

WegCenter/UniGraz Technical Report for FFG-ALR No. 2/2006

FFG-ALR Study:

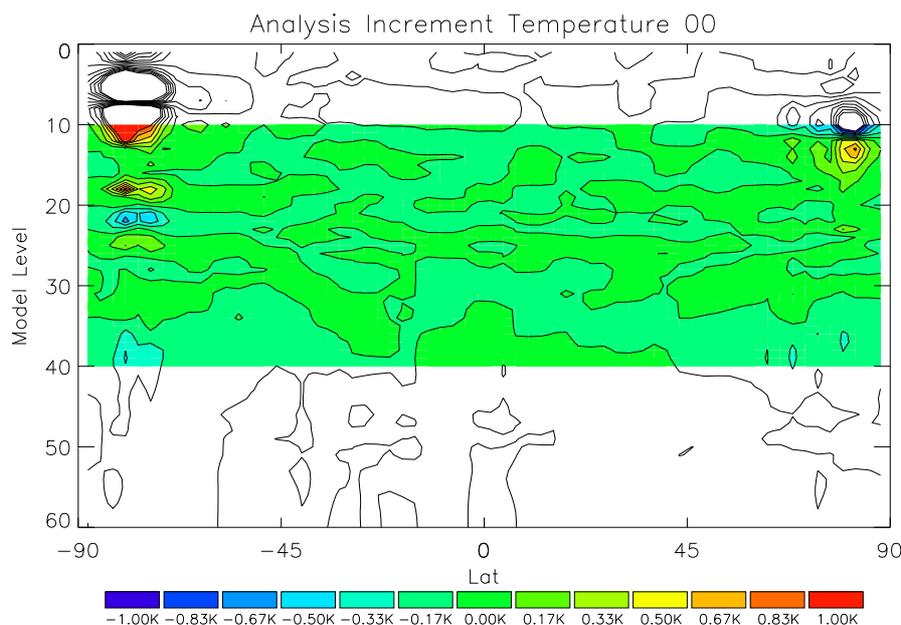
CHAMPCLIM – Radio Occultation Data Analysis Advancement and
Climate Change Monitoring Based on the CHAMP/GPS Experiment
[Contract No: ASAP-WV-203/05 – September 2005]

Assimilation Report
[WP4: MONITOR]

CHAMP Radio Occultation Data Assimilation into ECMWF Fields for Global Climate Analyses

A. Löscher, U. Foelsche, and G. Kirchengast

Wegener Center for Climate and Global Change (WegCenter),
University of Graz, Graz, Austria



February 2006

Contents

Introduction	1
1 Assimilation System Setup	3
1.1 Coordinate System	3
1.2 Temperature, Specific Humidity, and Surface Pressure Analysis	3
1.3 Variational Assimilation 3D-VAR	3
1.3.1 3D-VAR and Incremental 3D-VAR	3
1.4 Implementation Technique	4
1.4.1 Control Variables	4
1.4.2 Minimization	4
1.4.3 Preconditioning	4
1.4.4 The Adjoint Technique	4
1.5 Observation Operators	4
1.5.1 Interpolation Operator	4
1.5.2 Refractivity Operator	5
1.5.3 Vertical Coordinate Operator	7
1.6 Observation Error Covariance	8
1.6.1 Formulation of the Observation Error Covariance Matrix	9
1.7 Incest Problem	9
1.8 Background Error Covariance	10
1.8.1 Temperature, Specific Humidity, and Surface Pressure	10
1.8.2 Vertical and Horizontal ECMWF Temperature Error Correlations	10
1.9 Control Space Transformations	10
1.9.1 Vertical Control Variable Transform	15
1.9.2 Horizontal Control Variable Transform	17
1.9.3 Recursive Filters	17
1.10 Horizontal Background Error Covariances	20
1.11 The Minimization Algorithm	20
2 Data Preprocessing	23
3 Quality Control	23
4 Background Data	23
4.1 ECMWF Analysis Fields	24
4.2 Accuracy of ECMWF Analysis Fields	25
5 Radio Occultation Data	27
5.1 The RO Technique	27
5.1.1 Unique Advantages for Monitoring of Atmospheric Key Parameters	27
5.1.2 Characteristic Horizontal and Vertical Resolution of RO Measurements	28
5.2 Retrieval	29
5.2.1 General RO Retrieval	29
5.3 Data Products	31
5.3.1 Refractivity Profiles	31
5.3.2 Temperature Profiles	32
5.3.3 Humidity Profiles	32
5.3.4 Geopotential Height	33
5.3.5 Pressure	33

5.3.6	Total Electron Content	33
5.4	The CHAMP Satellite	33
6	Validation	35
6.1	Validation of Adjoint Code	35
6.2	System Test Runs	35
	Conclusions and Outlook	43
	Acknowledgments	43
A	Symbols	45
B	Constants	47
C	List of Acronyms	49
D	List of Figures	51
E	List of Tables	53
	References	55

Introduction

This document contains a description of the data assimilation system developed within the framework of the CHAMPCLIM-2 project (Contract No: ASAP-WV-203/05 – September 2005) in order to conduct climate studies using radio occultation (RO) data and forecast fields of the European Centre for Medium-range Weather Forecasts (ECMWF). It is a mature version of the system, which was presented in Loescher (2004) in detail.

The method of choice is a 3D-Var (Three Dimensional Variational Assimilation) system, which allows to merge information from different sources (in our case RO data and ECMWF short-range forecast fields) under the assumption of known error structures, in a statistically optimal way. This report is a technical description of the system and is organized as follows.

Section 1

Here the specific implementation of the temperature, specific humidity, and surface pressure (TQPsurf) 3D-Var system is described in detail. The realization of control space transformations and the pre-conditioning is explained, as well as the use of recursive filters within this framework. The observation and background covariance matrices are defined and the data thinning as part of the pre-processing of observations is described here. The observation operators are explained in detail and the minimization routine is introduced.

Section 2

The kind of data that is used for the assimilation experiments is introduced. On the one hand the ECMWF model data, which we used as background, and on the other hand the CHAMP RO data. Both sources of information are specified here in some detail, with a brief outlook on future model updates and RO missions.

Section 3

Within this section, the results of representative test runs with quasi operational CHAMP data from July 2003 are shown and the convergence behavior of that experimental assimilation setup is presented.

Conclusions and Outlook

The results of the first assimilation experiments and future perspectives are discussed. In the future the developed RO data assimilation system will be further fine-tuned and applied to create sequences of monthly climate analyses. These will contribute to systematic long-term monitoring of climate variability and change.

1 Assimilation System Setup

The system is implemented as a 3D-Var scheme using control space transformations and recursive filters. The dimensions of the background are flexible, but due to the specific application a GCM compliant Gaussian grid corresponding to T42L60, i.e., 64 latitudes \times 128 longitudes comprising 60 model levels is used (the system was tested with a T21L60, i.e., 32 latitudes \times 64 longitudes setup too).

1.1 Coordinate System

As mentioned above a GCM compliant Gaussian grid corresponding to T42L60, i.e., 64 latitudes \times 128 longitudes, comprising 60 model levels is used. The vertical coordinate system is derived from the surface pressure and the A and B vectors provided by ECMWF. This vertical grid comprises 60 hybrid levels. From this basic vertical coordinate system grids of geopotential height, and geometric height (over reference ellipsoid) can be derived. The assimilation scheme can be used either with geometric height or geopotential height. The necessary operators are discussed later within this chapter.

1.2 Temperature, Specific Humidity, and Surface Pressure Analysis

The assimilation scheme directly updates the temperature, specific humidity, and surface pressure input fields, what means that all fields are interpolated separately. At each iteration the new vertical fields (like pressure, geopotential, and height grid) have to be derived from the updated surface pressure field. That means that the whole vertical coordinate system is shifted up and down. That implies the necessity to calculate the interpolation coefficients for the background fields at observation location at every cycle (*simulation*) before the refractivity can be derived. After comparison of background refractivity (also denoted as *model observation*) and observation the gradients of the input fields and observations are calculated and a suitable correction is applied.

1.3 Variational Assimilation 3D-VAR

The solution of the minimization problem can be performed either in terms of full-fields $J(\mathbf{x}_a)$ or the analysis of increments $J(\mathbf{x}_a) = J(\delta\mathbf{x}_a = \mathbf{x}_a - \mathbf{x}_b)$ (Bouttier and Courtier, 1999). The latter solution method provides *optimal* analysis increments, which are added to the unmodified background field. This procedure has a number of advantages like the use of linearized control variable transforms which allow the straightforward use of adjoints to calculate the gradient of the cost function. Another advantage is that any imbalance introduced through the analysis procedure is limited to the small increments which are added to the balanced first guess.

1.3.1 3D-VAR and Incremental 3D-VAR

3D-VAR uses observations and background information in a statistical optimal way to derive combined analysis fields:

$$\mathbf{x}_a = \text{Arg min } J(\mathbf{x}) \quad (1)$$

$$J(\mathbf{x}) = J_b(\mathbf{x}) + J_o(\mathbf{x}) \quad (2)$$

$$J(\mathbf{x}) = \frac{1}{2} \left\{ (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + (\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y} - H(\mathbf{x})) \right\} \quad (3)$$

$$\nabla J_{\mathbf{x}_0} = \mathbf{B}_0^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) + \mathbf{H}^T \mathbf{R}^{-1} [H(\mathbf{x}_0) - \mathbf{y}] = 0. \quad (4)$$

The solution of the minimization problem requires the calculation of the gradient $\nabla J_{\mathbf{x}_0}$ and can be performed either in terms of full fields \mathbf{x} or in terms of an analysis of the increments $\delta\mathbf{x}$:

$$J(\delta\mathbf{x})=J(\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b) \quad (5)$$

$$J(\mathbf{x}) = \frac{1}{2} \{ \delta\mathbf{x}^T \mathbf{B}^{-1} \delta\mathbf{x} + (H(\delta\mathbf{x}) - \mathbf{d})^T \mathbf{R}^{-1} (H(\delta\mathbf{x}) - \mathbf{d}) \} \quad (6)$$

$$\mathbf{d} = \mathbf{y} - H(\mathbf{x}) \quad (7)$$

$$\nabla J_{\delta\mathbf{x}_0} = \mathbf{B}^{-1} \delta\mathbf{x} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}(\delta\mathbf{x}) - \mathbf{H}^T \mathbf{R}^{-1} \mathbf{d} = 0 \quad (8)$$

where the analysis is found by adding the final increment to the first guess:

$$\mathbf{x}_a = \mathbf{x}_b + \delta\mathbf{x}_a. \quad (9)$$

1.4 Implementation Technique

1.4.1 Control Variables

The control variables used in the analysis are temperature, specific humidity, and surface pressure. To avoid negative specific humidities in the analysis the field is checked for negative values at every iteration, which are corrected to a value of 10^{-6} , the corresponding gradients are set to 0. The cross correlations between the control variables are assumed to be small enough to be neglected. This assumption serves to effectively block-diagonalize the background error covariance matrix. For each control variable there still remains both, horizontal and vertical correlations. Those are assumed to be separable, which is a widely used assumption.

1.4.2 Minimization

The cost function is minimized by using an iterative descent algorithm, which is in our case the L-BFGS-B routine. The cost of the analysis is proportional to the number of evaluations of the cost function and of its gradient, denoted as *simulations*. If a new state \mathbf{x} is found, an *iteration* is performed, which means that to find a new \mathbf{x} , several simulations may be required (cf., Section 1.11).

1.4.3 Preconditioning

A preconditioning of the problem is performed as part of the control space transformations (cf., Section 1.9).

1.4.4 The Adjoint Technique

The adjoint technique allows to calculate the gradients, needed for the minimization procedure, at reasonable numerical cost (INRIA, 2002).

1.5 Observation Operators

1.5.1 Interpolation Operator

To calculate the background values at the spatial location of the measurements two bilinear horizontal and one logarithmic vertical interpolation are performed for each observation.

Horizontal Interpolation The bilinear interpolation consists of a weighted average of the four surrounding grid points to determine their interpolated value. Two linear interpolations on opposite sites are performed followed by a consecutive interpolation of these intermediate results. This horizontal interpolation is performed for the atmospheric layers above and below the observation

$$f(x, y) = (1 - u)(1 - v)f_{i,j} + u(1 - v)f_{i+1,j} + (1 - u)v f_{i,j} + uv f_{i+1,j+1} , \quad (10a)$$

$$u = \frac{(x - x_i)}{(x_{i+1} - x_i)} , \quad (10b)$$

$$v = \frac{(y - y_j)}{(y_{j+1} - y_j)} , \quad (10c)$$

where $x_i < x < x_{i+1}$ and $y_j < y < y_{j+1}$.

Vertical Interpolation Due to the fact of a globally non uniform vertical grid, the heights of the horizontal interpolated values are also calculated by bilinear interpolation from the vertical background grid.

Linear and Logarithmic Interpolation Given the background values above and below the spatial location of the observation either a logarithmic or a linear interpolation is performed to get the final value of the background at the location of the observation.

$$\text{Linear Interpolated Value} = Z_1 * \text{Weight}_1 + Z_2 * \text{Weight}_2 \quad (11)$$

$$\text{Logarithmic Interpolated Value} = e^{(\log(Z_1)*\text{Weight}_1)+(\log(Z_2)*\text{Weight}_2)} \quad (12)$$

where Z_1 and Z_2 denote the horizontal interpolated values of the layers above and below the observation which are weighted with Weight_1 and Weight_2 , calculated from the vertical distance between observation and Z_1 , Z_2 . Temperature is linear interpolated, specific humidity, and pressure are logarithmic interpolated.

1.5.2 Refractivity Operator

To calculate the refractivity at a given point, it is necessary to know the atmospheric variables absolute temperature, specific humidity, and pressure. There are two standard formulas used, which are the Thayer and the Smith-Weintraub formula. The Thayer Formula is the more accurate one, the Smith-Weintraub formula is basically the same, but assumes an ideal gas.

Thayer Formula

$$N = k_1 \cdot \frac{p_A}{T} \cdot \frac{1}{z_A} + k_2 \cdot \frac{e}{T} \cdot \frac{1}{z_W} + k_3 \cdot \frac{e}{T^2} \cdot \frac{1}{z_W} \quad (13)$$

Smith-Weintraub Formula

$$N = k_1 \cdot \frac{p_A}{T} + k_2 \cdot \frac{e}{T} + k_3 \cdot \frac{e}{T^2} \quad (14)$$

For further explanations of parameters see Appendix B.

Symbol	Parameter	Unit
N	= Refractivity	[N units]
T	= Absolute Temperature	[K]
e	= Partial pressure of water vapor	[hPa]
p_A	= Partial pressure of “dry air”	[hPa]
z_A	= Compressibility factor of “dry air”	[1]
z_W	= Compressibility factor of water vapor	[1]
k_1	= Empirical constant Thayer Formula	[K hPa ⁻¹]
k_2	= Empirical constant Thayer Formula	[K hPa ⁻¹]
k_3	= Empirical constant Thayer Formula	[K ² hPa ⁻¹]

Table 1: Parameters and variables used in Thayer and Smith-Weintraub formula.

Calculation of Refractivity Fields from ECMWF Fields Given the fields of temperature, surface pressure, and specific humidity (in our case forecast fields of ECMWF), we can calculate the field of refractivity using either the Thayer or Smith-Weintraub formula. A comparison of the two formulas shows no significant differences. A simplified version of the Smith-Weintraub formula was chosen as forward operator, to calculate refractivity from temperature, humidity, and surface pressure analysis fields.

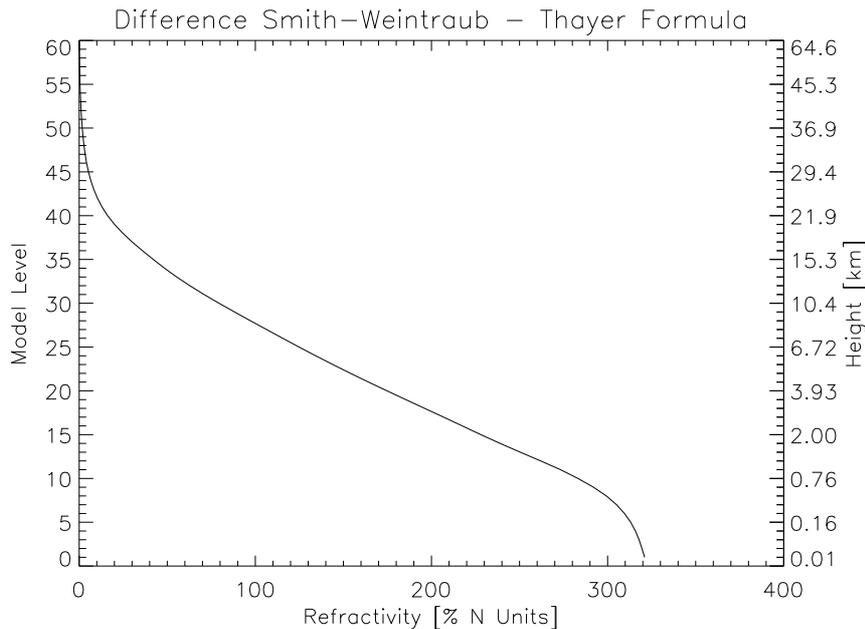


Figure 1: Mean global refractivity profile calculated from T42L60 ECMWF analysis fields.

Fig. 2 shows the negligible difference of the mean global refractivity profile calculated with the Thayer and the Smith Weintraub formula using an ECMWF T42L60 analysis field (Date: 2003.01.03; 12 UTC).

The use of a very simple formulation is justified for that application so the refractivity forward operator can be written as:

$$N = k_1 \cdot \frac{p_A}{T} + k_3 \cdot \frac{e}{T^2}. \quad (15)$$

This simple formulation is generally accepted and introduces no significant errors. Further we

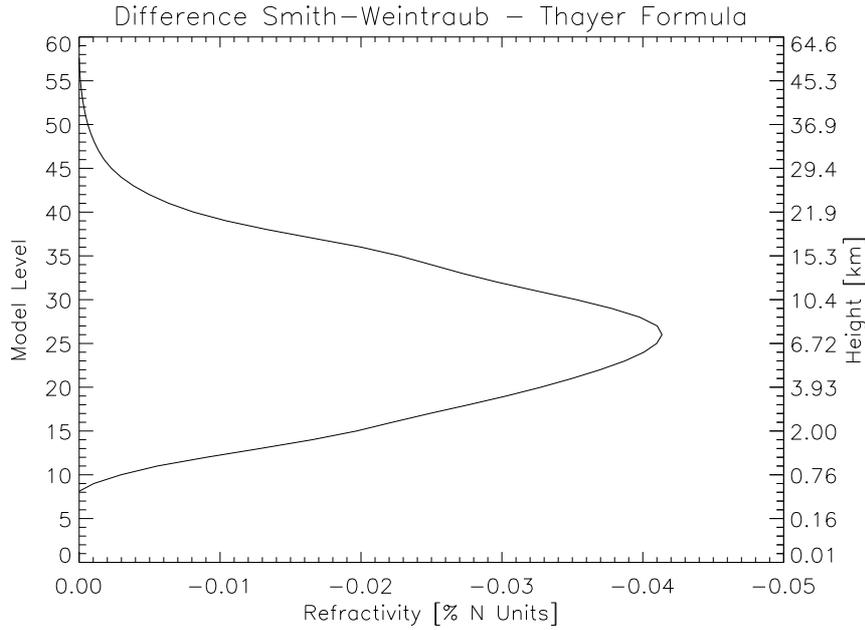


Figure 2: Difference in global mean refractivity calculated with Smith-Weintraub and Thayer Formula.

need the partial pressure of water vapor e which is given by

$$p_W = \frac{M_W}{M_D} \quad (16)$$

$$e = \frac{q \cdot p}{p_W + (1.0 - p_W) \cdot q}, \quad (17)$$

where e denotes the water vapor pressure, M_W , respectively M_D the molecular mass of water vapor and dry air.

1.5.3 Vertical Coordinate Operator

As can be seen in Eq. 13 and Eq. 14 the pressure at the location of the observation is needed to calculate the refractivity. ECMWF provides temperature, specific humidity, and surface pressure fields. The background pressure field is derived by a series of operators. This operators are also used to set up the vertical coordinate system of the assimilation scheme either as a vertical grid of geopotential heights or a vertical grid of geometric heights. Since the pressure field is also derived, pressure coordinates are possible but not implemented at the moment. The values of $T_{i,j,z}$ and specific humidity $q_{i,j,z}$, are given for the Gaussian grid of the latitudes φ_j and the homogenous grid of the longitudes λ_i , and an irregular spaced height grid $z_{i,j,z}$. For the T42L60 grid the index ranges are $i = 1 \dots 64$, $j = 1 \dots 128$, and $z = 1 \dots 60$ for full level quantities and $z = 0 \dots 60$ for half level quantities. Here and in other parts of this report z simply denotes the height coordinate whether if it is geometrical height or geopotential height, but further on in this section z denotes geometrical height and ϕ geopotential height. The vertical index of all quantities is always denoted as z . Coordinates are geocentric except denoted otherwise. The pressure corresponding to the z th half and full levels are calculated by the means of formulas,

(e.g., Roeckner *et al.*, 2003)

$$P_{i,j,z+\frac{1}{2}} = A_{z+\frac{1}{2}} + B_{z+\frac{1}{2}} \cdot P_{\text{Surf}_{i,j}} , \quad (18)$$

$$P_{i,j,z} = \frac{1}{2} \left(P_{i,j,z+\frac{1}{2}} + P_{i,j,z-\frac{1}{2}} \right) , \quad (19)$$

where $P_{\text{Surf}_{i,j}}$ denotes the surface pressure at the i th longitude and the j th latitude. The $A_{z+\frac{1}{2}}$ and $B_{z+\frac{1}{2}}$ are the vertical coordinate parameters provided by ECMWF. The calculation of the geopotential heights is based on the hydrostatic equation and on an interpolation between the half and the full levels (Gorbunov and Kornbluh, 2003)

$$\phi_{i,j,z+\frac{1}{2}} - \phi_{i,j,z-\frac{1}{2}} = R_{\text{Dry}} T_{v,i,j,z} \cdot \text{Ln} \left(\frac{P_{i,j,z+\frac{1}{2}}}{P_{i,j,z-\frac{1}{2}}} \right) , \quad (20)$$

$$\phi_{i,j,z_{\text{max}}+\frac{1}{2}} = \phi_{i,j,z_{\text{Surf}}} , \quad (21)$$

$$\phi_{i,j,z} = \phi_{i,j,z+\frac{1}{2}} + \alpha_{i,j,z} \cdot R_{\text{Dry}} T_{v,i,j,z} , \quad (22)$$

$$\alpha_{i,j,z} = \text{Ln}(2) \text{ for } z = 1 , \quad (23)$$

$$\alpha_{i,j,z} = 1 - \frac{P_{i,j,z-\frac{1}{2}}}{P_{i,j,z+\frac{1}{2}} - P_{i,j,z-\frac{1}{2}}} \cdot \text{Ln} \left(\frac{P_{i,j,z+\frac{1}{2}}}{P_{i,j,z-\frac{1}{2}}} \right) \text{ for } z > 1 , \quad (24)$$

where T_v denotes the virtual temperature as defined in Eq. 78 and ϕ_{Surf} is the surface geopotential which is equal the orography. The geometrical heights over reference ellipsoid were calculated from the geopotential heights by the approximate formula from the *US Standard Atmosphere*:

$$R_0 = \frac{2 \cdot 10^{-3} \cdot g_{\text{Surf}_{i,j}}}{3.085462 \cdot 10^{-6} + 2.27 \cdot 10^{-9} \cdot \cos(2\varphi_j) - 2 \cdot 10^{-12} \cdot \cos(4\varphi_j)} , \quad (25)$$

$$z_{i,j,z} = R_0 \cdot \frac{\phi_{i,j,z} \cdot (g_{\text{Mean}} - \phi_{i,j,z})}{g_{\text{Surf}_{i,j}} \cdot R_0} , \quad (26)$$

where R_0 denotes the effective Earth's radius, g_{Mean} the mean gravity acceleration and $g_{\text{Surf}_{i,j}}$ the local gravity acceleration on the surface. $g_{\text{Surf}_{i,j}}$ itself is calculated by the *International gravity formula plus standard z dependence*:

$$g_{\text{Equator}} = 9.7803 \text{ m/s}^2 \quad (27)$$

$$g_{\text{Surf}_{i,j}} = g_{\text{Equator}} \cdot \left(1 + 0.00531 \cdot \sin(\varphi_{\text{Geodetic}_j})^2 \right) . \quad (28)$$

Finally the standard height dependency of $g_{\text{Surf}_{i,j}}$ is calculated

$$R_{\text{Mean}} = 6371.0 \text{ km} , \quad (29)$$

$$f_{gz} = \left(\frac{R_{\text{Mean}}}{R_{\text{Mean}} + z} \right)^2 , \quad (30)$$

$$g_{\text{Surf}_{i,j,z}} = g_{\text{Surf}_{i,j}} \cdot f_{gz} , \quad (31)$$

g_{Equator} is the gravity acceleration at the equator and R_{Mean} the mean Earth's radius.

1.6 Observation Error Covariance

The observation covariance takes only vertical correlations into account. Due to the separation in space and time between the different RO events this simplification is justified.

1.6.1 Formulation of the Observation Error Covariance Matrix

A simple error covariance matrix formulation was deduced from the empirical estimated matrices (Steiner, 2004). A least square method was used to fit analytical functions to the relative standard deviation which shows a different behavior below and above the tropopause height. The empirical relative standard deviation can be approximated with an exponential increase above the upper troposphere/lower stratosphere region between about 14 km and 20 km, where it is closely constant, and with a decrease from near 14 km downwards proportional an inverse law. To be able to scale the error magnitude, which is receiver dependent, the standard deviation in the upper troposphere/lower stratosphere domain (s_{utls}) can be tuned. Eq. 32 gives the analytical functions for the relative standard deviation s_z over all altitude domains, where z denotes the height, z_{Tropotop} the top level of the “troposphere domain”, $z_{\text{Stratobot}}$ the bottom level of the “stratosphere domain” and H_{Strato} which is the scale height of the error increase over the stratosphere (Steiner and Kirchengast, 2004).

$$s_z = \begin{cases} s_{\text{utls}} + s_0 \cdot \left[\frac{1}{z^p} - \frac{1}{z_{\text{Tropotop}}^p} \right], & \text{for } 2 \text{ km} < z \leq z_{\text{Tropotop}} \\ s_{\text{utls}}, & \text{for } z_{\text{Tropotop}} < z \leq z_{\text{Stratobot}} \\ s_{\text{utls}} \cdot \exp \left[\frac{z - z_{\text{Stratobot}}}{H_{\text{Strato}}} \right], & \text{for } z_{\text{Stratobot}} < z \leq 50 \text{ km} \end{cases} \quad (32)$$

To be able to derive the error covariance matrix the correlation length $L(z)$ has to be determined. The best values for $L(z)$ are 2 km within the troposphere (up to approximately 15 km) and a linear decrease of $L(z)$ above the troposphere to 1 km at 50 km altitude. The observation error covariance matrix \mathbf{S} can now be described as:

$$\mathbf{S} = S_{ij} = s_i \cdot s_j \cdot \exp \left(-\frac{z_i - z_j}{L(z)} \right) \quad (33)$$

This formulation of the observation error covariance also accounts for the error of representativeness, so there is no additional specification within the assimilation framework necessary. The values which are currently used for CHAMP data are:

s_{utls} :	0.5%
s_0 :	4.5%
$z_{\text{Stratobot}}$:	20 km
z_{Tropotop} :	14 km
H_{Strato} :	15 km
p :	1.0
$L(z)$, $15 \text{ km} \leq z \leq 50 \text{ km}$:	Linear decrease to 1 km at 50 km
$L(z)$, $2 \text{ km} \leq z \leq 15 \text{ km}$:	2 km

1.7 Incest Problem

The assumption that there is no correlation between background and observation errors is usually justified, because the causes of the errors are supposed to be completely independent. However, one must be careful about observation preprocessing, like retrieval procedures, which use background information. These procedures can cause a bias of the observations towards the background. If observations, containing background information are used in an assimilation procedure, we are confronted with the so-called *incest problem*. The analysis is drawn closer to the background, which is caused by observations already containing background information, reducing the apparent background departures. If additional information is necessary in

the observation preprocessing procedure, one should carefully decide, which background will not influence the analysis result. The analysis is only optimal if the assumption of bias-free errors holds. In practice background and observations are often significantly biased. If the biases are known, they can be subtracted from the background and observations, which is in practice a delicate problem. Bias monitoring and removal are subjects of ongoing improvement and research.

1.8 Background Error Covariance

The used background fields of temperature, specific humidity, and surface pressure, are provided by ECMWF. So the basis of our considerations concerning the errors and correlations are based on ECMWF recommendations¹. Since this error structures describe analysis fields several assumptions concerning the used 24 and 30 hour forecast fields have to be made. Assuming a linear error growth in time the original standard deviations are multiplied by 1.5, the horizontal correlations have been multiplied by 2 for the 24 hour forecasts and by 1.625, respectively 2.25 for the 30 hour forecasts, with a limit which is effective for the uppermost levels, concerning the horizontal correlation length. This approach is highly empirical and should be studied further. The vertical correlations are not altered due to the fact that the resulting matrices are very sensitive and negative eigenvalues may occur, which suggests also further investigations.

1.8.1 Temperature, Specific Humidity, and Surface Pressure

Standard deviations for temperature, specific humidity, and surface pressure are provided by ECMWF as global means. The same applies to the vertical and horizontal correlations. The relative values of the standard deviations are derived with global mean temperature and specific humidity profiles.

Standard Deviation Temperature

1.8.2 Vertical and Horizontal ECMWF Temperature Error Correlations

Standard Deviation Specific Humidity

Vertical and Horizontal ECMWF Error Correlation Specific Humidity

Standard Deviation Surface Pressure The global mean standard deviation of surface pressure provided by ECMWF is 250 Pa.

Horizontal ECMWF Error Correlation Surface Pressure The correlation of the surface pressure error is shown in Fig. 11

1.9 Control Space Transformations

The principles of 3D-Var are briefly described in Section 1.3. For a model state \mathbf{x} with n degrees of freedom minimization of the cost function requires $O(n^2)$ calculations (Bouttier and Courtier, 1999), thus becomes prohibitively expensive for usual n 's. One practical solution to this problem is to perform the minimization in a control variable space \mathbf{v} given by

$$\mathbf{x} = \mathbf{U}\mathbf{v} .$$

¹Data provided by Mike Fisher, ECMWF, Reading U.K., 2004.

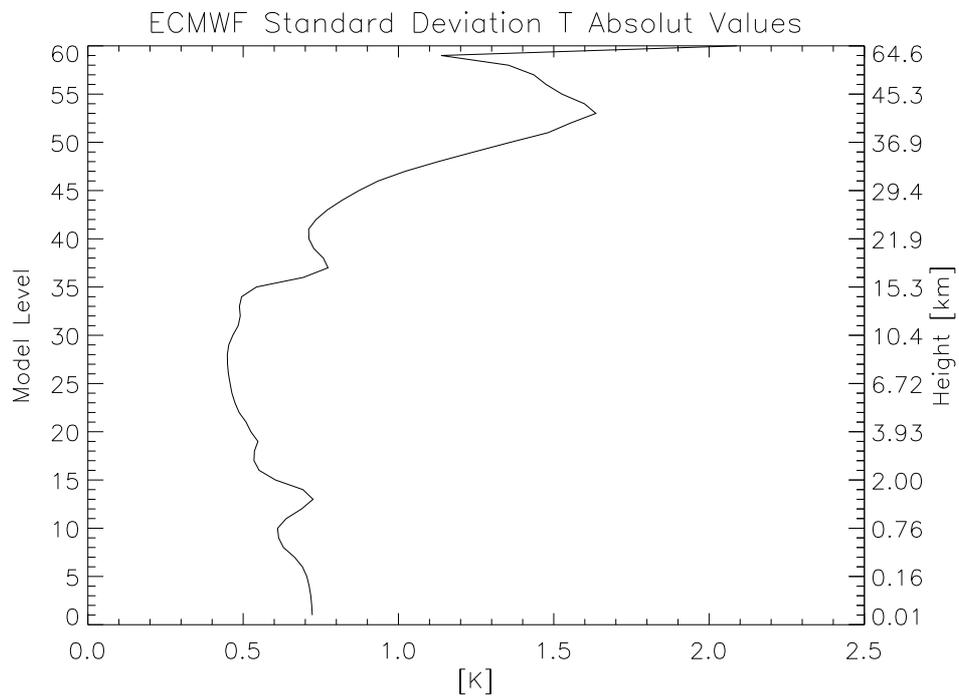


Figure 3: ECMWF global mean temperature standard deviation absolut values.

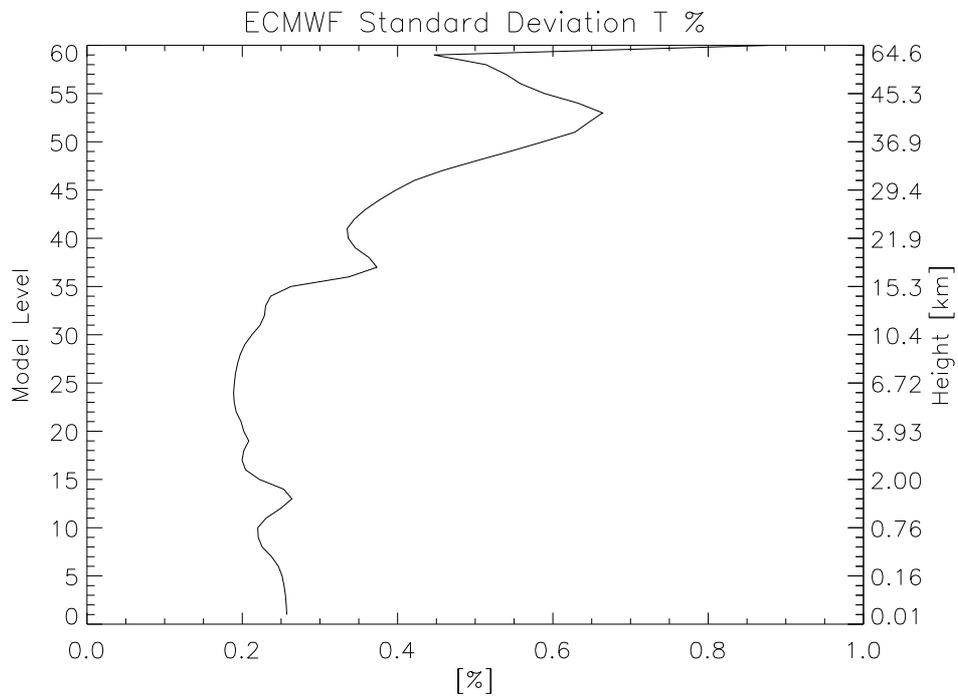


Figure 4: ECMWF global mean temperature standard deviation relative values.

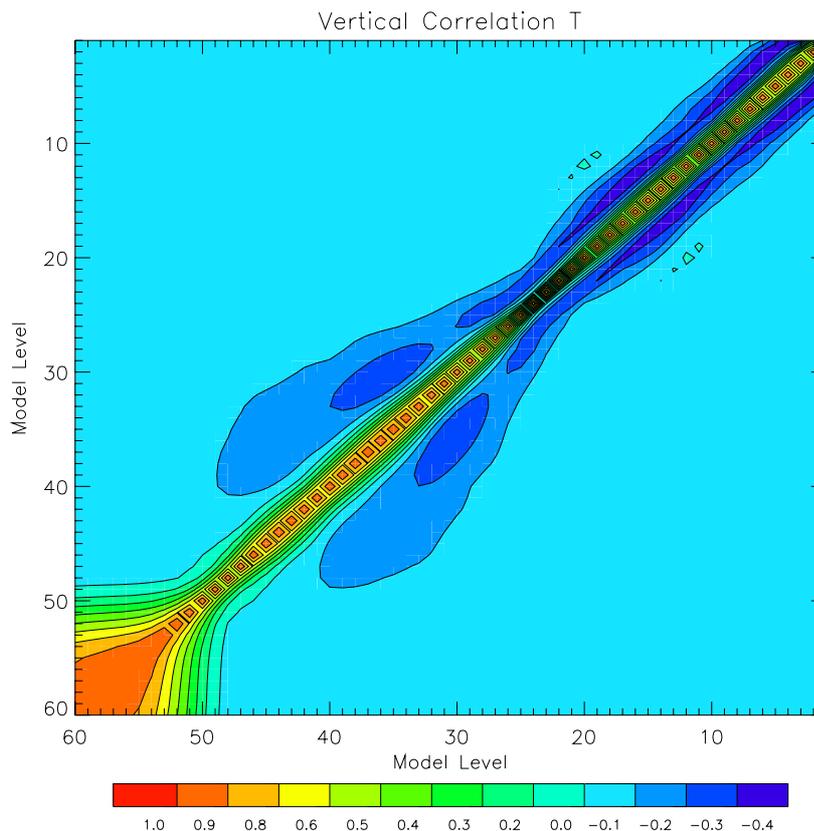


Figure 5: Global mean vertical error correlations of ECMWF temperature fields, L60 resolution.

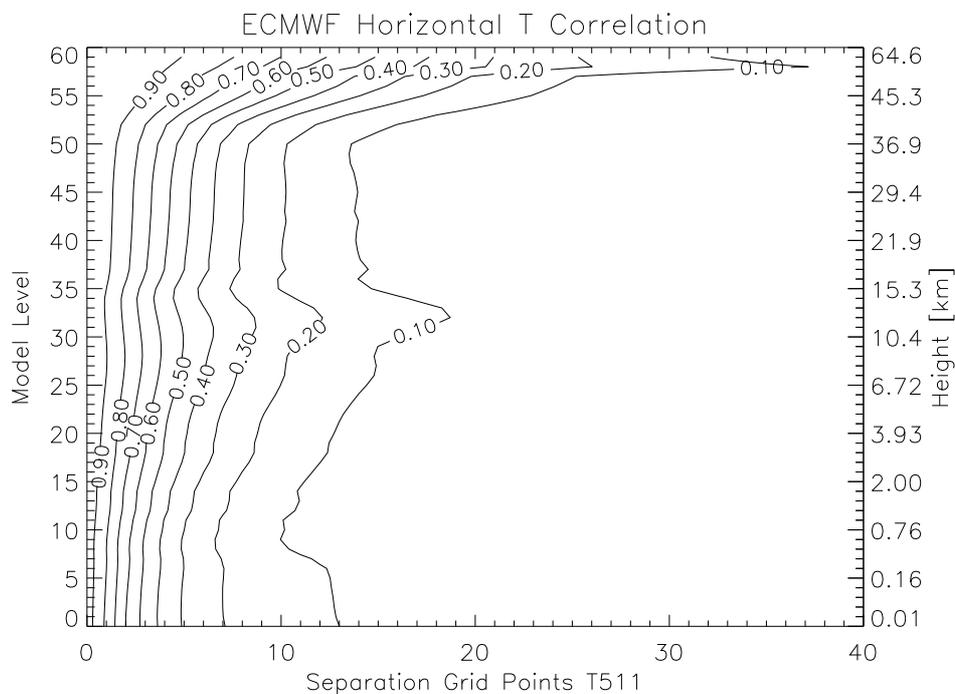


Figure 6: Global mean horizontal error correlations of ECMWF temperature fields.

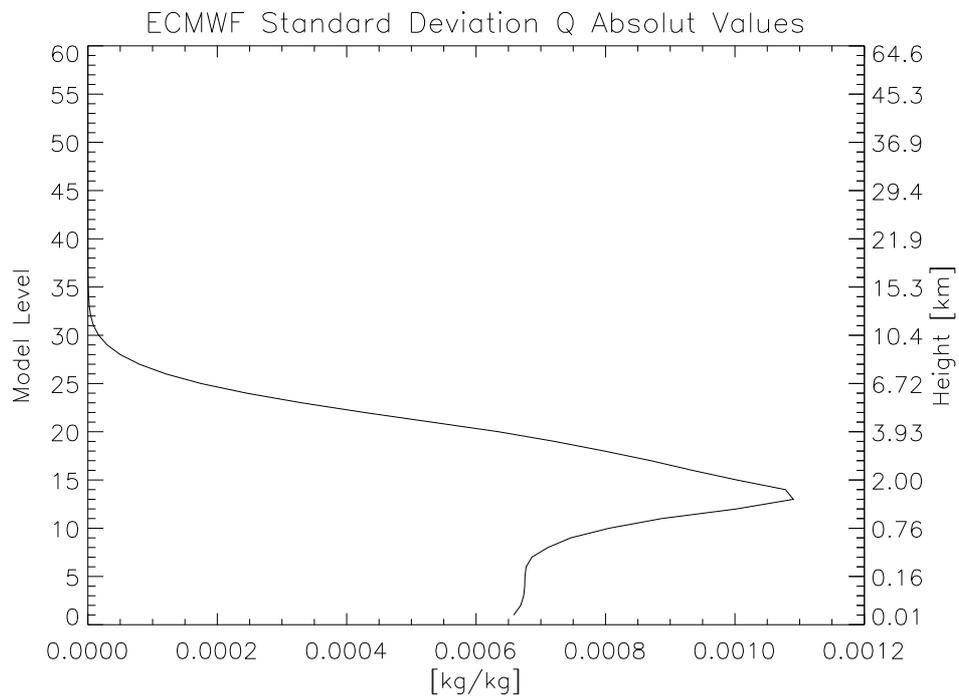


Figure 7: ECMWF global mean specific humidity standard deviation absolut values.

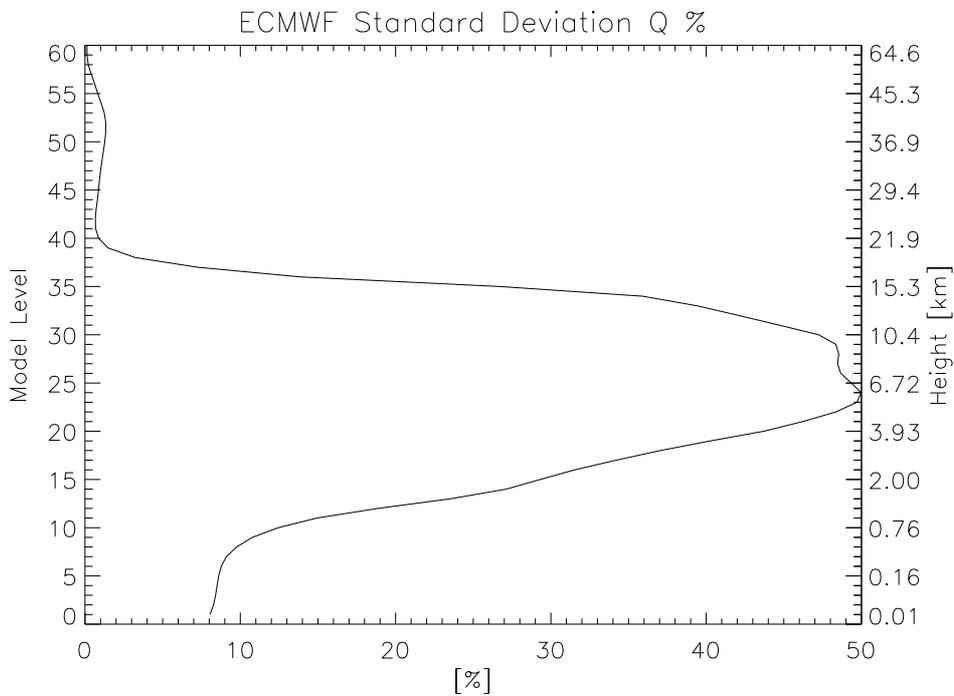


Figure 8: ECMWF global mean specific humidity standard deviation relative values.

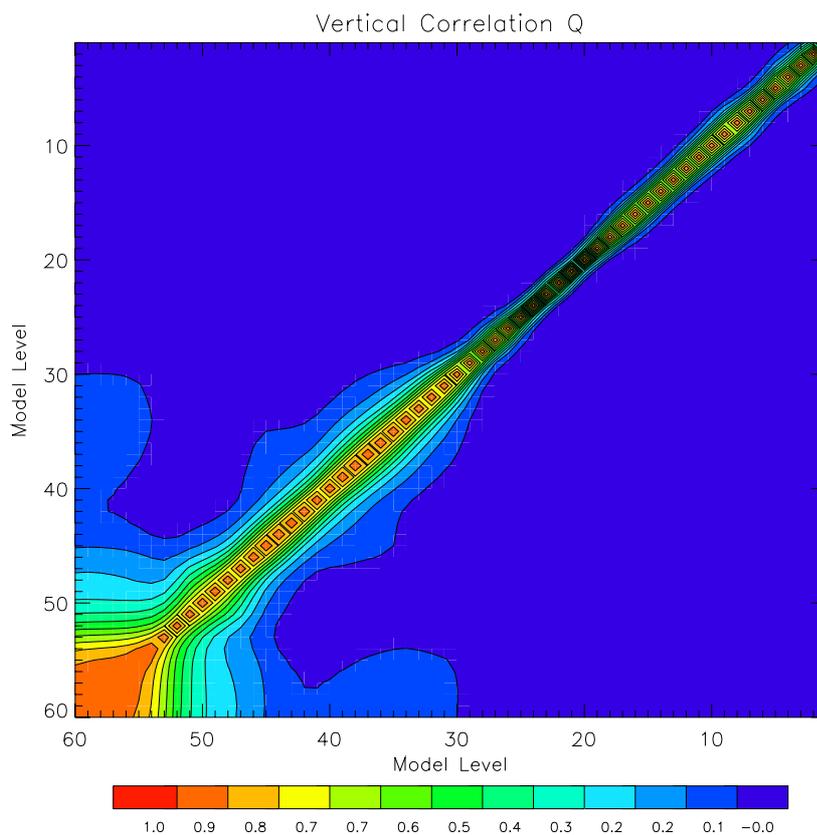


Figure 9: ECMWF vertical error correlation of specific humidity, L60 resolution.

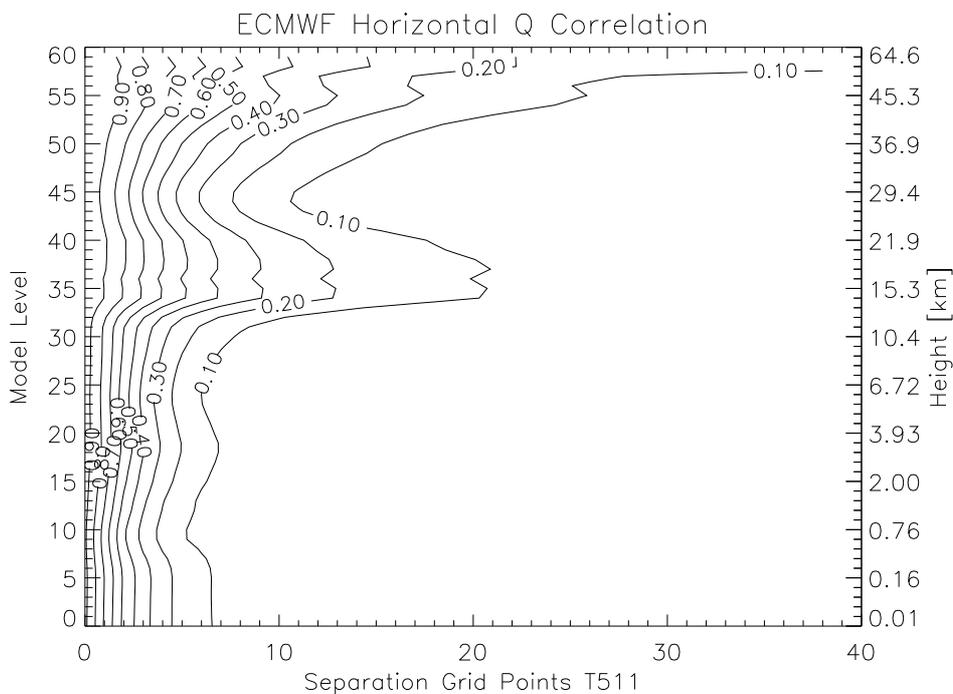


Figure 10: ECMWF horizontal error correlation of specific humidity.

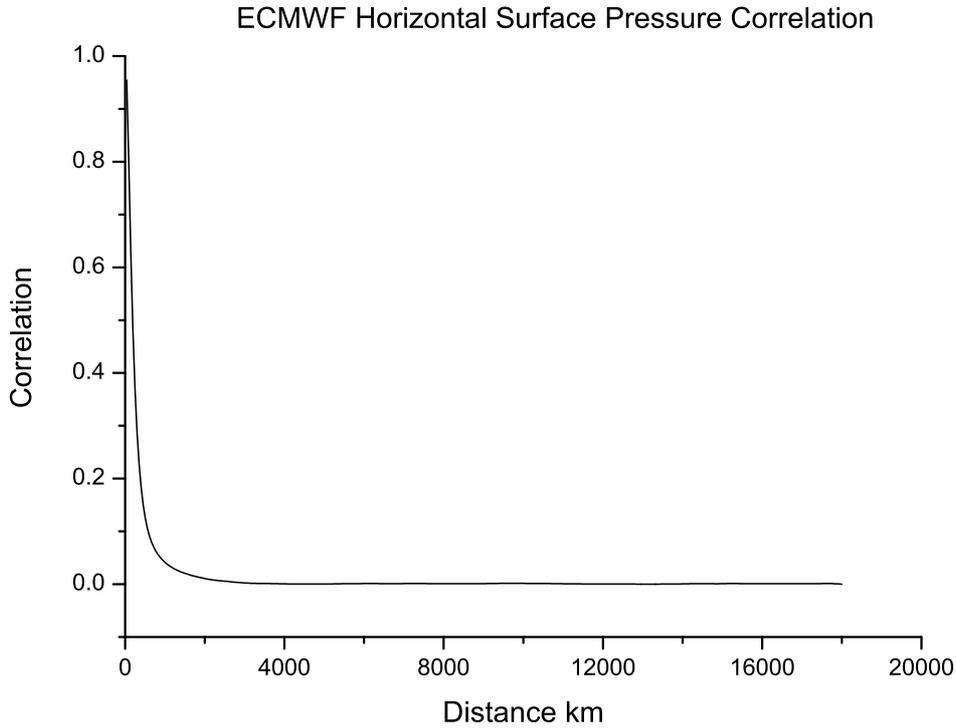


Figure 11: Global mean ECMWF horizontal error correlation of surface pressure.

The transform U has to be chosen in a way that

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T, \quad (34)$$

is approximately satisfied. In the control space \mathbf{v} the number of required minimization calculations is reduced from $O(n^2)$ to $O(n)$. Furthermore by using the transform Eq. 34, the background error covariance matrix becomes $\mathbf{B}_v = \mathbf{I}$, hence effectively preconditioning the problem. \mathbf{I} denotes the identity matrix, \mathbf{B}_v the vertical background covariance matrix. In terms of increments the control variable transform can be written as

$$\mathbf{x}' = \mathbf{U}\mathbf{v}.$$

The transformation

$$\mathbf{v} = \mathbf{U}^{-1}\mathbf{x}'$$

can be specified in different ways. The definition must provide a way to break down the atmospheric state \mathbf{x} into uncorrelated but physically realistic error modes which can be *penalized* in J_b according to their estimated error magnitude.

1.9.1 Vertical Control Variable Transform

The vertical transform serves to project control variables from model levels onto the weighted eigenvectors of the vertical component of the background error covariance matrix

$$\mathbf{B}_v = \overline{\epsilon_v \epsilon_v^T}. \quad (35)$$

Eq. 35 considers the vertical transformation of v at a single horizontal location. For practical reasons approximations must be made like the use of climatological eigenvectors and eigenvalues and the averaging over a geographical domain of these structures. At the moment global means are used, derived from error characteristics provided by ECMWF. (The global mean standard deviations are converted to relative values by the use of averaged absolute values from the used background fields). Since in the calculations relative values are used every single vertical covariance matrix is individual. Furthermore completely spatially dependent error formulations could be used if available, which is feasible using the T42L60 resolution.

General Formulation of U_v The vertical covariance matrix \mathbf{B}_v is given as $K \times K$ positive-definite symmetric matrix where K is equal the number of vertical levels. These are properties which allow to perform an eigendecomposition

$$\begin{aligned}\mathbf{B}_v &= \mathbf{P}^{-1} \mathbf{E} \mathbf{\Lambda} \mathbf{E}^T \mathbf{P}^{-1} , \\ &= \mathbf{P} \hat{\mathbf{B}}_v \mathbf{P}^{-1} .\end{aligned}\tag{36}$$

The inner product \mathbf{P} defines a weighted error $\hat{\epsilon}_v = \mathbf{P} \epsilon_v$, which may be used to allow for variable model level thickness or introduce synoptic dependencies. In the current version this option is not used. The columns of the matrix \mathbf{E} are the K eigenvectors $\mathbf{e}(m)$ of $\hat{\mathbf{B}}_v$ which obey the orthogonality relationship

$$\mathbf{E} \mathbf{E}^T = \mathbf{I} .$$

The diagonal matrix $\mathbf{\Lambda}$ contains the K eigenvalues $\lambda(m)$. With this standard theory it is possible to define a transform U_v between variables $v(k)$ on model levels k and their projection $v(m)$ onto vertical modes m defined by

$$\mathbf{B}_v = \mathbf{U}_v \mathbf{U}_v^T .\tag{37}$$

If one compares Eq. 36 and Eq. 37 it is possible to derive

$$\begin{aligned}\mathbf{v} &= \mathbf{U}_v \mathbf{v}_v \\ &= \mathbf{P}^{-1} \mathbf{E} \mathbf{\Lambda}^{\frac{1}{2}} \mathbf{v}_v .\end{aligned}\tag{38}$$

If Eq. 38 is inserted into the control variable space form of the background error cost function

$$J_b = \frac{1}{2} \mathbf{v}_v^T \mathbf{B}_v^{-1} \mathbf{v}_v ,$$

which gives

$$\begin{aligned}J_b &= \frac{1}{2} \mathbf{v}_v^T \mathbf{v}_v \\ &= \frac{1}{2} \sum_m \mathbf{v}_v(m)^2 ,\end{aligned}\tag{39}$$

and for the gradient

$$\nabla_v J_b = \mathbf{v}_v .\tag{40}$$

As can be seen easily there are several effects of the U_v transform.

- The projection onto uncorrelated eigenvectors of \mathbf{B}_v leads to very significant CPU savings as can be seen via Eq. 39 in the calculation of the background cost function and in its adjoint (gradient) calculations.
- The scaling by the square root of the eigenvalues $\lambda^{\frac{1}{2}}(m)$ serves as a preconditioner.

- The eigenvectors are ordered by the size of their respective eigenvalues what means $\lambda(1)$ is the dominant structure and $\lambda(k)$ essentially contains low amplitude noise. This ordering can be used to filter vertical grid scale noise which reduces CPU still further by neglecting small-scale eigenvalue structures, which contribute little to the total error.

Approximated Eigenstructures Assuming a single column model, with knowledge of the background covariance matrix and hence the eigenvectors and eigenvalues the U_v transform Eq. 38 is an efficient means of reducing CPU without any loss of information. In reality the background covariance matrix is not exactly known, so approximations have to be made. Furthermore our application is 3D-Var where averaging is necessary compared to the 1D-Var case. At the moment global mean error structures are used, resulting from the global mean error structures provided by ECMWF. Absolute temperature standard deviations are converted to relative values by the use of a global mean temperature profile calculated from the background. The same applies for specific humidity and surface pressure. The fractional errors applied in the calculation of the covariance matrix are the global means provided by ECMWF.

1.9.2 Horizontal Control Variable Transform

A recursive filter (RF) is used to represent the horizontal component of the background error covariance matrix. The implementation is based on the description of RFs in Lorenc (1992).

1.9.3 Recursive Filters

RF Basic Algorithm The basic algorithm for a recursive filter is quite simple. The RF is presented with an initial function A_j at gridpoints j where $1 \leq j \leq J$. A single pass of the RF consists of an initial smoothing from *left to right*

$$B_j = \alpha B_{j-1} + (1 - \alpha)A_j \text{ for } j = 1 \dots J , \quad (41)$$

followed by another pass from *right to left*

$$C_j = \alpha C_{j+1} + (1 - \alpha)B_j \text{ for } j = J \dots 1 . \quad (42)$$

The application of the RF in each direction is performed to ensure zero phase change. So a 1-pass filter is defined as a single application of Eq. 41 and Eq. 42. A N -pass RF is defined by N sequential applications.

RF Boundary Conditions Eq. 41 and Eq. 42 are used to compute recursively the RF response at all points $j = 2 : J - 1$ interior to the boundary. Explicit boundary conditions are required to specify the response at points $j = 1$ and J . If there is a limited area and thus a *real* boundary a method of Hayden & Purser (Hayden and Lorenc, 1995) can be used to specify boundary conditions which assume a given decay-tail outside the domain. This technique assures that the response to observations near the boundary is equivalent to the response within the center of the domain. The boundary conditions for B_1 and C_{J+1} depend on the particular number of passes p of the filter in *opposite directions*. Assuming no previous pass of the left moving filter ($p = 0$) we have

$$B_1 = (1 - \alpha)A_1 . \quad (43)$$

Following one pass of the filter in the opposite direction the $p = 1$ boundary condition is

$$(C_J, B_1) = \frac{1 - \alpha}{(1 - \alpha^2)^2} [(B_J, A_1) - \alpha^3(B_{J-1}, A_2)] . \quad (44)$$

Hayden & Purser (Hayden and Lorenc, 1995) suggest to use the $p = 2$ boundary condition also for $p > 2$. In our application there is no *real* boundary but the boundary conditions for B_1 and C_{J+1} still have to be defined

$$B_1 = \alpha A_J + (1 - \alpha)A_1, \quad (45)$$

and

$$C_J = \alpha B_1 + (1 - \alpha)B_J. \quad (46)$$

Matching of RF Output and Analytical Functions The smoothing operations performed by the RF algorithm are related to certain analytical functions. In particular, for $N = 2$, the RF output approximates a second order autoregressive (SOAR) function

$$\mu_s(r) = \left(1 + \frac{r}{s}\right) e^{-\frac{r}{s}}. \quad (47)$$

In the limit $N \rightarrow \infty$ it can be shown that the RF output tends to a Gaussian function

$$\mu_g(r) = \exp\left[-\frac{1}{2}\left(\frac{r}{2s}\right)^2\right] \quad (48)$$

where r is distance and s is a characteristic length scale. In this implementation a slightly modified version of Eq. 48 is used:

$$\mu_g(r) = \exp\left[-0.125 \cdot \left(\frac{r}{s}\right)^2\right]. \quad (49)$$

The α is calculated as

$$\frac{\alpha}{(1 - \alpha)^2} = \frac{1}{2E}, \quad (50)$$

where

$$E = \frac{N(\Delta x)^2}{4s^2}. \quad (51)$$

The definition of E is in this particular case the same for the SOAR and the Gaussian function. This arises from the particular scaling of the Gaussian function given by Eq. 48. Δx denotes the grid spacing, N and s are also known parameters, thus E can be calculated from Eq. 51. α can be calculated as follows

$$\alpha = 1 + E - \sqrt{E(E + 2)}. \quad (52)$$

This approach is matching the large-scale response of the RF that of a SOAR for $N = 2$ and approaches that of a Gaussian for increasing N . The matching of the large scale response to analytical SOAR and Gaussian functions serves the definition of α via Eq. 52. It is also required that the RF conserves the background error variance, for the zero distance case. The calculation of this scaling factor S is realized as the inverse of the zero distance response of a 1D N -pass RF to a delta function. A two dimensional N -pass RF is realized by performing N applications of multiple 1D RF's in one direction followed by the multiple applications of 1D RF's in the orthogonal direction. α and E are calculated in the same way as in the 1 dimensional case, however the RF output has to be scaled by S^2 instead of S which is defined as in the one dimensional case.

Handling of Border Conditions Since the filter runs from a grid point 1 to a grid point n and returns it is well suited for areas with defined borders. The border conditions can be specified and are a function of the number of filter pass. In our case we operate on a closed surface so information must be transferred from grid point n to grid point one and vice versa. The solution found doubles the number of filter operations, which can be handled without problems due to the numerical properties of the filter procedure. The approach can be explained with a simple graphical representation. Fig. 12 shows a latitude or longitude band divided into four equal elements.



Figure 12: Segments along one latitude or longitude band in original order.

To be able to transfer information from A to D and vice versa the filter procedure is applied a second time to a shifted arrangement of boxes:

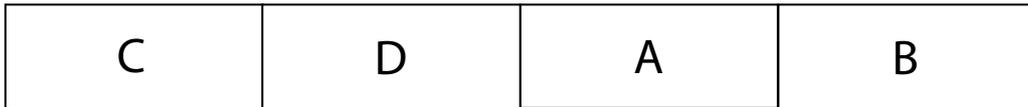


Figure 13: Shifted arrangement of segments along one latitude or longitude band.

After the second filtering process the original order of the boxes is reconstructed using the two middle segments from both filtering runs, respectively B and C from the original order and D and A from the shifted version. As can be seen, this approach ensures a smooth transition between the boxes A and D.

Transform to Non Dimensional Space Due to the fact that Δx is not uniform over the grid domain, the two dimensional field must be transformed to a non dimensional space prior to the filter procedure. This transform is realized as an inner product which is defined as

$$\frac{\text{Increment in Control Space}}{\sqrt{\text{Grid Box Area}}} . \quad (53)$$

The grid box area is calculated by subtraction of fractions of ellipsoid areas between the equator and the pole and subsequent division by the number of longitudes. The ellipsoid (WGS84) areas are calculated by

$$\begin{aligned} \text{Ellipsoid Area} = 2\pi (r_{\text{Equator}} + z)^2 \cdot (0.996647190 \cdot \sin(\text{Lat}) - 0.0011166660 & \quad (54) \\ \sin(3.0 \cdot \text{Lat}) + 1.68880838^{-6} \cdot \sin(5.0 \cdot \text{Lat}) - 2.70005436^{-9} & \\ \sin(7.0 \cdot \text{Lat}) + 4.41731436^{-12} \cdot \sin(9.0 \cdot \text{Lat})) , & \end{aligned}$$

where z denotes the height over the reverence ellipsoid (Lauf, 1983).

RF Representation of Background Error Covariances The control variable transform uses the identity

$$\mathbf{B} = \mathbf{U}\mathbf{U}^T, \quad (55)$$

to define a transform $\mathbf{x}' = \mathbf{U}\mathbf{v}$ which relates preconditioned control variables \mathbf{v} to analysis increments \mathbf{x}' in model space. The horizontal component \mathbf{U}_h defined by

$$\mathbf{B}_h = \mathbf{U}_h\mathbf{U}_h^T, \quad (56)$$

is realized by scaled recursive filters. The RF has to be applied in a non dimensional space

$$\hat{\mathbf{v}} = P_x^{\frac{1}{2}}\mathbf{v},$$

where the scaling factor P_x contains the grid box area as described above. The relation between model and non dimensional space background error covariance matrix $\hat{\mathbf{B}}$ is given as

$$\mathbf{B} = P_x^{-\frac{1}{2}}\hat{\mathbf{B}}P_x^{-\frac{1}{2}}. \quad (57)$$

The comparison between Eq. 55 and Eq. 57 indicates that the horizontal component of the control variable transform \mathbf{U}_h relating model space control variables \mathbf{v} to model space analysis variables \mathbf{x} via $\mathbf{x} = \mathbf{U}_h\mathbf{v}$ can be represented by using a recursive filter \hat{R} in non dimensional space as

$$\mathbf{x}' = \sigma_b P_x^{-\frac{1}{2}}\hat{R}P_x^{-\frac{1}{2}}\mathbf{v}. \quad (58)$$

When the two dimensional recursive filter \hat{R} is applied, only $\frac{N}{2}$ passes are performed, as indicated in Eq. 56 the other $\frac{N}{2}$ passes are performed by the adjoint transform (Barker, 1999).

1.10 Horizontal Background Error Covariances

As recursive filters approximate analytical functions it naturally occurs to be difficult to match them with statistically derived correlation functions. Nevertheless it is possible to archive a quite good agreement with the ECMWF provided horizontal correlations. To take the different grid point distances into account (nearly constant along the latitudes, getting smaller toward the poles along the longitudes), the filter coefficients are calculated separately along the latitude, longitude directions.

1.11 The Minimization Algorithm

The L-BFGS-B algorithm is a limited memory algorithm (L) for solving large nonlinear optimization problems subject to simple bounds on the variables. It is based on the Broyden-Fletcher-Goldfarb-Shanno Method (BFGS), which is from the class of Quasi-Newton methods, the most common. BFGS uses the following basic update for A_i

$$A_{i+1} = A_i + \frac{s_i s_i^T}{s_i^T v_i} + \frac{A_i v_i v_i^T}{v_i^T A_i v_i} + (v_i^T A_i v_i) \cdot u_i u_i^T, \quad (59)$$

with

$$u_i = \frac{s_i}{s_i^T} - \frac{A_i v_i}{v_i^T A_i v_i}, \quad (60)$$

where $s_i = x_{i+1}$ and $v_i = \nabla f_{i+1} - \nabla f_i$. For a symmetric positive definite matrix A_i the matrix A_{i+1} is also symmetric positive definite, and thus the Quasi-Newton condition is fulfilled.

This version was chosen to be able to apply simple bounds within the assimilation framework (denoted by the B). This option is currently not used. It is intended for problems in which information on the Hessian matrix is difficult to obtain or for large dense problems. L-BFGS-B can also be used for unconstrained problems, as currently in our application, and in this case performs similarly to its predecessor algorithm L-BFGS (Harwell routine VA15). The algorithm is implemented in Fortran 77 (Byrd *et al.*, 1994; Dong and Nocedal, 1989).

2 Data Preprocessing

Since the RO data products consist of profiles with a vertical resolution (300 to 400 observations within the interesting altitude domain), which exceeds vertical resolution of the used sigma level grid (60 vertical levels) by far, a data thinning procedure has to be applied prior to the use of the data. Several studies, also indicate that CHAMP observations have the best quality between 5 km and 35 km, so only observations within this altitude domain are used. Within the preprocessing step the quality flags of the observations (WegCenter processed CHAMP data) are utilized to reject suspicious data.

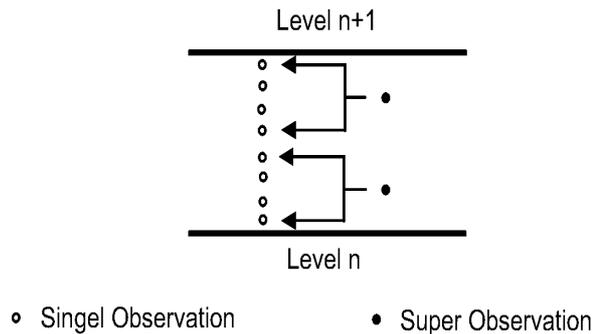


Figure 14: Interpolated measurement distribution using two super observations.

During the preprocessing the number of measurements is reduced by averaging, taking the background grid into account. The linear averaging procedure in *LOG* space, takes the spacing of the background levels into account which is derived from the mean global vertical grid (Loescher, 2004). Fig. 14 depicts the chosen averaging pattern, giving two super observations between two sigma levels. The observation error assumptions (see Section 1.6) can be seen as very conservative taking the fact of the observation averaging procedure within the pre-preprocessing step into account.

3 Quality Control

The assimilation system allows for additional quality control steps and observation screening, (e.g., certain criteria to refuse observation too far from the background values, taking the error characteristic from both, background and observation into account), which are not implemented at the moment. As a “quasi” quality control step the observations are screened for cases below the model orography (cf., model orography 15), which might happen if the profiles are used within their full vertical domain (cf., horizontal interpolation of height grid, Eq. 10a).

4 Background Data

In general, background data provided by ECMWF, is used, due to its high standards, accuracy and availability. The ECMWF, spectral model (IFS) runs currently with a T511L60 resolution, which is foreseen to be upgraded to T511L90 during 2005. After increasing the vertical resolution from 60 to 90 model levels, the next scheduled step is an upgrade of the horizontal resolution to T711 (ECMWF, 2004).

4.1 ECMWF Analysis Fields

As background for the assimilation procedure ECMWF T42L60 analysis fields, corresponding to a Gaussian grid composed of 128×64 geographic areas, and 60 standard model levels up to a height of ~ 64 km are used. The resolution of the vertical grid is highest in the planetary boundary layer and lowest in the stratosphere and lower mesosphere. These vertical levels are realized as hybrid levels, which means they are composed as σ -levels which follow the Earth's surface in the lower and mid troposphere, but are surfaces of constant pressure (pressure levels) in the upper stratosphere and mesosphere with a smooth transition between these types of levels. ECMWF uses two different numerical representations for the horizontal grid. A *spectral* method, based on a spherical harmonic expansion, truncated at total wavenumber 511, for the representation of upper air fields and the computation of the horizontal derivatives. Apart from the T511L60 operational model a T255L40 model is run for ensemble prediction, a T159L40 model for the 4D-Var assimilation and a T63L31 model for seasonal forecast. In addition, there is a grid point representation used for computing dynamic tendencies and the diabatic physical parametrization. This so called Gaussian grid, is regular in longitude and almost regular in latitude. Due to the convergence of the longitudes towards the poles, the east-west distance between the grid points decreases polewards. To avoid some numerical problems around the poles and most importantly, to save computing time, a reduced Gaussian grid was introduced by reducing the number of grid points along the shorter latitude lines near the poles, so as to keep the east-west separation between points on different latitudes almost constant. With the current resolution the grid is identical to a regular Gaussian grid between 24°N and 24°S . The model surface is logically divided into sea and land points, by using a *land-sea-mask*. The representation of the *orography* uses the mean orography and is significantly smoother than reality (Fig. 15).

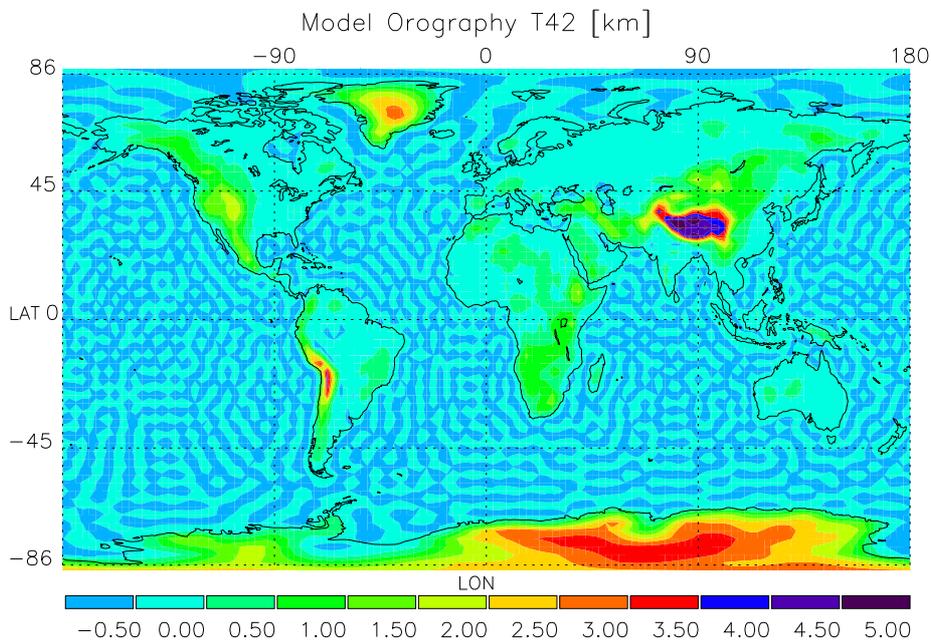


Figure 15: Model orography for the T42 horizontal resolution.

At ECMWF four global analyses per day are produced at 00, 06, 12 and 18 UTC. These are obtained by two 4D-Var minimization cycles running from 03 UTC to 15 UTC and from 15 UTC to 03 UTC. The analysis is performed by comparing the observations directly with a

very short forecast, using exactly the same model as the operational medium-range forecast. The differences between the observed values and the equivalent values predicted by the short-range forecast are used to make a correction to the first guess field in order to produce the atmospheric analysis. The use of these analysis times is partly traditional, since 4D-Var operates quasi continuously. Forecasts are run twice daily at ECMWF, starting from 00 UTC and 12 UTC. For our climatological application the T42L60 resolution of the 24 respectively 30 hour forecast fields was chosen due to computational constraints and some unique advantages. Since there are only two forecast runs started per day, we use for the 00 and 12 time slices 24 hour forecasts and for the 06 and 18 time slices the corresponding 30 hour forecasts (with appropriate adaption of the error structures) as first guess data. As discussed below the spatial characteristics of Radio Occultation data (moderate horizontal, high vertical resolution) fits quite well to this background grid spacing. ECMWF offers a broad spectrum of analysis and forecast products, for our application global temperature, specific humidity, and surface pressure fields are used (Person, 2003; Gobiet and Kirchengast, 2004).

4.2 Accuracy of ECMWF Analysis Fields

For climatological applications the knowledge of error bounds is crucial. For the operational daily analyses, information about the standard deviations and correlations of the atmospheric parameters are available. It is possible to derive them with statistical methods like the NMC method, where the error characteristics are derived by averaged forecast differences. The error characteristics of monthly and seasonal means are still unknown and believed to be bias driven (Mike Fisher ECMWF Reading, U.K., pers. communication 2003). The error characteristics of the used background fields are discussed in detail in Section 1.8.

5 Radio Occultation Data

With the successful launch of the CHAMP satellite in summer 2000 and the start of its GPS Radio Occultation experiment in February 2001, the number of available RO-based atmospheric profiles increased in a way that long term climatological studies become feasible. In addition to CHAMP, also the RO experiments on the Argentinean SAC-C satellite contributes data (SAC-C data is fragmentary and limited to certain periods, at the moment further data from SAC-C is questionable), and the GRACE mission is expected to constitute data in 2004. The first successful processed GRACE RO profiles were published by JPL on 29 of July 2004. Furthermore, a RO receiver (GRAS) (EUMETSAT, 2003) is part of the payload of the MetOp series of polar-orbiting, operational meteorological satellites currently prepared by EUMETSAT and ESA. Other RO missions are scheduled like COSMIC (US-Taiwan) and ACE+ (ESA) whose further implementation was unfortunately stopped recently. The global coverage, all-weather capability, high vertical resolution, accuracy and long term stability of RO data makes them an ideal candidate to build global climatologies of fundamental variables such as temperature, geopotential height, and water vapor (Kirchengast *et al.*, 2004; Gobiet and Kirchengast, 2004).

5.1 The RO Technique

Radio Occultation (RO) is a novel active limb sounding technique to derive atmospheric key parameters. The measurement setup comprises a receiver mounted on a low Earth orbit (LEO) satellite, which tracks the signal of a global navigation system (GPS) satellite positioned in a medium Earth orbit (MEO), in an occultation geometry. The challenges from a technical point of view imply the necessity of an extremely high frequency stability in the signal and the positions and velocities of transmitter and receiver must be known to very high accuracy. The concept was successfully proven on-board the Micro Lab 1 satellite (GPS/MET experiment) (Kursinski *et al.*, 1996) and is now quasi operationally implemented as part of the CHAMP mission. Fig. 16 illustrates the concept of the RO technique, which is the interaction of electromagnetic waves (GPS signals) and the (in our application) terrestrial atmosphere. An electromagnetic ray passing through the atmosphere is bent and retarded due to the ionosphere and the Earth's refractivity field. In our application the signal must be corrected for the influence of the ionosphere which is accomplished by a differential approach. For other applications this part of the signal is used to derive maps of the ionosphere and the total electron content (TEC). The effect of the atmosphere onto the electromagnetic waves can be characterized by a total bending angle (α) as a function of the impact parameter (a). The impact parameter is defined, assuming spherical symmetry, as the perpendicular distance between the center of local curvature at the perigee of the occultation ray and the ray asymptote at the GPS or LEO satellite.

5.1.1 Unique Advantages for Monitoring of Atmospheric Key Parameters

Due to the measurement principle, radio occultation features some preferable characteristics which makes it an ideal technique for a long term monitoring of atmospheric key parameters. Its long term stability and self-calibrating concept makes it an ideal candidate for climate studies.

- The atmospheric profiles are not derived from absolute intensities or phase delays.
- The profiles are derived from transmissions (normalized intensities) and the Doppler shift (phase change) profiles (intrinsic self-calibration).
- Only short-term stability is necessary during the occultation event.

Geometry of GPS limb sounding with CHAMP

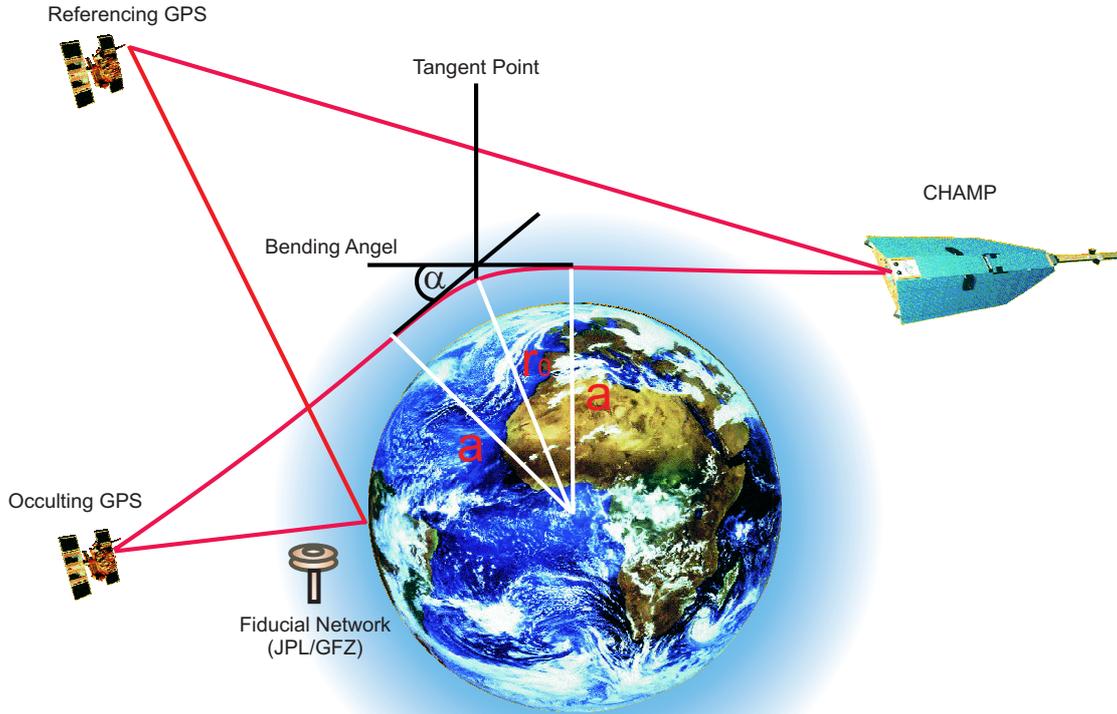


Figure 16: Basic geometry of a radio occultation measurement including LEO satellite (CHAMP), GPS satellites and fiducial ground network.

5.1.2 Characteristic Horizontal and Vertical Resolution of RO Measurements

One of the characteristics of RO measurement is its high vertical (Δz) and moderate to low horizontal (ΔL) spatial resolution. The following relation is valid

$$\Delta L = 2 \cdot \sqrt{2 \cdot R \cdot \Delta z} , \quad (61)$$

where ΔL denotes the chord, which is defined by the tangent of the inner of two concentric circles with radii which differ about Δz . R is the radius of the inner circle, which is in fact the atmospheric radius at the tangent point of the ray path. If geometric optics is applied, the vertical resolution is limited by the diameter of the first Fresnel Zone d_F . For occultation geometry, this can be calculated by negligible atmosphere (stratosphere) as follows

$$d_F = 2 \cdot \sqrt{\lambda \cdot D} ,$$

where λ denotes the wavelength of the GPS signal and D the distance between the GPS receiver on-board CHAMP and the tangent point. With $\lambda = 19$ cm and $D = 2600$ km (orbit height of 500 km) for the diameter of the first Fresnel Zone follows 1.4 km. Using Eq. 61 a horizontal resolution ΔL of ~ 270 km can be calculated. Caused by the exponential increase of the refractivity towards the Earth's surface (troposphere) d_F becomes smaller and reaches close to the Earth's surface a value of 0.5 km (Kursinski *et al.*, 1997) which translates into a horizontal resolution of about ~ 80 km. It is possible to enhance the vertical resolution by using methods which take diffraction effects into account (Gorbunov and Gurvich, 1998; Wickert, 2002).

5.2 Retrieval

5.2.1 General RO Retrieval

As a basic idea the radio signals emitted by the GNSS satellites can be treated as rays, which means a geometric optics assumption. This is a valid simplification from the mid-troposphere upwards. However, below some 5 km wave optics methods, which can cope with complex signal structures in the presence of strong refractivity gradients enhance the retrieval performance significantly (Gorbunov, 2002; Hocke *et al.*, 1999; Sokolovskiy, 2003). At the moment the Weg-Center retrieval for CHAMP data uses only the geometric optics approach. The use of a retrieval procedure which blends geometric optics assumption from the mid-troposphere upwards with wave optics methods derived data in the lower regions is foreseen to be used in the near future.

$$\begin{aligned} \text{GPS Frequencies } L_1 &= 1575.42 \text{ MHz} \\ L_2 &= 1227.60 \text{ MHz} \end{aligned}$$

The primary observables are the phase delays of GNSS signals, resulting from the deceleration of the electromagnetic wave's phase velocity by the atmosphere. The Doppler shifts and total bending angles α as function of the ray's impact parameter a can be deduced from the phase delays (Kursinski *et al.*, 1997). The basis to derive α is the Doppler-Shift equation (Gorbunov *et al.*, 1996):

$$f_d = f_c \left(\frac{c - \vec{v}_2 \vec{m}_2 n_2}{c - \vec{v}_1 \vec{m}_1 n_1} - 1 \right), \quad (62)$$

where \vec{v}_1 and \vec{v}_2 denote the velocity vectors of the occulting GPS and CHAMP satellites, \vec{m}_1 and \vec{m}_2 are the unit vectors of the wave vector of the transmitted and received signals, n_1 and n_2 are the refractivities at the corresponding satellite positions. The Doppler shift f_d corresponds to the measured phase as

$$f_d = -\frac{f_c}{c} \frac{dL}{dt}, \quad (63)$$

with the carrier frequency f_c and the vacuum speed of light c . It is possible to decompose L as

$$L = L_0 + dA_{L_0}, \quad (64)$$

so it is possible to decompose the Doppler shift

$$f_d = f_{d0} + f_{dA}. \quad (65)$$

The first term describes the frequency shift without atmospheric influence, caused by the relative motion of the satellites and can be calculated from precise orbit data. The second term is the time derivative of calibrated atmospheric induced signal delay of the occultation link which is composed of an ionospheric part and a part of the neutral atmosphere. The angle of refraction α can be derived as follows

$$\alpha = \Phi_1 + \Phi_2 + \Theta - \pi. \quad (66)$$

The angles Φ_1 and Φ_2 are the only unknowns in Eq. 66, after solving the scalar product and the introduction of the measured Doppler shift in Eq. 62 (Kursinski *et al.*, 1997), and can be derived under the assumption of local spherical symmetry of the refractivity $n = n(r)$ using Snells law

$$r_1 n(r_1) \sin \Phi_1 = r_2 n(r_2) \sin \Phi_2 = a. \quad (67)$$

Eq. 62 and Eq. 67 are a non linear system which cannot be solved analytically, but with a simple iterative procedure (Gorbunov *et al.*, 1996). Eq. 67 also provides the impact parameter

a. Furthermore an ellipsoid and an ionospheric correction have to be applied. As a next step the refractivity index n can be derived via an inverse Abel transform (Fjeldbo *et al.*, 1971)

$$n(a) = \exp \left[\frac{1}{\pi} \cdot \int_a^{\infty} \frac{\alpha(a')}{\sqrt{a'^2 - a^2}} da' \right] . \quad (68)$$

The refractivity as a function of height $N(a)$ is obtained via Eq. 68:

$$\begin{aligned} N(a) &= 10^6 \cdot (n(a) - 1) \\ z(a) &= \frac{a}{n(a)} - R_c \end{aligned}$$

where R_c denotes the local radius of curvature of the Earth's ellipsoid at the occultation location. Bending angles above ~ 45 km are dominated by ionospheric effects (Hocke, 1997). Since the ionosphere is a dispersive medium and thus causes different L_1 and L_2 phase delays, it is possible to remove this effects to first order by linear combination of this two signals. The method of linear correction of bending angles Eq. 69 has been applied most successfully (Vorobev and Krasnilnikova, 1994) by

$$\alpha_{LC}(a) = \frac{f_1^2 \alpha_1(a) - f_2^2 \alpha_2(a)}{f_1^2 - f_2^2} , \quad (69)$$

where α_{LC} denotes the ionosphere corrected bending angle, α_1 and α_2 the uncorrected bending angles of the L_1 and L_2 signals. Still, retrieval results above 20 km to 30 km are sensitive to residual ionospheric noise (resulting from higher order terms, which are not corrected by Eq. 69) and other errors like receiver noise, residual clock errors, local multipath and orbit uncertainties. Since the upper integration limit of the inverse Able transform Eq. 68 ranges to infinity it needs in practice some kind of high altitude initialization to avoid downward propagation of errors via the Abel transform itself and subsequently via the hydrostatic integration Eq. 84. To minimize these errors the concept of statistical optimization is applied (Sokolovskiy and Hunt, 1996). The best linear unbiased estimator (BLUE Eq. 70;(Loescher, 2004))

$$\mathbf{x}_a - \mathbf{x}_b + \mathbf{K} (\mathbf{y} - H(\mathbf{x}_b)) \quad (70)$$

α_{opt} is derived from an observed α_O and a background α_B bending angle profile under the assumption of unbiased Gaussian errors. The \mathbf{O} and \mathbf{B} are the observation and background error covariance matrices, respectively. The α_{opt} is derived by

$$\alpha_{opt} = \alpha_B + \mathbf{B} \cdot (\mathbf{B} + \mathbf{O})^{-1} \cdot (\alpha_O - \alpha_B) , \quad (71)$$

where α_{opt} is a fused bending angle profile dominated by the background in the upper part and by the observation in the lower part. The WegCenter retrieval schemes integrate background information only at one point of the retrieval (at bending angle level), so that the results have well defined error characteristics. One has to be careful if background information is used in a retrieval procedure, if the retrieved data is used in a consecutive assimilation framework. If the assimilation framework uses the same background data as the retrieval we end up with a so-called *incest problem*². The analysis in the assimilation procedure is artificially drawn to the background. At WegCenter, statistical optimization is implemented in two ways, both relying on Eq. 71, but using different sources of background information and different ways of preprocessing of this information. WegCenter/MSIS uses bending angle profiles extracted from the MSIS-90 climatology (Hedin, 1991) and applies best fit profile library search and bias

²cf. Section 1.7.

correction procedures (Gobiet *et al.*, 2004a) in order to diminish known biases in the climatology (Randel *et al.*, 2002). WegCenter/ECMWF uses bending angle profiles derived from ECMWF operational analyses. For assimilation purposes only data derived with the WegCenter/MSIS framework was used to avoid the incest problem.

	WegCenter/MSIS	WegCenter/ECMWF
Outlier Rejection and Smoothing	3σ outlier rejection on phase delays and smoothing using regularisation	Like WegCenter/MSIS
Ionospheric Correction	Linear combination of bending angles. Correction is applied to low-pass filtered bending angles (1 km sliding average), L_1 high-pass contribution is added after correction. L_2 bending angles < 15 km derived via $L_1 - L_2$ extrapolation.	Like WegCenter/MSIS
Bending Angle Initialisation	Statistical optimisation of bending angles 30 – 120 km. Vertical correlated background (corr. length $L = 6$ km) and observation ($L = 1$ km) errors. Obs. error estimated from obs. profile > 65 km. Background error: 15%. Background information: MSIS-90 best fit-profile, bias corrected.	Like WegCenter/MSIS, but co-located bending angle profile derived from ECMWF operational analysis (above ~ 60 km: MSISE-90) as background information. No further processing.
Hydrostat. Integral Init.	At 120 km: pressure = $p(\text{MSISE-90})$.	Like WegCenter/MSIS
Quality Control	Refractivity 5 – 35 km $\frac{\Delta N}{N} < 10\%$; Temperature 8 – 25 km: $\Delta T < 20$ K; Reference: ECMWF analysis.	Like WegCenter/MSIS

Table 2: Overview of WegCenter CHAMP RO retrieval schemes (EGOPS/CCR Version 2).

5.3 Data Products

From the phase delay measurements a variety of atmospheric parameters can be derived. In theory, some parameters (e.g., dry temperature) could be derived without any background information, but in practice, as mentioned above, the retrieval procedure has to be initialized. For a detailed description of the refractivity formulas (Smith-Weintraub and Thayer formula) cf., Section 1.5.2.

5.3.1 Refractivity Profiles

Refractivity profiles are derived as described above from the statistical optimized bending angle α . This is the retrieval product which is used within the assimilation framework. To avoid the

so called *incest problem* (and data quality reasons), which is mentioned in more detail in Section 1.7, only data processed with the WegCenter/ECMWF retrieval scheme below 35 km is used. At that altitude the influence of the initialisation on the retrieved refractivity should be small enough, to pose no major problem.

5.3.2 Temperature Profiles

The dry temperature T_{Dry} can then be derived from Eq. 13 or Eq. 14 by neglecting the effect of water vapor (ignoring the wet terms, k_1 empirical constant cf., Section 1.5.2)

$$N = k_1 \cdot \frac{p_A}{T_{\text{Dry}}}, \quad (72)$$

which is valid in the mid- to upper troposphere and stratosphere. If the temperature is below 250 K, the temperature error caused by a 50% error of the water vapor climatology is less than 1 K (Kursinski *et al.*, 1996). So the assumption of a *dry atmosphere* can be expanded down to the ground at high latitudes beginning from the subpolar regions on. If this assumption does not hold *a priori* information about the humidity below ~ 6 km is necessary to solve the ambiguity. Assuming a *dry atmosphere*, using Eq. 72 and introducing the ideal gas law

$$p_A = \frac{\rho_A T_{\text{Dry}} R_{\text{Dry}}}{m_A}, \quad (73)$$

where ρ_A denotes the dry air density, T_{Dry} the *dry air temperature*, R_{Dry} the universal gas constant for dry air, p_A the *dry air pressure* and m_A the mean molar mass of dry air, it follows with the use of Eq. 72

$$\rho_A = \frac{m_A}{k_1 R_{\text{Dry}}} \cdot N, \quad (74)$$

what means that the density of air is directly proportional to the refractivity and thus can be derived directly. If the vertical air density $\rho(z)$ is known the vertical pressure can be derived using the equation of hydrostatic equilibrium

$$dp_A(z) = -g(z)\rho_A(z)dz, \quad (75)$$

and integration over z

$$p_A(z) = \int_z^\infty g(z')\rho_A(z')dz'.$$

A second application of Eq. 73 allows to derive the vertical profile of the dry temperature T_{Dry}

$$T_{\text{Dry}} = k_1 \frac{p_A(z)}{N(z)}. \quad (76)$$

5.3.3 Humidity Profiles

To derive humidity profiles *a priori* information about the humidity is necessary to resolve the ambiguity. An iterative procedure to calculate specific humidity profiles works as follows:

1. Assumption of dry atmosphere

$$q(z) = 0. \quad (77)$$

2. Calculation of the virtual temperature profile

$$T_v(z) = T(z) \cdot (1 + 0.608 \cdot q(z)). \quad (78)$$

3. Calculation of the pressure profile as in Eq. 84

$$p(z) = \frac{T(z)^2}{c_2} \cdot \left(n(z) - 1 - c_1 \cdot \frac{p(z)}{T(z)} \right), \quad (79)$$

4. Calculation of the specific humidity profile

$$q(z) = \frac{0.622 \cdot p_w(z)}{(p(z) - 0.378 \cdot p(z))}. \quad (80)$$

With the calculated $q(z)$ the iteration starts again at step 2, the procedure converges fast. An other approach would be a 1D-Var procedure to determine the most likely state of the atmosphere taking background information into account (Gorbunov and Sokolovskiy, 1993).

5.3.4 Geopotential Height

The geopotential height profile can be computed corresponding to a given geometric height profile. The (geodetic) latitude dependence of gravitation weighted by $\frac{g_{\text{Equ}}}{g_0}$ is needed and calculated as factor

$$g_{\text{fact}} = \frac{g_{\text{Equ}}}{g_0} + \frac{0.00531}{g_0} \cdot \sin(\varphi_j)^2. \quad (81)$$

To calculate the geopotential height the relation

$$dZ = \frac{g}{g_0} dz \quad (82)$$

is used. The geopotential height is calculated from the geometric height by integration

$$Z(z) = \int_{z_1}^{z_n} \left(\frac{R_{\text{Mean}}}{R_{\text{Mean}} + \frac{1}{2} \cdot h(z')} \right)^2 \cdot g_{\text{fact}} \cdot h(z') dz'. \quad (83)$$

5.3.5 Pressure

Eq. 84 describes the calculation of dry pressure $p_A(z)$ which is equal to the atmospheric pressure if humidity $p_W(z)$ can be neglected, i.e., everywhere above the lower to middle troposphere

$$p_A(z) = \frac{M_d}{k_1 R} \cdot \int_z^{\infty} g(z') \cdot N(z') dz'. \quad (84)$$

5.3.6 Total Electron Content

For the ionosphere, phase changes measured with a dual band GPS receiver can be used to calculate electron density profiles. This specific data product is especially valuable for space weather applications (Jakowski *et al.*, 2004).

5.4 The CHAMP Satellite

The CHAMP satellite was launched from the COSMODROM at Plesetzsk, ~ 800 km north of Moscow, at the 15 July 2000, 12:00 UTC on-board a COSMOS-3B launcher. The initial orbit was nearly circular ($\varepsilon = 0.004$) at a height of 454 km and an inclination of 87.3° . Due to the atmospheric drag the orbit height is decreasing approximately 50 m/day to 200 m/day, depending on the solar activity. The TRSR-2 (*Black Jack*) receiver used for the RO experiment is a key component of the science payload and serves several purposes onboard CHAMP (Reigber *et al.*, 1995; Wickert *et al.*, 2002, 2001; Reigber *et al.*, 2003).



Figure 17: Artist view of the CHAMP satellite in orbit (courtesy GFZ Potsdam, NASA picture archive, 2004).

6 Validation

6.1 Validation of Adjoint Code

Beside general validation strategies like test-bed setups, adjoint code can be verified with numerical methods,

$$\langle TL(x), TL(x) \rangle = \langle (x), AD(TL(x)) \rangle. \quad (85)$$

The identity expressed in Eq. 85 should hold up to machine accuracy, where TL , denotes the tangent linear, AD , the adjoint and x the original input. These tests have been performed for the observation operators adjoints (interpolation, forward model and filter). This procedure can be applied to single do loops as well as to whole subroutines or even larger sections of the code. It is to mention that this methodology verifies the adjoint code with respect to the forward model. It will not show any bugs in the forward model (the adjoint code can be correct with respect to a erroneous forward model). In our case, single or sets of subroutines as a whole have been tested depending on the setup of the individual operators.

6.2 System Test Runs

The efficiency of an assimilation system can only be verified by real runs, using original data sets. As a system test-bed July 2003 was chosen, using 24 hour forecasts for the 00 and 12 hours time slices respectively 30 hour forecasts for the 06 and 18 hours time slices. The following sample plots are showing the results of test runs comprising 30 simulations within the iterative procedure. The colored part of the figures covers the altitude domain predominately influenced directly by the presence of observations. Given the selected data cut-off heights of 35 km respectively 5 km, the increments above and below these limits are caused by the information spreading within the assimilation system by the background covariance matrices. Model level 20 corresponds approximately to 4 km, model level 50 to 37 km, the contour lines show the increment structures over the whole vertical domain. This interval was chosen to be able to compare with RO only derived temperature maps.

These first validation results agree, in their nature, very well with the first RO only derived climatologies (cf., Gobiet *et al.*, 2004b), and are shown on the following pages. Increments are plotted for temperature and specific humidity which are based on zonal means. The surface pressure increment plots illustrate the horizontal spread of information due to the filter procedure, furthermore they show the global distribution of the observed profiles.

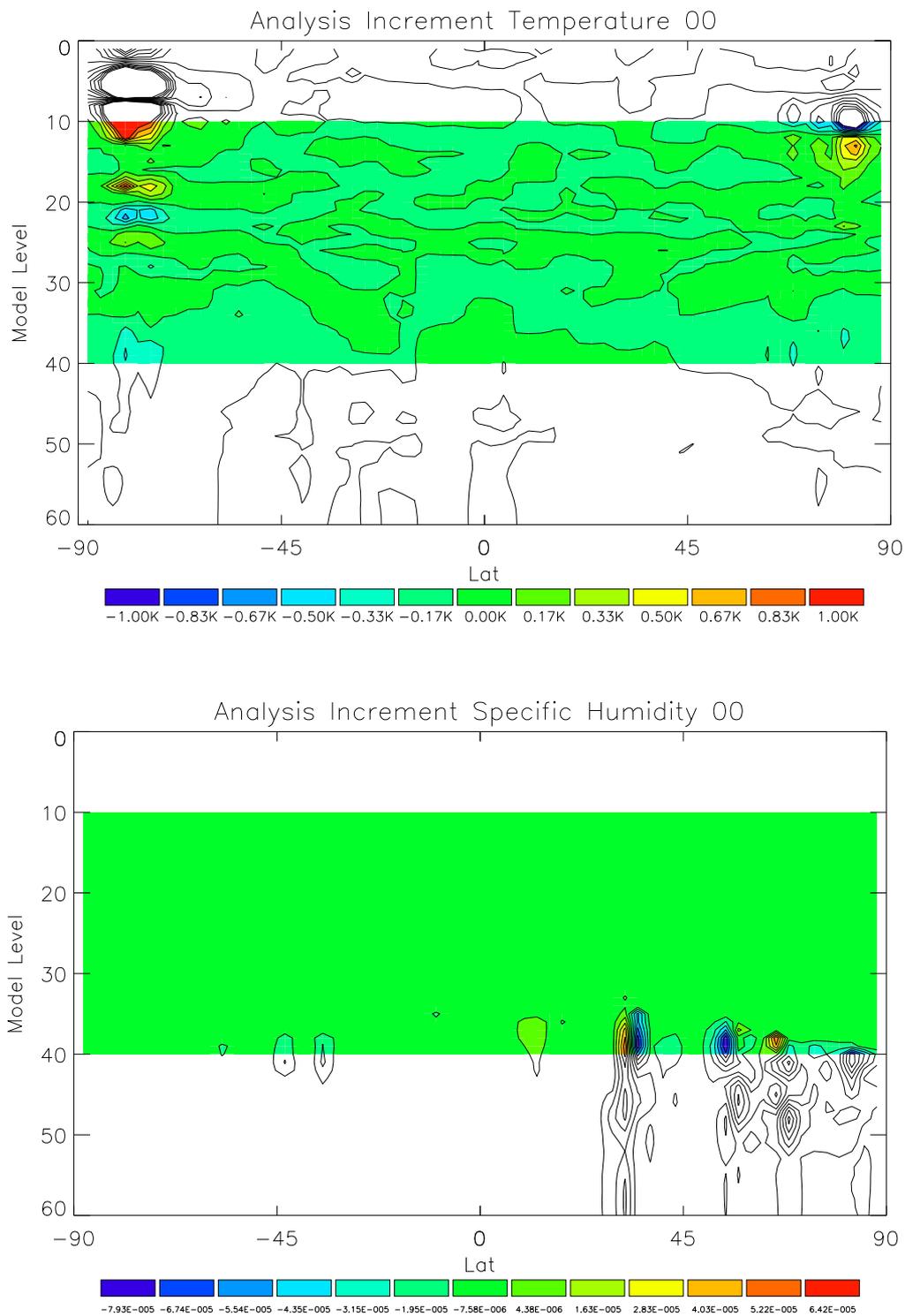


Figure 18: Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 00 UTC time layer.

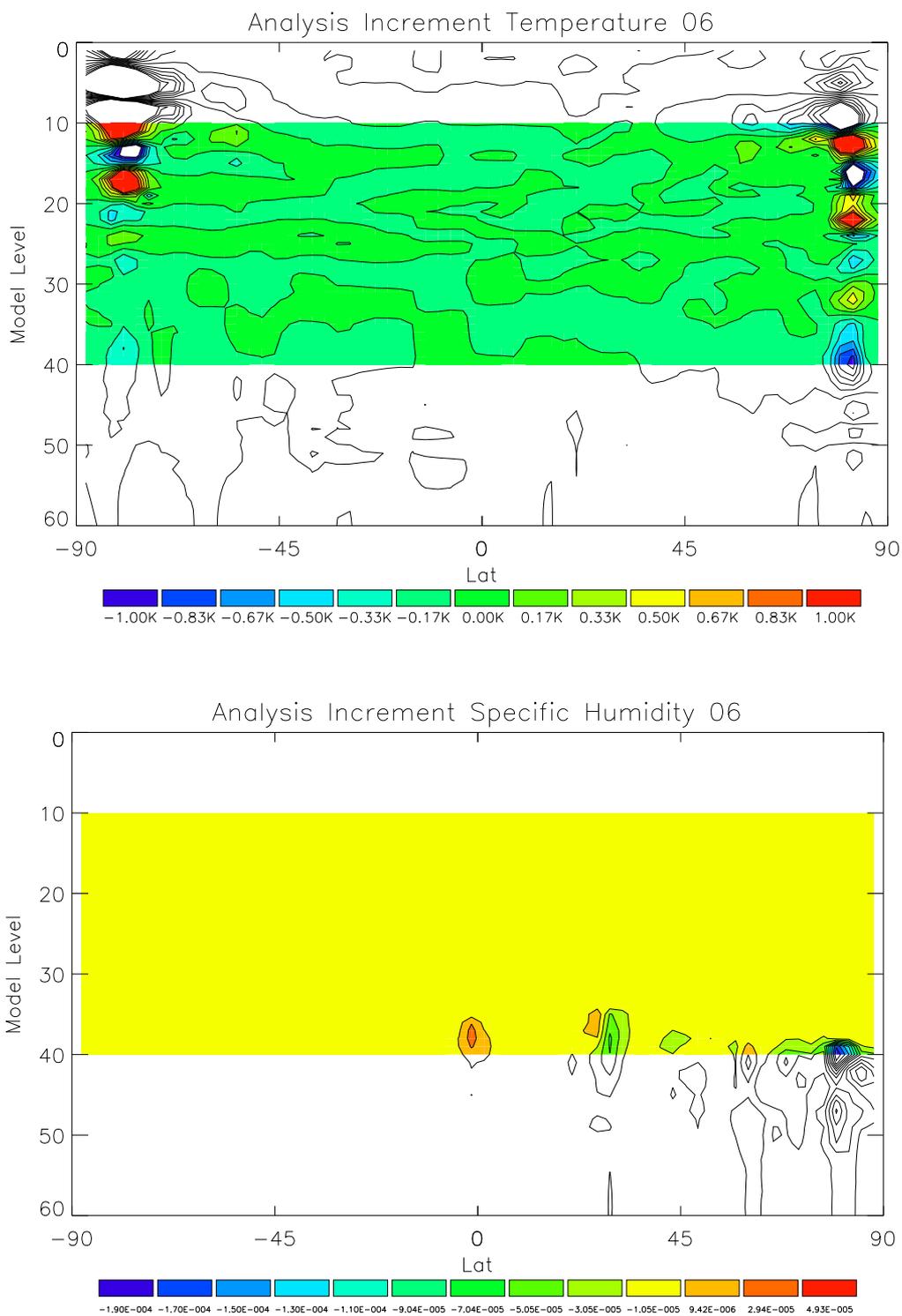


Figure 19: Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 06 UTC time layer.

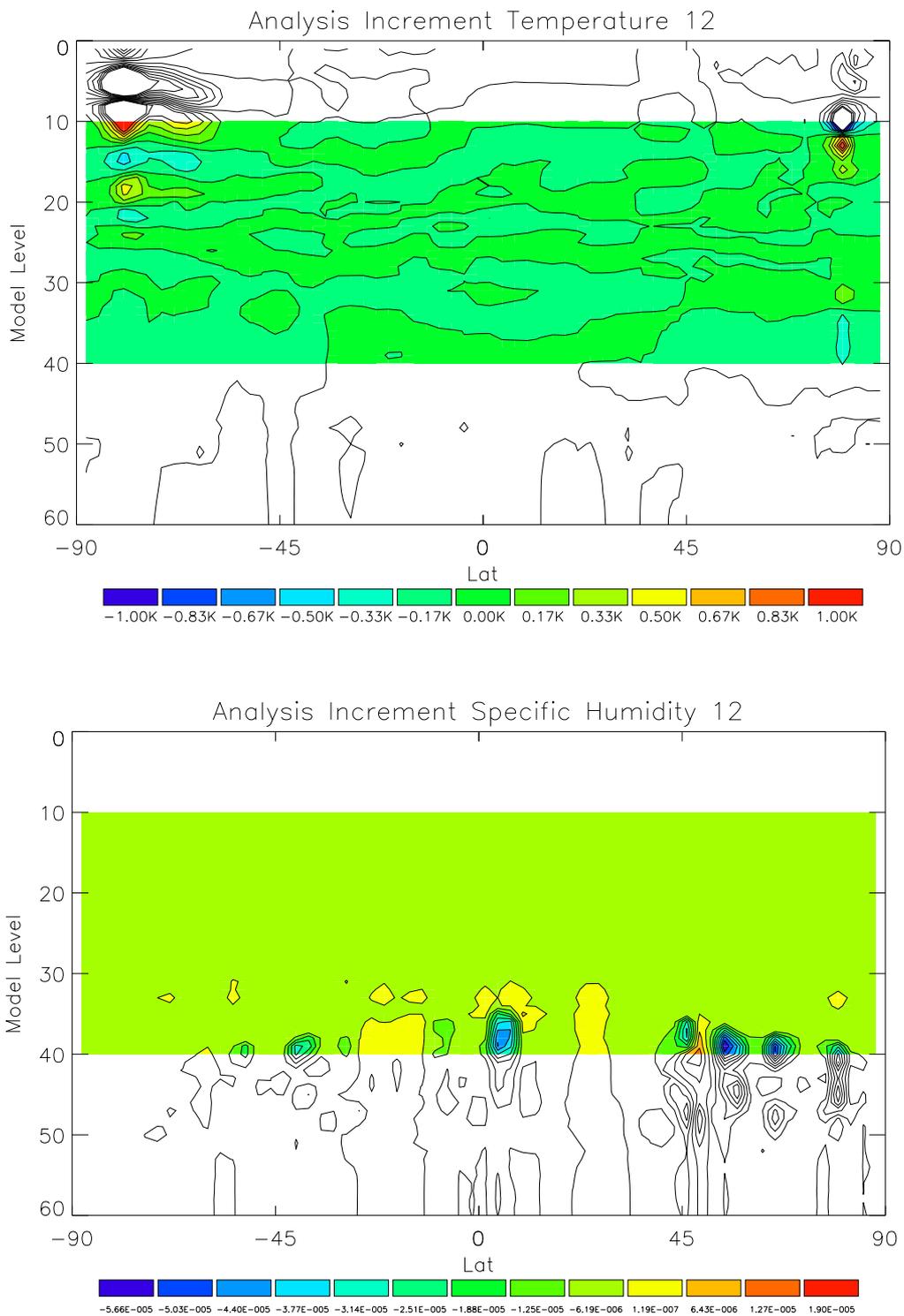


Figure 20: Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 12 UTC time layer.

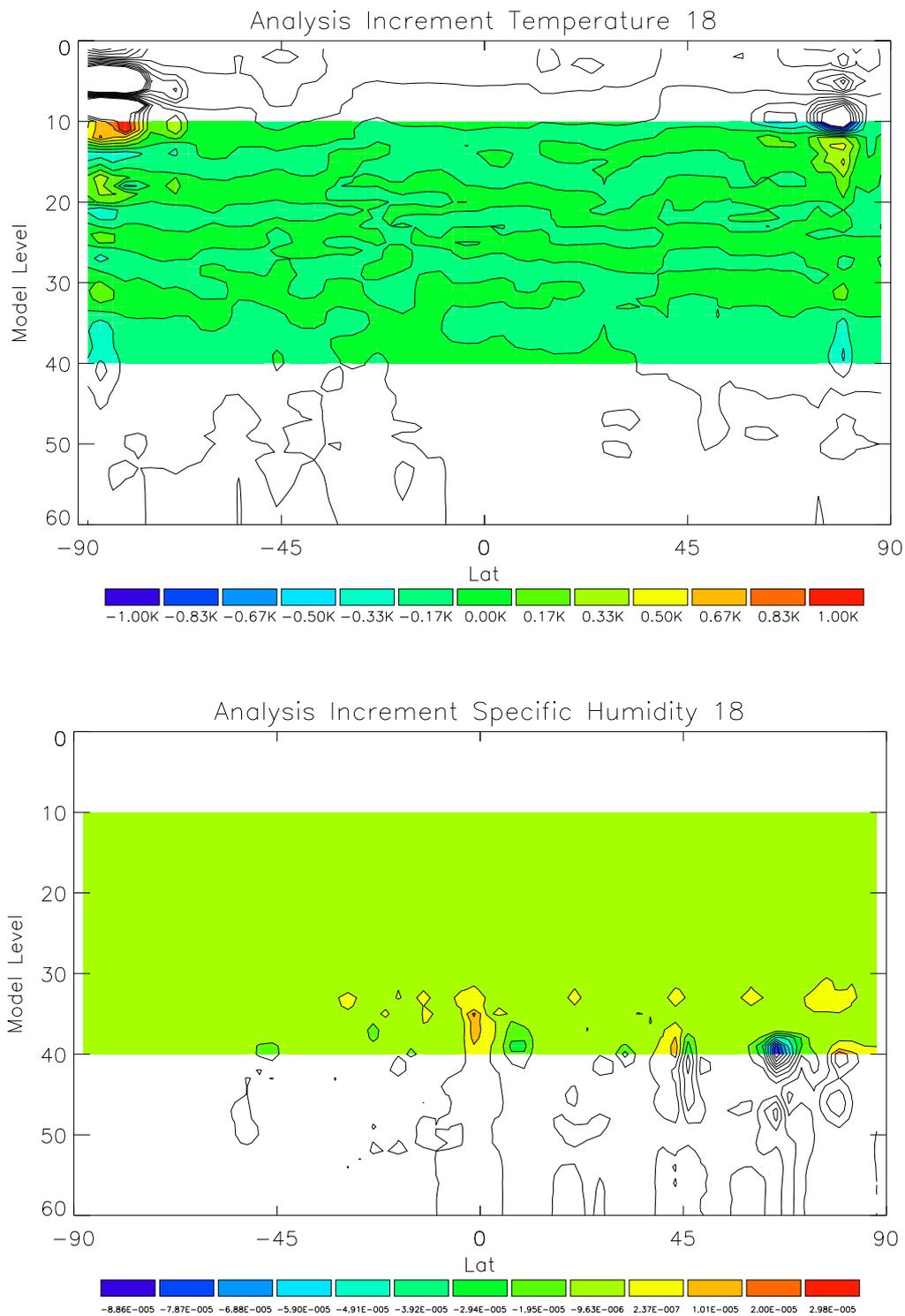


Figure 21: Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 18 UTC time layer.

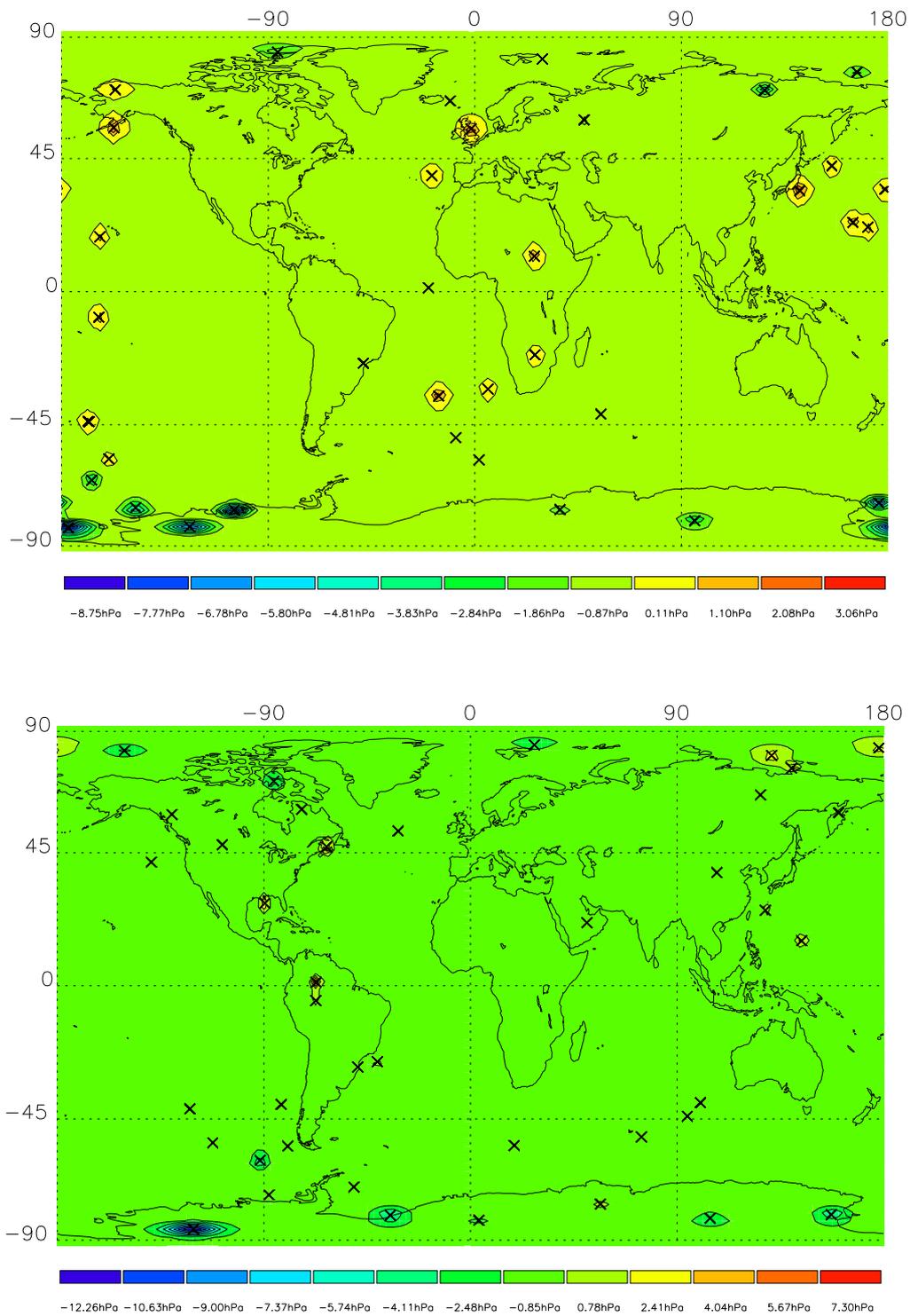


Figure 22: Surface pressure analysis increment 1 July 2003, 00 UTC time layer (top) and 06 UTC time layer (bottom).

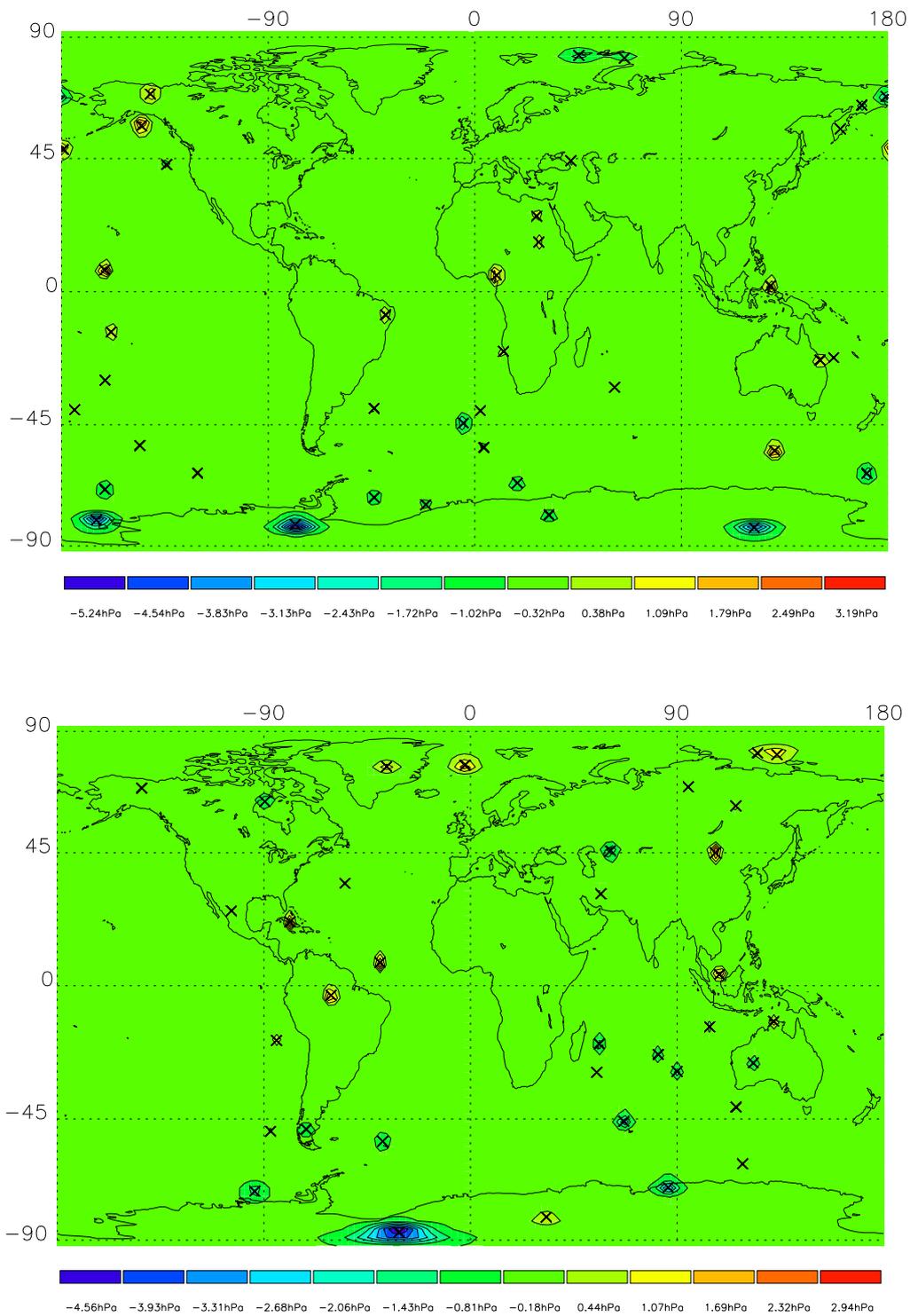


Figure 23: Surface pressure analysis increment 1 July 2003, 12 UTC time layer (top) and 18 UTC time layer (bottom).

Conclusions and Outlook

As shown in the previous section, the first results are promising. The use of variational techniques is not yet widespread within the climate community, mostly due to the strong requirement to use data as *pure* and *uncontaminated* as possible. Fact is that most remote sensing data, since the measurements are in general indirect, do not fulfill these user requirements.

Although RO data is closer to that *ideal observation* than most other observations, a variational approach is promising, since the *bias free* requirement (paramount to variational methods, and also important for climate applications) might be essentially fulfilled for RO data in the near future. Advanced retrieval techniques like wave optics are able to cope with the problems occurring in the lower troposphere and hence are superior to geometric optics below about 5 km (Gorbunov, 2002; Hocke *et al.*, 1999; Sokolovskiy, 2003; Jensen *et al.*, 2003; Beyerle *et al.*, 2003). Future RO data products will take advantage of the improved retrieval techniques and thus deliver high quality observations down into the lower troposphere at least to the top of the planetary boundary layer.

One next step could be the implementation of a bending angle operator, following another principle of data assimilation, using observations as raw and unprocessed as possible. This approach mitigates introduction of errors due to processing and auxiliary data (*incest problem*). It allows for a simpler observation error characterization, avoiding correlations caused by the data processing steps itself.

Furthermore, the chosen concept can be expanded to other observations in future to get multi-instrument analyses. The derivation of first guess fields from, e.g., ERA40 data should be considered to be completely independent from model changes. This methodology offers the unique opportunity to derive climatological fields to complement the traditional climate products, which are in general based on interpolation techniques. Assimilation technique can combine different sources of information in an optimal and consistent way, a clear advantage to interpolation techniques which show their weakness when it comes to merging different sources of information.

In the future the developed RO data assimilation system will be further fine-tuned and applied to create sequences of monthly climate analyses. These will contribute to systematic long-term monitoring of climate variability and change.

Acknowledgments

The authors gratefully acknowledge the GFZ Potsdam, Germany, for the provision of CHAMP RO data, in particular J. Wickert and T. Schmidt for advice and technical support. Thanks go to ECMWF Reading, UK, for providing the analysis/forecast data and especially to M. Fisher and E. Anderson for providing the error specifications for ECMWF fields. Thanks are also due to INRIA for their very useful adjoint compiler. Within the WegCenter/UniGraz A. Steiner is thanked for providing the error characteristics of the CHAMP data, M. Schwärz, J. Fritzer and C. Retscher for helping solve programming and computational issues and for scientific-technical discussions and A. Gobiet and M. Borsche for support concerning the CHAMP data and discussions on the CHAMP based binning-climatologies.

Financial support for the work was received from the START research award of G. Kirchengast funded by the Austrian Ministry for Education, Science, and Culture and managed under Program No. Y103-N03 of the Austrian Science Fund (FWF) and from the CHAMPCLIM project funded by the Austrian Ministry for Traffic, Innovation, and Technology and managed under Contract No. ASAP-WV-203/05 of the Austrian Aeronautics and Space Agency (FFG-ALR).

A Symbols

\mathbf{x}_t	: True State of the Atmosphere	Dimension n
\mathbf{x}_b	: Background Model State	Dimension n
\mathbf{x}_a	: Analysis Model State	Dimension n
\mathbf{y}	: Observation Vector	Dimension p
H	: Observation Operator	Dimension $n \rightarrow p$
\mathbf{H}	: Linear Observation Operator	Dimension $n \rightarrow p$
\mathbf{B}	: Background Covariance Matrix	Dimension $n \times n$
\mathbf{R}	: Observation Covariance Matrix	Dimension $p \times p$
\mathbf{A}	: Analysis Covariance Matrix	Dimension $n \times n$
\mathbf{K}	: Gain Matrix	Dimension $n \times n$
\mathbf{I}	: Identity Matrix	
J_b	: Background Cost Function	
J_o	: Observation Cost Function	
J	: Total Cost Function	
\mathbf{v}	: Background State Vector in Control Space	Dimension n
\mathbf{E}	: Matrix Containing Eigenvectors (Columns)	Dimension $n \times n$
U	: Control Space Transformation Operator	
λ	: Eigenvalues	
P	: Inner Product	
z	: Geometric Height	
ϕ	: Geopotential Height	
φ	: Latitude	
λ	: Longitude	
σ	: Standard Deviation	

B Constants

k_1	=	77.60	[K hPa ⁻¹]	Empirical Constant Thayer Formula & Smith Weintraub
k_2	=	70.40	[K hPa ⁻¹]	Empirical Constant Thayer Formula & Smith Weintraub
k_3	=	373900.00	[K ² hPa ⁻¹]	Empirical Constant Thayer Formula & Smith Weintraub
A	=	6.02214×10^{23}	[mol ⁻¹]	Avogadro Number
R	=	8.3145	[Pa·m ³ /K·mol]	Universal Gas Constant
R_{Dry}	=	287.06	[J K ⁻¹ kg ⁻¹]	Dry Air Gas Constant
R_{WatVap}	=	461.52	[J K ⁻¹ kg ⁻¹]	Water Vapor Gas Constant
m_A	=	28.964	[kg kmol ⁻¹]	Molar Mass of Dry Air
m_W	=	18.015	[kg kmol ⁻¹]	Molar Mass of Water Vapor
g_{Mean}	=	9.80665	[m s ⁻¹]	Mean Acceleration of Gravity
r_{Mean}	=	6371.0	[km]	Mean Radius of Earth
r_{Pol}	=	6356.752314	[km]	Polar Radius of Earth
r_{Equ}	=	6378.137	[km]	Equatorial Radius of Earth
M^*	=	6.022140×10^{26}	[kmol ⁻¹]	Kilo Mol
g_{Equ}	=	9.7803	[m s ⁻¹]	Acceleration of Gravity at Equator

C List of Acronyms

AD	Adjoint
BFGS	Broyden-Fletcher-Goldfarb-Shanno Method
BLUE	Best Linear Unbiased Estimator
CHAMP	Challenging Mini Satellite Payload
COSMIC	Constellation Observing System for Meteorology, Ionosphere & Climate
DMI	Danish Meteorological Institute
ECMWF	European Centre for Medium-Range Weather Forecasts
ERA40	ECMWF Re-Analysis 40
ESA	European Space Agency
GCM	Global Circulation Model
GFZ	Geo Forschungszentrum Potsdam
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GRAS	GNSS Receiver for Atmospheric Sounding
IFS	ECMWF Integrated Forecast System
INRIA	Institut National de Recherche en Informatique et en Automatique
IPCC	Intergovernmental Panel on Climate Change
LEO	Low Earth Orbit Satellite
MEO	Medium Earth Orbit Satellite
METOP	Meteorological Operational Satellite
MSISE	Mass Spectrometry Incoherent Scatter (Extended)
NASA	National Aeronautics and Space Administration
RF	Recursive Filter
RO	Radio Occultation
SOAR	Second Order Autoregressive Function
TEC	Total Electron Content
TL	Tangent Linear
WMO	World Meteorological Organization
3D-VAR	Three Dimensional Variational Data Assimilation
4D-VAR	Four Dimensional Variational Data Assimilation

List of Figures

1	Mean global refractivity profile calculated from T42L60 ECMWF analysis fields.	6
2	Difference in global mean refractivity calculated with Smith-Weintraub and Thayer Formula.	7
3	ECMWF global mean temperature standard deviation absolut values.	11
4	ECMWF global mean temperature standard deviation relative values.	11
5	Global mean vertical error correlations of ECMWF temperature fields, L60 resolution.	12
6	Global mean horizontal error correlations of ECMWF temperature fields.	12
7	ECMWF global mean specific humidity standard deviation absolut values.	13
8	ECMWF global mean specific humidity standard deviation relative values.	13
9	ECMWF vertical error correlation of specific humidity, L60 resolution.	14
10	ECMWF horizontal error correlation of specific humidity.	14
11	Global mean ECMWF horizontal error correlation of surface pressure.	15
12	Segments along one latitude or longitude band in original order.	19
13	Shifted arrangement of segments along one latitude or longitude band.	19
14	Interpolated measurement distribution using two super observations.	23
15	Model orography for the T42 horizontal resolution.	24
16	Basic geometry of a radio occultation measurement including LEO satellite (CHAMP), GPS satellites and fiducial ground network.	28
17	Artist view of the CHAMP satellite in orbit (courtesy GFZ Potsdam, NASA picture archive, 2004).	34
18	Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 00 UTC time layer.	36
19	Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 06 UTC time layer.	37
20	Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 12 UTC time layer.	38
21	Temperature (top) and specific humidity (bottom) analysis increment 1 July 2003, 18 UTC time layer.	39
22	Surface pressure analysis increment 1 July 2003, 00 UTC time layer (top) and 06 UTC time layer (bottom).	40
23	Surface pressure analysis increment 1 July 2003, 12 UTC time layer (top) and 18 UTC time layer (bottom).	41

List of Tables

- 1 Parameters and variables used in Thayer and Smith-Weintraub formula. 6
- 2 Overview of WegCenter CHAMP RO retrieval schemes (EGOPS/CCR Version 2). 31

References

- Barker, D. (1999). A general formulation for 3dvar control variables.
- Beyerle, G., Wickert, J., Schmidt, T., and Reigber, C. (2003). Atmospheric sounding by gnss radio occultation: An analysis of negative refractivity bias using champ observations. *J. Geophys. Res.*, (doi:10.1029/2002JD003922).
- Bouttier, F. and Courtier, P. (1999). Data assimilation concepts and methodes, meteorological training course lecture series. In *European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, RG2 9AX, England*.
- Byrd, R. H., Peihuang, L., Nocedal, J., and Ciyou, Z. (1994). A limited memory algorithm for bound constrained optimization. Technical report, Northwestern University.
- Dong, L. C. and Nocedal, J. (1989). On the limited memory bfgs methode for large scale optimisation. *Mathematical Programming*, **5**, 503–528.
- ECMWF (2004). Ifs documentation, cy28r1. Technical report, ECMWF.
- EUMETSAT, editor (2003). *2nd GRAS SAF User Workshop, Workshop Proceedings*, volume EUM P 40. EUMETSAT.
- Fjeldbo, G. F., Eshleman, V. R., and Kliore, A. J. (1971). The neutral atmosphere of venus as studied with the mariner v radio occultation experiments. *Astron. J.*, **76**, 123–140.
- Gobiet, A. and Kirchengast, G. (2004). Advancements of global navigation satellite system radion occultation retrieval in the upper stratosphere for optimal climate monitoring utility. *J. Geophys. Res.*, **109**(doi:10.1029/2004JD005117).
- Gobiet, A., Steiner, A. K., Retscher, C., Foelsche, U., and Kirchengast, G. (2004a). Algorithms validation based on champ/gps data. Technical report, Inst. for Geophys., Asrophys., and Meteorol., Univ. of Graz, Austria.
- Gobiet, A., Foelsche, U., Steiner, A. K., Borsche, M., Kirchengast, G., and Wickert, J. (2004b). Climatological validation of stratospheric temperatures in ecmwf operational analysis with champ radio occultation data. *J. Geophys. Res.*, **32**(doi:10.1029/2005GL022617).
- Gorbunov, M. E. (2002). Canonical transform methode for processing radio occultation data in the lower troposphere. *Radio Sci.*, **37**(5)(doi: 10.1029/2000RS002592), 1057.
- Gorbunov, M. E. and Gurvich, A. S. (1998). Microlab-1 experiment: Multipath effects in the lower troposphere. *J. Geophys. Res.*, **103**, D12, 13819–13826.
- Gorbunov, M. E. and Kornblueh, L. (2003). Principles of variational assimilation of gnss radio occultation data. Technical Report Report No.350, Max-Planck-Institut fr Meteorologie, Hamburg.
- Gorbunov, M. E. and Sokolovskiy, S. V. (1993). Remote sensing of refractivity from space for global observations of atmospheric parameters. Technical report, Max Planck Institute for Meteorology, Hamburg.
- Gorbunov, M. E., Sokolovskiy, S., and Bengtsson, L. (1996). Space refractivity tomography of the atmosphere: Modeling of direct and invers problems. Technical report, Max Planck Institute for Meteorology, Hamburg.

- Hayden, C. M. and Lorenc, A. C. (1995). Recursive filter for objective analysis of meteorological fields. applications to nesdis operational processing. *J. Appl. Meteor.*, **34**, 3–15.
- Hedin, A. E. (1991). Extension of the msis thermosphere model into the middle and lower atmosphere. *J. Geophys. Res.*, **96**, 1159–1172.
- Hocke, K. (1997). Inversion of gps meteorology data. *Annales. Geophysicae*, **15**, 443–450.
- Hocke, K., Pavelyev, A., Yakovlev, O. I., Barthes, L., and Jakowski, N. (1999). Radio occultation data analysis by the radioholographic methode. *J. Atmos. Terr. Phys.*, **61**, 1169–1177.
- INRIA (2002). Software tapenade inria 2002, version 2.0. Technical report, Domaine de Voluceau, Rocquencourt - BP 105, 78153 Le Chesnay Cedex, FRANCE.
- Jakowski, N., Heise, S., Wehrenpfennig, A., and Tsybulya, K. (2004). *Ionospheric Radio Occultation Measurements and Space Weather*, pages 383–392. In Kirchengast *et al.* (2004).
- Jensen, A. S., Benzon, H.-H., and Nielsen, A. (2003). Full spectrum inversion of radio occultation signals. *Radio Sci.*, **18(3)**(doi:10.1029/2002RS002763), 1159.
- Kirchengast, G., Foelsche, U., and Steiner, A., editors (2004). *Occultations for Probing Atmosphere and Climate*. Springer.
- Kursinski, R. E., Hajj, G. A., Bertiger, W. I., Leroy, S. S., Meehan, T. K., Romans, L. J., Schofield, J. T., McCleese, D. J., Melbourne, W. G., Thorton, C. L., Yunck, T. P., Eyre, J. R., and Nagatani, R. N. (1996). Initial results of radio occultation observations of earth’s atmosphere using the global positioning system. *Science*, **271**, 1107–1110.
- Kursinski, R. E., Hajj, G. A., Schofield, J. T., Linfield, R. P., and Hardy, K. R. (1997). Observing earth’s atmosphere with radio occultation measurements using the global positioning system. *J. Geophys. Res.*, **120**, 23429–23465.
- Lauf, G. B. (1983). Geodesy and map projections. *TAFE Publications Unit*.
- Loescher, A. (2004). Assimilation of gnss radio occultation data into gcm fields for global climate analysis. Technical report, Inst. for Geophys., Asrophys., and Meteorol., Univ. of Graz, Austria.
- Lorenc, A. C. (1992). Analysis methodes for numerical weather prediction. *Qart. J. Ray. Meteor. Soc.*, **112**, 1177–1194.
- Person, A. (2003). *User Guide to ECMWF Forecast Products*.
- Randel, W., Chanic, M. L., and Michaut, C. (2002). Sparc intercomparison of middel atmosphere climatologies. Technical report, WCRP 116.
- Reigber, C., Schwinger, P., and Kohlhas, A. (1995). Champ - a challenging micro satellite payload for geophysical research and application, feasibility study for dara, final report. Technical report, Geo Forschungs Zentrum Potsdam, Germany.
- Reigber, C., Lhr, H., and Scheintzer, P., editors (2003). *First CHAMP Mission Results for Gravity, Magnetic and Atmospheric Studies*. Springer.
- Roeckner, E., Buml, G., Bonaventure, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A. (2003). The atmospheric general circulation model echam5 parti. Technical Report Report No.349, Max-Plank-Institut fr Meteorologie, Hamburg.

- Sokolovskiy, S. V. (2003). Effects of superrefraction on inversions of radio occultation signals in the lower troposphere. *Radio Sci.*, **38(3)**(doi: 10.1029/2002RS002728), 1058.
- Sokolovskiy, S. V. and Hunt, D. (1996). Statistical optimisation approach for gps/met data inversions, paper presented at the ursi gps/met workshop, tucson, arizona. In *URSI GPS/Met Workshop*.
- Steiner, A. K. (2004). Error analyses of refractivity profiles retrieved from champ radio occultation data. Technical Report Scientific Report 04-02, Danish Meteorological Institute.
- Steiner, A. K. and Kirchengast, G. (2004). Error analysis for gnss radio occultation data based on ensembles of profiles from end-to-end simulations. *J. Geophys. Res.*, **110**(doi:10.1029/2004JD005251).
- Vorobev, V. V. and Krasnilnikova, T. G. (1994). Estimation of the accuracy of the atmospheric refractive index recovery from doppler shift measurements at frequencies used in the navstar system. *Phys. Atmos. Ocean*, **29**, 602–609.
- Wickert, J. (2002). Das champ-radiookkultationsexperiment: Algorithmen, prozessierungssystem und erste ergebnisse. Technical Report Scientific Technical Report STR02/07, Geoforschungszentrum Potsdam, Germany.
- Wickert, J., Reigber, C., Beyerle, G., Knig, R., Marquardt, C., Schmidt, T., Grunwald, L., Galas, R., Mehan, T., Melbourne, W. G., and Hocke, K. (2001). Atmosphere sounding by gps radio occultation: First results from champ. *J. Geophys. Res.*, **28**, 3263–3266.
- Wickert, J., Schmidt, T., Marquardt, C., Reigber, C., Neumayer, K. H., Beyerle, G., Galas, R., and Grunwald, L. (2002). Gps radio occultation with champ: First results and status of the experiment, proc. of iag scientific assambly, 2.-7. september, budapest. In *Proc. of IAG Scientific Assambly*, volume 125.

– end of document –