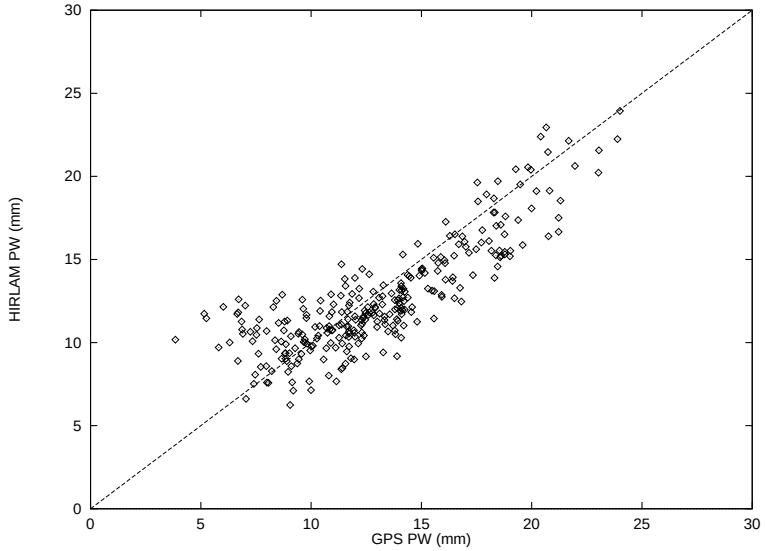


Figure 4.8: Comparison of PW estimates against PW modeled for all the network and for all the campaign. The dashed line shows the results of perfect correlation.

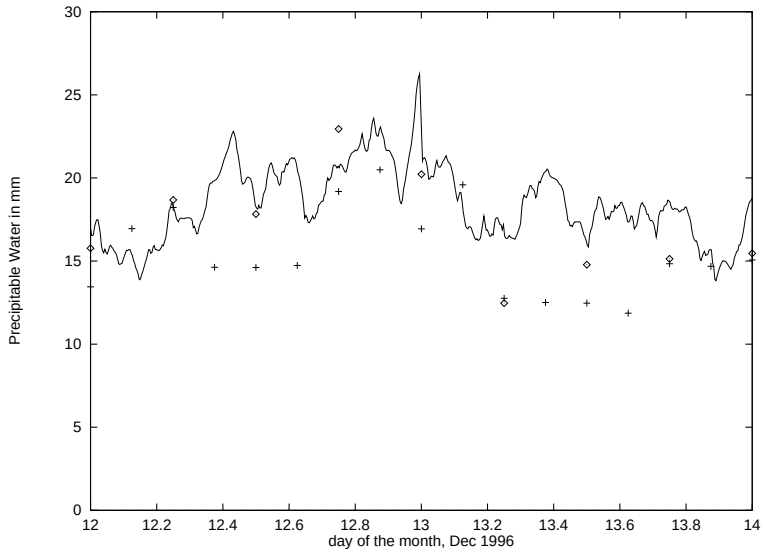


Figure 4.9: The Robledo station PW estimates (continuous line) vs PW modeled with the HIRLAM/A (diamond) and HIRLAM/F (cross).

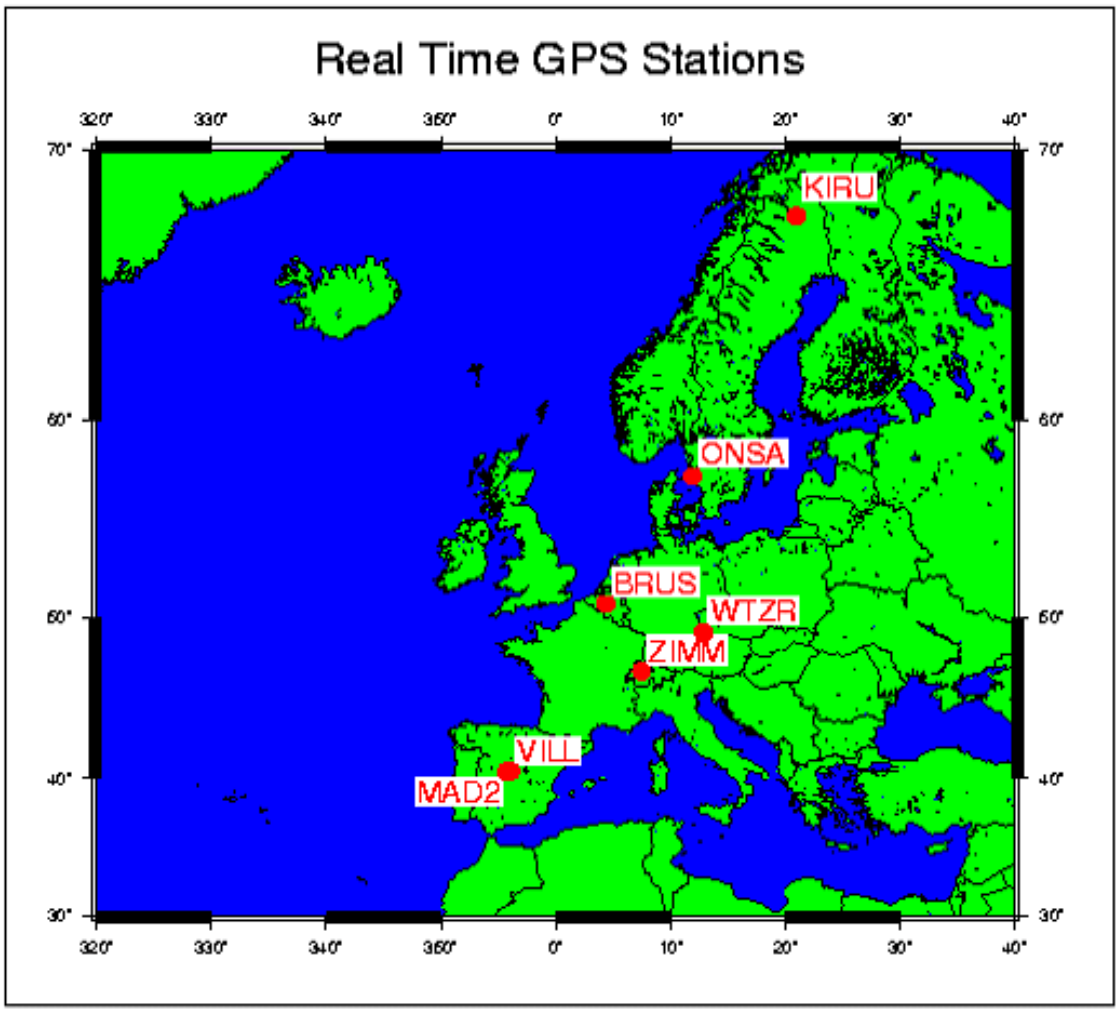


Figure 4.10: Location of the near-real time GPS sites.

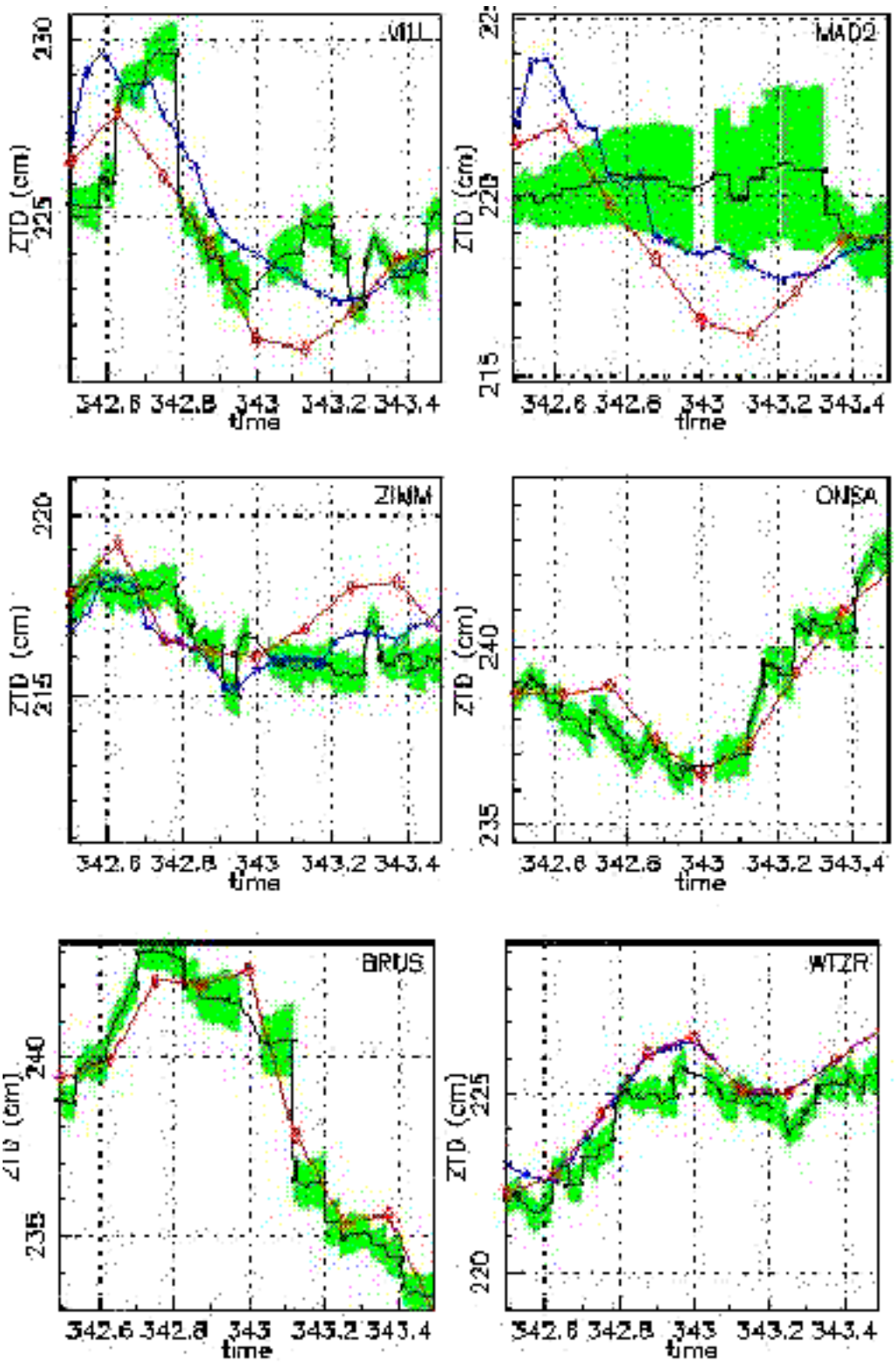


Figure 4.11: The GPS ZTD estimates (black with error bar in green) vs the ZTD modeled with the HIR (blue) and OPR (red) modes of the HIRLAM model as a function of time (day of the year).

4.2 The MM5 model: September 1999

In the previous section, we have performed a comparative study of the spatial and temporal distribution of the water vapor content as obtained with GPS and modeled with HIRLAM on a small scale (5-50 km). However, a low resolution hydrostatic model was used and the focus was mainly addressed to analyze the GPS-derived PW in synoptic situations. It is therefore necessary to further investigate if these conclusions found in the previous work still apply when high resolution models are used, i.e. non-hydrostatic models (Cucurull and Vandenberghe 1999), and for situations where trigger mechanisms such as surface heating or orographic uplift can lead to sudden changes in the water vapor distribution (Cucurull et al. 2001b).

The western Mediterranean is frequently affected by situations connected with heavy rainfalls over localized areas (less than 50 km) and during a short time (less than two hours). These events are mostly the results of mesoscale convective systems (Llasat and Puigcerver 1992; Ramis et al. 1994; Codina et al. 1996; Romero et al. 1998) which are closely related to the land surface heat flux conditions and the topography of the area.

The aim of this section is to use precise and continuous measurements of the water vapor column by means of the GPS technique to study the evolution of a mesoscale convective system at the northeast coast of the Iberian Peninsula. This region has complex orography and the surface conditions are very heterogenous (from heavily urbanized areas to forest and bare soil). The GPS observations are compared to the ZTD values calculated by means of fine-scale modeling for the same situation by using the MM5 Modeling System (Anthes and Warner 1978; Dudhia 1993; Grell et al. 1994).

The case study is carried out on 14 September 1999 during the evolution of a mesoscale convective system which produced a large amount of precipitation in the area. In order to asses the dependence of the GPS data on the meteorological situation, an additional day with absence of precipitation and low moisture variability was selected (10 September 1999).

A continuous monitoring of the ZTD is carried out by five GPS receivers located at several sites in the NW Mediterranean region. The contributions to the ZTD fluctuations are analyzed in terms of the two components (Davis et al. 1985; Bevis et al. 1992): the ZHD and the ZWD. The ZHD is the largest term and can be accurately calculated if measurements of surface pressure are available (Saastamoinen 1972). The ZWD is associated with the atmospheric water vapor and is very difficult to model

because this variable is highly variable in space and time (Emardson et al. 1999). From the ZWD, one can derive the PW variable (1 cm of ZWD converts to around 0.15 cm of PW) (see Section 3).

4.2.1 Meteorological description

The meteorological situation under study was the result of the interaction of two atmospheric phenomena. On 14 September 1999 around noon a thermal low was well developed above the center of the Iberian Peninsula. The Meteosat imagery at 12 UTC 14 September 1999 is shown in Figure (4.12). The origin of the thermal low is a synoptic situation characterized by a low pressure area over the North of Africa, Mediterranean sea, and eastern coast of the Iberian Peninsula which is reinforced by the intense heating of the land surface in the previous days. During the afternoon this system moved eastward due to the passage of a sharp and deep trough which crossed the Iberian Peninsula with a NW-SE jet stream behind the trough axis. This trough simultaneously intensified the cyclogenesis over eastern Iberian Peninsula and the Mediterranean sea. At 18 UTC the low center was situated above the eastern coast (see Figure 4.13). As a result, moist and warm air was advected from the Mediterranean into the NE of the Iberian Peninsula and the Gulf of Lion. The radiosonde measurements showed at 00 UTC on September 15 a nearly saturated atmosphere from above 850 hPa up to 200 hPa. The synergism of the surface phenomena, low level advection of moist and warm air, and upper conditions, through transporting cold air and intensifying the cyclogenesis and convection over the area, was the origin of the active mesoscale convective system over NE Iberian Peninsula.

The rainfall precipitation measured was 80 mm in the East of Catalonia between 00 UTC 14 September and 00 UTC 15 September. At several stations this precipitation fell in a short time and with high intensity (for instance, in the city of Barcelona 55 mm within 40 min with 3 cm hails). The intense rainfall that produced the mesoscale convective system over Catalonia from 00 UTC 14 September to 00 UTC 15 September is depicted in Figure (4.14) [obtained from the Servei de Meteorologia de Catalunya (SMC), Spain]. The map is produced from a Cressman objective analysis of 81 meteorological stations available in the area of interest. The minimum, average and maximum distances between these sites are 0.02°, 0.12° and 0.33°, respectively. The analysis uses a grid size of 0.1° and seven iterations with radius of 7,6,5,4,3,2 and 1 times the grid size value (0.1°), respectively.

To compare the GPS observations with the MM5 results under different meteoro-

logical conditions, we have selected a second meteorological event characterized by a cloudless sky, high temperature, and a high pressure situation. The selected dry and relative low values of water vapor period corresponds to 10 September 1999.

4.2.2 GPS data processing

The GPS network consists of five Trimble 400SSI GPS receivers operated by the Institut Cartogràfic de Catalunya (ICC, Spain). The names and heights above sea level of these stations are shown in Table (4.3). These GPS sites form baselines ranging in length from about 100 to 350 km with maximum altitude difference between GPS sites of about 2400 m. The geographical location of the GPS sites covers from around 0° to 4°E and from 40°N to 43°N (see Figure 4.15).

We use the GPS precise orbits and clocks as well as consistent earth-rotation parameters provided by the International GPS Service (IGS) together with the GIPSY/OASIS-II (version 4) software package (Webb and Zumberge 1993) to estimate ZTD (taken every 15 min) at the five GPS sites with a formal error of 0.5 cm. This software uses a stochastic filter to provide time-dependent estimates of the atmospheric delays for each site.

4.2.3 Model simulation

The NCAR/Penn State MM5 Modeling System is used to simulate the ZTD variable. The MM5 is a primitive equation, finite-difference based non-hydrostatic mesoscale model. It uses a sigma vertical coordinate defined in terms of a time independent reference pressure p_0,

$$\sigma = \frac{p_0 - p_t}{p^\star}; \quad p^\star = p_s - p_t \tag{4.2}$$

where p_s and p_t are the surface and top reference pressures of the model, respectively. MM5 variables are coupled with $p^\star$ which is a time-independent 2-dim model constant.

We set up three (2-way nested) domains with grid distance ranging from 54 km down to 6 km. At the finest domain the grid dimensions are 82 grid points in the north-south direction, 97 in the east-west direction, and 24 vertical sigma levels (see Figure 4.13). A 5 min topographic source is used for the third domain. To investigate the impact of modeling with finer grid resolution and more accurate topography source, we have defined five additional nested domains, each one centered at the

location of a GPS receiver. In these fine domains, a grid resolution of 2 km is prescribed with a mesh of 52 x 52 x 24, and a 30 sec topography source. The physical options used are: the high-resolution Blackadar parameterization of the planetary boundary layer (PBL), multi-layer soil model, the simple scheme of Dudhia (1993) for explicit moisture parameterization, and the clouds are explicitly solved for the smaller domains (grid resolutions of 6 km and 2 km).

The initial and boundary conditions are provided by the ECMWF analyses at 00 UTC 10 and 14 September 1999 in order to study the two different meteorological situations. Both simulations are integrated for a 24 h period. ZTD values are calculated (15-min interval) at all grid points by adding the simulated ZHD and ZWD components. A bilinear interpolation from the four closest grid point values is used in the horizontal in order to estimate the simulated ZTD at the GPS sites.

4.2.4 ZTD differences due to the meteorological situation

The development and evolution of a mesoscale convective system is studied in terms of the ZTD differences at a maritime (EBRE) and a mountain (LLIV) stations. First, we have used the model resolution of 6 km and a topography source of 5 min to simulate the ZTD variable at these GPS sites. In the simulations, the same parameterization of physical processes is used.

Figure (4.16) shows the 3-h interval rms fluctuations of ZTD and the PW modeled by means of MM5 for 14 September 1999. (All the rms values are calculated around the mean value). The increase in the rms value at selected sites during the second half of day 14 (more than 1 cm of ZTD) reveals a high variability of the ZTD variable during this period which corresponds to the activity of the mesoscale convective system. The fluctuations of the ZTD are mainly the ZWD contribution due to the variations in the water vapor content. This can be also observed in the figure where the modeled PW variable has been represented at the same sites and for the same period. The rise and high variability of the PW starting at about 12 UTC 14 September correlates with the increase of the rms value and with the large precipitation recorded in the area. During 10 September (clear day), there is almost no variation in the ZTD fluctuation, and the average rms values for LLIV and EBRE sites are 0.3 cm of ZTD. A similar behavior is found for the rest of the GPS stations.

We next turn our attention to compare the differences between the GPS-derived ZTD values and the ZTD simulations using MM5 with the grid resolution of 6 km. The frequency distribution of these differences for 14 September are presented in

Figure (4.17a). The histogram for LLIV station (grey) shows its maximum frequency between 5 and 6 cm of ZTD difference with an average bias of 5.5 cm (observations values are higher than the MM5 results) and a rms value of 1.4 cm. The coastal site error distribution (white) presents an average bias of 1.0 cm of ZTD (observations are higher than the model simulations) and a rms of 2.2 cm of ZTD. In that case, the data distribution is more centered at zero but the rms of the histogram is higher than for the mountain station.

In order to study the impact of the meteorological conditions on those frequency distributions, Figure (4.17b) shows the same analyses of distribution but for 10 September 1999. Although there is still some positive bias between the observations and the modeled ZTD (average bias of 4.5 cm of ZTD at LLIV and 0.8 cm of ZTD at EBRE), the data distribution is more centered around its mean (rms of 1.0 cm at LLIV and 1.6 cm at EBRE) which corresponds to a lower variability of the ZTD differences between the model and the observations. This is reasonable since we have found that in clear days the evolution of ZTD shows lower variability than in a stormy period. Similar results are found for the other stations treated in the study.

As Figure (4.17) shows, the mountain station has a large bias in 14 September, while a lower bias is obtained for the coastal site. When the clear day is analyzed the average bias decreases by around 20 % of its original value at both GPS stations. This reduction may be due to a decrease of the differences between the observed and modeled ZHD or it may be attributed to a more accurate modeling of the PW variable during the clear day. Measurements of surface pressure at GPS sites are needed to evaluate the former in order to compare the observed pressure data with the values predicted by the meteorological model and thus evaluate the ZHD term. The only ground-based receiver from our GPS network that operates a barometer is CREU station. For this site, the average bias between the observed pressure and the values obtained with the model accounts for around 2 cm of ZHD (modeled values higher) in both meteorological situations. However, there is a slightly decrease in the average PW difference during 10 September 1999. Therefore, the decrease of the average ZTD bias is due to the water vapor contribution term (ZWD), while the hydrostatic component remains unchanged. As a consequence, the reduction of the average ZTD bias accounts for about 1.5 mm and 0.3 mm of PW at LLIV and EBRE stations, respectively.

The variability of the water vapor content slightly reduces the differences between the ZTD observations and the modeled values, but it is not the main source for such

a bias. In the next section, we analyze this average bias by simulating the mesoscale convective system with a finer grid and topography database.

4.2.5 ZTD differences due to the topography and model resolutions

We have selected two model configurations to analyze the impact of the topographic and modeling resolutions on the ZTD differences between observations and modeled values. We have used a topography source of 5 min for the lower model resolution tested (6 km) and the fine data set of 30 sec for the higher grid resolution (2 km). The physical processes were parameterized in the same way for both cases.

The elevations from the different topographic sources are shown in Table (4.3). The mountain stations ESCO and BELL are better represented by the 30 sec resolution source when compared to the GPS heights, but we do not find the same situation for the coastal sites and LLIV. Both, the 5 min and 30 sec landuse categories misrepresent CREU station, which is described as a water body. It should be mentioned that this station is situated on a cliff. The lower resolution topography results are more appropriate than the 30 sec source to model the height of EBRE station. For LLIV, the difference between the station height and the elevation modeled using the 5 min and 30 sec topography sources is the same. Therefore, we can not expect an improvement on the ZTD simulation with the use of a finer terrain database.

Figure (4.18) shows the frequency distribution of the ZTD differences between observations and modeling simulations at ESCO station for 14 September. The grey histogram corresponds to the lower topographic resolution, whereas the white histogram depicts the ZTD differences calculated with the finest data set. The average bias of these ZTD fluctuations is largely reduced when the 30 sec topography source is used. Although the frequency of the distribution is similar in both histograms, the ZTD values simulated using the finest topographic data agree better with the observations than when the coarse data resolution is used. This means that the impact of the topography is to reduce (for the fine source) or to increase (for the low resolution source) the average ZTD bias (mainly through the ZHD contribution). However, it does not affect the variability (rms) of these differences.

This result shows that the ZTD value strongly depends on the height of the GPS station, but its variability is mainly driven by the meteorological conditions as was shown in the previous section. We have found similar results for all the GPS sites analyzed in this study (Table 4.3). The table shows the average bias

and rms values for the two topographic sources tested. The average bias decreases largely at mountain stations when the fine topography is used. The resolution of the selected source slightly affects the mean difference at CREU, which is consistent with the inaccurate height given by the topography data source. Although the lower topography resolution source simulates slightly better the orography of EBRE site, a smaller reduced bias is obtained with the fine data set. This is mainly caused by a more appropriate vegetation category from the 30 sec source and the use of a higher model grid resolution.

The ZTD variable depends on the atmosphere layer depth between the GPS satellite and the receiver. This value increases for denser atmospheres and longer trajectories of the signal. As a consequence, for a given atmospheric profile, the ZTD variable will be larger at the receivers located at the sea level in comparison to those situated on mountain tops. From the table, all the stations which have a model height below (above) the GPS sites present a larger (smaller) simulated ZTD than the observed values. For instance, ESCO has a real value of 2458 m compared to the 2310 m given by the 30 sec data source. As a result of this difference, the ZTD bias is negative (-7.1 cm).

The increase of the ZTD differences between observations and modeled values with the use of the 5 min topography resolution when compared to the finer data set may be explained in two ways: (a) it may be caused by a bias in the station pressure entering in the ZHD calculation due to the low terrain resolution or (b) it may produced by innacurate predictions of the PW content. In order to analyze these two different contributions, the average ZTD difference between the use of the low and high topography resolutions are shown in Table (4.4) for all the stations.

The table also includes the hydrostatic contribution to the ZTD bias due to innacurate modeling of the surface pressure variable due to the use of the low terrain data set. The highest differences in the ZHD variable between both topographic sources are found for the mountain stations, which are the stations with higher differences between 5 min and 30 sec derived elevations. Since the elevation of CREU is always inaccurately given by both data sets, no differences are found for the simulated ground pressure variable. As opposite to this coastal site, all the ZTD bias is attributed to the hydrostatic term at EBRE site.

The remaining average difference of ZTD when using different topography and model resolutions can be attributed to the moisture content of the atmosphere. The average difference between the PW simulations with the use of a low and fine terrain

data sets is also shown in Table (4.4). Generally, the lower model resolution results in an overestimation of the PW variable compared to the finer simulation during 14 September. The opposite situation is only found at LLIV station, which also showed a negative biased ZTD difference. We also note that the average PW difference between using the lower and higher topography sources is larger at mountain stations and ranges between 1 mm and 2 mm of PW.

We next investigate the hydrostatic and moisture contributions to the average difference between the GPS-derived ZTD and the values simulated with MM5 (see Table 4.3).

4.2.6 Corrections to remove the ZHD bias

We have already shown the bias in the ZTD differences that is caused by using topographic sources of different resolutions. The major part of this bias comes from the inaccurate calculation of the hydrostatic component (ZHD). In the following, we propose a procedure to correct for this bias.

The modeling of the ZHD using the coarse and fine topography data sources is used to infer the hydrostatic component of the average ZTD difference between observations and model simulations. This component is estimated by analyzing the differences between the pressure values predicted for the locations of the GPS sites as a function of the height increment. The pressure calculation is mainly governed by the hydrostatic equation (non-hydrostatic effects are of second order) and depends on the physical options selected in the model, on the latitude, and on the weather situation. The correction to remove the ZTD difference bias due to the ZHD contribution can be easily applied to other meteorological models.

Figure (4.19) compares the average ZTD difference between the observations and the simulated values against the difference between the stations and modeled heights for the topographic sources under study. A straight-line fit to these data yields a slope of -0.024 ± 0.003 (for the 5 min source, in asterisk), and -0.013 ± 0.010 (for the 30 sec source, in diamond). The χ^2 (per degree of freedom) are 1.3, and 0.7, respectively. In the figure, the dashed line fits the average ZHD differences between the model simulations with the use of the 5 min and 30 sec topography sources as a function of the altitude increment Δz (in cross) which are summarized in Table (4.4). This line yields a slope of -0.0225 ± 0.0004 and gives the ZTD variation as a function of the altitude increment assuming that the atmosphere is in hydrostatic equilibrium. For instance, ESCO has an elevation of 2458 m compared to the 2097

m given by the 5 min topography source. As a consequence, $\Delta z = 361$ m and the expected average ZHD difference between the observations and the model simulations is 361 x $(-0.0225) = -8.1$ cm.

The main difference between the ZTD derived from GPS and MM5 arise by the inaccurate values of the topography. The contribution to the average ZTD difference by assuming hydrostatic equilibrium is summarized in Table (4.5) for all the stations. The largest values of the derived ZHD differences are found at mountain stations for all the topographic sources tested in this study. As expected, no differences are found at CREU station between both terrain data sets. The value obtained at this coastal site (-1.9 cm of ZHD) is consistent with the average ZHD derived from the differences between the modeled and observed surface pressure values (see Section 4.1). This good agreement confirms the validity of the slope of -0.0225 used to infer the average ZHD difference from estimates of ZTD differences.

The remaining average bias of ZTD may be explained in terms of the high PW fluctuations due to development and evolution of the mesoscale convective system. This average bias on the moisture field ranges from -1.4 to 3.7 mm of PW for the low resolution terrain database, and from -0.6 to 2.9 mm of PW for the fine topographic data set. In general, the simulations of PW underestimate the moisture content of the atmosphere when compared to the derived observations. The model only predicts a surplus of humidity field at CREU station. The largest PW differences are found at mountain stations, which also showed a higher average ZTD bias. A distinctive element of the table is that although the elevation of the GPS stations is in general better modeled with the use of the fine topography source, the simulation of the PW variable is not always improved by using a higher terrain resolution because it depends on the physical parameterization. The coastal and LLIV stations reduce their average PW bias when using the 30 sec topography source, while ESCO and BELL sites increase the average difference between observed and modeled PW with the fine database by around 1 mm of PW.

4.2.7 Conclusions

The zenith total delay observed and modeled during the occurrence of a mesoscale convective system is studied. The emphasis is placed to analyze the hydrostatic and moisture contributions to the differences between the ZTD observed and calculated in a situation with high water vapor variability. The impact of the topography and model resolutions is studied by doing simulations with different resolutions.

The area under study is the NW Mediterranean region which is characterized by complex orography and heterogenous land surface conditions. The case selected was 14 September 1999 due to the large precipitation recorded in the area and the high variability of the water vapor content. In order to compare the dependence of the ZTD results on the meteorological situation we also analyzed a clear day with high temperature and absence of precipitation.

Significant average bias between the observed and modeled ZTD values are found for the stormy meteorological situation at mountain stations using the lower model resolution tested in the analysis (6 km). Such biases are largely reduced when a fine 30 sec topography source and a higher grid resolution are used. However, the rms values are large for all topographic sources which indicates an increase of the variability of the ZTD differences between observations and modeled values during periods of high variability in the water vapor content. For the clear day, the average bias is only slightly reduced and this is found to be associated to the moisture content of the atmosphere. The rms errors largely decrease at all GPS stations.

Important differences are found in the average ZTD bias when different topography data sets are tested. Probability distributions of the ZTD differences show reasonable data distributions with large biases at mountain stations with the 5 min topography source and the low resolution model. These bias are mainly due to the large height differences between the GPS station elevation and the model topography. The use of a finer topographic data source and a higher grid resolution, results in a reduction of these ZTD differences. This reduction is due to a better simulation of the ZHD since the topography is better represented. We have proposed a correction for the topography-derived error contribution to the ZTD differences between observations and modeled simulations.

The remaining average bias is attributed to innacurate water vapor values given by the representations of the physical processes in the model. It is also found that the model tends to underestimate the moisture content of the atmosphere during the stormy situation analyzed in this study regardless of the topography source used.

Table 4.3: Altitude of the GPS stations and elevation from different topographic sources, together with the average bias and rms (in brackets) between GPS-derived ZTD and MM5 modeled values for all the stations. M: mountain site, C: coastal site.

Source	ESCO (M)	LLIV (M)	BELL (M)	CREU (C)	EBRE (C)
station height (m)	2458	1418	803	83	58
5 min height (m)	2097	1553	652	0	71
ZTD bias and rms (cm)	-7.1 (0.7)	5.5 (1.4)	-2.4 (1.8)	-2.8 (1.3)	1.0 (2.2)
30 sec height (m)	2310	1273	726	0	14
ZTD bias and rms (cm)	-1.4 (0.7)	-2.3 (1.3)	-0.1 (1.7)	-2.6 (1.1)	-0.4 (2.1)

Table 4.4: Average difference and rms (in brackets) between the model simulations of ZTD, ZHD and PW variables with the use of a 5 min topography source and the values obtained with a 30 sec finer terrain database.

station	modeled ZTD bias (rms) (cm)	modeled ZHD bias (rms) (cm)	modeled PW bias (rms) (mm)
ESCO (M)	5.7 (0.5)	4.5 (0.1)	1.8 (0.8)
LLIV (M)	-7.8 (1.0)	-6.5 (0.1)	-2.1 (1.6)
BELL (M)	2.3 (1.0)	1.7 (0.1)	0.9 (1.5)
CREU (C)	0.2 (0.6)	0.0 (0.1)	0.3 (0.9)
EBRE (C)	-1.4 (0.8)	-1.5 (0.1)	0.1 (1.3)

Table 4.5: Contribution of the surface pressure (related to ZHD) and moisture variable (related to PW) to the average ZTD difference between the GPS-derived observations and model simulations with the use of 5 min and 30 sec topography sources.

Source	ESCO (M)	LLIV (M)	BELL (M)	CREU (C)	EBRE (C)
5 min					
ZHD bias (cm)	-8.1	3.0	-3.4	-1.9	0.3
PW bias (mm)	1.4	3.7	1.5	-1.4	1.1
30 sec					
ZHD bias (cm)	-3.3	-3.3	-1.7	-1.9	-1.0
PW bias (mm)	2.9	1.5	2.5	-0.6	0.9

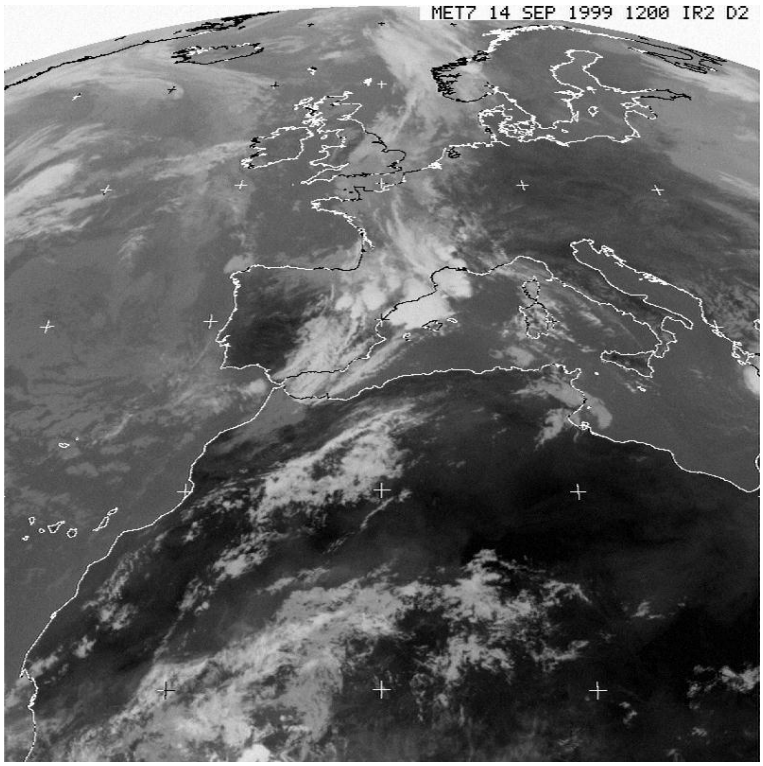


Figure 4.12: Meteosat imagery at 12 UTC 14 September 1999.

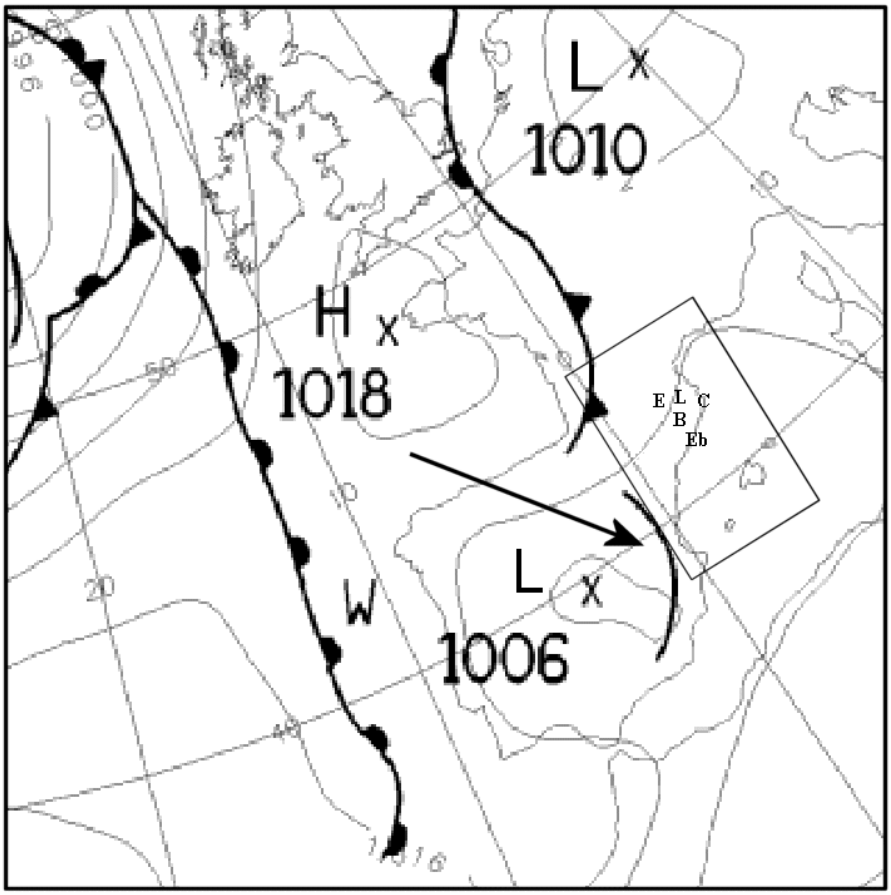


Figure 4.13: Surface analysis (from The Met. Office, UK) of atmospheric flow for 18 UTC 14 September 1999 (the large arrow shows the location of the trough axis), geographical location of the GPS sites (E for ESCO, L for LLIV, C for CREU, B for BELL, Eb for EBRE), and the 6-km grid resolution domain simulated with the MM5 model.

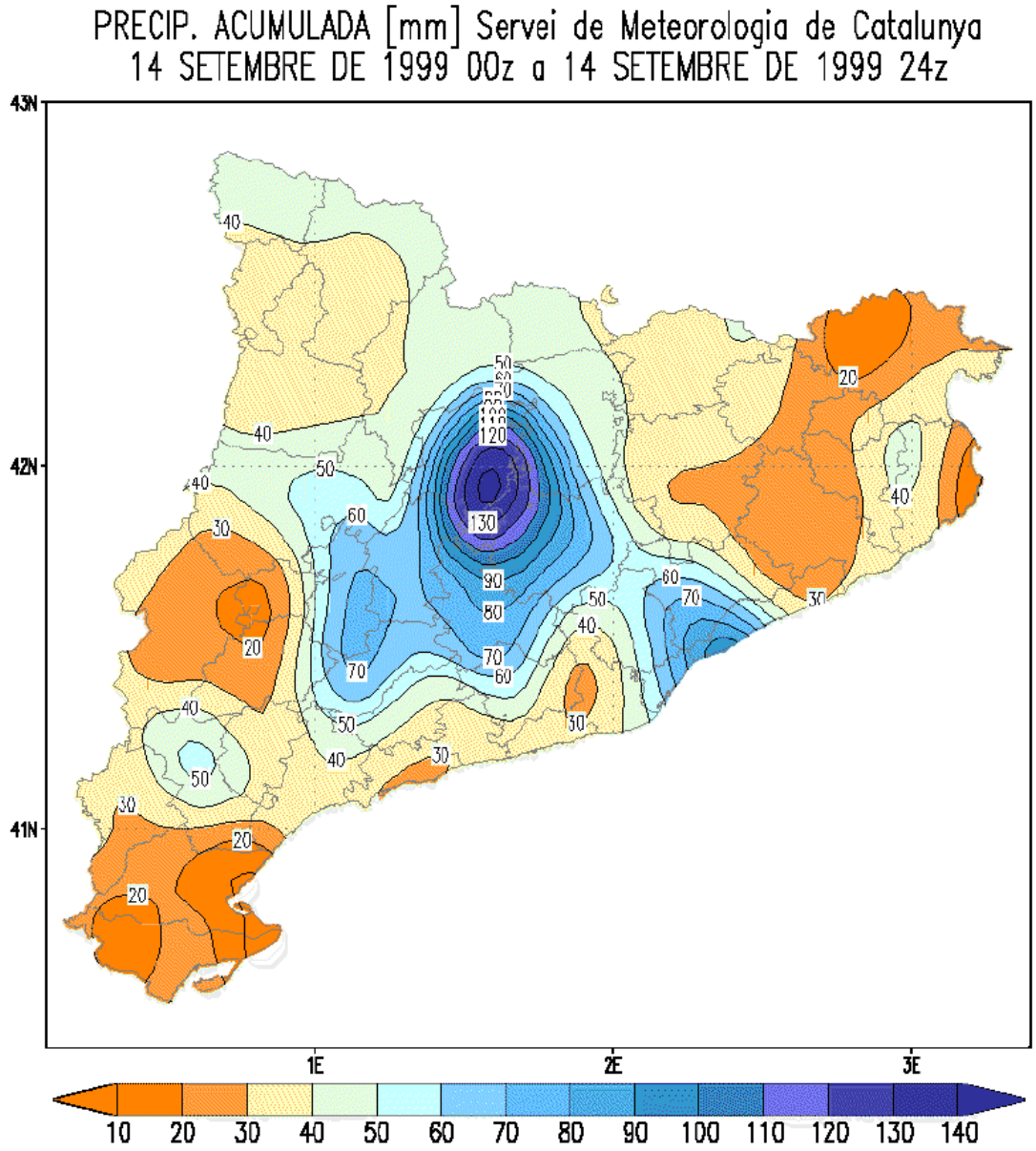


Figure 4.14: Observed accumulated precipitation (mm) over the 24-h period ending at 00 UTC 15 September 1999.

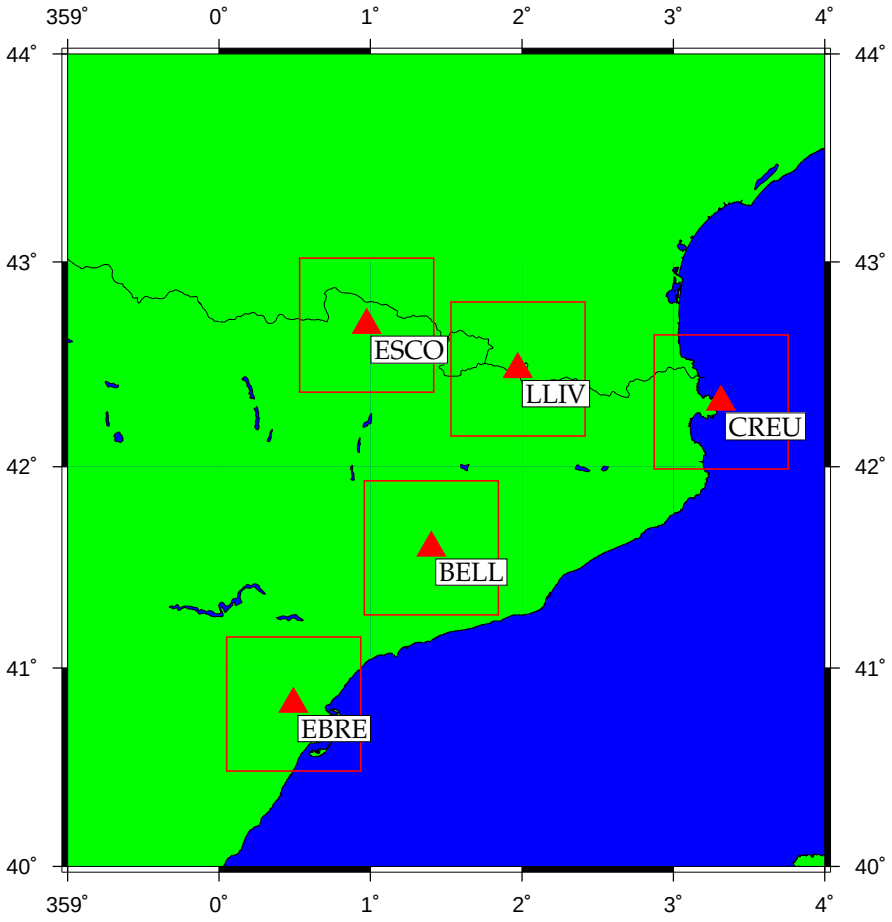


Figure 4.15: Geographical location of the GPS sites (triangles) used in the experiment and the (2-way nested) inner domains simulated with the MM5 model. The geoid altitude of the sites is the following: BELL (803 m a.s.l.), CREU (83 m a.s.l.), EBRE (58 m a.s.l.), ESCO (2458 m a.s.l.) and LLIV (1418 m a.s.l.).

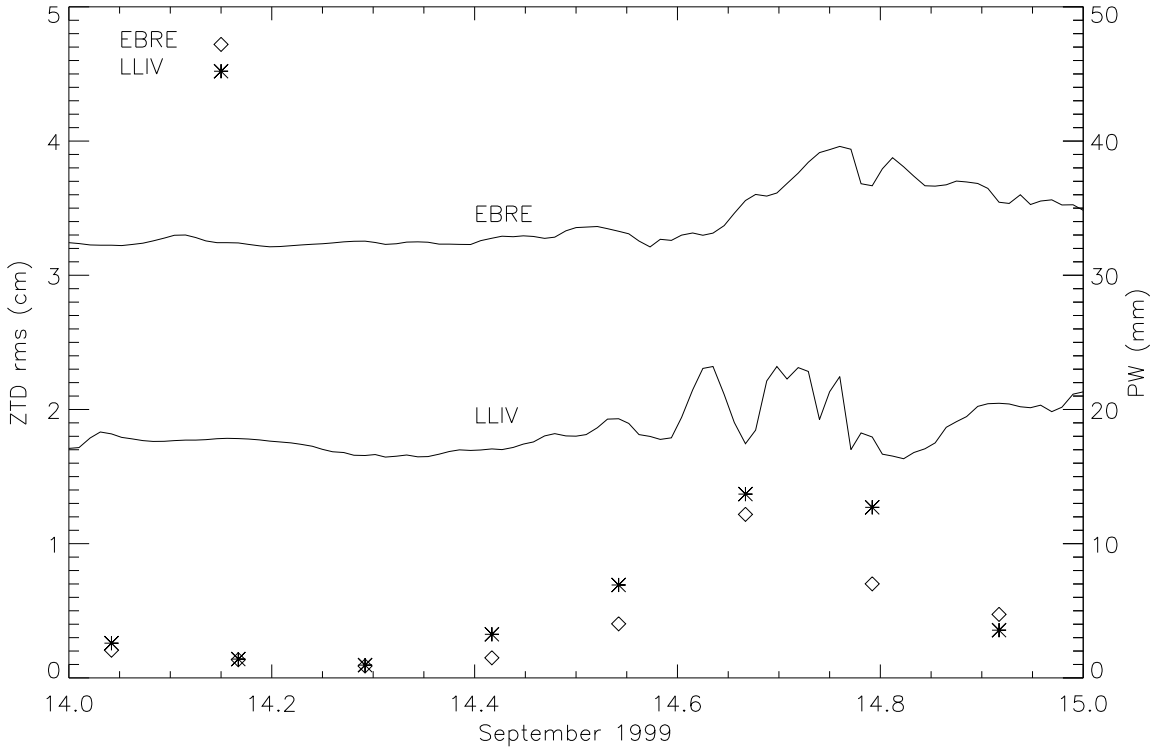


Figure 4.16: 3-hour interval rms ZTD errors ($*$ LLIV, $\diamond$ EBRE) and modeled precipitable water (continuous line) as a function of time at LLIV and EBRE stations.

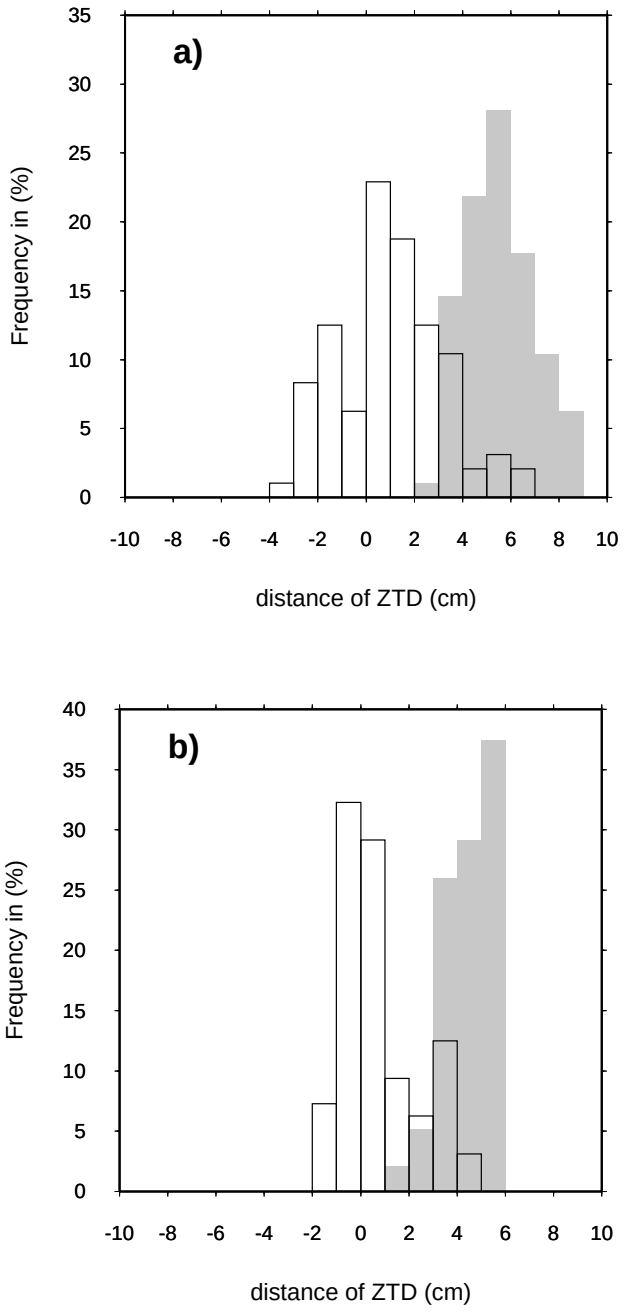


Figure 4.17: Frequency distribution of the differences between the GPS-derived ZTD observations and the modeled values (low grid and topography resolutions) at LLIV (grey) and EBRE (white) stations during (a) 14 September 1999 and (b) 10 September 1999.

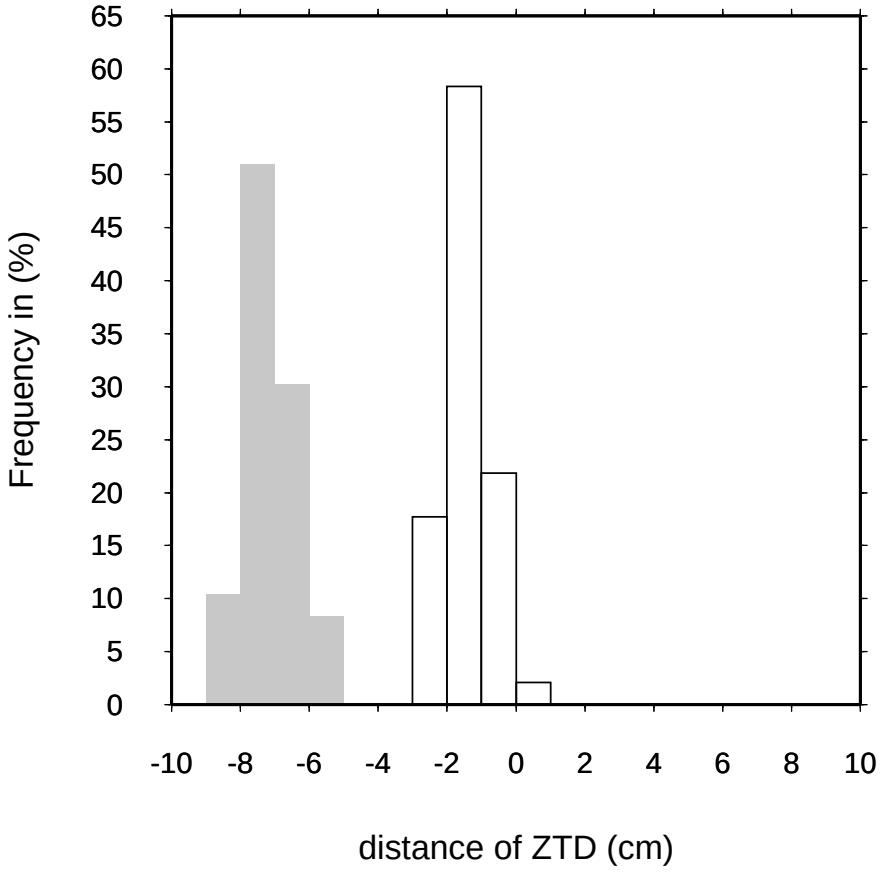


Figure 4.18: Frequency distribution of the differences between the GPS-derived ZTD observations and the modeled values at ESCO station during 14 September 1999 using a 5 min topographic and land-use sources (grey) and 30 sec topographic and land-use data sets (white).

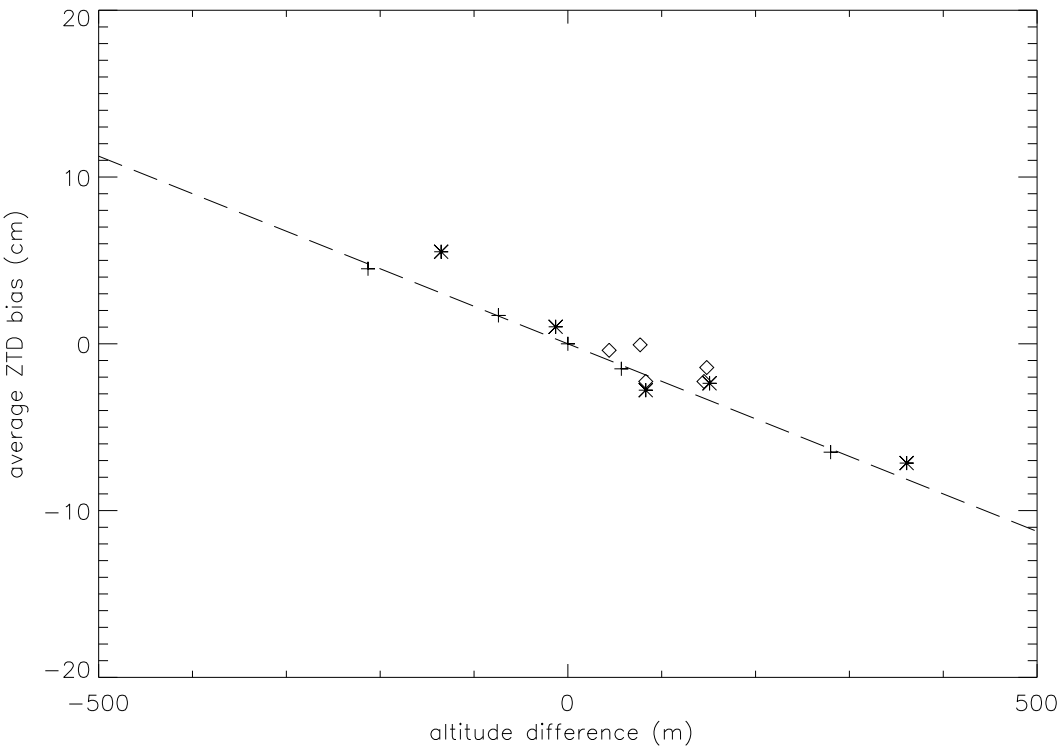


Figure 4.19: Comparison of the ZTD differences between the GPS estimates and MM5 modeled values as a function of the altitude increment (Δz) for 5 min (asterisk), and 30 sec (diamond) topographic sources for 14 September 1999. The dashed line (-0.022 Δz - 0.134) fits the average ZHD differences between the model simulations with the 5 min and 30 sec topography sources as a function of the altitude increment, and shows the ZTD variation assuming hydrostatic equilibrium in the atmosphere.

4.3 The MASS model: September 1998

In this section, we briefly describe the recent implementation of the comparison of the GPS-derived observations with the hydrostatic MASS model (Kaplan et al. 1982; Kaplan and Karyampudi, 1992), which is operational at the SMC. This model includes a multi-nest capability and incorporates a high resolution Blackadar type planetary boundary layer and detailed surface energy and moisture budgets.

At the SMC, the MASS model is run at two different horizontal resolutions, 55 km by 55 km (low resolution run, 55 x 55 grid point distribution), and 10 km by 10 km (high resolution run, 55 x 55 grid point distribution), both with the same 20 levels in the vertical. The coarse domain covers SW Europe and North Africa while the smaller domain is configured to represent the Catalonian area. The higher resolution domain is 1-way nested into the coarser domain. The initial and boundary conditions fort the largest domain are provided by the U.S. National Meteorological Center's Medium Range Forecast model (MRF), and their predictions are used as a first guess and boundary condition for the finer simulation. See Codina et al. (1997) for a description of the implementation of this model in Catalonia.

We have selected a 3-day period to compare the GPS-derived ZTD values estimated at two GPS stations with the values calculated by means of the MASS higher resolution model. The GPS data is processed following the same procedure as described in Sections (4.1) and (4.2). The modeled ZTD values are derived from the PW and surface pressure variables (see Chapter 3). A bilinear interpolation is carried out in the horizontal direction to estimate the model variables at the location of the GPS sites.

The nested simulations were initialized at 00 UTC on 21, 22, 23, and 24 September 1998, and all runs were executed for a period of 21 hours. Large amount of precipitation was recorded during 23 September 1998 due to the passage of a frontal system over Catalonia. Figure (4.20) shows the Meteost imagery at 12 UTC 23 September 1998 and a surface analysis of the atmospheric flow for this day at 12 UTC is depicted in Figure (4.21). The ZTD variable was output from the model at 3-h interval and these values were then interpolated to the ground-based receivers location. In addition, GPS data was available at two GPS stations for days 22, 23, and 24 September 1998.

The results of this intercomparison are displayed in Figure (4.22) for EBRE and BCN stations. In general the model reproduces reasonably well the main trends of the observations. However, the NWP model tends to overestimate the ZTD content

at BCN when compared to GPS data. This feature is also observed at EBRE station on 23 September 1998. Another distinctive element of the figure is the strong decrease of zenith total delay at the beginning of 23 September. The MASS model does not drop the total ZTD amount so sharply. Indeed, the model overestimates this value with respect to what is observed with the GPS data.

These are preliminary results found with the MASS for the Catalonian region. Future work should be addressed to improve the methodology developed with the Universitat de Barcelona (UB, Spain) in order to monitor the MASS short-range forecast of precipitation during the passage of frontal systems.

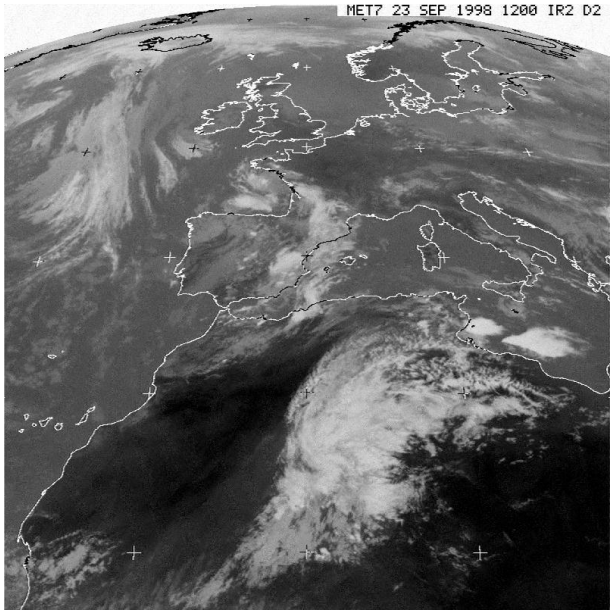


Figure 4.20: Meteosat imagery at 12 UTC 23 September 1998.

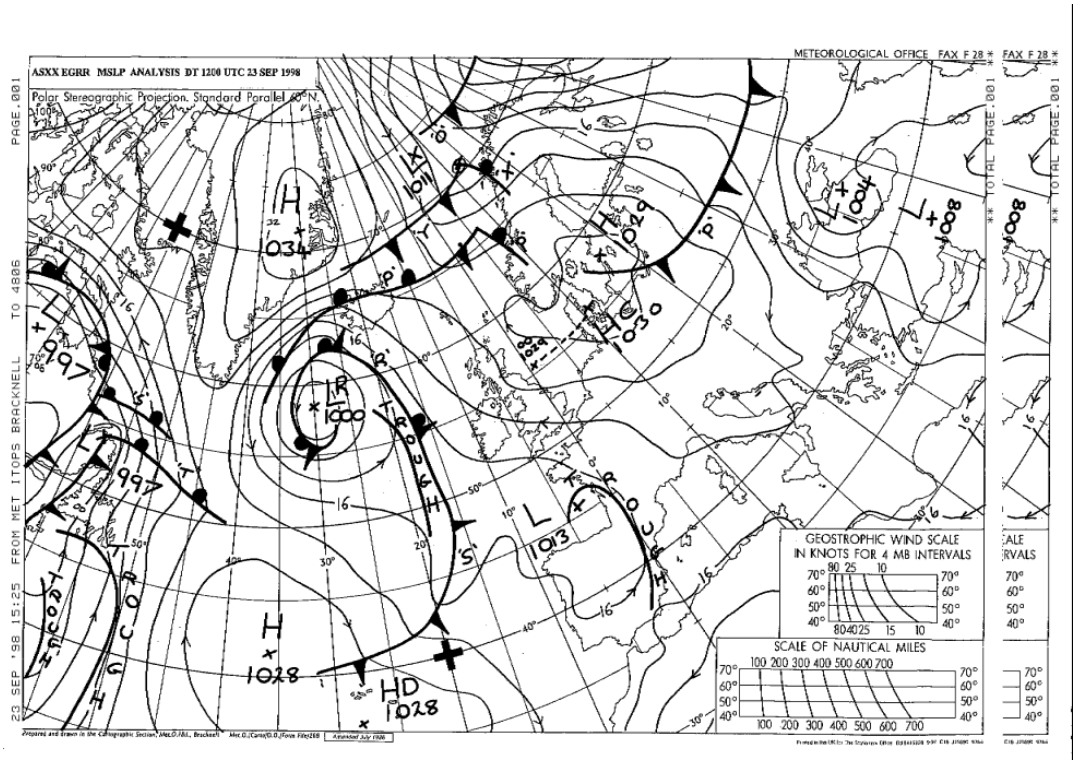
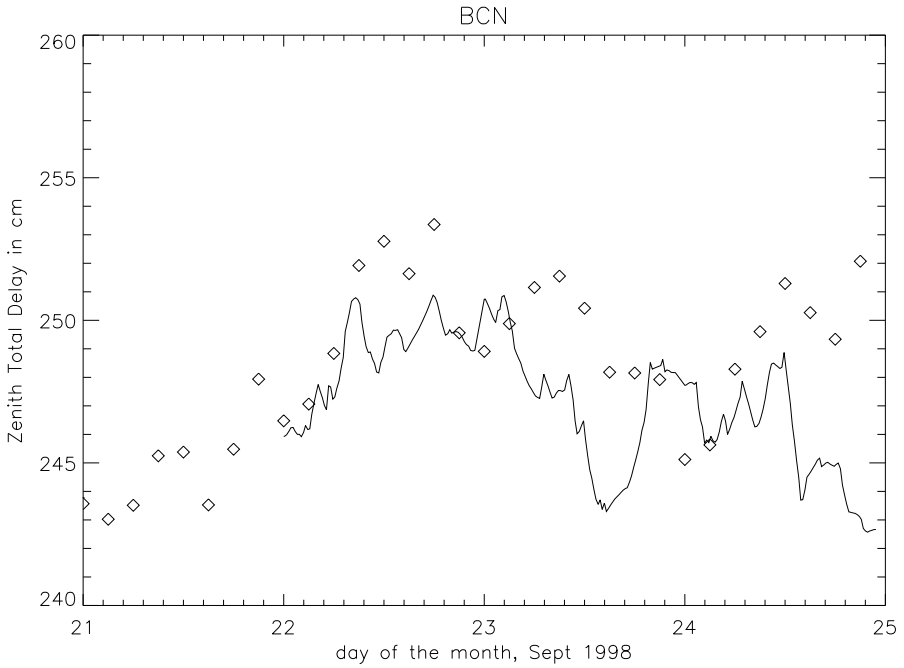
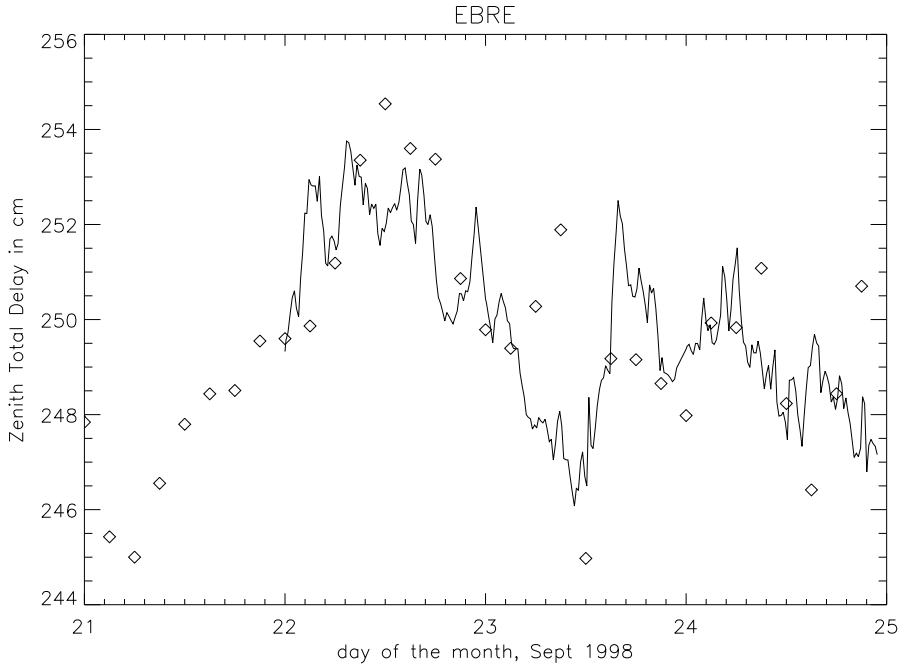


Figure 4.21: Surface analysis (from The Met. Office, UK) of atmospheric flow for 12 UTC 23 September 1998.



a)



b)

Figure 4.22: Estimates of ZTD from GPS measurements (continuous line) and from the MASS model (diamond) at (a) BCN and (b) EBRE stations.

Chapter 5

Variational assimilation of GPS-derived observations: 1D-VAR

We now analyze the impact of the variational assimilation of GPS-derived observations on moisture fields.

As a very useful example, in this chapter we first investigate the 1-dim variational (1D-VAR) retrieval problem by performing a local analysis of one atmospheric column at the location of the GPS ground-based receiver. The procedure is based on the 1D-VAR methodology developed at The Met. Office (Healy 1999)

The 1D-VAR assimilation of ZTD will be extended to a global 4D-VAR assimilation problem in next chapter.

5.1 1D-VAR assimilation of zenith total delay and surface pressure

The PW content of the atmosphere is usually derived from measurements of the zenith total delay, the surface pressure P_s, and the mean temperature T_m defined in (3.8).

The relationship between the PW variable and the derived ZWD trough (3.7) is a function of T_m, which depends on the specific area and season. However, as was seen in Chapter (3), the T_m temperature has been empirically found to be well correlated with surface temperature. For example, Bevis et al. (1992) derived a linear relation

between the surface temperature and T_m, and Emardson and Derks (1999) derived a relation between the zenith wet delay and the precipitable water based on surface temperature, site latitude, and day of the year.

These methods work well for simple models of the atmosphere (Saastamoinen 1972; Davis et al. 1985) and statistical analysis of meteorological radiosonde observations (Emardson and Derks 1999). For instance, Askne and Nordius (1986) presented a formulation in terms of a lapse rate α for the temperature profile and a parameter λ for the water vapor decrease in the atmosphere. They concluded that their formulation was consistent with the Saastamoinen model for a lapse rate of α=6.2 K/km and an exponent λ=3. However, when complicated meteorological situations are analyzed, we expect that more precise information on the atmospheric state is needed. In the following, we study the use of 1D-VAR assimilation technique to retrieve the moisture content from the GPS-derived ZTD.

The objective of the problem is to determine the analyzed state $\mathbf{x_a}$ of the atmosphere which minimizes the functional (2.18). Here, H($\mathbf{x}$) is the nonlinear forward operator which maps the vector defining the state of the atmosphere, in the vertical direction, into the space of the observables.

The non linearity of the problem is treated using a Newton descent method. The required optimal state is found after a few iterations.

5.2 Evaluation of the assimilation of ZTD observations

The expression for the total refractivity of the atmosphere (3.3) can be approximated by

$$N = k_1 \frac{P}{T} + k_4 \frac{P_w}{T^2} \tag{5.1}$$

where k_4 = 3.82 x 10^5 $K^2 hPa^{-1}$ and P is the total pressure. The ZTD is then

$$\mathrm{ZTD} = 10^{-6} \int_0^\infty N dz = 10^{-6} \int_0^\infty (\frac{k_1 P}{T} + \frac{k_4 P_w}{T^2}) \cdot dz \tag{5.2}$$

and applying the hydrostatic equation, this integral can be expressed as

$$\mathrm{ZTD} = 10^{-6} \int_{P_s}^0 (\frac{k_1 P}{T} + \frac{k_4 P_w}{T^2}) \cdot (\frac{R_d T^V}{Pg}) \cdot dP \tag{5.3}$$

where T^V is the virtual temperature in K, g is the gravity acceleration in (m/s^2), R_d is the dry gas constant and P_s is the surface pressure in hPa. The forward operator H defined in (2.18) is then the discrete form of the expression (5.3)

$$\mathrm{H} = 10^{-6} \sum_i (\frac{k_1 P_i}{T_i} + \frac{k_4 P_{wi}}{T_i^2}) \cdot (\frac{R_d T_i^V}{P_i g_m}) \cdot \Delta P_i \tag{5.4}$$

The parameters P_i, T_i, P_{wi} and T_i^V are model layer averages, which have to be related to the model variables. The model variables involved in the computation of the H operator for this case under study are the temperature and the specific humidity at the model pressure levels, and the surface pressure. The water vapor pressure and the virtual temperature are then calculated using the following approximations:

$$P_{wi} = \frac{P_i \cdot q_i}{\epsilon} \tag{5.5}$$

$$T_i^V = T_i(1 + ct \cdot q_i) \tag{5.6}$$

where q_i is the average specific humidity at layer i in (gr/kg), ϵ is the ratio of the molecular weights of water vapor and dry air (ϵ = 621.9 gr/kg) and ct = ($\frac{1}{\epsilon}$ - 1) = 0.608 x 10^{-3}.

In (5.4) g_m is the acceleration due to gravity at the center mass of the vertical column. The value at this point is (Saastamoinen 1972)

$$g_m = 9.784(1 - 0.00266cos(2\lambda) - 0.00028H)\mathrm{m}s^{-2} \tag{5.7}$$

where λ is the latitude of the site and H is the height of the surface above the ellipsoid in km.

The minimization of the cost function (2.18) requires the derivative of the H operator with respect to the model variables (2.19). The vector ($\frac{\partial \mathrm{H}}{\partial x}$) $\equiv \mathbf{H}$ is given by

$$\mathbf{H}^T = (\frac{\partial \mathrm{H}}{\partial T_1}, ..., \frac{\partial \mathrm{H}}{\partial T_n}, \frac{\partial \mathrm{H}}{\partial q_1}, ..., \frac{\partial \mathrm{H}}{\partial q_n}, \frac{\partial \mathrm{H}}{\partial P_{z0}}) \tag{5.8}$$

where P_{z0} is the model variable for P_s and n is the number of vertical layers. The previous derivatives are computed and simplified.

derivative with respect to T_i:

$$\frac{\partial \mathrm{H}}{\partial T_i} = -10^{-6} \cdot (1 + ct \cdot q_i)(\frac{R_d q_i k_4}{\epsilon g_m T_i^2}) \cdot \Delta P_i \tag{5.9}$$

derivative with respect to q_i:

$$\frac{\partial \mathrm{H}}{\partial q_i} = 10^{-6} \cdot [(1 + ct \cdot q_i)\frac{k_4 R_d}{\epsilon g_m T_i} + (k_1 + \frac{q_i k_4}{T_i \epsilon})\frac{R_d ct}{g_m}] \cdot \Delta P_i \tag{5.10}$$

derivative with respect to P_{z0}:

$$\frac{\partial \mathrm{H}}{\partial P_{z0}} = 10^{-6} \cdot (1 + ct \cdot q_s)(k_1 + \frac{q_s k_4}{T_s \epsilon})\frac{R_d}{g_m} \tag{5.11}$$

Above, T_s and q_s are the surface temperature and specific humidity, respectively.

5.3 Assimilation of simulated ZTD observations

As a test case, a simulated ZTD observation is assimilated to investigate the dependence of the model variables on the assimilation of the GPS-derived observables.

For this case under study, we have selected a standard background profile of the atmosphere. The $\mathbf{x_b}$ is described by a 46-dim vector composed of 30 temperature values (at 30 decreasing pressure levels), 15 specific humidity layers (at the first 15 decreasing pressure levels) and the surface pressure.

In order to simulate a ZTD observation, a model state ($\mathbf{x^t}$) is first generated from the background profile by adding a perturbation profile,

$$\mathbf{x^t} = \mathbf{x_b} + \sigma\lambda \tag{5.12}$$

where σ is the rms error for $\mathbf{x_b}$ and λ is a unit variance Gaussian deviate vector.

The observation is then simulated by adding another perturbation to $\mathrm{H}(\mathbf{x_t})$.

$$\mathbf{y} = \mathrm{H}(\mathbf{x^t}) + \sigma_y \lambda_y \tag{5.13}$$

where once again λ_y is a unit variance Gaussian deviate and σ_y is the ZTD rms error.

Once the simulated observation of ZTD is assimilated, the optimal state is computed. This assimilation procedure is repeated a statistically significant number of times (a few hundred).

We analyzed the obtained innovation vector (i.e. the difference between the background and the optimal state). In addition, different simulated background covariance

matrices were tested (diagonal and weakly not diagonal). For this case under study, we did not assume any topography bias between the observations and the modeled ZTD values. That means that the model topography represents accurately the elevation of the observational site. Therefore, the surface parameters (P_s, T_s, q_s) are represented in the model by the first pressure level variables. Such approximation preserves from any bias between the observation and the modeled ZTD value caused by an inaccurate model topography.

For instance, when a diagonal B matrix is considered with $\sigma_i = 1$, for $i = (1...30)$ (we assume 1K error for temperature), $\sigma_i = 0.05 * \mathbf{x_b}(i)$ for $i = (31...45)$ (we assume a 5% of the specific humidity value for moisture error) and $\sigma_i = 0.01$, for $i = 46$ (we assume 0.01 hPa error for surface pressure); and ZTD observations are assimilated with a rms value of $\sigma_y = 5$ mm, the temperature values of $\mathbf{x_b}$ remain nearly unchanged (relative increment of less than 1 per thousand). However, the humidity variable is largely modified due to the ZTD assimilation (relative increment of around 30 %). This seems reasonable since the surface pressure, which is assumed very precise, is believed to be weakly correlated to the wet component of the atmosphere but strongly related to the hydrostatic contribution term.

5.4 Evaluation of the assimilation of ZTD and P_S observations

We next analyze the assimilation of the ZTD variable together with the surface pressure measurements. In addition, and as a more realistic case, we now consider a bias between the surface pressure measured at the observational site (i.e. the variable P_s) and the ground pressure provided by the NWP model.

Let P_{lev} be the lowest model pressure level above surface. In the meteorological model parameterization, the *surface* is represented by the *lev* layer in the vertical. Accordingly, the forward operator has to be modified with the $(P_s - P_{lev})$ contribution as follows:

$$\mathrm{H} = 10^{-6}[\sum_{P_{lev}}^{0}(\frac{k_1 P_i}{T_i} + \frac{k_4 P_{wi}}{T_i^2})(\frac{R_d T_i^V}{P_i g_m}) \cdot \Delta P_i + A(lev)(P_{z0} - P_{lev})] \tag{5.14}$$

where $A(lev)$ is given by

$$A(lev) \approx (\frac{k_1 P_{lev}}{T_{lev}} + \frac{k_4 P_{wlev}}{T_{lev}^2})(\frac{R_d T_{lev}^V}{P_{lev} g_m}) \tag{5.15}$$

This approximation is justified because the bias between the surface pressure (which is represented by the P_{z0} variable in the meteorological model) and P_{lev} is assumed small. The water vapor pressure, temperature and virtual temperature at the lev level are represented by P_{wlev}, T_{lev}, and T_{lev}^V, respectively.

To assimilate pressure observations, the forward operator (5.14) is modified as follows:

$$\mathrm{H}^T = (\mathrm{H}_{ZTD}, \mathrm{H}_{P_{z0}}), \tag{5.16}$$

where H_{ZTD} and $\mathrm{H}_{P_{z0}}$ are the operators which map the model state to the space of the observations. This can be described as follows:

$$\mathrm{H}_{ZTD} : \mathbf{x} \longrightarrow \mathrm{ZTD} \tag{5.17}$$

$$\mathrm{H}_{P_{z0}} : \mathbf{x} \longrightarrow P_{z0} \tag{5.18}$$

The H_{ZTD} is the H operator defined in (5.14) whereas $\mathrm{H}_{P_{z0}}$ has the linear representation

$$\mathrm{H}_{P_{z0}} = (0, 0, ...0, 1)^T \tag{5.19}$$

Since both observations are decoupled, the J cost function and its gradient ∇J can be rewritten as follows:

$$J(\mathbf{x}) = [\mathbf{x} - \mathbf{x_b}]^T B^{-1} [\mathbf{x} - \mathbf{x_b}] + \tag{5.20}$$

$$[\mathrm{ZTD}_{obs} - \mathrm{H}_{ZTD}(\mathbf{x})]^T (R_{ZTD})^{-1} [\mathrm{ZTD}_{obs} - \mathrm{H}_{ZTD}(\mathbf{x})]) + \tag{5.21}$$

$$[P_s - \mathrm{H}_{P_{z0}}(\mathbf{x})]^T (R_{P_s})^{-1} [P_s - \mathrm{H}_{P_{z0}}(\mathbf{x})] \tag{5.22}$$

$$\nabla J(\mathbf{x}) = 2 \cdot (B^{-1} [\mathbf{x} - \mathbf{x_b}] - \tag{5.23}$$

$$(\frac{\partial \mathrm{H}_{ZTD}}{\partial \mathbf{x}})^T (R_{ZTD})^{-1} [\mathrm{ZTD}_{obs} - \mathrm{H}_{ZTD}(\mathbf{x})] - \tag{5.24}$$

$$(\frac{\partial \mathrm{H}_{P_{z0}}}{\partial \mathbf{x}})^T (R_{P_s})^{-1} [P_s - \mathrm{H}_{P_{z0}}(\mathbf{x})]) \tag{5.25}$$

In (5.21) and (5.24) the $(R_{ZTD})^{-1}$ and $(R_{P_{z0}})^{-1}$ values are the inverse of the ZTD and surface pressure observations covariance matrices, respectively. In this case, these matrices are reduced to scalars. Above, the zenith total delay observation is represented by ZTD_{obs}.

In the following, we investigate the impact of the assimilation of zenith total delay and ground pressure observations at one selected GPS site and for a period of two weeks.

5.5 Assimilation of real ZTD and $\mathbf{P_S}$ observations

In this section, the assimilation of real observations of ZTD and surface pressure variables is carried out at Robledo site for the Madrid campaign, which is extensively described in Chapter (4.1).

Estimates of ZTD were derived from GPS data using a Trimble 400SSE receiver during 2-15 December 1996. In addition, meteorological data were collected at the GPS site. As part of the meteorological package, the station operated a well calibrated barometer. ZTD estimates for the whole campaign every 150 sec are shown in Figure (5.1) and surface pressure observations are depicted in Figure (5.2). Since ZTD and P_s observations were available during the campaign, both data sets will be assimilated by means of the 1D-VAR procedure.

HIRLAM analysis (available every 6 h) are used to define the background state. We define a control vector in a 63-dim space to parameterize the state of the atmosphere, where the first 31 components are the temperature values at decreasing pressure levels. Since the HIRLAM model uses the specific humidity as the water vapor parameter, the variables from components 32 to 62 are the specific humidity at the 31 pressure levels, and the 63th component is the pressure at surface.

Figure (5.3) shows the temperature profiles derived from the HIRLAM model for all the campaign. At each pressure level there are (14 x 4) 56 values for the temperature variable. The two solid lines correspond to the profiles proposed by Askne and Nordius (1986) with a lapse rate of α=6.2 K/km and an exponent λ=3. These profiles correspond to a standard atmosphere. It is evident from the picture that the simple model do not represent correctly the actual state of the atmosphere analyzed in this study at several pressure levels. We thus expect better results when using the 1D-VAR algorithm to retrieve the amount of moisture.

We next calculate the modeled values of the precipitable water and the zenith wet and hydrostatic delays with the HIRLAM background state.

(1) The PW content of the atmosphere is estimated from the HIRLAM state using

$$\mathrm{PW} = (\frac{1}{g_m \rho_v}) \sum_i q_i \cdot \Delta P_i \tag{5.26}$$

The bias between the station surface pressure and the model surface pressure yields

$$\mathrm{PW} = (\frac{1}{g_m \rho_v}) \cdot [\sum_{P_{lev}}^{0} q_i \cdot \Delta P_i + q_{lev} \cdot (P_{z0} - P_{lev})] \tag{5.27}$$

(2) The zenith wet delay is the second term in (5.3) and its discrete form is given by

$$\begin{aligned} \mathrm{ZWD} &= 10^{-6} \cdot [\textstyle\sum_i (\frac{k_4 P_{wi}}{T_i^2}) \cdot (\frac{R_d T_i^V}{P_i g_m}) \cdot \Delta P_i + & (5.28)\\ &\quad + A_{zwd}(lev) \cdot (P_{z0} - P_{lev})] & (5.29) \end{aligned}$$

with

$$A_{zwd}(lev) \approx (\frac{k_4 e_{lev}}{T_{lev}^2}) \cdot (\frac{R_d T_{lev}^V}{P_{lev} g_m}) \tag{5.30}$$

(3) Simulations of the zenith hydrostatic delay are estimated from the first term in (5.3):

$$\begin{aligned} \mathrm{ZHD} &= 10^{-6} \cdot (\tfrac{R_d k_1}{g_m}) \cdot [\textstyle\sum_i (\frac{T_i^V}{T_i}) \cdot \Delta P_i + & (5.31)\\ &\quad + A_{zhd}(lev) \cdot (P_{z0} - P_{lev})] & (5.32) \end{aligned}$$

with

$$A_{zhd}(lev) \approx (\frac{T_{level}^V}{T_{level}}) \tag{5.33}$$

The Π function defined in (3) can be derived from the ZWD and the PW calculations by applying (3.7). In addition, this function can also be estimated from the modeled mean temperature (3.8), which is given by the following expression:

$$T_m = \frac{\sum_i (\frac{P_i q_i}{T_i}) \cdot (\frac{R_d T_i^V}{P_i g_m}) \cdot \Delta P_i + (\frac{P_{lev} q_{lev}}{T_{lev}}) \cdot (\frac{R_d T_{lev}^V}{P_{lev} g_m}) \cdot (P_{z0} - P_{lev})}{\sum_i (\frac{P_i q_i}{T_i^2}) \cdot (\frac{R_d T_i^V}{P_i g_m}) \cdot \Delta P_i + (\frac{P_{lev} q_{lev}}{T_{lev}^2}) \cdot (\frac{R_d T_{lev}^V}{P_{lev} g_m}) \cdot (P_{z0} - P_{lev})} \tag{5.34}$$

In (5.29), (5.32), and (5.34) the sum is taken from P_{lev} to zero pressure.

The modeled PW content of the atmosphere and the ZWD variable are estimated with the integration of (5.27) and (5.29). The ratio Π^{-1}=(ZWD/PW) obtained with the HIRLAM background profile is shown in Figure (5.4) as a function of the surface temperature for all the campaign. In the picture, continuous lines are the contour lines provided by the integration of a large number of radiosonde profiles whereas the dashed line represents the best fit to the data using the polynomial model for the Mediterranean area (Emardson and Derks 1999). From the figure, most of the values obtained using the numerical model fit the dashed line. However, some profiles that correspond to surface temperature values around T_s= 280 K yield lower Π^{-1} values than when those are compared to the polynomial model fit.

The results shown in Figures (5.3) and (5.4), and the availability of the GPS-derived observations, motivated us to use the 1-dim variational technique instead of the standard procedure to retrieve the moisture content of the atmosphere. For this purpose, observations of ZTD and P_s were assimilated every 6 h for all 15-day campaign to get the optimal state of the atmosphere. Simulations of PW, ZWD, and ZHD were estimated from the optimal profiles following (5.27), (5.29), and (5.32), respectively. The zenith total delay was given by the sum of the ZWD and ZHD variables.

The difference between the observations of zenith total delay, derived with the GPS technique, and the values estimated with the HIRLAM profiles prior to data assimilation is represented in Figure (5.5) as a function of time. We find an average bias of 3.4 mm of ZTD (GPS-derived ZTD values are higher than those modeled with HIRLAM). This small bias may be caused by two contributions: the hydrostatic component, which depends basically on the surface pressure, and the wet term, which is estimated from temperature and humidity profiles. Since the bias between the modeled and the observed surface pressure ranges between 0.1 to 1.2 hPa (which converts to about 1.4 mm of ZHD), the contribution of the wet component to the ZTD bias should account for around 2 mm.

The Π^{-1} function is obtained from the ratio ZWD/PW and is then compared with the value derived from the estimation of the mean temperature T_m. The results show that both methods agree at 1 % level.

In the following, we compare the derived moisture content obtained with the 1D-VAR algorithm with the PW value retrieved using other procedures. We analyze two different situations depending on the precision of the observations to be assimilated.

As a first case study, we start with the assimilation of observations which are assumed very precise ($\sigma_{ZTD} = 3.5$ mm and $\sigma_{P_s} = 0.01$ hPa).

Temperature profiles are reasonably well reproduced by NWP models. However, this is not the case for the moisture component, which is much more difficult to model. Consequently, if we assume small variances for the temperature variable and higher values for the moisture profile, the assimilation procedure will modify the humidity values (and therefore the PW content of the atmosphere) but the temperature profile will remain almost unchanged. According to that, we chose for the background covariance matrix (assumed diagonal) an error of 0.1 K for the temperature variable and a humidity error which varies from 1 gr/kg to 0.01 gr/kg depending on the vertical pressure level. The error for the HIRLAM surface pressure is set to 1.5 hPa.

Results of the 1D-VAR assimilation procedure are displayed in Figure (5.6). The figure shows the difference of the PW content of the atmosphere analyzed by the HIRLAM model with and without the variational assimilation of ZTD and surface pressure observations. This difference results in an average bias of about 0.5 mm of PW (precipitable water from the optimal state is slightly higher than when it is calculated from the background state) and a higher rms value of around 4 mm of PW.

In Section (4.1), the ZHD was calculated from surface pressure measurements using (3.6). This quantity was then removed from the observed ZTD to derive the ZWD term which was converted to PW using a Π^{-1} value of 6.6. It is interesting to compare these results found in the previous chapter with the PW estimated by using the 1D-VAR procedure. This intercomparison is shown in Figure (5.7) as a function of time. The difference between both methods results in a bias value of less than 0.5 mm of PW. As a consequence, we can conclude from our results, that for the used covariance matrices and for this analyzed campaign, the use of $\Pi^{-1} = 6.6$ is a good approximation.

In the following, we investigate if the results found for the assimilation of very precise observations are modified with the use of more relaxed observational and background covariance matrices. For this purpose, we increase the error associated to the observations by assuming now an error of 6 mm for the GPS-derived ZTD and

1 hPa barometer error for the surface pressure. The background covariance matrix is still assumed diagonal (i.e. model variables are uncorrelated), with 1 K error for the temperature variable and a humidity error of $\sigma(q) \approx q$. This error configuration corresponds to a more realistic situation.

The difference between the PW variable estimated by the integration of the HIRLAM background profile and the PW obtained from the optimal state for this case under study, is shown in Figure (5.8) as a function of time. There is not much variation from the results obtained in Figure (5.6) where the observations were assumed more precise. We find now again an average bias of 0.5 mm of PW (the results obtained after the assimilation are higher than the modeled values) and a slightly higher rms value of around 4.5 mm of PW.

Figure (5.9) compares the PW estimates from the numerical weather prediction model against the values obtained after the assimilation procedure. A straight-line fit to this data yields a slope of 0.68 ± 0.06, and the χ^2 (per degree of freedom) is 3.1. A distinctive element of the figure is that the HIRLAM model underestimates the moisture content with respect to the optimal state of the atmosphere around values of about 15 mm. In contrast, the model tends to overestimate the moisture content of the atmosphere below 10 mm of precipitable water. The same results were found in the previous chapter when PW variable modeled with HIRLAM was compared with the GPS-derived PW values.

We finally compare the results obtained with the variational technique against the PW values estimated with the polynomial model proposed by Emardson and Derks (1999) for the Mediterranean area. In this model, the Π^{-1} function is given by

$$\Pi^{-1} = a_0 + a_1 \cdot \Delta T + a_2 \cdot \Delta T^2 \tag{5.35}$$

where ΔT^2 is the surface temperature minus the mean surface temperature T_0 for the area in K. For the Mediterranean area, $a_0 = 6.324$, a_1 = -1.77 x 10^{-2}K^{-1}, $a_2 = 7.5$ x 10^{-5} K^{-2} and $T_0 = 289.76$ K.

The difference between the PW variable obtained using the polynomial model from above and the value obtained with the 1D-VAR algorithm is shown in Figure (5.10) for all the campaign. From the figure, we note that the derived moisture content does not depend on the used technique. The average bias of the comparison is 0.3 mm of PW (the polynomial model yields a slightly higher PW variable than the variational assimilation technique) and a rms value of only 0.2 mm of PW. This can be also shown in Figure (5.11) where the PW variable retrieved using the polynomial

model is compared with the value obtained from the optimal state of the atmosphere. A straight-line fit to this data yields a slope of 1.02 ± 0.01, and the χ^2 (per degree of freedom) is 0.06.

We can conclude from our results that the PW variable retrieved with the 1D-VAR assimilation technique is equivalent to the moisture content derived with the standard procedures. This is probably caused because the meteorological conditions for the Madrid campaign can be represented accurately enough by simple models of the atmosphere. We should look for more interesting meteorological situations in order to get an improvement in the retrieval of the PW using the 1-dim variational assimilation technique.

In the following chapter, we are going to extend the variational assimilation of ZTD in a one-dimensional space to a global four-dimensional variational assimilation problem.

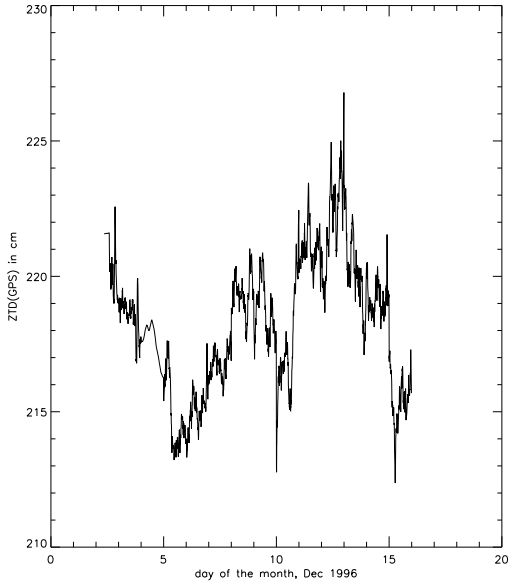


Figure 5.1: GPS zenith total delay at Robledo site as a function of time.

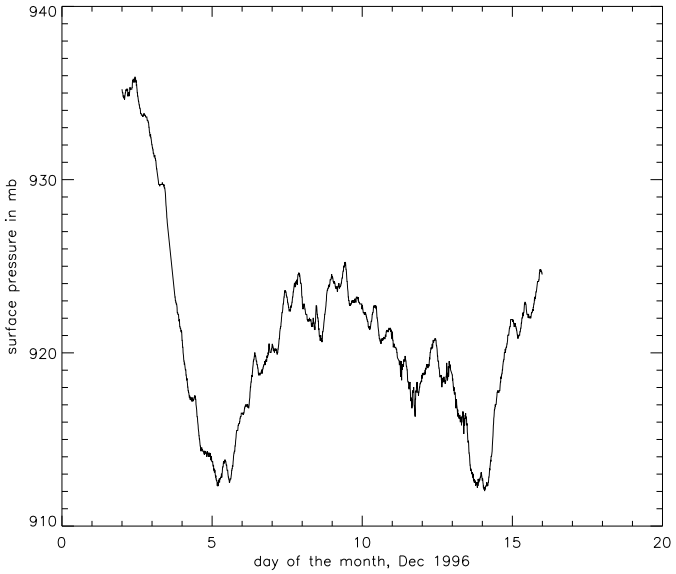


Figure 5.2: Surface pressure at Robledo site as a function of time.

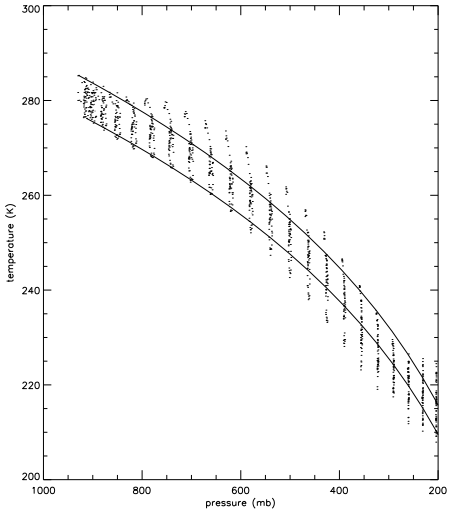


Figure 5.3: Temperature profiles derived from the HIRLAM model for all the campaign (dashed line) and Saastamoinen profiles corresponding to the minimum (lower solid line) and maximum (upper solid line) surface temperature of the campaign.

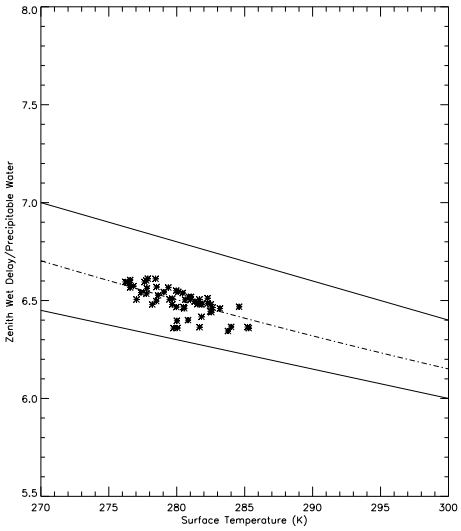


Figure 5.4: The (ZWD/PW) ratio as a function of the surface temperature for all the campaign using the HIRLAM model (dots), the integration of radiosonde profiles (solid lines) and the polynomial model (dashed line).

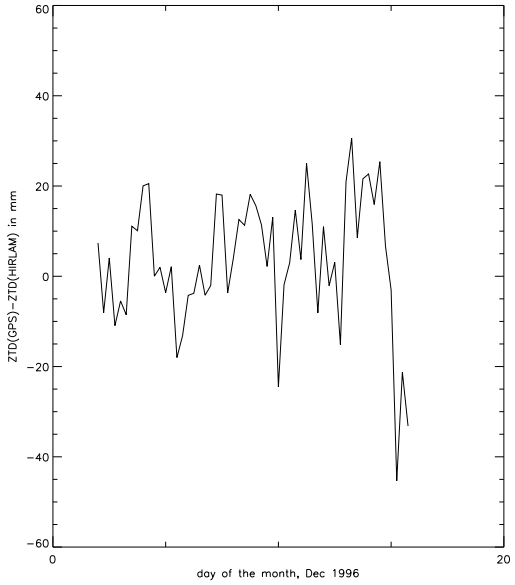


Figure 5.5: Difference between the ZTD derived from GPS data and the ZTD simulated with the HIRLAM model as a function of time.

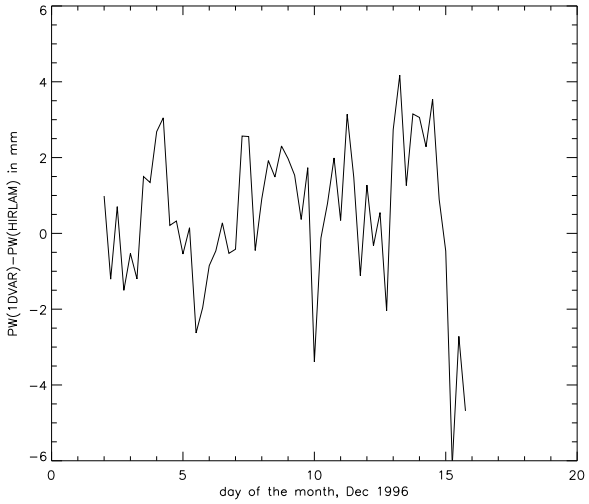


Figure 5.6: Difference between the PW variable obtained with the 1D-VAR procedure and the PW modeled with the HIRLAM profiles for the 15-day campaign. For the variational case, the observations are assumed very precise.

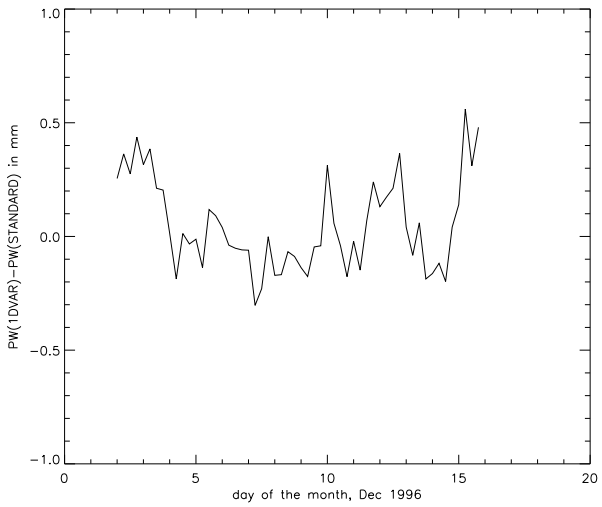


Figure 5.7: Difference between the PW variable obtained with the 1D-VAR procedure and using a standard Π^{-1} value of 6.6 for the 15-day campaign. For the variational case, the observations are assumed very precise.

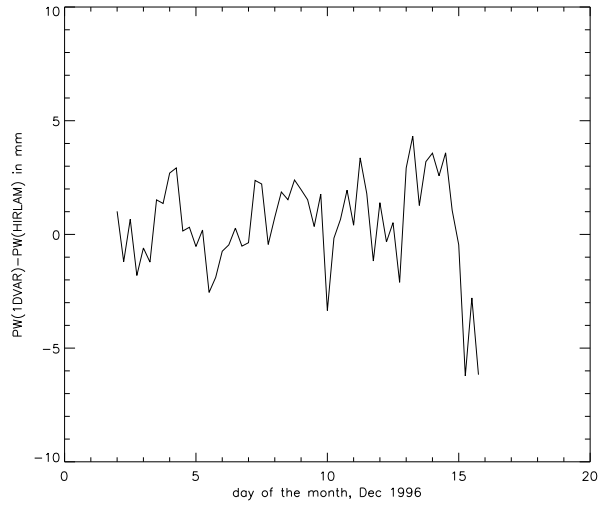


Figure 5.8: Difference between the PW variable obtained with the 1D-VAR assimilation technique and the PW derived from the HIRLAM model as a function of time.

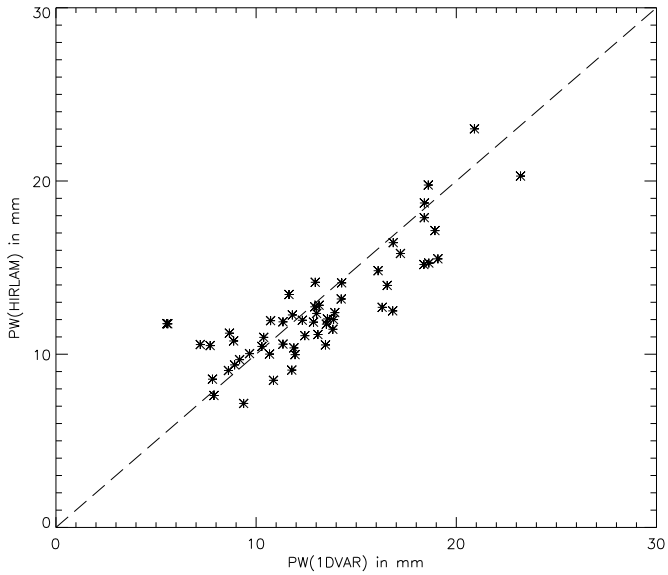


Figure 5.9: Comparison of the PW variable modeled by HIRLAM against the PW value obtained with the optimal atmospheric profile for all the campaign. The dashed line shows the results of perfect correlation.

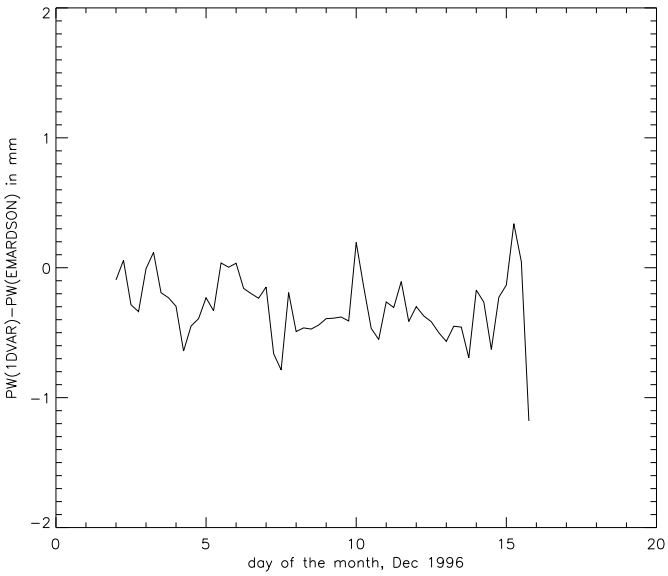


Figure 5.10: Difference between the PW variable obtained with the optimal profile and the polynomial model from Emardson for all the campaign.

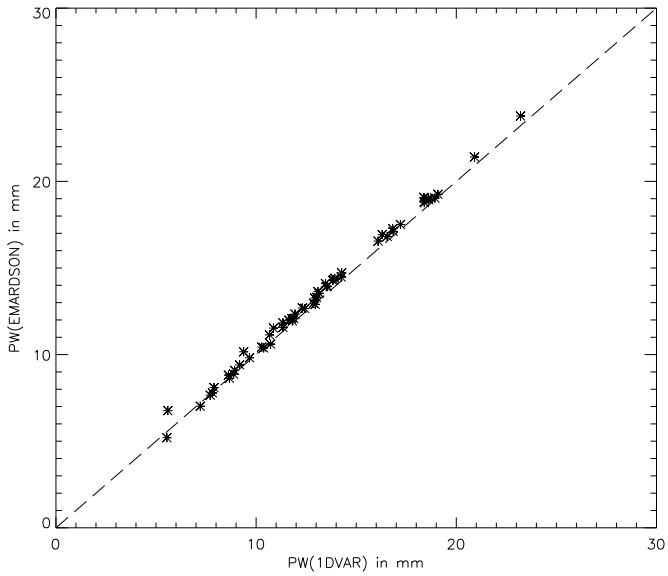


Figure 5.11: Comparison of PW values derived with the polynomial model from Emardson against the optimal PW estimates for all the campaign. The dashed line shows the results of perfect correlation.

Chapter 6

Variational assimilation of GPS-derived observations: 4D-VAR

In this chapter we analyze the impact of the four-dimensional variational assimilation of GPS-derived observations on moisture and precipitation fields.

The assimilation of a large number of PW observations is analyzed in Section (6.1), while the ZTD variable is assimilated in Section (6.2) in a smaller GPS network.

6.1 4D-VAR assimilation of simulated precipitable water into a mesoscale model

In this section we examine the impact of PW assimilation into a mesoscale model on moisture and precipitation fields.

We conduct an *observing system simulation experiment* (OSSE) to simulate a set of PW measurements from a hypothetical network of GPS receivers over a selected area. This idealized network consists of one receiver located at each grid point of the domain chosen for the study. In addition, we consider that every GPS site operates a precise barometer. Therefore, since ZTD and surface pressure are available at all grid points, the precipitable water content of the atmosphere can be derived at all GPS sites following the procedure described in Section (3).

The ground-based GPS observations are simulated by means of the MM5 meteorological model every 30 min and for a period of 3 h. We have selected the 4D-VAR

assimilation technique to investigate the impact of the assimilation in order to make use of the large amount of (simulated) observational data in time.

This study will give us a measure of the impact of the gridded GPS-derived PW assimilation into MM5 for a selected domain. Note that in our idealized network, we have a large number of ground-based receivers. The experiment under study will provide us some theoretical guidance on the impact of the GPS-derived data assimilation.

6.1.1 Simulation of the PW observations

We have used the MM5 mesoscale model to investigate the impact of the assimilation of the PW variable, which is extensively described in Section (4.2). The version of MM5 used in this study includes a high-resolution Blackadar parameterization of the planetary boundary layer (PBL) processes, the USGS topography and land-use sources of 10 min, a non-explicit treatment for the moisture content, and explicitly solves the clouds.

The retrieval experiment makes use of the MM5 Adjoint Model (Zou et al. 1995; Zou et al. 1997). The adjoint model of MM5 is a tool which effectively computes the gradient of any MM5 forecast aspect with respect to the model input control variables. In this work under study, the control variables are the wind field (u, v, w), the temperature (T), the specific humidity (q) and the perturbation pressure (i.e. the difference between the actual and the reference pressure) (p').

The domain selected for the assimilation study is the north-east of the Iberian Peninsula (see Figure 6.1). The grid distance of the domain is 18 km and the dimensions are 27 x 32 x 15, i.e. 27 grid points in the north-south direction, 32 in the east-wet direction, and 15 vertical sigma layers. These layers are defined at $\sigma = 1.0$, 0.96, 0.93, 0.89, 0.85, 0.80, 0.75, 0.7, 0.65, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 and 0.0.

We have chosen a 3-h period from day 14 September 1999 to carry out this sensitivity analysis. As it was seen in Section (4.2), this period was characterized by large amount of precipitation in the area under study. The initial conditions are provided by the ECMWF analysis. The lateral boundary condition is obtained from the analysis.

The PW variable is estimated at 30-min interval from 18 UTC 14 September 1999 to 21 UTC 14 September 1999 by means of the MM5 modeling system. The precipitable water content is obtained from the vertical integration of the specific humidity

$$\mathrm{PW} = \frac{p^*}{g} \sum_{1}^{KP} q(k) \cdot \Delta\sigma(k) \tag{6.1}$$

where p^* is defined as $(p_s - p_t)$, p_s is the surface pressure and p_t is the pressure at the top of the model. The $q(k)$ variable is the specific humidity at the k-th layer, $\Delta\sigma(k)$ is the thickness of the k-th layer, and KP is the total number of layers.

The analyzed PW values at the beginning of the experiment are depicted in Figure (6.2a). Large PW values are observed over all the domain up to around 40 mm along the coast. Values of precipitable water analyzed with MM5 at particular sites during this stormy period are described in Section (4.2).

Figure (6.2b) shows the PW predicted by means of MM5 at the end of the simulation window. There is an important advection of moisture over the coast at the ending time, basically along the northern and southern coastal areas.

The 3-h accumulated precipitation predicted by the MM5 model for this interval time is represented in Figure (6.3). Rainfall amounts about a few millimeters (with a maximum value of around 7 mm) are found in the northern part of the domain.

The simulation run described above (hereafter, the *control run*) is used to generate the "observations" of PW from 18 UTC 14 September 1999 to 21 UTC September 1999. Three different cases are considered:

1. The PW values obtained from the control run are increased by 10 %. We thus consider that the ground-based receivers measure higher moisture content than the value predicted by MM5.

2. The PW variable from the control run is reduced by 25 % and therefore the GPS network gives lower humidity values than the meteorological model.

3. For one selected area, we modify its PW content. For this purpose, we divide the domain along the east-west direction into two regions and then we replace the moisture content of the southern sub-domain by the PW values measured in the north of the domain. The PW variable is not modified for the northern sub-domain.

We simulate three different observational data sets in order to carry out a sensitivity analysis of the impact of the assimilation of the PW variable. We modify the moisture content of the atmosphere from the control run by slightly increasing or reducing its predicted value over all the domain in experiments (1) and (2) and we modify the spatial distribution of the humidity variable in experiment (3).

6.1.2 4D-VAR and statistical issues

The value of the cost function described in (2.18) consists of the background and the observational terms. The latter requires the calculation of the adjoint of the observation operator and the covariance matrix of the observations that will be ingested into the model. In our case, the observational term can be written as follows:

$$J_o(\mathbf{x}) = \sum_{0}^{n} (\mathbf{y}_t - \mathrm{PW}_t[\mathbf{x}_t])^T R_t^{-1} (\mathbf{y}_t - \mathrm{PW}_t[\mathbf{x}_t]) \tag{6.2}$$

where $\mathbf{y}_t$ are the PW observations available at time t and PW_t is the forward operator which calculates the precipitable water from the model variables at all grid points at time t. The accuracy of the simulated GPS observations of precipitable water is given at the millimeter level which is a realistic assumption (see Sections 3 and 4.1). The observations are assumed uncorrelated.

The *a priori* information of the atmosphere is introduced in the first term of (2.18). The background state of the atmosphere and its statistics is crucial to ensure successful data assimilation, especially when sufficient observations are not available. From a theoretical point of view, this *a priori* information is not needed if the number of observations is larger than the degrees of freedom of the control variables.

A proper background term (J_b) for the MM5 model is not yet available since the adjoint model does not have the necessary error statistics. However, we can use some approximations to obtain a simple background information. We use the ECMWF analysis as initial condition (background) and the B^{-1} matrix written in (2.18) is a diagonal matrix obtained by taking the inverse of the square of the maximum difference of the model variables between the beginning and the ending time of the assimilation window (see Navon et al. 1992). This approximation is justified because, in the 4D-VAR approach, the information on model variables is spread from one grid point to another one by the covariance matrix and the strong constraint (2.25). This is not the case for the 3D-VAR procedure, where the B^{-1} matrix is the only available source for propagating the information over the domain and thus a more accurate covariance matrix should be used. The error associated to the PW variable using this covariance matrix is around 3 mm of precipitable water.

The method used to minimize the cost function is the limited-memory quasi-Newton method (Liu and Nocedal 1989) and the first guess conditions for the minimization procedure are given by the initial model conditions.

For all experiments, the assimilation of the PW observations is carried in the same domain as in the control run, and at all grid points. The observations are ingested

into the model from 18 UTC 14 September 1999 to 21 UTC 14 September 1999 every 30-min. This is a realistic assumption since GPS-derived data are given at even shorter intervals.

The PW field is the only variable that is assimilated into the model because we are only interested in the impact of the GPS-derived precipitable water and not in the assimilation of other model variables (such as wind and temperature) which are not derived from ground-based GPS data.

6.1.3 Results and conclusions

We focus on the integrated water vapor and the short-range forecast of precipitation in order to investigate if the assimilation of a large number of PW observations, generated at this hypothetical GPS network, drives the forecast of the MM5 model towards an important modification of the meteorological variables. The former is the variable that is assimilated into the mesoscale model and the latter is a useful field to examine because it is not assimilated and its accuracy is sensitive to various dynamical and thermodynamical processes of the model.

The gradient of the cost function is obtained by integrating the adjoint model from the final to the initial time of the assimilation window. The minimization procedure gives the optimal state of the atmosphere at 18 UTC 14 September 1999 (see Section 2.2). The number of allowed iterations is 30 for all the experiments analyzed in this section.

Experiment 1: assimilation of higher moisture content

Figure (6.4) shows the cost function value (J) and the norm of its gradient (∇J) as a function of the iteration level. The cost function decreases to around 20 % of its original value and the norm of the gradient decreases one order of magnitude during the minimization procedure.

The PW field derived from the optimal state is shown in Figure (6.5) for the initial time. As expected, we find an overall increase of the precipitable water with respect the initial guess PW (Figure 6.2a) when high values of precipitable water are assimilated into the model. The correlation coefficient between the optimal PW and the control value of precipitable water is 0.99 and the rms value is 0.06 cm of PW. The average increase of the moisture content of the atmosphere due to the assimilation of the precise observations is 0.31 cm of PW.

The minimum, maximum and average values of the increment of the precipitable water obtained after the assimilation procedure with respect to the initial guess of PW are 108 %, 111 %, and 109 %, respectively. This means that, in average, the 4D-VAR assimilation increases the moisture content of the atmosphere by the same factor used to generate the observations of PW. This is caused by the assimilation of observations of PW which are assumed to be very precise and forces the model to reproduce them.

The increase of the PW field at all grid points after the minimization procedure is due to the large number of GPS receivers available in the domain (one per grid point!). We expect less impact on the PW field for a smaller GPS network. Studies on the impact in the PW fields using a reduced number of ground-based GPS receivers is analyzed in the following section in terms of the ZTD variable. The error associated to the the optimal precipitable water value for this case under study is around 1 mm of PW. As expected, the variational assimilation increases the accuracy of the meteorological fields.

The assimilation of higher values of PW also produces an increase of the relative humidity field.

Figure (6.6) shows the moisture content at 1000 hPa from (a) the control run and (b) the optimal run (i.e. the model simulation with the optimal initial conditions) at 18 UTC 14 September 1999. Prior to data assimilation, the relative humidity variable varies between 60 % and 90 % in the most part of the domain. However, when the GPS-derived observations are assimilated into the model, there is an increase of this variable to above 90 %.

The increment of moisture fields after the assimilation procedure is also observed at upper levels. Figure (6.7a) displays the relative humidity at 850 hPa pressure level from the optimal run at 18 UTC 14 September 1999. At 850 hPa, the relative humidity field is above 80 % over all the domain. This variable is slightly reduced when it is compared with the results obtained from the control run. The relative humidity field at the same pressure level and for the same time from the control simulation is shown in Figure (6.7b). From the figure, the moisture field varies from 70 % to slightly above 90 % at 850 hPa.

We next examine the impact of these PW simulations on precipitation forecast, the second field tested in our study. In order to investigate how this predicted field is modified after the minimization procedure, we carry out a 3-h forward run starting from the optimal initial conditions.

Figure (6.8) shows the 3-h accumulated precipitation ending at 21 UTC 14 September 1999. The pattern is now completely different from the one obtained with the first guess initial conditions. The assimilation of the high PW values simulated at the GPS sites results in an overall increase of rainfall. When compared to the short-range forecast of precipitation from the control run (Figure 6.3), weak and moderate rains are collected over the whole area after the assimilation of the higher PW values. The correlation coefficient between Figures (6.3) and (6.8) is 0.49 with a rms value of 1.2 mm of rainfall.

For the selected area and under the meteorological conditions of 14 September 1999, the assimilation of high PW observations at the hypothetical GPS network results in an increase of the moisture content of the atmosphere and intensifies the 3-h accumulated precipitation over the whole area.

Experiment 2: assimilation of lower moisture content

We next investigate the impact of the assimilation of lower PW values into the mesoscale model. The evolution of the cost function and the norm of its gradient are shown in Figure (6.9) as a function of the iteration number. The function J reaches its minimum after the first ten iterations and decreases to around 20 % of its original value. The gradient of the cost function decreases two orders of magnitude during the minimization procedure.

The assimilation of low GPS-derived observations of PW gives a lower moisture content of the atmosphere. Figure (6.10) shows the PW field from the optimal state of the atmosphere for experiment (2) at initial time. There is a general reduction of the modeled PW variable once the observations are assimilated. The correlation coefficient between this pattern and the control PW field is 0.99 and the rms value is 0.16 cm of PW. The average difference between the precipitable water from the initial conditions and the value obtained from the optimal condition is 0.7 cm of PW, which converts to around 5 cm of ZTD.

The ratio between the optimal PW and the value obtained from the initial guess varies between 0.70 and 0.84 with an average value of 0.78. As we found for experiment (1), the model tends to reproduce the observed PW values once they are correctly assimilated into the model.

The assimilation of lower PW values also modifies the short-range forecast of precipitation. Figure (6.11) displays the 3-h accumulated rainfall ending at 21 UTC 14 September 1999 from the optimal conditions. The model predicts now a different

distribution of precipitation when comparing with the results from the control run (Figure 6.3). There is a reduction of the accumulated rainfall on the north coast of Catalonia and in the western part of the Pyrenees. The maximum precipitation peak covers now a major area and shifts slightly to the east. In addition, several precipitation areas develop along the center of the domain. The correlation coefficient between Figures (6.3) and (6.11) is 0.60 with a rms value of 0.8 mm of rainfall.

The assimilation of lower moisture content does not result in a reduction of the forecast of precipitation. This should not come a surprise since the equations that drive the evolution of the atmosphere are non-linear and consequently the prediction of rainfall does not depend on the integrated moisture content only (see, e.g., Abbs 1999).

Experiment 3:

We now analyze the impact of the assimilation of the observations generated in experiment (3). To generate these observations, the spatial distribution of the precipitable water is modified by replacing the PW variable of the southern domain by the PW moisture content of the northern region. We do not perturb the precipitable water modeled for the northern part of the area under study.

The cost function and the norm of its gradient are displayed in Figure (6.12) as a function of the iteration number. The value of J decreases to around 11 % of its original value after the assimilation of the observations. The norm of its gradient is reduced by two orders of magnitude during the minimization procedure.

Figure (6.13) shows the precipitable water field from the optimal run at 18 UTC 14 September 1999. As we have found for experiments (1) and (2), the assimilation of the precise observations forces the model to reproduce the observed moisture field. From the figure, if we divide the domain along the east-west direction, the southern region reproduces the PW observations (i.e. northern part of Figure 6.2a) and, consequently, both sub-domains obtained from the optimal run have the same spatial distribution of the PW variable. The correlation coefficient between the control and optimal precipitable water fields is 0.81 with a rms value of 0.48 cm of PW. The average difference between both data sets its 0.29 cm of PW (values from the control run higher).

The reduction of PW in the south of Catalonia and the lower elevation of this area when comparing with the Pyrenees, result in a drastic reduction of the relative humidity field in that region. This is shown in Figure (6.14) where we display the

relative humidity variable at 850 hPa pressure level from the optimal initial conditions. The main characteristic of this pattern when comparing with Figure (6.13) is that although the PW content is the same for both sub-domains, the relative humidity is higher for the north of Catalonia because of the higher elevation of the terrain, which forces the model to allocate the same water vapor content in a thiner atmospheric layer. Similar results are found at upper pressure levels (not shown).

Another effect of this assimilation experiment, is found in the direction of the wind field. Prior to the assimilation, the wind is from the SE, advecting warm and humid air towards the coast of Catalonia as is shown in Figure (6.15) at 850 hPa pressure level. This produces indeed the high moisture content of the atmosphere for the period under study.

The assimilation of lower PW values in the southern domain modifies the wind direction in this area in order to decrease its moisture content. This is shown in Figure (6.14) where we display the winds at 850 hPa for 18 UTC 14 September 1999. From the figure, the winds on the southern coast of the domain turn to the southwest after the assimilation with higher speeds than the control ones. As a result of the assimilation procedure, the model modifies the winds accordingly to the observed amount of precipitable water measured at the virtual GPS network. Since there is a large area with lower observed moisture content than the nearby, the assimilation generates a divergence of water vapor to reproduce the observations of PW in that region.

We now examine the forecasted precipitation between 18 UTC and 21 UTC 14 September 1999. Figure (6.16) shows the 3-h accumulated rainfall from the optimal run for experiment (3). Rainfall of different amounts are now collected at several areas in both sub-domains. The correlation coefficient between Figure (6.16) and the rainfall field prior to data assimilation is 0.62 with a rms value of 0.9 mm of rainfall.

The new precipitation areas developed with the assimilation of the observations are coherent with the modifications found for the wind direction. The amount of water vapor expelled from the southwestern area in Figure (6.14) increases the precipitation recorded in its mountainous neighborhood.

We have shown that the moisture field and the short-range forecast of precipitation are sensitive to the 4D-VAR assimilation of PW observations which are assumed very

precise. Although there is not a linear dependence between PW and precipitation, the amount of rainfall and its distribution in the atmosphere are very sensitive to this variable (PW). This suggests the assimilation of precipitable water into mesoscale models when GPS-derived observations are available. The results found in this section give us confidence on the positive impact of the GPS estimates once they are correctly assimilated. However, in our study we have used a large number of simulated GPS receivers which is not the actual situation in the north-east of the Iberian Peninsula. The results that we present are obtained from an idealized situation. In addition, we have assumed that each GPS site was provided with a well calibrated barometer. At present, in most GPS networks measurements of the surface pressure are not available. All these reasons motivated us to investigate the impact of the ZTD assimilation (instead of PW) measured in a reduced (and more realistic) GPS network over the same domain. This is analyzed in the following section.

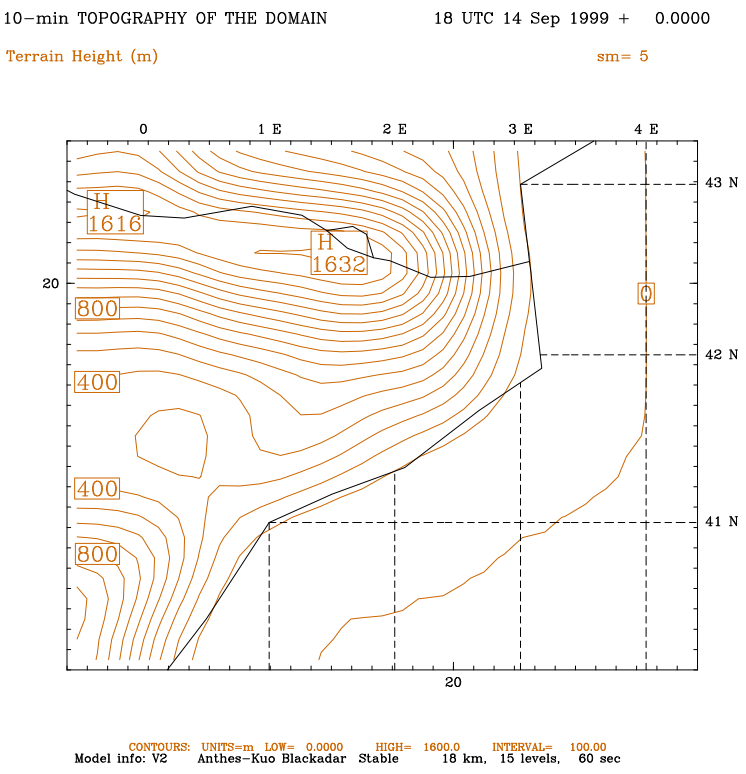
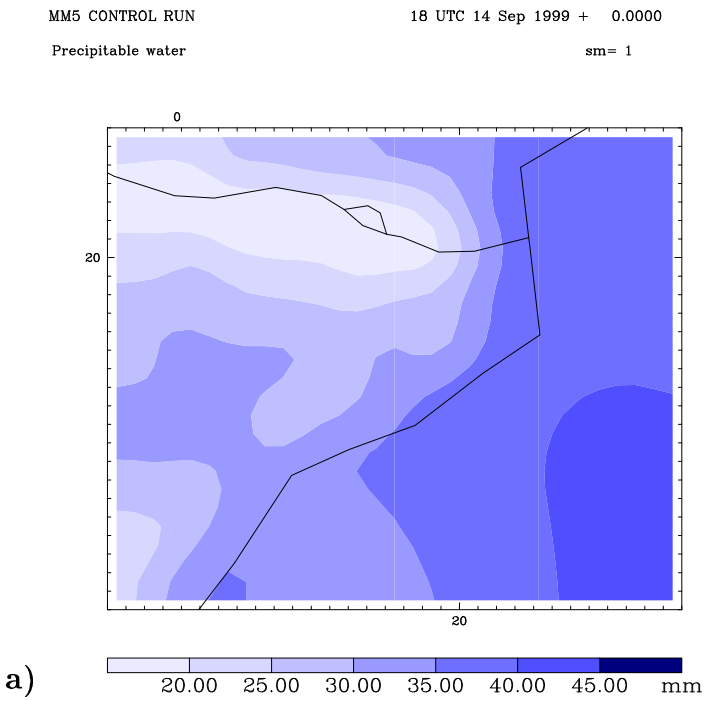
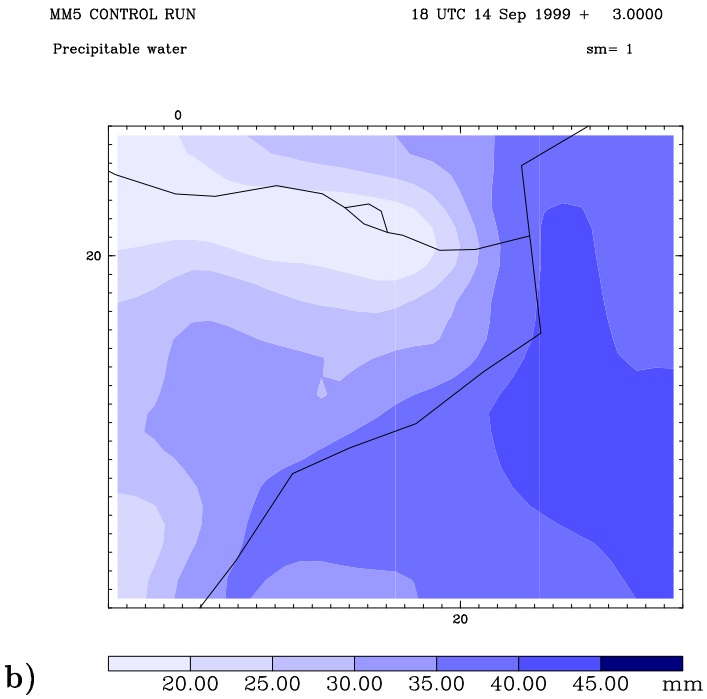


Figure 6.1: 10-min resolution topography of the domain.



a)



b)

Figure 6.2: Precipitable water fields for 14 September 1999 at (a) starting time and (b) ending time of the simulation window.

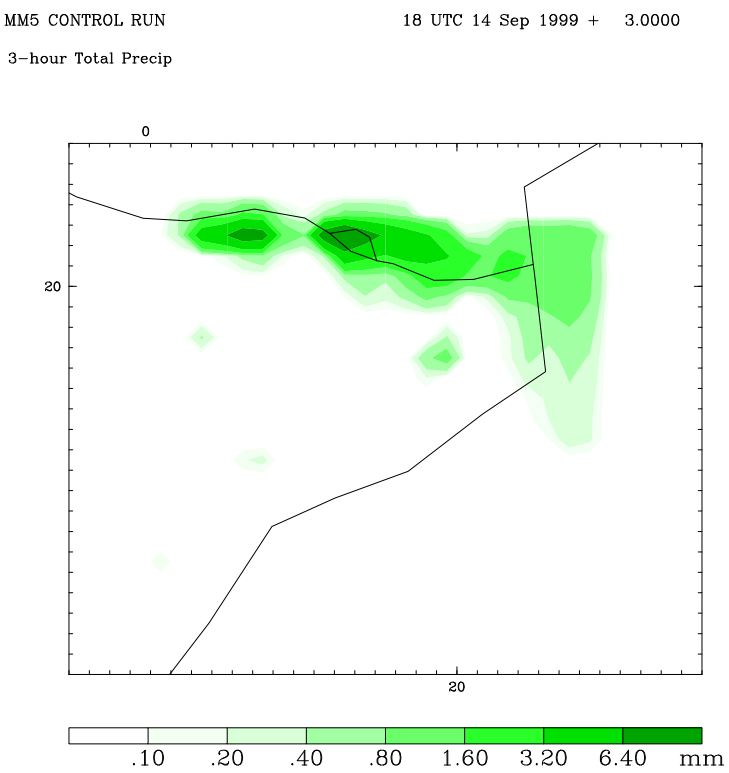


Figure 6.3: 3-hour accumulated precipitation from the control run ending at 21 UTC 14 September 1999.

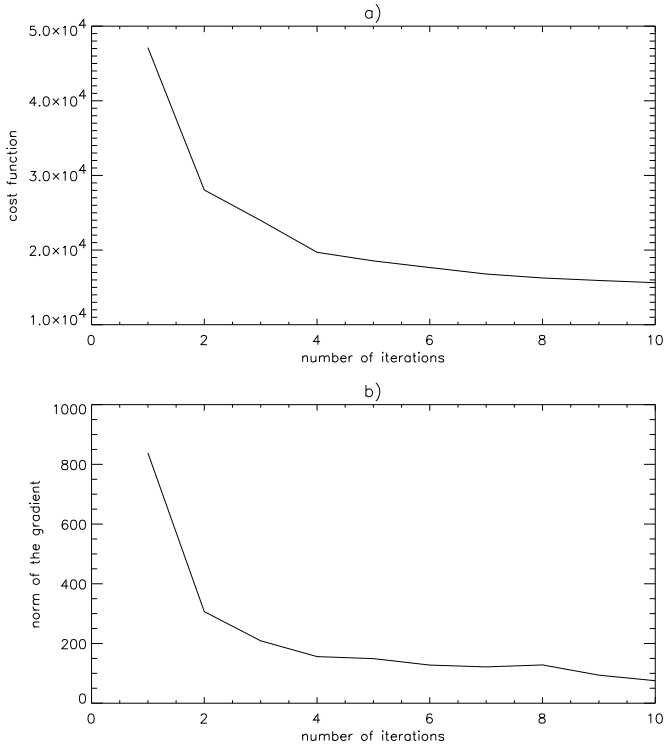


Figure 6.4: Values of (a) the cost function and (b) the norm of its gradient for experiment (1) as a function of the iteration number.

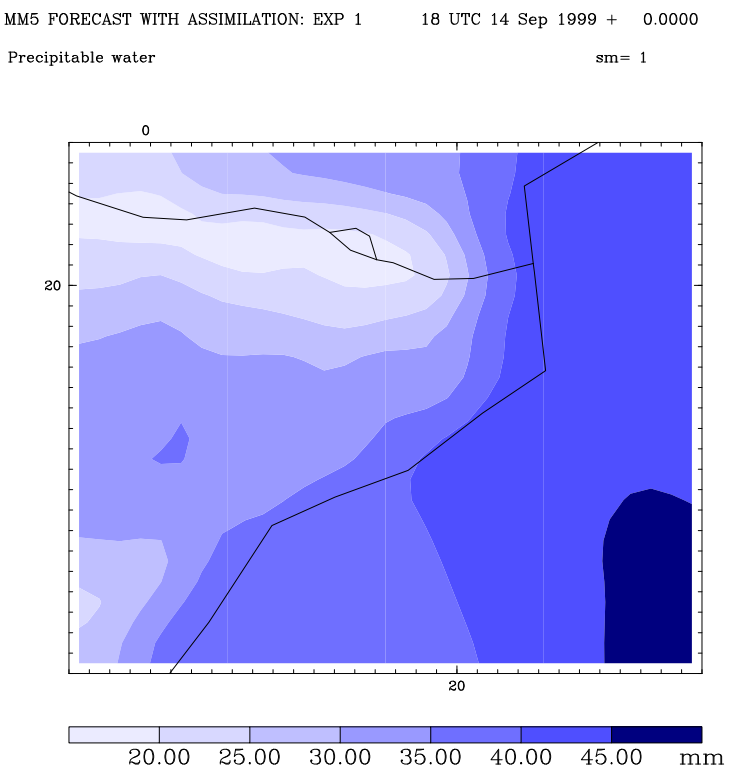
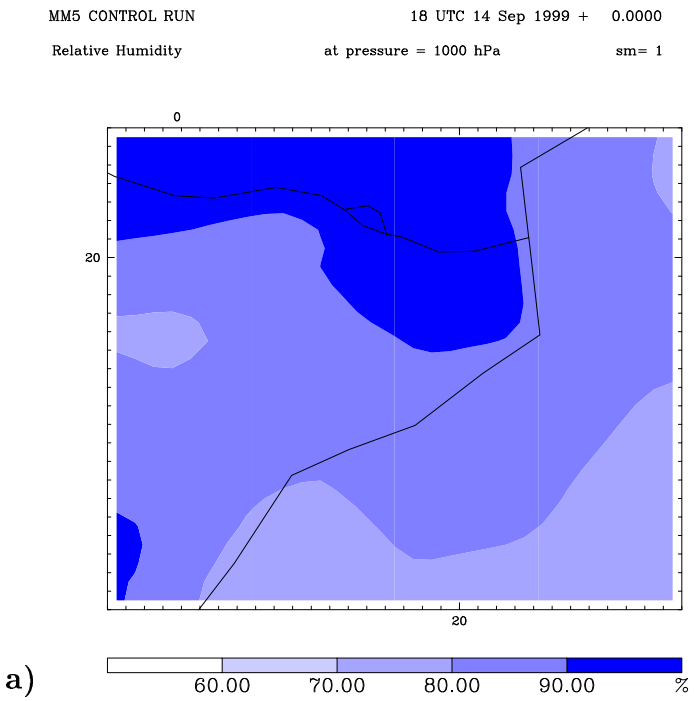
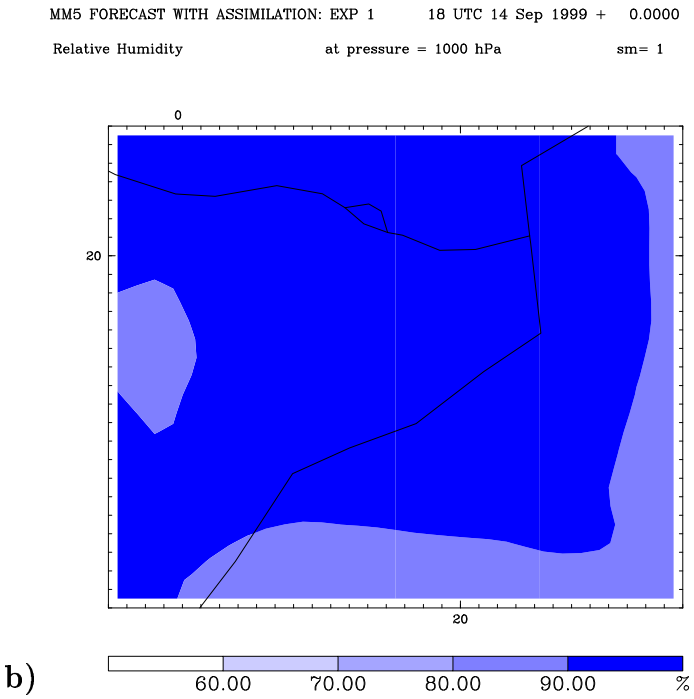


Figure 6.5: Precipitable water field from the optimal run at 18 UTC 14 September 1999 for experiment (1).

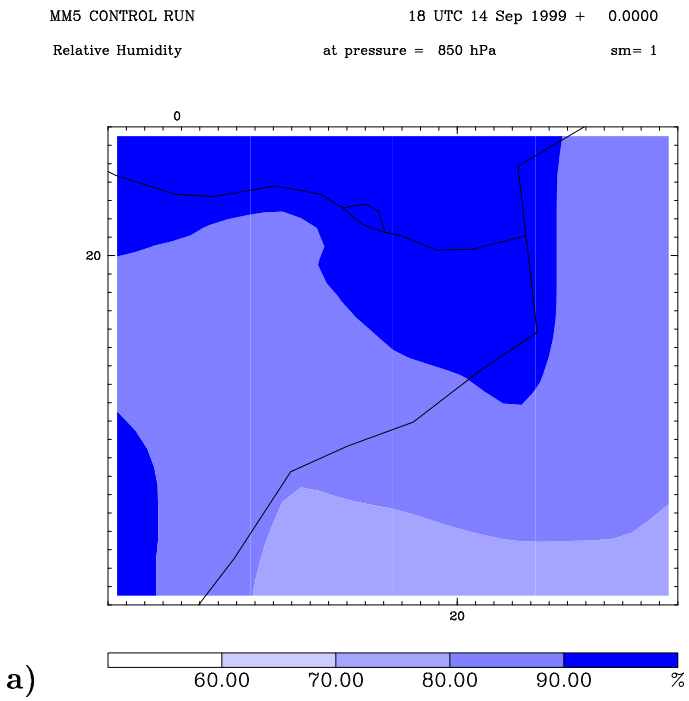


a)

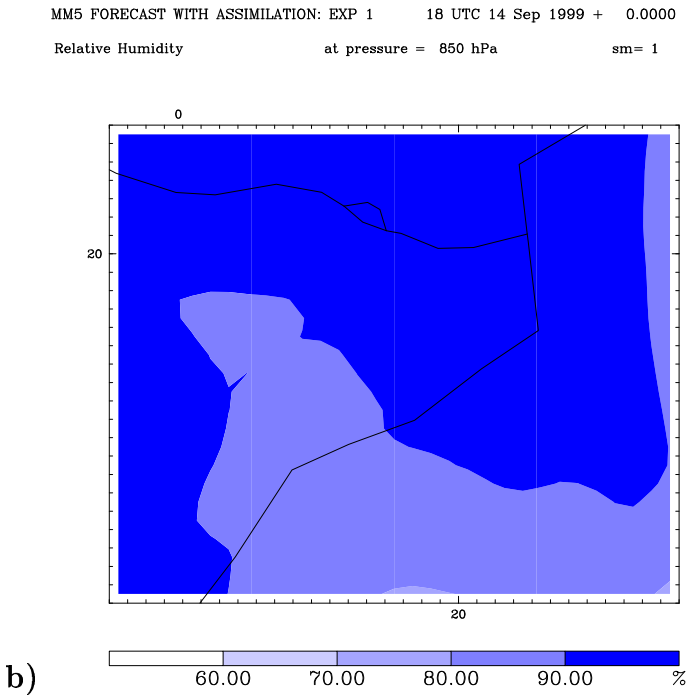


b)

Figure 6.6: Relative humidity at 1000 hPa from (a) the control and (b) optimal runs at 18 UTC 14 September 1999. The results are for experiment (1).



a)



b)

Figure 6.7: Relative humidity at 850 hPa from (a) the control and (b) the optimal runs at 18 UTC 14 September 1999. The results are for experiment (1).

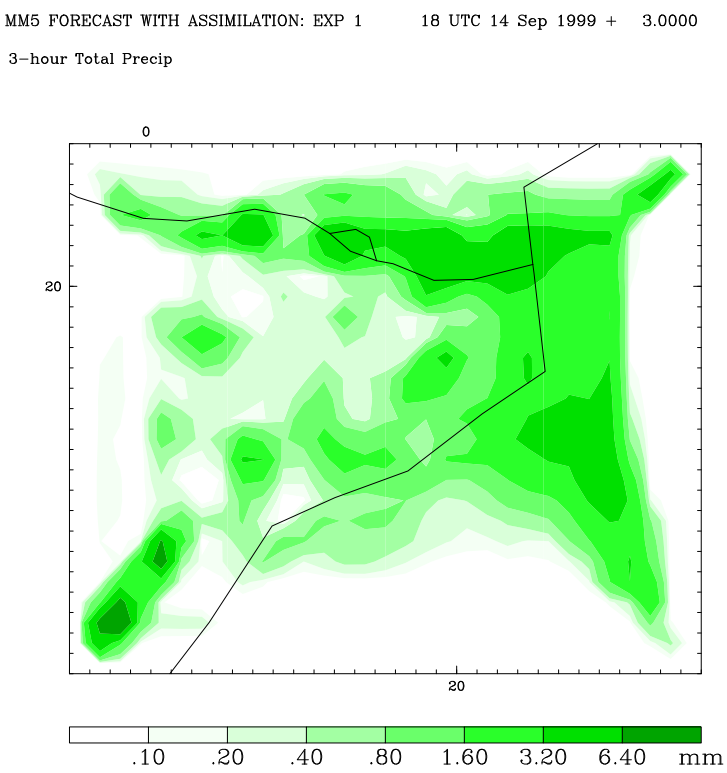


Figure 6.8: 3-hour accumulated precipitation from the optimal run ending at 21 UTC 14 September 1999. The results are for experiment (1).

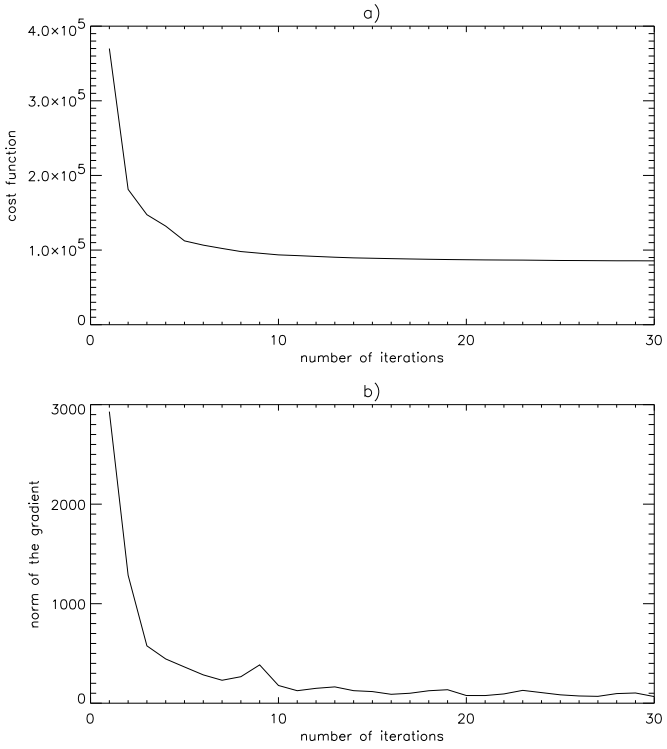


Figure 6.9: Values of (a) the cost function and (b) the norm of its gradient for experiment (2) as a function of the iteration number.

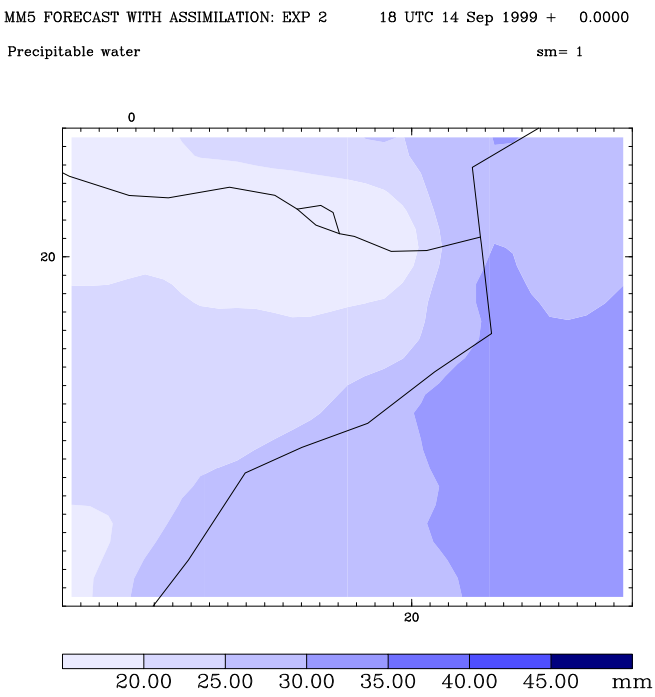


Figure 6.10: Precipitable water field from the optimal run at 18 UTC 14 September 1999 for experiment (2).

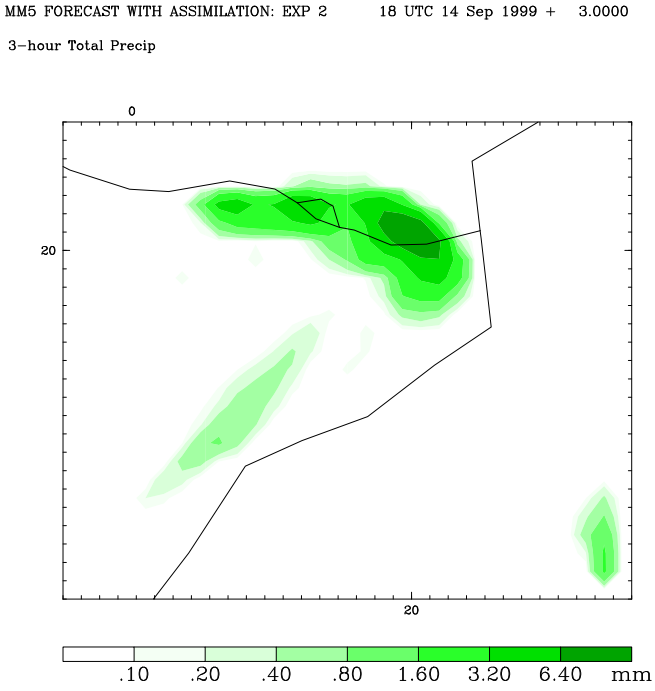


Figure 6.11: 3-hour accumulated precipitation from the optimal run ending at 21 UTC 14 September 1999. The results are for experiment (2).

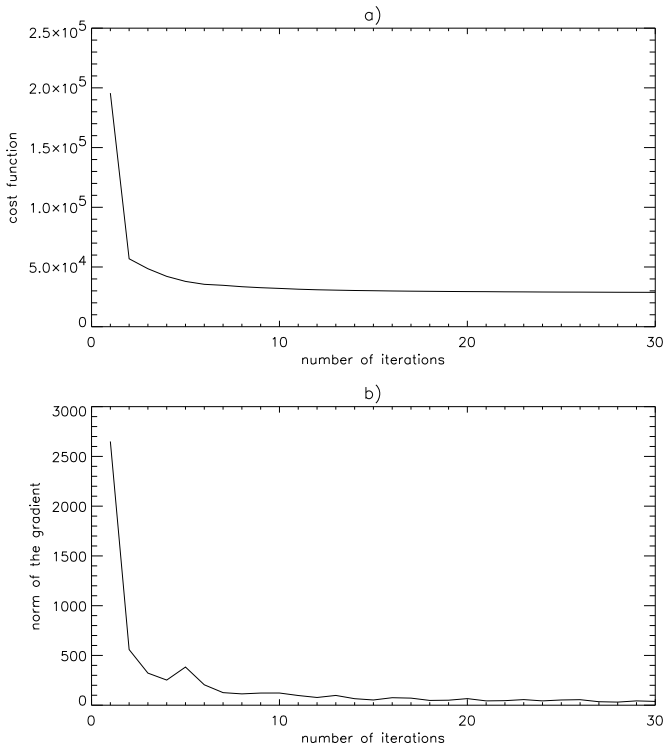


Figure 6.12: Values of (a) the cost function and (b) the norm of its gradient for experiment (3) as a function of the iteration number.

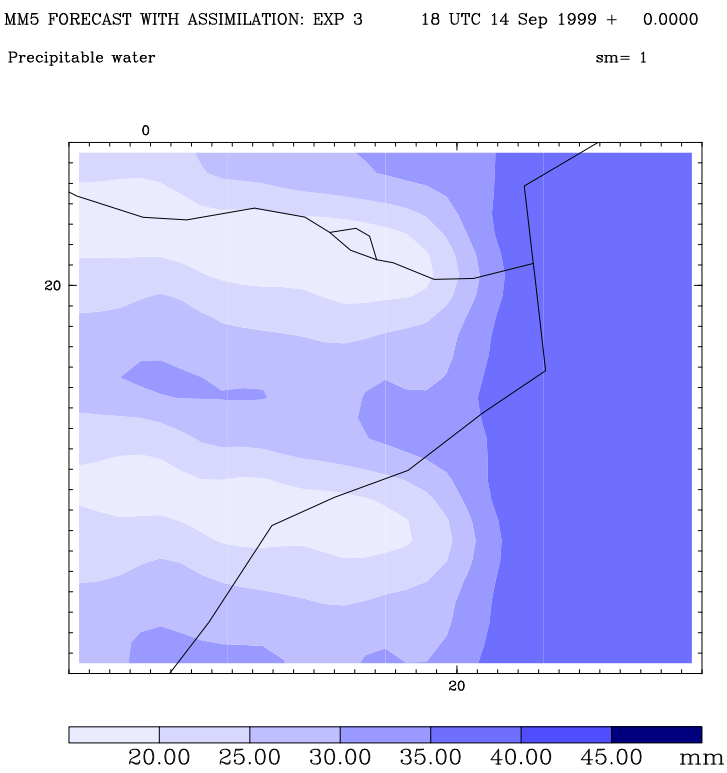


Figure 6.13: Precipitable water field from the optimal run at 18 UTC 14 September 1999 for experiment (3).

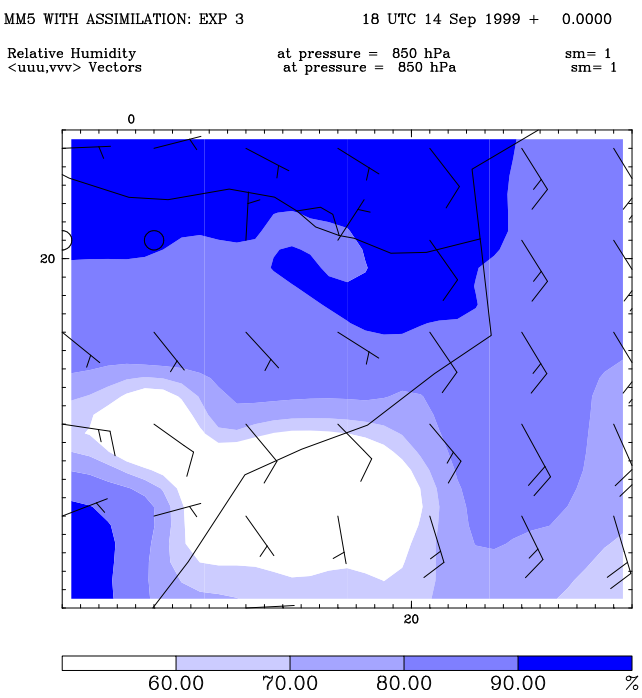


Figure 6.14: Relative humidity and wind at 850 hPa from the optimal run at 18 UTC 14 September 1999. The results are for experiment (3).

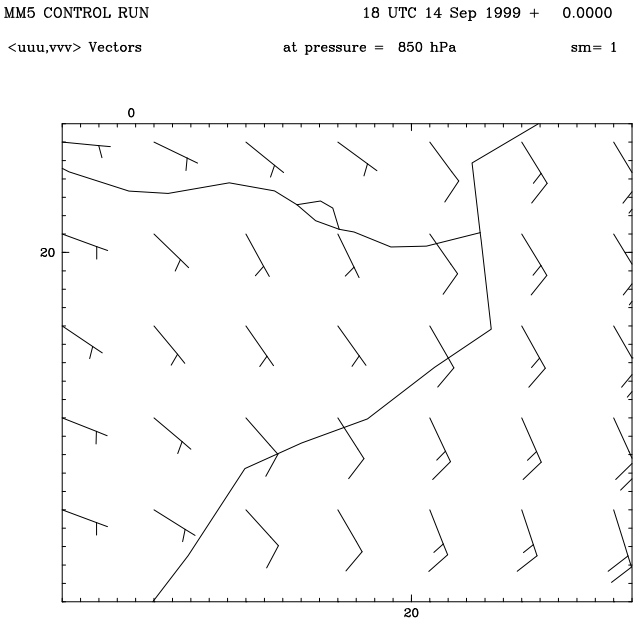


Figure 6.15: Wind at 850 hPa from the control run at 18 UTC 14 September 1999.

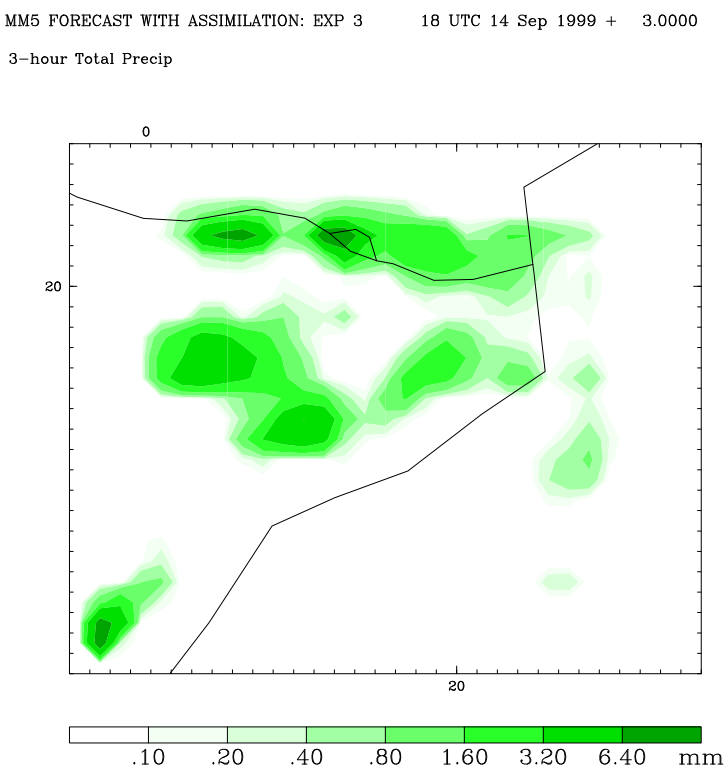


Figure 6.16: 3-hour accumulated precipitation from the optimal run ending at 21 UTC 14 September 1999. The results are for experiment (3).

6.2 4D-VAR assimilation of zenith total delay observations into a mesoscale model

We finally investigate the impact of the assimilation of GPS data gathered at five ground-based receivers operated by the ICC in the Catalonian area. We use the zenith total delay as the observation variable instead of the precipitable water field because pressure measurements are not available in this GPS network.

The domain selected to carry out this assimilation experiment is the same region as in the previous section and the location of the GPS stations on the area under study is shown in Figure (4.15). The distance between the sites ranges from 100 to 350 km and the maximum altitude difference between sites is around 2400 m. The complex topography of the area suggests the use of a non-hydrostatic mesoscale model such as MM5 to better resolve the orographic effects. We make use of the MM5 Adjoint Model (see the previous section) to carry out this assimilation study.

This work analyzes the 6-h period from 18 UTC to 24 UTC 14 September 1999. The assimilation of the ZTD values is a 3-h window experiment, starting at 18 UTC 14 September 1999.

6.2.1 Case description

We select the stormy episode of 14 September 1999 in this assimilation analysis to explore the capability of the 4D-VAR procedure to improve the forecasted fields with the use of GPS-derived data in a rainfall event.

The meteorological variable associated with the ZTD data, and weakly modeled by NWP systems, is the precipitable water. As a consequence, we expect interesting results with the 4D-VAR assimilation technique for periods characterized by large amount of precipitation, which is strongly related to the moisture content of the atmosphere.

Section (4.2) describes in detail the atmospheric conditions for 14 September 1999 in the domain under study. In this assimilation study, we concentrate on the second half of the day because higher values of PW were found (see Figure 4.16 in Section 4.2) and we expect that the impact of the assimilation of GPS data on the moisture content of the atmosphere will be higher.

The grid resolution of the domain selected for this retrieval experiment is the same as defined in Section (6.1) for the assimilation of simulated PW values (i.e. 27 x 32 x 15).

The forecast run is initialized at 18 UTC 14 September 1999. The initial and lateral boundary conditions are provided by the ECMWF and interpolated to the domain under study. The model uses a high-resolution Blackadar parameterization of the boundary layer, explicitly solves the clouds, and includes no explicit moisture physics. This control run is a 6-h forecast run.

Figure (6.17) shows the precipitable water field prior to data assimilation at 18 and 21 UTC September 1999. Large amount of moisture is found over the domain, with higher values along the coast and the Mediterranean sea.

The forecast of precipitation is depicted in Figure (6.18). The figure shows the 3-h accumulated precipitation ending at 21 and 24 UTC 14 September. The passage of the frontal system produced rainfall over the Pyrenees from 18 to 21 UTC covering ESCO, LLIV and CREU stations. This control run does not predict rainfall at BELL and EBRE GPS sites. Later, areas of moderate precipitation cover the most part of the domain from 21 to 24 UTC as is shown in Figure (6.18b).

The precipitation recorded at the GPS sites during the same period is found in Table (6.1). Large values of rain are found for the first 3-h period at all GPS stations except for CREU, which did not measure any precipitation. This situation is not reproduced correctly by MM5 when we compare the observations from the table with Figure (6.18). Forecasts prior to data assimilation overestimates the actual amount of rainfall at CREU station and underestimates the observed accumulated precipitation at BELL site from 18 to 21 UTC. For the second 3-h period the model tends to overestimate the precipitation at most of the observational sites.

6.2.2 4D-VAR data assimilation of the ZTD variable

In this 4D-VAR exercise we minimize the cost function J defined in (2.18) that measures the misfit between the observations and their modeled values. This functional introduces a priori state of the atmosphere,

$$J_b(\mathbf{x}) = (\mathbf{x} - \mathbf{x_b})^T B^{-1} (\mathbf{x} - \mathbf{x_b}). \tag{6.3}$$

The background state is provided by the ECMWF analysis and interpolated to the domain under study. The inverse of the covariance matrix of this a priori information (B^{-1}) is assumed diagonal. The inverse of the square of the maximum difference of the model variables between the beginning and the ending time of the assimilation window are the elements of this matrix. The error associated to the ZTD variable

using this covariance matrix is around 2.4 cm of ZTD (or around 3 mm of precipitable water).

The observational term of the cost function in (2.18) is written now as follows:

$$J_o(\mathbf{x}) = \sum_{0}^{n} (\mathbf{y}_t - \mathrm{ZTD}_t[\mathbf{x}_t])^T R_t^{-1} (\mathbf{y}_t - \mathrm{ZTD}_t[\mathbf{x}_t]) \tag{6.4}$$

where $\mathbf{y}_t$ are the observations of the ZTD variable measured at the five ground-based receivers at time t. The ZTD_t value is the forward operator which estimates the zenith total delay from the model variables ($\mathbf{x}_t$) at the location of the GPS sites at t. This operator can be written as the sum of two terms: the ZHD and the ZWD forward operators:

$$\mathrm{ZTD}_t[\mathbf{x}_t] = \mathrm{ZHD}_t[\mathbf{x}_t] + \mathrm{ZWD}_t[\mathbf{x}_t] \tag{6.5}$$

As it was seen in Section (3) the hydrostatic contribution can be estimated from surface pressure measurements. Therefore, the ZHD operator in (6.5) basically estimates the surface pressure at the GPS sites from the model variables. We use the bilinear interpolation in the horizontal direction to interpolate the ground pressure values from the grid points of the domain to the location of the GPS sites.

A more accurate treatment is needed for the interpolation in the vertical. We use the methodology developed by the de Pondeca and Zou (2001) to interpolate the model surface pressure to the elevation of the GPS sites. We thus consider three different cases:

- station elevation higher than the model topography: the surface pressure of the GPS site is found from the two sigma levels above and below the station. The interpolation assumes an exponential dependence for pressure as a function of height.

- station elevation lower than the model topography: the ground pressure at the GPS site is interpolated from the sea level pressure and the model ground pressure.

- station elevation coincides with the model topography: the surface pressure at the ground-based receiver is provided by the model pressure.

De Pondeca and Zou (2001) found that this method used to interpolate the surface pressure from the model to the observational sites yields average errors of the order

of 1.0 to 1.5 hPa, which converts to around 3 mm of ZTD. However, these authors obtained values as large as 3 hPa at some locations of the domain centered around the Los Angeles Basin from 12 UTC 9 December 1996 to 24 UTC 9 December 1996.

We note the need for a carefully designed operator in the vertical direction if we want to avoid a bias between the zenith total delay measured at the ground-based receiver and the modeled ZTD value caused by orographic effects. This is because the surface pressure strongly depends on the height of the GPS sites and these are not modeled correctly as is shown in Table (6.2). The table shows the elevation of the ground-based receivers and their modeled values by means of MM5. We find differences as large as 500 m which corresponds to around 13 cm of ZHD. The bias between the observations and the meteorological model has to be removed prior to data assimilation in order to get the optimal state of the atmosphere.

The evaluation of the hydrostatic contribution in the zenith total delay operator is computed from the ground pressure data using (3.6). The zenith wet delay value at GPS sites is given by the temperature and water vapor pressure variables (see equations 3.7 and 3.8 in Section 3).

Figure (6.19) shows the time series of the differences between the model and the GPS zenith total delay prior to data assimilation. The mean-values of the differences between the observations and the MM5 simulation of the ZTD are summarized in Table (6.3) for all GPS sites. In most of the cases the observed values are higher than the modeled results at ESCO and LLIV stations and we find the opposite situation for CREU and EBRE sites. The atmospheric delay gathered at these sites is over-estimated by the meteorological model. An error as large as around 3 cm of ZTD is found for BELL station at the initial time of the period under study.

The R_t^{-1} matrix in (6.4) is the inverse of the covariance matrix associated to the observations. The zenith total delay observations are assimilated every 15-min with a precision that varies in time. The errors of the observations are assumed uncorrelated.

The ECMWF analysis interpolated to the domain under study is taken as first guess for this assimilation experiment.

The control variables for this case under study are the three components of the wind field, the temperature, the pressure perturbation and the specific humidity at all grid points. We expect a reduction of the ZTD differences between the MM5 model and the observed values and thus improve the model initial conditions.

6.2.3 Results and conclusions

The minimization procedure gives the optimal state of the atmosphere at the initial time after the 35 allowed iterations.

The values of the cost function and the norm of its gradient are shown in Figure (6.20) as a function of the iteration number.

The functional J decreases by around 90 % of its original value during the minimization procedure, with the largest drop occurring during the first 10 iterations. At the beginning of the minimization, the cost function is entirely due to the observational term since the first guess condition coincides with the background state of the atmosphere. The background contribution to J increases with the iteration level, while the observational term decreases in order to minimize the cost function. The norm of the gradient decreases one order of magnitude during the minimization procedure.

The accuracy of the ZTD value obtained from the optimal state of the atmosphere decreases from 2.4 cm (prior to data assimilation) to around 0.5 cm of ZTD after the assimilation of the precise observations.

Figure (6.21) shows the ZTD differences between the GPS observations and the values obtained from the optimal initial conditions. As expected, the assimilation of the zenith total delay reduces the mean error of the model at all sites. The average differences of ZTD between observations and MM5 are summarized in Table (6.3). We find that EBRE, CREU and BELL reduce their absolute ZTD error by more than 90 % and LLIV and ESCO by above 70 %.

The reduction of the ZTD error can provide from a reduction of two different error sources:

- error associated to the hydrostatic contribution
- error caused by the zenith wet delay term

Measurements of surface pressure at GPS sites are needed to evaluate the first source of error in order to compare the observed pressure data with the values predicted by the meteorological model. The only ground-based receiver from our GPS network that operates a barometer is CREU station. The time evolution of the surface pressure at CREU is shown in Figure (6.22). The figure also includes the modeled ground pressure from initial and optimal states of the atmosphere at the time when ZTD observations are available. The average bias between the observed pressure and

the values obtained from the first guess initial state is around 1.6 hPa. This value is in agreement with the pressure errors found by de Pondeca and Zou (2001) associated with the estimation of the ground pressure.

The assimilation of GPS-derived ZTD data does not reduce this pressure error at CREU. The difference between the a priori pressure at this site and the value obtained after the assimilation procedure is basically not changed because of the accuracy of the initial surface pressure which is given at around 0.7 hPa with the B^{-1} matrix defined in (6.3). This error on the ground pressure error only accounts for 1.5 mm of ZTD.

We find similar results for the modeled pressure at surface with and without assimilation at all GPS stations. As a consequence, all the improvement achieved with the assimilation of the zenith total delay is found in the zenith wet delay contribution, i.e. the moisture content of the atmosphere.

In the following, we are going to analyze the impact of the ZTD assimilation in terms of the atmospheric moisture. We select the precipitable water field and the short-range forecast of precipitation to carry out such analysis.

Figure (6.23) shows the PW variable at 18 and 21 UTC 14 September 1999 from the optimal state of the atmosphere after the 35 iterations allowed in the minimization procedure. The figure reproduces the main features obtained from the control run in Figure (6.17). The correlation coefficient between both plots at the beginning and ending times of the assimilation window is 99 % (average rms value of 5 mm of PW). However, some differences are found in very localized areas. The control run overestimates the moisture content at CREU station at the end of the assimilation window when compared to the optimal PW field. Similar tendencies are found around EBRE and BELL sites. The mountain stations (ESCO and LLIV) located on the Pyrenees do not modify their PW content after the assimilation of ZTD because the largest part of the moisture content of the atmosphere is located below 3000 m and these sites have elevations of around 1500-2000 m. We expect that the impact of the assimilation of ZTD will be higher for lower GPS heights.

The decrease of the precipitable water field at CREU station with the assimilation of GPS data reduces the forecast of rainfall at this site. This is shown in Figure (6.24a) where we display the 3-h accumulated precipitation ending at 21 UTC 14 September 1999. Another characteristic of this pattern when it is compared with Figure (6.18) is the increase of produced rainfall at the area covering BELL station.

When these results are compared with the observations at the GPS stations (see

Table 6.1) we find that the optimal run reproduces more accurately the precipitation measured at CREU, BELL, and LLIV sites than the optimal run. However, the model does not predict the intense peaks of rain measured at BELL and ESCO sites. We expect better results on the prediction of precipitation when other variables such as wind, temperature and surface humidity, which have direct influence on the precipitation predictions, are also assimilated in a 4D-VAR procedure along with the PW field.

We find less differences between the 3-h accumulated precipitation ending at 24 UTC when comparing the control and optimal runs. This is probably caused by the "contamination" of the boundary condition which is not optimized by the 4D-VAR assimilation technique. The main effect of the minimization for this second 3-h window is found at EBRE site. The run from the optimal conditions reduces the predicted amount of rainfall over this area, which agrees with the observational data at this site.

In this section, we have analyzed the impact of the assimilation of ZTD at five selected grid points. The results yield an important improvement of the short-range precipitation forecast around the location of the GPS sites. This influence, in general, will depend on the model resolution and on the area under study.

This study provides results for this particular case and under these meteorological conditions. These are preliminary results on the impact of the ingestion of GPS data into a mesoscale model.

The use of the MM5 Adjoint Model to get the optimal state of the atmosphere requires different data sources. A schematic diagram of the procedure used in this exercise is presented in Figure (6.25). We first need to select the area where to carry out the assimilation study and get the topography and land-use data sets accordingly to the selected grid size. We have used the 10-min resolution source. In this mesoscale model, this is done in the TERRAIN package. Analysis given by ECMWF are used to provide initial and boundary conditions for the domain under study in the REGRID module. In case that multiple nesting is defined in our domain, we can use the INTERP program to interpolate the initial conditions to the nested domains. In addition, this program handles the data transformation (vertical and horizontal interpolation, diagnostic computation and simple data reformatting).

At this point, we are ready to run the meteorological model (i.e. the MM5 program) and to analyze the predicted fields. We have called this run the *control run*.

The availability of the GPS-derived data, gathered at the ICC and analyzed at the

IEEC, every 15-min interval, allows us to use the 4D-VAR procedure to assimilate the zenith total delay and get the optimal state of the atmosphere by using the adjoint version of the MM5 meteorological model.

Once the optimal state of the atmosphere is achieved, we run again the MM5 program to analyze the improved meteorological predicted fields. Then, we compare the state of the atmosphere obtained from the initial and optimal runs with the available observations over the domain to analyze the impact of the assimilation of the observations.

Table 6.1: 3-hour accumulated precipitation (mm) ending at 21 and 24 UTC 14 September 1999 measured at the five GPS stations.

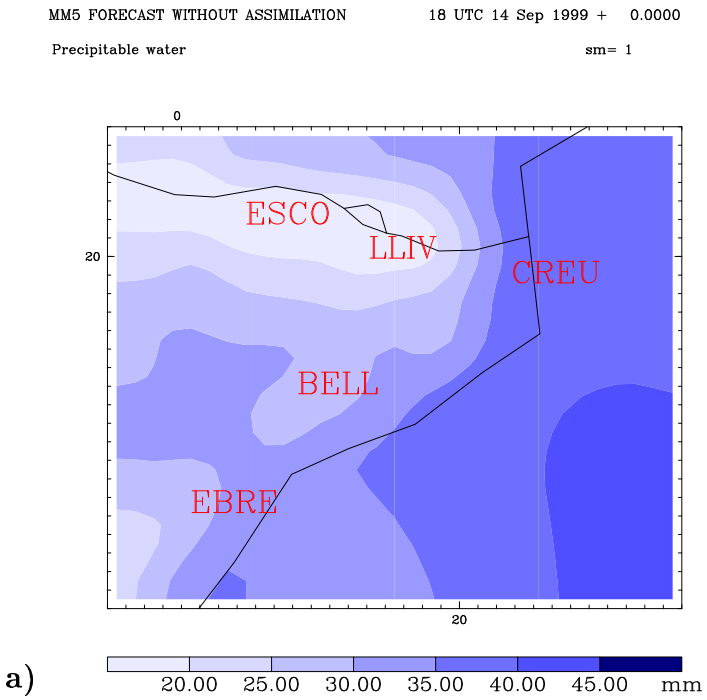
station	18-21 UTC 14 September	21-24 UTC 14 September
CREU	0.0 mm	0.0 mm
EBRE	1.2 mm	0.0 mm
BELL	8.7 mm	0.0 mm
LLIV	8.6 mm	5.6 mm
ESCO	7.0 mm	5.0 mm

Table 6.2: Altitude of the GPS stations and elevation from the 10-min topography source. All the heights are in meters above sea level.

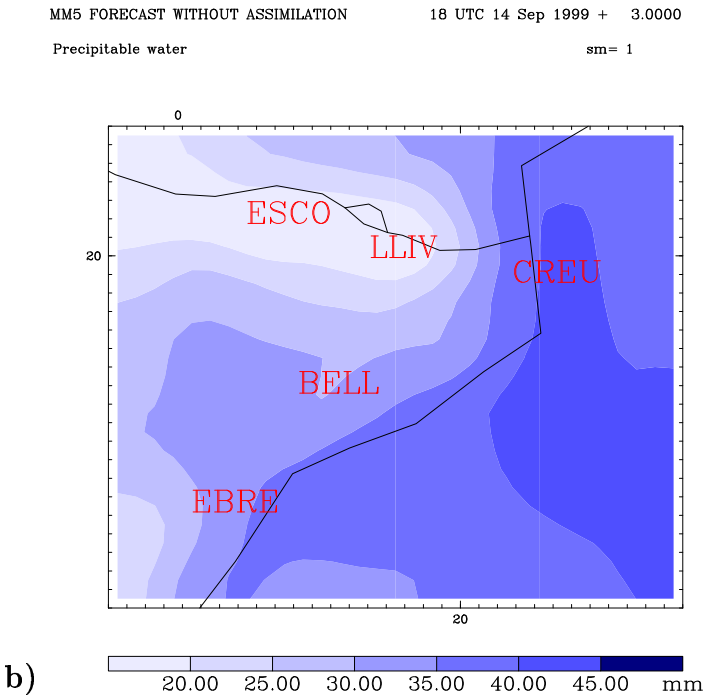
station	model height	GPS elevation
CREU	0 m	83 m
EBRE	204 m	58 m
BELL	624 m	803 m
ESCO	1971 m	2458 m
LLIV	1823 m	1418 m

Table 6.3: Average bias between GPS-derived ZTD and MM5 modeled values from initial and optimal conditions for all the stations. All values are expressed in cm of ZTD.

station	Initial Conditions	Optimal Conditions
BELL	1.19 cm	0.01 cm
CREU	-1.93 cm	-0.04 cm
EBRE	-2.17 cm	-0.08 cm
ESCO	0.74 cm	0.21 cm
LLIV	2.06 cm	0.34 cm

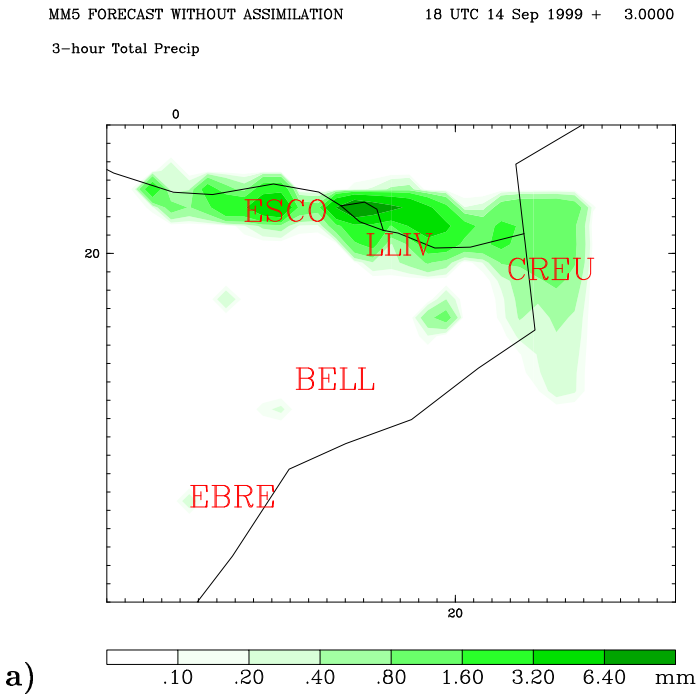


a)

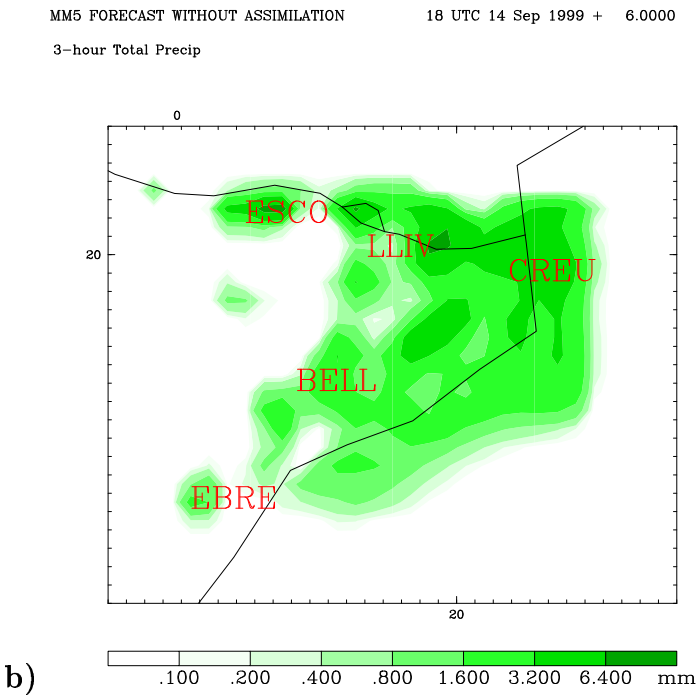


b)

Figure 6.17: Precipitable water fields for 14 September 1999 at (a) 18 UTC and (b) 21 UTC from the control run.



a)



b)

Figure 6.18: 3-hour accumulated precipitation from the control run ending at (a) 21 UTC and (b) 24 UTC 14 September 1999.

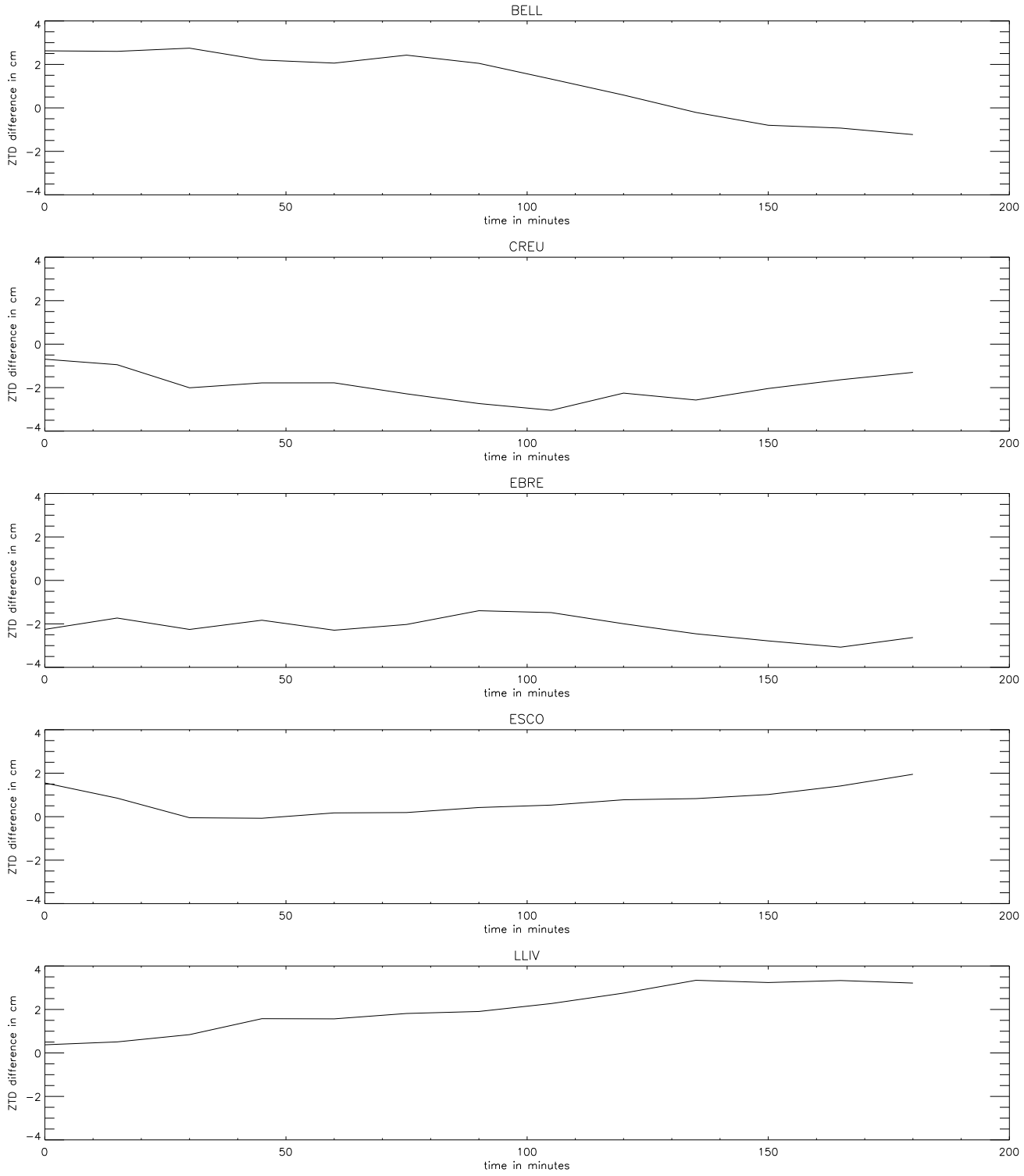


Figure 6.19: Evolution of the ZTD differences prior to data assimilation between the GPS observations and the MM5 modeled values as a function of time for all GPS sites.

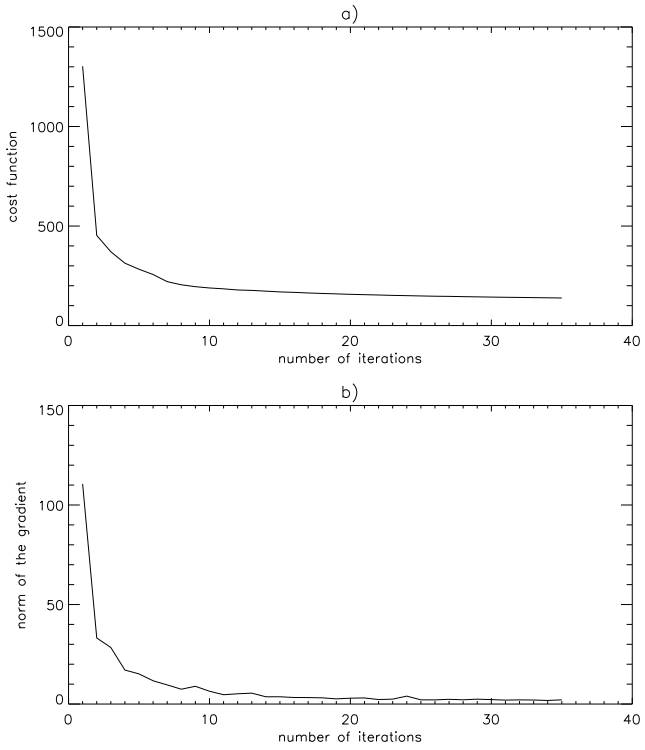


Figure 6.20: Values of the cost function and the norm of its gradient as a function of the iteration number.

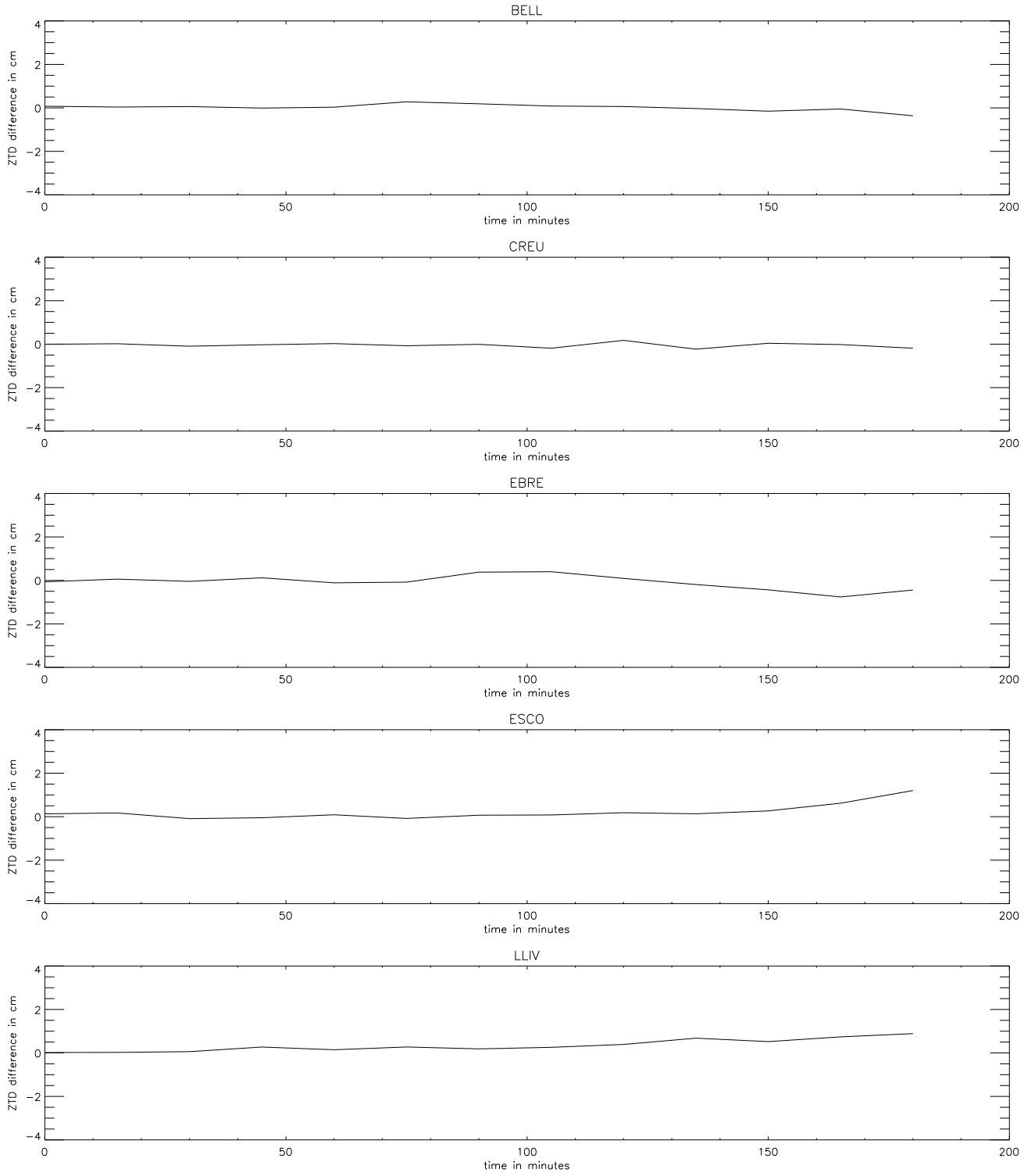


Figure 6.21: Evolution of the ZTD differences from optimal conditions between the GPS observations and the MM5 modeled values as a function of time for all GPS sites.

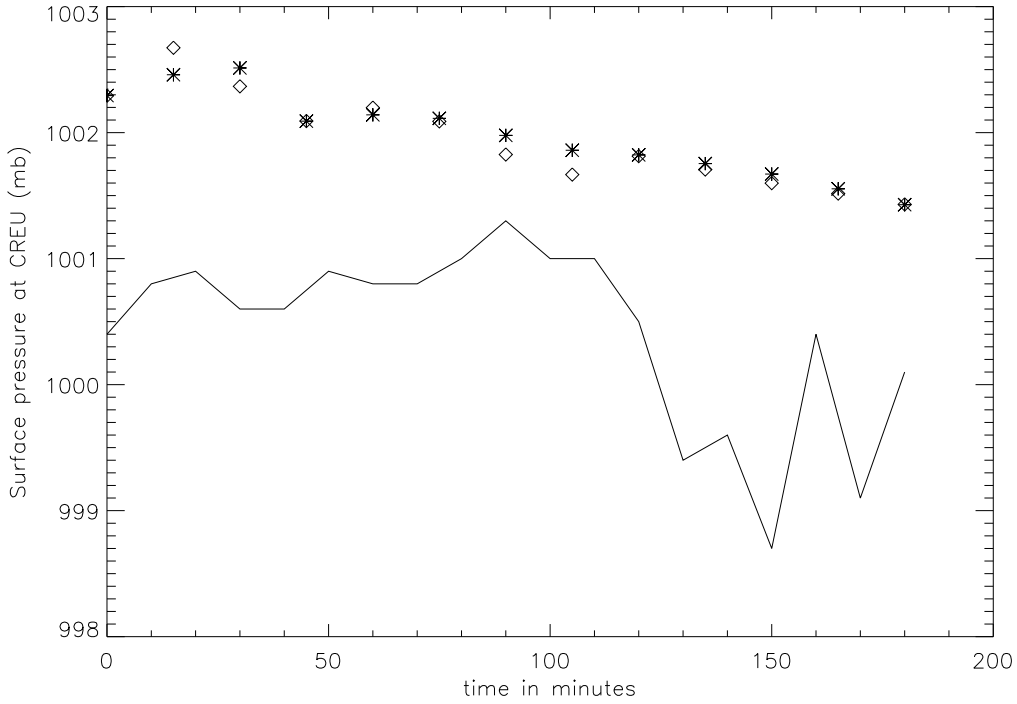
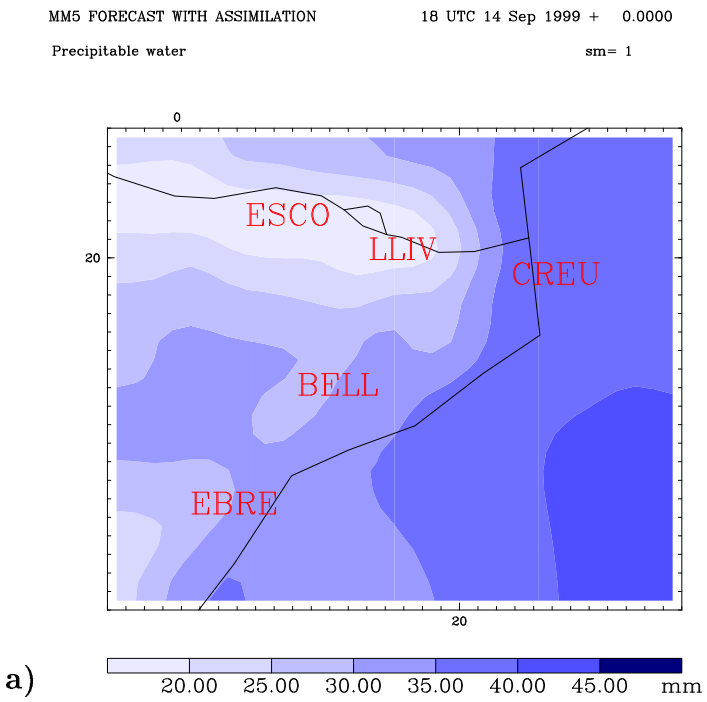
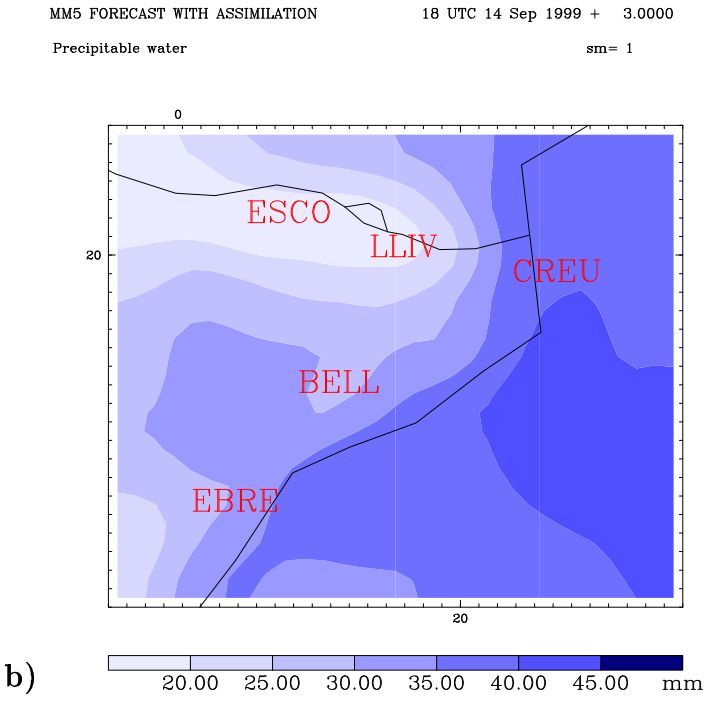


Figure 6.22: Surface pressure at CREU station as a function of time from the observed values (continuous line) and the pressure variable simulated with the model from the initial conditions ($*$), and from the optimal state of the atmosphere ($\diamond$).

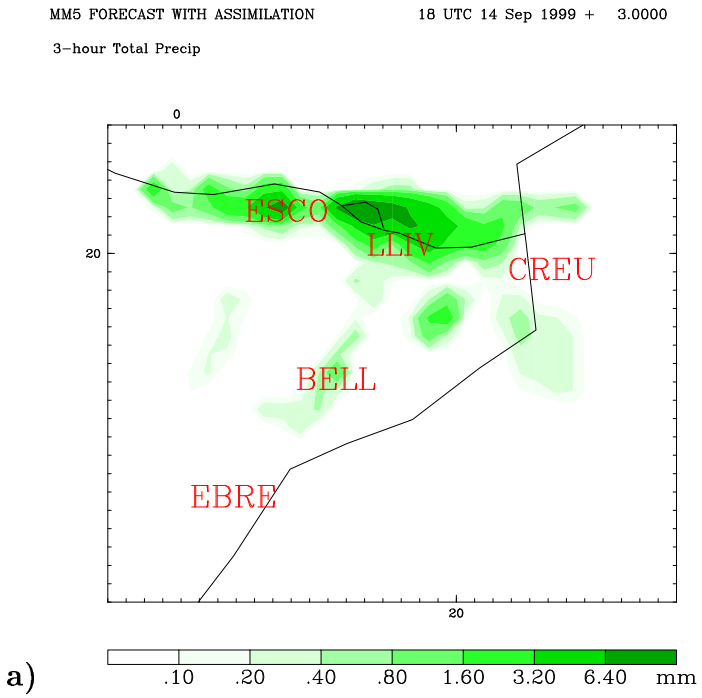


a)

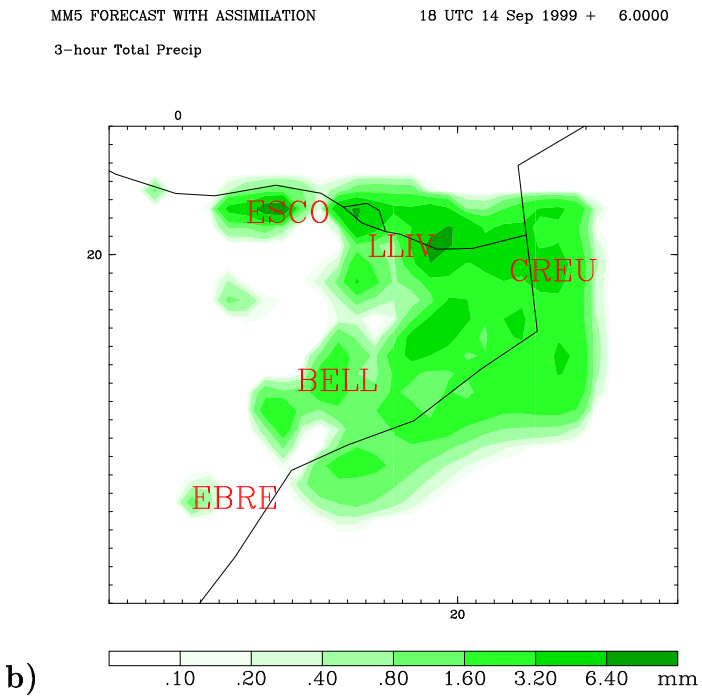


b)

Figure 6.23: Precipitable water fields for 14 September 1999 at (a) 18 UTC and (b) 21 UTC from the optimal initial conditions.



a)



b)

Figure 6.24: 3-hour accumulated precipitation from the optimal run ending at (a) 21 UTC and (b) 24 UTC 14 September 1999.

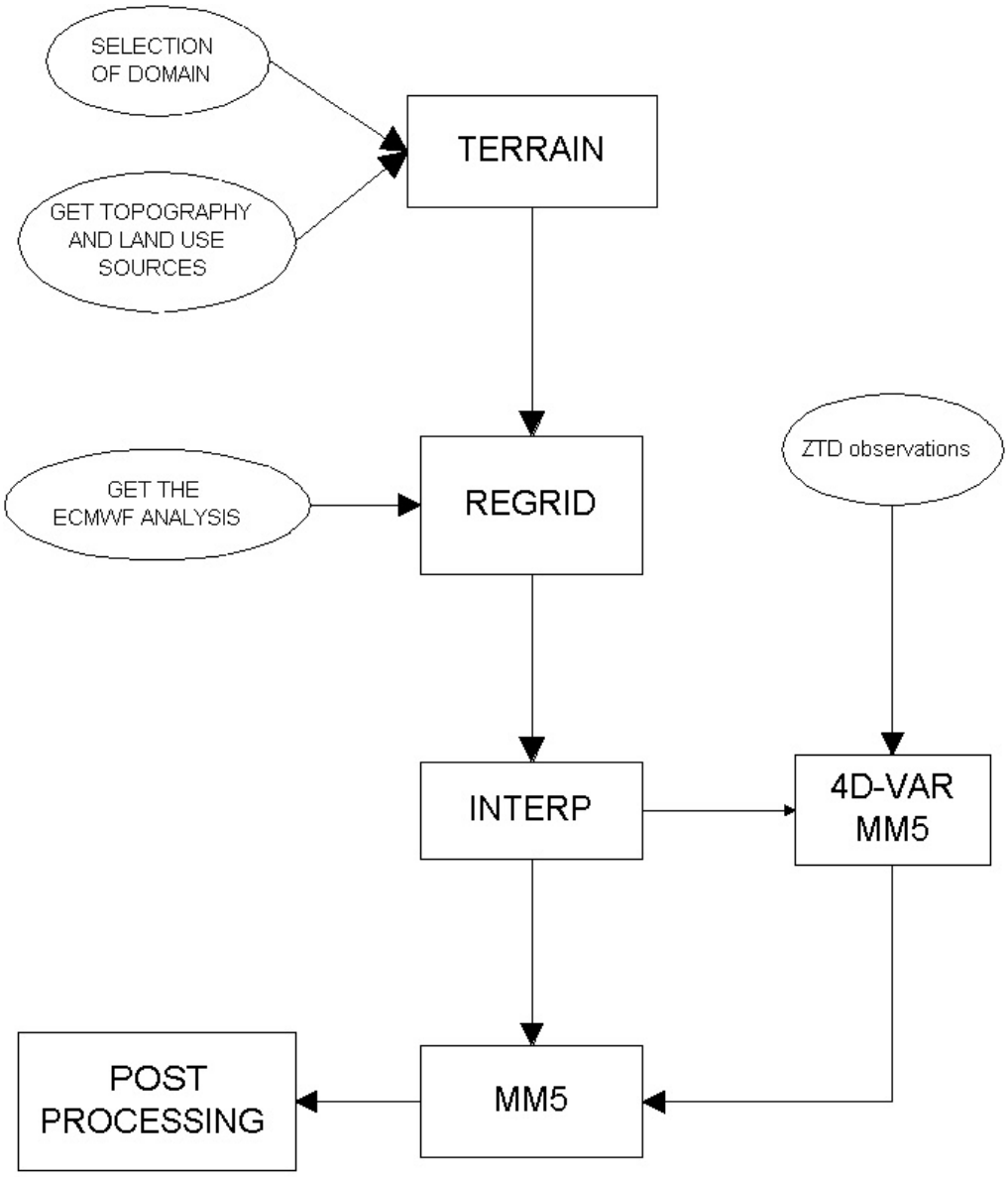


Figure 6.25: Flow chart of the sources and programs used in the 4D-VAR assimilation of the GPS-derived observations.

Chapter 7

Conclusions

The measurement and modeling of water vapor field is still associated with large uncertainties and is one of major error sources in short term forecast of precipitation. New observing systems have been available during the last years. Some of them can be used to derive the moisture content of the atmosphere and thus improve the short-range forecasts of rainfall. This has been the case of the GPS system.

In this thesis we have made use of this new instrumental technique to monitor meteorological models and to justify its assimilation in NWP models.

The most suitable meteorological application of this satellite technique is perhaps the assimilation of the GPS estimates into NWP models in a variational way. However, as a first step and previous to the assimilation, we have determined how GPS-derived observations compare with the values estimated by means of NWP models.

We have investigated the comparison between the GPS-derived observations and the values analyzed and forecasted with three mesoscale numerical models and under different meteorological situations. GPS data has been converted to atmospheric delays mapped to the zenith direction or to a more meteorological variable, such as precipitable water, depending on the availability of surface pressure measurements. The good agreement found among all the methods tested in this thesis (GPS, NWP models, and radiosondes) is very encouraging for the possible use of GPS atmospheric products in NWP. The availability of such measurements would potentially be useful for studying the distribution of the water vapor variability on small spatial scales. We have shown that GPS measurements can detect small (temporal) scale fluctuations and therefore can be used to evaluate meteorological models with fine resolution.

We have then explored the capability of the 4-dim variational assimilation tech-

nique to evaluate the impact of the GPS-derived observations in NWP models. This data assimilation method takes into account the high rate of the GPS data retrievals.

A global impact on humidity and precipitation fields is found when a large number of GPS-derived observations of PW are assimilated into the model. We have shown that the moisture field and the short-range forecast of precipitation are sensitive to the 4D-VAR assimilation of PW observations which are assumed very precise. Although there is not a linear dependence between PW and precipitation, the amount of rainfall and its distribution in the atmosphere are very sensitive to this variable (PW). The impact of the assimilation of the GPS data decreases for a smaller number of observations.

We have also analyzed the influence of the assimilation of the atmospheric delay instead of the precipitable water in a small GPS network. The largest impact of the assimilation of the GPS observations is found for the moisture component of the atmosphere. The results yielded an important improvement of the moisture field and precipitation forecast around the location of the GPS sites. This encourages the use of dense GPS networks. In general, this influence will depend on the topography and model resolutions, and on the area and the meteorological situation under study.

This thesis provides preliminary results on the impact of the GPS data into NWP models for selected areas and under particular meteorological situations. Some of the results have been published in international papers and others are in the process to be published. These are, at least, the results of the efforts from the last years.

There are many issues regarding this subject that should be further analyzed in the future. For example, it is necessary to develop and implement better covariance matrices of NWP models, and is still missing an observational covariance matrix for spatial correlations between GPS observations. The use of GPS-derived observations into meteorological models in an operational environment also requires the availability of the GPS data in near real time and the implementation of a good quality control check. All these topics should be investigated before the system could be operational in the weather services.

Appendix

Atmospheric delays from radiosonde profiles and propagation of statistical errors.

To estimate the zenith hydrostatic and wet delays from radiosonde profiles in Section (4.1) we used as input variables the geopotential height, the pressure, the temperature and the dew point temperature values at each sounding level.

Vapor pressure profiles were calculated from the dew point temperature using the Clausius-Clapeyron equation,

$$\frac{dP_{wet}^{sat}}{dT} = \frac{LP_{wet}^{sat}M_{rv}}{RT^2} \tag{7.1}$$

where P_{wet}^{sat} is the saturated vapor pressure, M_{rv} is the molecular weight of the water vapor, R is the universal gas constant and L is the latent heat. The vapor pressure of the atmosphere was obtained by using the dew point temperature instead of the atmospheric temperature in the expression above. The approximation used for (7.1) in the exercise described in Section (4.1) lets a latent heat which varies with temperature.

The dry pressure was then calculated by subtracting the vapor pressure the total pressure. The estimates of the ZWD and ZHD atmospheric delays, as well as their propagated errors, where calculated by integrating the refractivity profiles.

The errors associated to the radiosonde atmospheric delays where obtained from the propagation of the statistical errors associated to the variables involved in the ZWD and ZHD calculations. These errors were calculated as follows:

Let E be a function which depends on parameters v_i. The propagation of the errors associated with v_i to the function E is given by

$$(\delta E)^2 = \Sigma_i \left[\left(\frac{\delta E}{\delta v_i} \right)^2 \cdot (\delta v_i)^2 \right] \tag{7.2}$$

In our case, the function E is given by

$$E = (10^{-6})\Sigma_i N(z) \cdot \Delta z_i \tag{7.3}$$

where N is the neutral (wet or hydrostatic) atmospheric refractivity. The variables that contribute to the total error are the dry pressure, the water pressure, the temperature, the refractivity constants and the height increments. The errors associated with these variables were provided by Vaisala. We did not assume any error for the compressibility terms in (3.3) because they differ from unity by a few parts per thousand.

The error associated to the ZTD variable varied from 10 to 13 mm. The estimates of ZWD and ZHD values at the IGNE station are shown in the following table for all the campaign:

day	ZWD_S (mm)	ZHD_S (mm)	ZTD_S (mm)	$\sigma(ZWD_S)$ (mm)	$\sigma(ZHD_S)$ (mm)	$\sigma(ZTD_S)$ (mm)	ZTD_{GPS}) (mm)	ΔZTD (mm)
2.5	71	2146	2217	9	9	13	2225	-8
3.0	72	2140	2213	6	9	10	2216	-3
3.5	58	2126	2184	3	9	10	2200	-16
4.0	79	2112	2192	4	9	10	2193	-1
4.5	98	2103	2202	4	8	9	2202	0
5.0	91	2098	2189	5	7	9	2190	-1
5.5	69	2099	2168	2	8	8	2160	8
6.0	47	2103	2151	1	9	9	2154	-3
7.5	48	2115	2163	2	8	9	2180	-17
8.0	77	2123	2200	3	9	9	2200	0
8.5	80	2118	2199	4	8	9	2205	-6
9.	58	2120	2178	4	8	9	2193	-15
9.5	68	2115	2183	2	9	9	2194	-11
10.	62	2117	2180	3	9	9	2186	-6
11.5	109	2102	2211	6	9	11	2242	-31
12.	103	2110	2213	5	8	10	2227	-14
12.5	127	2133	2261	6	9	11	2229	32
13.	130	2111	2241	7	8	10	2262	-21
13.5	98	2107	2205	4	9	10	2205	0
14.	81	2099	2180	4	9	10	2203	-23
14.5	84	2107	2191	4	8	9	2209	-18
15.0	79	2118	2197	5	8	10	2197	0
15.5	52	2119	2171	3	9	9	2165	6
16.0	63	2124	2187	2	10	10	2162	25

In the table, ZTD_S and ZTD_{GPS} account for the radiosonde and GPS atmospheric delays, respectively, and ΔZTD is the difference between them. The χ^2 value for this data set is given by

$$\chi^2 = \frac{1}{N} \sum \frac{\Delta ZTD^2}{\sigma_S^2 + 5^2} = 1.7 \tag{7.4}$$

In the expression above, σ_S is the σ associated to the radiosonde calculations of ZTD. We chose a mean value of 5 mm for the σ of the GPS-derived ZTD. From the χ^2 value

obtained in the equation above, we should increment the error associated to the GPS observations in a 1.3 factor (or a 30 % of its value) to get χ^2=1.

Bibliography

Abbs, D. J., 1999: A numerical modeling study to investigate the assumptions used in the calculation of probable maximum precipitation. *Water Resour. Res.*, **35**, 785-796.

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.

Arakawa, A., 1966: Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. Part I. *J. Comp. Phys.*, **1**, 119.

Arakawa, A., 1984: The use of integral constraints in designing finite-difference schemes for the two-dimensional advection equation. *ECMWF Seminar 1983 Numerical Methods for Weather Prediction*, **1**, 159.

Askne, J. and H. Nordius, 1986: Estimation of tropospheric delay for microwaves from surface weather data. *Radio Sci.*, **22**, 379-386.

Behrend, D., L. Cucurull, J. Vilà, and R. Hass, 2000: An Inter-Comparison Study to Estimate Zenith Wet Delays Using VLBI, GPS, and NWP Models, 2000. *Earth, Plantes and Space*, **52**, 10, 691-694.

Bertin J. J., 1996: *Global Positioning System: Theory and Applications.* Volume 1, edited by B. W. Parkinson and J. J. Spilker Jr., Progress in Astronautics and Aeronautics, 517-546.

Bevis, M., S. Businger, T. A. Herring, C. Rocken, R. A. Anthes, and R. H. 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, S. R. Chiswell, T. A. Herring, R. A. Anthes, C. Rocken, and R. H. Ware, 1994: GPS Meteorology: mapping zenith wet delays onto Precipitable Water. *J. Appl. Meteorol.*, **33**, 379-386.

Bouttier, F., and P. Courtier, 1998: *Data assimilation concepts and methods.* Meteorological Training Course. Internal report from the ECMWF, 64 pp.

Businger S., S. R. Chiswell, M. Bevis, J. Duan, R. A. Anthes, C. Rocken, R. H. Ware, T. Exner, T. Van Hove, and F. S. Solheim, 1996: The promise of GPS in atmospheric monitoring. *Bull. Am. Meteorol. Soc.*, **77**, 5-18.

Carlsson, T. R., G. Elgered, and J. M. Johansson, 1996: A Quality Assessment of the Wet Path Delay Estimated From GPS Data. Proc. of the 11th Working Meeting on European VLBI for Geodesy and Astrometry, Research Report 177, Ed. G. Elgered, Onsala Space Observatory, Chalmers Univ. of Tech., 89-95.

Codina, B., M. Aran, S. Young, and A. Redaño, 1997: Prediction of a Mesoscale Convective System over Catalonia (Northeastern Spain) with a Nested Numerical Model. *Meteor. Atmos. Phys.*, **62**, 9-22.

Cucurull, L., P. Sedó, D. Behrend, E. Cardellach, and A. Rius, 2001a: Integrating NWP products into the analysis of GPS observables. *Submitted to Phys. Chem. Earth.*

Cucurull, L., J. Vilà, and A. Rius, 2001b: Zenith Total Delay study of a Mesoscale Convective System: GPS observations and Fine-Scale Modeling . *Submitted to Tellus A.*

Cucurull, L., B. Navascués, G. Ruffini, P. Elósegui, A. Rius, and J. Vilà, 2000: The use of GPS to validate NWP systems: the HIRLAM model. *J. Atmos. and Ocean. Tech.*, **17**,6, 773-787.

Cucurull, L., F. Vandenberghe, G. Ruffini, J. Vilà, A. Flores, A. Rius and B. Navascués, 1999: The GPS system as a tool to validate NPW models. *Geophysical Research Abstracts*, **1**.

Cucurull, L., and F.Vandenberghe, 1999: Comparison of PW estimated from MM5 and GPS data. *Proc. of the 9th PSU/NCAR Mesoscale Model User's Workshop*, 75-78.

Cucurull, L., J. Vidal, A. Rius, and G. Ruffini, 1998: Asimilación de datos GPS en modelos numéricos de mesoescala. *1st Asamblea Hispano Portuguesa de Geodesia y Geofísica Abstracts*, **1**, 338.

Chen and Herring, 1997: Effects of atmospheric azimuthal asymetry on the analysis of space geodetic data. *J. Geophys. Res.*, **102**, 20489-20502.

Davis, J. L., T. A. Herring, I. I. Shapiro, A. E. Rogers, and G. Elgered, 1985: Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline lengths. *Radio Sci.*, **20**, 1593-1607.

de Pondeca, M.S.F.V. and X. Zou, 2001: Moisture retrievals from simulated zenith delay "observations" and their impact on short-range precipitation forecasts. *Tellus*, **53 A**.

Dixon, T. H., and S. Kornreich Wolf, 1990: Some tests of wet tropospheric calibration for the CASA Uno Global Positioning System experiment. *Geophys. Res. Lett.*, **17**, 203-206.

Duan, J., M. Bevis, P. Fang, Y. Bock, S. R. Chiswell, S. Businger, C. Rocken, F. S. Solheim, T. Van Hove, R. H. Ware, S. McClusky, T. A. Herring, and R. W. King, 1996: GPS Meteorology: Direct Estimation of the Absolute Value of Precipitable Water. *J. Appl. Meteorol.*, **35**, 830-838.

Dudhia, J., 1993: A non-hydrostatic version of the Penn State/NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**, 1493-1513.

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

Elósegui, P., A. Rius, J. L. Davis, G. Ruffini, S. J. Keihm, B. Bürki, and L. P. Kruse, 1998: An Experiment for Estimation of the Spatial and Temporal Variations of Water Vapor Using GPS Data. *Phys. Chem. Earth*, **23**, 125-130.

Emardson, T. R., G. Elgered, and J. M. Johansson, 1998: Three months of continuous monitoring of atmospheric water vapor. *J. Geophys. Res.*, **103**, 1807-1820.

Emardson, T. R. and H. J. P. Derks, 1999: On the relation between the wet delay and the integrated precipitable water vapor in the European atmosphere. *Meteorol. Appl.*, **6**, 1-12.

Flores, A., J. Vilà, L. Gradinarsky, and A. Rius, 2001: Tomography of the Lower Troposphere Using a Small Dense Network of GPS Receivers. *IEEE Trans. Geosci. Remote Sensing*, **39**,2.

Flores, A., A.Rius, and G. Ruffini, 2000a: 4D tropospheric tomography using GPS slant wet delays. *Ann. Geophysicae*, **18**, 223-224.

Flores, A., A. Escudero, M.J. Sedó, and A. Rius, 2000b: A near real time system for tropospheric monitoring using IGS hourly data. *Earth, Planets and Space*, **52**, 681-684.

Flores, A., 1999: *Atmospheric Tomography Using Satellite Radio Signals.* PhD thesis, Insitut d'Estudis Espacials de Catalunya.

Grell, G.A., J. Duhia, and D.R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). NCAR Tech. Note, NCAR/TN-398+STR, National Center for Atmospheric Research, Boulder, CO, 138 pp.

Healy, S. and L. Eyre, 1999: Retrieving temperature, water vapor and surface pressure information from refractive index profiles derived by radio occultation: a simulation study. *Quart. J. R. Met. Soc.* Accepted.

Herring T. A., J. L. Davis, and I. I. Shapiro, 1990: Geodesy by radio interferometry: The application of Kalman filtering to the analysis of very long baseline interferometry data, *J. Geophys. Res.*, **95**, 561-581.

Herring T. A., 1999: Geodetic applications of GPS. *Proceedings of the IEEE*, **87**,1.

Holton, J. A., 1992: An introduction to dynamic meteorology, 3rd. ed., edited by Academic Press Limited, 500 pp.

Jarrau, M., and A. J. Simmons, 1984: The spectral technique. ECMWF Seminar 1983. Numerical Methods for Weather Prediction, 2, 1.

Källen, E., 1996: HIRLAM documentation manual, System 2.5. Edited by E. Källen,

120 pp. Available from SMHI, S-601 79 Norrkoping, Sweden, attention Erland Källen.

Kaplan, M. L., and V. M. Karyampudi, 1992: Meso-beta scale numerical simulations of terrain drag-induced frontogenesis. Part I: Midtropospheric frontogenesis. *J. Meteorol. Atmos. Phys.*, **49**, 133-156.

Kaplan, M. L., J. W. Zack, V. C. Wong, and J. J. Tuccillo, 1982: Initital results from a mesoscale atmospheric simulation system and comparsions with an AVE-SESAME I data set. *J. Mon. Wea. Rev.*, **110**, 1564-1590.

Kuo, Y.-H., Y.-R. Guo, and E. R. Westwater, 1993: Assimilation of precipitable water into mesoscale numerical model. *Mon. Wea. Rev.*, **121**, 1215-1238.

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

Kursinski, E. R., 1997: *The GPS Radio Occultation Concept: Theoretical Performance and Initial Results.* PhD thesis, California Insitute of Technology.

Lilly, D. K., 1964: Numerical solutions for the shape-preserving two-dimensional thermal convection element. *J. Atmos. Sci.*, **21**, 83

Liu, D. C. and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimitaztion. *Math. Program.*, **45**, 503-528.

Llasat, M. C., and M. Puigcerver, 1992: Pluies extremes in Catalogne: Influence orographique et caracteriques synoptiques. *Hydrol. Continen.*, **71**/2, 99-115.

Lönnberg, P., and D. Shaw (Editors), 1987: ECMWF Data Assimilation Scientific Documentation. Research Manual 1. 3/92 3rd Edition. *ECMWF Meteorol. Bull.*, 100 pp. Available from ECMWF library, Shinfield Park, Reading, Berkshire, RG2 9AX, UK.

Lorenc, A.C., 1986: Analysis methods for numerical weather prediction. *Quart. J. R. Met. Soc.* **112**, 1177-1194.

Machenhauer, B., 1979: The spectral method. Numerical Methods used in Atmospheric Models. *GARP Publ. Ser.* **17**, 2.

Navon, I. M., X. Zou, J. Derber, and J. Sela, 1992: Variational data assimilation with and adiabatic version of the NMC spectral model. *Mon. Wea. Rev.*, **120**, 1433-1446.

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

Owens, J. C., 1967: Optical refractive index of air: dependence on pressure, temperature and composition. *Appl Opt.*, **6**, 51-58.

Ramis, C., M. C. Llasat, A. Genovés, and A. Jansà, 1994: The October-1987 floods in Catalonia: synoptic and mesoscale mechanism. *Met. Apps.*, **1**, 337-350.

Rapp, R.H., Y.M. Wang, and N.K. Pavlis, 1991: The Ohio State 1991 Geopotential and Sea Surface Topography Harmonic Coefficient Models. Report No. 410, Department of Geodetic Science and Surveying, The Ohio State University, 100 pp. Available from the Ohio State University, Columbus, Ohio 43210.

Rius, A., G. Ruffini, and L. Cucurull, 1997: Improving the vertical resolution of ionospheric tomography with GPS occultations. *Geophy. Res. Lett.*, **24**, 2291.

Rius, A., L. Cucurull, M.J. Sedó, and B. Navascués, 2000: On the use of GPS Zenith Total Delays to monitor NWP Systems. *Proceedings of the Americ. Meteorol. Soc.*, 224-228.

Rocken, C., R. H. Ware, T. Van Hove, F. S. Solheim, C. Alber, J. Johnson, M. Bevis, and S. Businger, 1993: Sensing Atmospheric Water Vapor with the Global Positioning System. *Geophy. Res. Lett.*, **20**, 2631-2634.

Rocken, C.,T. Van Hove, J. Johnson, F. S. Solheim, R. H. Ware, M. Bevis, and S. R. Chiswell, 1995: GPS/STORM - GPS Sensing of Atmospheric Water Vapor for Meteorology. *J. Atmos. and Ocean. Tech.*, **12**, 468-478.

Rocken, C., T. Van Hove, and R. H. Ware, 1997: Near real-time GPS sensing of atmospheric water vapor. *Geophy. Res. Lett.*, **24**, 3221-3224.

Romero, R., C. Ramis, S. Alonso, C.A. Doswell III, and D.J. Stensrud, 1998: Mesoscale model simulation of three heavy precipitation events in the western Mediterranean region. *Mon. Wea. Rev.*, **126**, 1859-1881.

Rowland, S and I. S. A. Isaksen, 1988: The Changing atmosphere. Edited by Wiley, London.

Ruffini, G., L. P. Kruse, A. Rius, B. Bürki, L. Cucurull, and A. Flores, 1999a: Estimation of Tropospheric Zenith Delay Gradients over the Madrid area using GPS and WVR data. *Geophy. Res. Lett.*, **26**, 447-450.

Ruffini, G., L. Cucurull, A. Flores and A. Rius, 1999b: A PIM-aided Kalman Filter for GPS Tomography of Ionospheric Electron Content, 1999. *Phys. Chem. Earth*, **24**, 365-369.

Ruffini, G., E. Cardellach, A. Flores and L. Cucurull, 1998: Ionospheric Calibration of Radar Altimeters Using GPS Tomography. *Geophy. Res. Lett.*, **25**, 20.

Saastamoinen, J., 1972: Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites. *The Use of Artificial Satellites for Geodesy, Geophys. Monogr. Se.*, **15**, eds. S.W. Henriksen, A. Mancini, and B. H. Chovitz, 247-251, AGU, Washington, D.C., U.S.A.

Sadourny, R., 1975: The dynamics of finite-difference models of the shallow-water equations. *J. Atmos. Sci.*, **32**, 680.

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

Talagrand, O., 1991: *Four-dimensional variational assimilation.* Meteorological Training Course. Internal Report from ECMWF, 30 pp.

Thayer, D. M., 1974: An improved equation for the radio refractive index of air. *Radio Sci.*, **9**, 803-807.

Thompson, A. R., J. M. Moran, and G. W. Swenson, Jr., 1986: Interferometry and Synthesis in Radio Astronomy. John Wiley and Sons, 534 pp.

Tregoning, P., R. Boers, D. O'Brien, and M. Hendy, 1998: Accuracy of absolute precipitable water vapor estimates from GPS observations. *J. Geophys. Res.*, **103**, 28701-28710.

Ware, R. H., C. Alber, C. Rocken, and F. S. Solheim, 1997: Sensing integrated water

vapor along GPS ray paths. *Geophy. Res. Lett.*, **24**, 417-420.

Webb, F. H., and J. F. Zumberge, 1993: An introduction to the GIPSY-OASIS-II. *JPL Publ.*, D-11088, 300 pp.

Yang, X., B. H. Sass, G. Elgered, J. M. Johansson, and T. R. Emardson, 1999: A comparison of the integrated water vapor estimation by a NWP simulation and GPS observations. *J. Appl. Meteorol.*, **38**, 941-956.

Yuan, L., R. A. Anthes, R. H. Ware, C. Rocken, W. D. Bonner, M. Bevis, and S. Businger, 1993: Sensing climate change using the Global Positioning System. *J. Geophys. Res.*, **98**, 14925-14937.

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

Zou, X. and Y.-H. Kuo, 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.*, **124**, 2859-2882.

Zou, F. Vandenberghe, M. Pondeca, and Y.-H. Kuo, 1997: Introduction to adjoint techniques and the MM5 Adjoint Modeling System. *NCAR Technical Note*, NCAR/TN-435-STR, 110 pp.