Four-dimensional variational data assimilation for estimation of the atmospheric chemical state from the tropopause to the lower mesosphere

I n a u g u r a l – D i s s e r t a t i o n
zur
Erlangung des Doktorgrades
der Mathematisch–Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von
Jörg Schwinger
aus Bonn

Köln 2006

Berichterstatter: Prof. Dr. A. Ebel
Prof. Dr. M. Kerschgens

Tag der mündlichen Prüfung: 30.10.2006

Abstract

A four-dimensional variational data assimilation system for stratospheric trace gas observations has been developed and operated. The system offers the flexibility to make use of data from different instruments and it was designed (1) to enforce chemical consistency by constraining the analyses to a state of the art stratospheric model, (2) to provide realistic estimates of anisotropic and inhomogeneous background error covariances, and (3) to be sufficiently efficient for application in near real time. The assimilating model has been assembled from scratch to allow for a couple of novel features: The meteorological fields are computed online by the global forecast model GME of German Weather Service, leading to an improved numerical representation of wind fields compared to traditional offline chemistry-transport models, which use spatially and temporally interpolated meteorological analyses. A number of 155 photolysis, gas phase, and heterogeneous reaction of 41 stratospheric trace gases is considered by the chemistry module. Since spatial correlations between background errors evolve according to the atmospheric flow, a flow dependent formulation of the background error covariance matrix has been devised by means of a diffusion approach. It can be shown that this measure considerably improves the analysis quality particularly in regions where large gradients of potential vorticity prevail. The governing equations are discretised on an icosahedral grid, as this significantly reduces the computational cost. Therefore, it is possible to operate the model with a relatively fine spatial resolution without violating the near real time constraint. A comprehensive set of case studies has been conducted in order to test and evaluate the new system. Trace gas profiles derived from measurements of the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) have been assimilated. Comparison with independent (not assimilated) control data sets and statistical evaluation demonstrates an excellent performance of the new assimilation system.

Kurzzusammenfassung

Ein System zur Assimilation stratosphärischer Spurengasmessungen basierend auf der vierdimensionalen variationellen Methode wurde entwickelt und angewandt. Dieses neue System bietet die Möglichkeit, Messdaten verschiedenster Sensoren zu verwenden, und zeichnet sich durch (1) chemische Konsistenz der analysierten Felder im Sinne eines umfassenden stratosphärischen Modells sowie (2) realistische Modellierung von anisotropen und inhomogenen Hintergrundfehler-Kovarianzen aus und ermöglicht (3) aufgrund seiner hohen nummerische Effizienz einen operationellen Einsatz. Das für die Assimilation verwendete Modell wurde neu entwickelt und weist eine Reihe von vorteilhaften Eigenschaften auf. Die meteorologischen Felder werden mithilfe des globalen Wettervorhersagemodells GME des Deutschen Wetterdienstes direkt erzeugt. Dadurch wird, im Vergleich zu herkömmlichen Chemie-Transport-Modellen, die zeitlich und räumlich interpolierte meteorologische Analysen verwenden, die Darstellung der Windfelder im Modell entscheidend verbessert. Das Chemiemodul berücksichtigt 155 Photolyse-, Gasphasen-, und heterogene Reaktionen zwischen 41 stratosphärischen Spurengasen. Da die räumlichen Korrelationen von Hintergrundfehlern an die Dynamik der atmosphärischen Strömung gekoppelt sind, wurde eine strömungsabhängige Formulierung der Hintergrundfehler-Kovarianzmatrix mithilfe eines Diffusionsansatzes realisiert. Es lässt sich zeigen, dass dadurch die Qualität der Analysen deutlich verbessern werden kann, insbesondere in Gebieten, in denen die potentielle Vorticity starke Gradienten aufweist. Die Lösung der Modellgleichungen erfolgt auf einem Ikosaedergitter, durch dessen Verwendung der Rechenaufwand signifikant verringert wird. Aus diesem Grund ist es möglich, das Modell auch im operationellen Einsatz mit einer relativ hohen räumlichen Auflösung zu betreiben. Anhand von umfangreichen Fallstudien konnte das neue System getestet und evaluiert werden. Dazu wurden Spurengasprofile, abgeleitet aus Messungen des Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), assimiliert. Vergleiche mit unabhängigen (nicht assimilierten) Beobachtungen und statistische Auswertungen zeigen, dass das neue Assimilationssystem hervorragend arbeitet.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

The term *data assimilation* originates from meteorology. With the advent of numerical weather prediction in the 1950s it became necessary to provide initial conditions for the computer models in order to obtain a numerical forecast. To this end, an *objective analysis* of the atmospheric state on regular model grids had to be derived from a set of sparse and spatially scattered weather observations. It soon became apparent that the numerical model can contribute valuable information to the analysis process, as it includes the physical laws governing the atmospheric flow. According to *Talagrand* [1997],

> *assimilation of meteorological or oceanographical observations can be described as the process through which all the available information is used in order to estimate as accurately as possible the state of the atmospheric or oceanic flow. The available information essentially consists of the* observation *proper, and of the physical laws which govern the evolution of the flow. The latter are available in practice under the form of a* numerical model.

Assimilation of atmospheric trace gas observations is closely related to meteorological data assimilation, as the numerical models are similar in terms of complexity and the large dimension of the state vector.

Stratospheric processes have been recognised to be of utmost importance for the earth system in recent years. The discovery of the ozone hole in the 1980s and, later on, the question how human activity influences the earth's

climate system has led to an increased interest in a thourough understanding of the complex interaction between chemistry, radiation, and dynamics of the stratosphere. Stratospheric trace gases play a key role in the radiation budget of the earth and their distribution influences the atmospheric circulation. Conversely, atmospheric conditions determine the rate of chemical reactions as well as the redistribution of trace gases by advection and turbulent diffusion (*World Meteorological Organization* [1999], *Intergovernmental Panel on Climate Change* [2001]).

For a long time only sparse stratospheric observational data was available. The lower stratosphere up to a height of approximately 32 km is being observed by a more or less dense network of ozone sondes, while for larger altitudes instruments mounted on research rockets were the only source of in situ observational data. With the advent of space borne remote sounding devices, information about stratospheric trace gases can be derived from emitted, scattered, or transmitted radiation, which is recorded by these instruments. In March 2002 the European research satellite EnviSat was launched into a polar orbit carrying the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) and the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) aboard, which are delivering an unprecedented wealth of observations of a number of stratospheric trace gases with global coverage.

However, satellite observations are, by nature, scattered in space and time, whereas the vast majority of applications require spatially and temporally uniform and consistent fields of atmospheric constituents. These applications include operational weather forcasting (e. g. *Derber and Wu* [1998], *Geer et al.* [2006]), ozone forecasting (*Eskes et al.* [2002], *Eskes et al.* [2004]), process studies (e. g. *Hoffmann and Riese* [2004]), and initialisation of climate models. Hence, advanced data assimilation techniques, must be applied to draw full advantage of stratospheric remote sounding data. For stratospheric constituent assimilation these techniques should ideally be able

1. to combine *all* available observational data from different sensors in an optimal way according to their error statistics, thereby
2. producing a comprehensive and chemically consistent picture of the atmospheric state, and
3. to reveal information about unobserved species, which are chemically coupled with observed constituents.

Furthermore, data assimilation can be used to identify inconsistencies between atmospheric models and observational data. It should be emphasised

that this task *cannot* be accomplished by comparing an unconstrained model integration with measurements, as even a perfect model is not expected to reproduce observations, unless adequate initial conditions are supplied. Hence, advancements in the understanding of atmospheric processes require a confrontation of models with data by employing assimilation techniques, which are rigorously based on statistical theory. In this spirit, data assimilation has been used, for example, to test chemical theories (*Lary et al.* [2003]) and to monitor observation errors (*Stajner et al.* [2004]).

Data assimilation of stratospheric constituents is a complex and computationally expensive endeavour. The elaborate model is only one part of the whole assimilation system and often considerable simplifications have been introduced in oder to reduce the computational cost. In this regard, the implementation of the background error covariance matrix (BECM) is a crucial issue: Since an optimal analysis requires a realistic representation of error statistics, the BECM constitutes a core element of the assimilation system. Especially if data is sparse, the analysis quality largely depends on the background error correlations encoded in the BECM. However, the dimension of this this matrix is enormous for comprehensive three-dimensional models and, consequently, it can only be handled in parameterised form. As the background field is invariably given by a short range forecast in present assimilation systems, it is the error statistics of this forecast that is to be approximated. Several approaches to this problem with differing degrees of sophistication exist (see *Riishøjgaard* [1998] for a detailed discussion). Despite the necessary simplicity, a skillful parameterisation should be capable of representing the relevant structures of the background error covariances. This includes the possibility of modelling inhomogeneous (varying with location) and anisotropic (varying with direction) correlations length scales. Until now all approaches that have been adopted for stratospheric constituent assimilation suffer from the fact that they are either overly simplified or prohibitively expensive in conjunction with three-dimensional models.

Among the algorithms to be considered for assimilation of stratospheric constituents there are only two candidates, which are able to comply with all of the three requirements mentioned above, namely the Kalman filter (*Kalman* [1960], *Cohn* [1997]) and the four-dimensional variational (4D-var) method (*Talagrand and Courtier* [1987], *Elbern et al.* [1997]). The former is a sequential method, i. e. the model state is corrected at times when observations are encountered. It possesses the theoretical advantage that the background error covariances are evolved according to the model dynamics and analysis error covariances are provided together with the analysed fields. However, the implementation of the full Kalman filter algorithm is not feasible for

large problems. Moreover, the costly analysis step cannot, in practise, be performed every time observations are encountered. Instead, measurements falling into a certain time interval, the *assimilation window*, are collected and assimilated together. This particularly hampers the assimilation of photochemically active species, which show a diurnal cycle. Unlike the Kalman filter, the 4D-var algorithm acts as a smoother, as it adjusts the initial values of the assimilating model, such that differences between observations and model state within a predefined time interval are minimised in a root-mean-square sense. The 4D-var method is sufficiently efficient to be implemented without serious simplifications. A drawback, however, is that there is no simple strategy to derive an analysis error estimate.

A first application of the 4D-var method to a small stratospheric reaction mechanism in connection with a trajectory model has been presented by *Fisher and Lary* [1995]. This was the first study considering the assimilation of chemically active stratospheric constituents. In the following years a number of studies on stratospheric trace gas assimilation has been published. *Lyster et al.* [1997] presented an implementation of the full Kalman filter atop a two-dimensional isentropic tracer transport model. The same system was used by *Ménard et al.* [2000*a*] and *Ménard et al.* [2000*b*] to assimilate methane observations. These studies together with the work presented by *El Serafy et al.* [2002] have been the only considering a rigorous treatment of the BECM including the time evolution of the error variances and covariances. This, however, was only possible at the cost of employing a model with strongly reduced complexity. *Levelt et al.* [1998] and, later on, *Khattatov et al.* [2000] and *Chipperfield et al.* [2002] assimilated observations of long-lived tracers into three-dimensional chemistry transport models (CTMs) with a detailed representation of stratospheric chemistry. A simplified (suboptimal) Kalman filter was employed together with an isotropic and homogeneous parameterisation of the BECM and the models were operated with relatively low resolution. The first application of the 4D-var algorithm together with a comprehensive stratospheric CTM was presented by *Errera and Fonteyn* [2001]. By virtue of the relative numerical efficiency of 4D-var these authors have been able to assimilate a large number of trace gas profiles obtained by the Cryogenic Infrared Spectrometers and Telescopes (CRISTA). Besides ozone, the set of assimilated species comprised the chemically more active constituents HNO_3 and $ClONO_2$. The CRISTA instrument provided a global coverage within 24 hours with exception of the polar regions, which remained unobserved. Therefore, it proved acceptable to implement the BECM as a diagonal matrix, neglecting background error correlations, allthough this approach is not expected to work well in general. The benefit of an aniso-

tropic *flow dependent* background error covariance formulation, which was proposed by *Riishøjgaard* [1998], was demonstrated by *Fierli et al.* [2002]. In conjunction with a high-resolution two-dimensional tracer transport model it was possible to preserve fine scale structures in the analysed ozone field.

Without exception, all models employed in the studies mentioned above use external analyses to provide the meteorological fields for the solution of the reaction-advection equation. While this reduces the complexity of the model, the meteorological fields have to be interpolated in space and time, leading to a poor representation of wind fields, especially of the vertical wind. An alternative approach has been taken in recent years by numerical weather prediction centres: Ozone assimilation capabilities are being added to their operational data assimilation schemes (*Derber and Wu* [1998], *Struthers et al.* [2002], *Geer et al.* [2006], *Bormann et al.* [2005]). This is motivated by the fact that a realistic distribution of ozone must be known to improve the derivation of temperature information from space borne remote sounding data. Furthermore, it can be hoped that a better representation of ozone will improve medium- and long-range weather forecasts, since the feedback between atmospheric dynamics and ozone distribution can be accounted for. With the latter solution the consistent, high quality wind fields of the meteorological models are used for the transport of ozone. However, as it is not intended to deliver a comprehensive analysis of atmospheric trace gases, no or only a linearised chemistry schemes are included. A model with full coupling of radiation, chemistry, and dynamics –such models are usually called *General Circulation Models* or GCMs, (*Lahoz* [2003])– has been recently presented by *Polavarapu et al.* [2005]. A three-dimensional variational assimilation scheme has been operated with this model and it is foreseen to add ozone assimilation capabilities. A GCM is the most complex among the models describing the atmosphere, and a number of benefits can be expected from this approach. Advanced stratospheric constituent assimilation with a full GCM, however, is still a distant prospect.

The work presented here addresses some of the shortcomings of present assimilation systems for stratospheric constituents. The SACADA (**S**ynoptic **A**nalysis of **C**hemical Constituents by **A**dvanced **D**ata **A**ssimilation) system has been assembled from scratch to allow for a couple of novel features. The four-dimensional variational approach, which is outlined in Chapter 2 along with an introduction to data assimilation theory, has been selected for the new system. A key issue was the development of a sophisticated background error covariance parameterisation. It was decided to follow a diffusion approach proposed by *Weaver and Courtier* [2001], which is, in some detail, described in Chapter 3. The resulting backround error covariance operator

is well suited for the application with large models and allows for anisotropic and inhomogeneous background error correlations, a feature that was utilised to devise a *flow dependent* formulation of the BECM.

The chemistry-transport module has been build atop the global weather forecast model GME of German Weather Service, as described in Chapter 4. By virtue of this concept, the use of numerically consistent wind fields for the solution of the reaction-advection equation is guaranteed. Furthermore, this can be seen as a first step towards a full General Circulation Model, as future extensions of the model aiming to capture the feedback between chemistry, radiation and dynamics are conceivable. The *icosahedral model grid*, which has been adopted from GME, reduces the computational effort for the chemistry-transport calculation by about 25% compared to traditional CTMs with comparable spatial resolution. Aditionally, the assimilation software has been designed for parallel compute environments, leading to an excellent efficiency of the new system. Therefore, it possible to operate the model with a relatively high spatial resolution, such that –particularly in combination with the flow dependent BECM formulation– filament structures of constituent fields can be analysed with the new system.

A comprehensive set of case studies has been accomplished to evaluate and test the SACADA assimilation system. Profiles of various stratospheric trace gases derived from MIPAS spectra have been assimilated. Observational data from the Stratospheric Aerosol and Gas Experiment II (SAGE II), the Halogen Occultation Experiment (HALOE), and SCIAMACHY served as independent control data sets. Chapter 5 provides an overview about these instruments and the respective data products. The results presented in Chapter 6 confirm that the new SACADA system is able to efficiently deliver high quality analyses of stratospheric constituents. A statistical evaluation of analysed, background, and observational profiles is performed, which proves the fully satisfying skill of the new system. It is demonstrated that ozone analyses can be employed to *cross validate* ozone retrievals from different sensors. The benefit of a flow dependent background error covariance formulation could be objectively quantified for the first time in the context of trace gas assimilation. A summary of the present work and a discussion of results is given in Chapter 7.

CHAPTER 2

Data assimilation and the 4D-var technique

Estimation theory provides the theoretical foundations of advanced data assimilation methods that are nowadays used in atmospheric modelling. This theory deals with the question how to estimate a quantity in a manner that is optimal in some sense, using information from different sources with different error characteristics. As all available information on the atmospheric state is imperfect, that is, afflicted with errors, an estimate of the true atmospheric state will be stochastic in nature. In this context, the data assimilation problem can be formulated as follows (see *Cohn* [1997] for a comprehensive discussion): Assume that a probability density function (PDF) $p(\boldsymbol{x})$ of the discrete atmospheric state $\boldsymbol{x}$ is available. This PDF may be based on a climatology, but usually it is derived from a short term forecast of a computer model initialised with previous analyses. Further assume that additional information is provided by observations $\boldsymbol{y}$. If the error characteristics of $\boldsymbol{y}$ is known, it is possible to formulate a PDF $p(\boldsymbol{y}|\boldsymbol{x})$, which describes the probability of taking a measurement $\boldsymbol{y}$ under a given atmospheric state $\boldsymbol{x}$. Data assimilations aims to update the *a priori* PDF $p(\boldsymbol{x})$ with incoming new information to yield the *a posteriori* PDF $p(\boldsymbol{x}|\boldsymbol{y})$ of the atmospheric state, given the observations $\boldsymbol{y}$. The solution can be obtained by applying Bayes' theorem:

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{x})\, p(\boldsymbol{x})}{\int p(\boldsymbol{y}|\boldsymbol{x})\, p(\boldsymbol{x})\, d\boldsymbol{x}}\,. \tag{2.1}$$

This theorem is very general and no constraints on the nature of the involved PDFs are imposed. Note however, that, as the dimension of $\boldsymbol{x}$ is as large

as 10^6–10^7 for state of the art atmospheric models, a complete numerical representation of the these PDFs is not possible, allthough the a posteriori PDF may be sampled by means of an ensemble integration, as recently shown by *van Leeuwen* [2003]. This technique is not yet very well established, but it is a promising approach for problems with strongly non-linear model dynamics and in cases where no simple approximation of $p(\boldsymbol{x})$ and $p(\boldsymbol{y}|\boldsymbol{x})$ can be devised.

2.1 Maximum likelihood and minimum variance estimates

Usually it is not possible to obtain $p(\boldsymbol{x}|\boldsymbol{y})$ or a suitable approximation thereof, and hence, an atmospheric state $\boldsymbol{x}^a$ may be selected, which represents some optimum, for example the state having the maximum likelihood or the minimum variance. If it is justified to approximate $p(\boldsymbol{x})$ and $p(\boldsymbol{y}|\boldsymbol{x})$ by Gaussian PDFs, and if observations are *not scattered* in time, the a posteriori probability density reads

$$\begin{aligned} p(\boldsymbol{x}|\boldsymbol{y}) = A \exp\left\{-\frac{1}{2}\left[\boldsymbol{y}-H(\boldsymbol{x})\right]^T \mathbf{R}^{-1}\left[\boldsymbol{y}-H(\boldsymbol{x})\right]\right\} \times \\ \exp\left\{-\frac{1}{2}\left[\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}}\right]^T \mathbf{B}^{-1}\left[\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}}\right]\right\}, \end{aligned} \tag{2.2}$$

where A is a normalisation factor arising from the denominator in (2.1). The *background* state $\boldsymbol{x}^{\mathrm{b}}$ is the expected value of the a priori PDF $p(\boldsymbol{x})$. $\mathbf{R}$ and $\mathbf{B}$ are the observation and background error covariance matrices, respectively, and H is an operator which maps from the n-dimensional space of model variables into the p-dimensional observation space. In the simplest case H performs a linear interpolation of model values from the grid point domain to the location of observations, but H also may contain radiative transfer calculations, for example, if the observed quantity is atmospheric radiance recorded by a remote sounding instrument. The maximum of the a posteriori PDF can be found by minimising the *cost function*

$$\begin{aligned} J(\boldsymbol{x}) := \frac{1}{2}\left[\boldsymbol{y}-H(\boldsymbol{x})\right]^T \mathbf{R}^{-1}\left[\boldsymbol{y}-H(\boldsymbol{x})\right] + \\ \frac{1}{2}\left[\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}}\right]^T \mathbf{B}^{-1}\left[\boldsymbol{x}-\boldsymbol{x}^{\mathrm{b}}\right]. \end{aligned} \tag{2.3}$$

If the observation operator is linear ($H = \mathbf{H}$), the minimum of J can be easily obtained (see e. g. *Rodgers* [2000], *Kalnay* [2003]) by

$$\boldsymbol{x}^{\mathrm{a}} = \boldsymbol{x}^{\mathrm{b}} + \left[\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1} \mathbf{H}^T\mathbf{R}^{-1}\left(\boldsymbol{y} - \mathbf{H}\boldsymbol{x}^{\mathrm{b}}\right), \tag{2.4}$$

and the error covariance matrix of the analysis state is given by

$$\mathbf{P}^{\mathrm{a}} = \left[\mathbf{B}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1} . \tag{2.5}$$

Equation (2.4) can be derived in a different way, in the context of linear optimal estimation. Thereby, no assumptions have to be made about the error statistics encoded in $\mathbf{R}$ and $\mathbf{B}$ apart from the requirement that $\boldsymbol{x}^{\mathrm{b}}$ and $\boldsymbol{y}$ have to be unbiased against the true state $\boldsymbol{x}^{\mathrm{t}}$. Then, $\boldsymbol{x}^{\mathrm{a}}$ turns out to be the minimum variance estimate (see e. g. *Talagrand* [1997]). This estimate is also called *Best Linear Unbiased Estimate* (BLUE), which, as we have seen, is equivalent to the maximum likelihood estimate in the linear case with Gaussian error statistics. Equation (2.4) is the basis of the optimal interpolation algorithm and it constitutes the analysis step of the Kalman filter[1].

In the non-linear case, the formal solution of the problem of minimising the cost function (2.3) can be found by applying a Gauss-Newton iteration (see *Rodgers* [2000] for details)

$$\begin{aligned} \boldsymbol{x}_{i+1} &= \boldsymbol{x}^{\mathrm{b}} + \mathbf{W}_i \left(\boldsymbol{y} - H(\boldsymbol{x}_i) + \mathbf{H}_i'(\boldsymbol{x}_i - \boldsymbol{x}^{\mathrm{b}})\right) , \\ \text{with} \quad \mathbf{W}_i &= \left[\mathbf{B}^{-1} + \mathbf{H}_i'^T\mathbf{R}^{-1}\mathbf{H}_i'\right]^{-1} \mathbf{H}_i'^T\mathbf{R}^{-1} , \end{aligned} \tag{2.6}$$

and $\mathbf{H}_i'$ being the linearisation of the observation operator around the state $\boldsymbol{x}_i$. The analysis error covariance matrix now contains the linearisation $\mathbf{H}_a'$ of H at the analysis point:

$$\mathbf{P}^{\mathrm{a}} = \left[\mathbf{B}^{-1} + \mathbf{H}_a'^T\mathbf{R}^{-1}\mathbf{H}_a' +\right]^{-1} . \tag{2.7}$$

Note that (2.7) is a close approximation to the Hessian matrix of J at $\boldsymbol{x} = \boldsymbol{x}^{\mathrm{a}}$. A wide range of applications in different areas of geosciences is based on (2.6), for example, the retrival of atmospheric trace gas profiles from spectral data recorded by remote sounding instruments, as described in Chapter 5.

For the data assimilation problem however, neither (2.4) nor (2.6) can be applied without simplifications, due to the large dimension of the state vector $\boldsymbol{x}$. The three-dimensional variational method for example, circumvents the direct computation of the expression $\left[\mathbf{B}^{-1} + \mathbf{H}_i'^T\mathbf{R}^{-1}\mathbf{H}_i'\right]^{-1}$ by employing minimisation algorithms which rely on the gradient of J only (and possibly on some coarse approximation of the Hessian). Variants of optimal estimation, which have been used in operational data assimilation schemes are described in *Daley* [1991] and *Kalnay* [2003].

[1] The Kalman filter analysis step is usually written using the equivalent formulation $\boldsymbol{x}^{\mathrm{a}} = \boldsymbol{x}^{\mathrm{b}} + \mathbf{B}\mathbf{H}^T \left[\mathbf{H}\mathbf{B}\mathbf{H}^T + R\right]^{-1} (\boldsymbol{y} - \mathbf{H}\boldsymbol{x}^{\mathrm{b}})$

2.2 Four-dimensional variational data assimilation

Observations of atmospheric quantities are usually distributed in time and Equation (2.2) can be generalised such that the evolution of the atmospheric state is accounted for. Let M_i denote the non-linear model operator that propagates the atmospheric state $\boldsymbol{x}_0$ at time t_0 to time t_i. An a posteriori PDF considering all observations within the time interval $[t_0, t_N]$ can be constructed by defining

$$p(\boldsymbol{x}_0|\boldsymbol{y}) = A \exp\left\{-\frac{1}{2}\sum_{i=0}^{N}\left[\boldsymbol{y}_i - H\left(M_i(\boldsymbol{x}_0)\right)\right]^T \mathbf{R}^{-1}\left[\boldsymbol{y}_i - H\left(M_i(\boldsymbol{x}_0)\right)\right]\right\} \times \exp\left\{-\frac{1}{2}\left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}\right]^T \mathbf{B}^{-1}\left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}\right]\right\}. \tag{2.8}$$

Here, $\boldsymbol{y}_i$ is the vector of observations available within time step i. Equation (2.8) implements the model M as a *strong constraint*, that is, the model is assumed to reproduce the evolution of the atmosphere without errors (*perfect model assumption*). Weak constraint formulations of (2.8) are possible, as outlined by *van Leeuwen and Evensen* [1996], however, the computational effort needed to find the maximum of the a posteriori PDF increases considerably.

The four-dimensional variational method finds the maximum of the PDF (2.8) by minimising the following cost function:

$$\begin{aligned} J(\boldsymbol{x}_0) = J^{\mathrm{b}} + J^{\mathrm{o}} = \\ \frac{1}{2}\left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}\right]^T \mathbf{B}^{-1}\left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}\right] + \\ \frac{1}{2}\sum_{i=0}^{N}\left[H\left(M_i(\boldsymbol{x}_0)\right) - \boldsymbol{y}_i\right]^T \mathbf{R}^{-1}\left[H\left(M_i(\boldsymbol{x}_0)\right) - \boldsymbol{y}_i\right]. \end{aligned} \tag{2.9}$$

Efficient minimisation algorithms like quasi-Newton or Conjugate-Gradient methods require the gradient of the cost function with respect to the control variables $\boldsymbol{x}_0$ in order to find the minimum of J. The gradient of the background portion J^{b} of the cost function is readily obtained by

$$\nabla_{\boldsymbol{x}_0} J^{\mathrm{b}} = \mathbf{B}^{-1}\left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}\right], \tag{2.10}$$

but the gradient of J^{o} with respect to the initial model values is more difficult to calculate. Allthough it is easy to express the gradient of J^{o} with respect

to the model variables at time t_i, namely

$$\nabla_{\boldsymbol{x}_i} J^{\mathrm{o}} = \mathbf{H}'^T \mathbf{R}^{-1} \left[H(\boldsymbol{x}_i) - \boldsymbol{y}_i \right] , \tag{2.11}$$

it is the calculation of $\nabla_{\boldsymbol{x}_0} J^{\mathrm{o}}$ that is the computationally most demanding task of 4D-var data assimilation. As the number of control variables in atmospheric models is of the order 10^6–10^7, the only feasible strategy to accomplish this calculation is given by utilising the adjoint model operator.

Let $\langle \cdot \, , \, \cdot \rangle$ be the canonical scalar product. Then, the variation of a scalar function $f : \mathbb{R}^n \to \mathbb{R}$ in response to a small variation $\delta\boldsymbol{x}$ about $\boldsymbol{x}$ can be approximated to the first order by

$$\delta f \approx \langle \nabla_{\boldsymbol{x}} f \ , \ \delta\boldsymbol{x} \rangle \ .$$

Consequently, due to the linearity of the scalar product, the variation of J^{o} is given by

$$\delta J^{\mathrm{o}} \approx \sum_{i=0}^{N} \langle \nabla_{\boldsymbol{x}_i} J^{\mathrm{o}} \ , \ \delta\boldsymbol{x}_i \rangle \ , \tag{2.12}$$

$$\text{where} \qquad \delta\boldsymbol{x}_i := M_i(\boldsymbol{x}_0 + \delta\boldsymbol{x}_0) - M_i(\boldsymbol{x}_0) \approx \mathbf{M}'_i \, \delta\boldsymbol{x}_0 .$$

In other words, $\delta\boldsymbol{x}_i$ is linked to the variation of the initial model values $\delta\boldsymbol{x}_0$ by the tangent linear model $\mathbf{M}'$, that is, the Jacobian of the model operator M. Using (2.12), the variation of the cost function can be expressed as

$$\delta J^{\mathrm{o}} \approx \sum_{i=0}^{N} \langle \nabla_{\boldsymbol{x}_i} J^{\mathrm{o}} \ , \ \mathbf{M}'_i \, \delta\boldsymbol{x}_0 \rangle = \sum_{i=0}^{N} \langle \mathbf{M}_i^* \nabla_{\boldsymbol{x}_i} J^{\mathrm{o}} \ , \ \delta\boldsymbol{x}_0 \rangle \ . \tag{2.13}$$

$\mathbf{M}^*$ is the *adjoint model operator*, which is the transpose of the tangent linear $\mathbf{M}'$ (see *Talagrand and Courtier* [1987] for a detailed discussion). From (2.13) and considering (2.11), it can be concluded that the gradient of J^{o} with respect to the initial model values $\boldsymbol{x}_0$ is given by

$$\nabla_{\boldsymbol{x}_0} J^{\mathrm{o}} = \sum_{i=0}^{N} \mathbf{M}_i^* \nabla_{\boldsymbol{x}_i} J^{\mathrm{o}} = \sum_{i=0}^{N} \mathbf{M}_i^* \mathbf{H}'^T \mathbf{R}^{-1} \left[H(\boldsymbol{x}_i) - \boldsymbol{y}_i \right] .$$

Hence, the complete gradient of the cost function with respect to the control variables $\boldsymbol{x}_0$ can be written as

$$\nabla_{\boldsymbol{x}_0} J = \mathbf{B}^{-1} \left[\boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}} \right] \ + \ \sum_{i=0}^{N} \mathbf{M}_i^* \mathbf{H}^T \mathbf{R}^{-1} \left[H(\boldsymbol{x}_i) - \boldsymbol{y}_i \right] . \tag{2.14}$$

2.2.1 Properties of the adjoint model

A numerical model integration over a time interval $[t_0, t_i]$ is subdivided into a number of timesteps:

$$\boldsymbol{x}_i = M_{i,i-1} \circ \cdots \circ M_{2,1} \circ M_{1,0}(\boldsymbol{x}_0)\,.$$

Consequently, the tangent linear $\mathbf{M}'_i$ of this sequence of operators is given by

$$\mathbf{M}'_i = \mathbf{M}'_{i,i-1} \cdots \mathbf{M}'_{2,1}\mathbf{M}'_{1,0}\,.$$

As the model is non-linear, each of the linearised operators $\mathbf{M}'_{l,l-1}$ explicitly depends on the current atmospheric state $\boldsymbol{x}_{l-1}$. By forming the transpose of the tangent linear in order to obtain the adjoint, the sequence of operators is reversed:

$$\mathbf{M}^*_i = \mathbf{M}^*_{1,0} \cdots \mathbf{M}^*_{i-1,i-2}\mathbf{M}^*_{i,i-1}\,.$$

Hence, the adjoint model operator $\mathbf{M}^*_i$ propagates the gradient of the cost function with respect to $\boldsymbol{x}_i$ backwards in time, to deliver the gradient of the cost function with respect to the vector of control variables $\boldsymbol{x}_0$. Note that, as each adjoint operator $\mathbf{M}^*_{l,l-1}$ depends on $\boldsymbol{x}_{l-1}$, the sequence of atmospheric states must be available in reverse order. To this end, all intermediate model states $\boldsymbol{x}_l$ must be stored for $l = 0, \cdots, i$ during a forward integration of the model M, or, alternatively, they have to be recomputed during the course of adjoint integration.

The adjoint model can be created from the computer code implementing the model M.[2] A short example should illustrate this approach: Suppose the statement `x = y**2+a*z` is given in your source code, where `a` is an arbitrary constant. This instruction can be interpreted as a function

$$F : \mathbb{R}^3 \to \mathbb{R}^3, \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} y^2 + az \\ y \\ z \end{pmatrix}\,.$$

The transpose Jacobian of F and the corresponding adjoint mapping are given by

$$\mathbf{M}^* = \begin{pmatrix} 0 & 0 & 0 \\ 2y & 1 & 0 \\ a & 0 & 1 \end{pmatrix} \quad F^* : \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} \mapsto \mathbf{M}^* \begin{pmatrix} x^* \\ y^* \\ z^* \end{pmatrix} = \begin{pmatrix} 0 \\ 2yx^* + y^* \\ ax^* + z^* \end{pmatrix}\,.$$

[2] There are other possible approaches. See *Giering and Kaminski* [1998] and references therein for details.

F^* can be implemented as the corresponding piece of adjoint code taking care that the instruction `x*` = `0` is the one to come last, of course. By examining the whole program unit in reverse order the adjoint code can be constructed statement by statement. A detailed description of this technique is given in *Giering and Kaminski* [1998]. Clearly this approach would not be suited very well for comprehensive atmospheric models, which consist of thousands of lines of code. However, the extensive and cumbersome coding work can be alleviated, because the task can be partly automated by using adjoint compilers like TAMC (*Giering* [1999]) or *Oðyssée* (*Faure and Papegay* [1998]).

CHAPTER 3

Background error covariances

Given a single observation at a certain location, an assimilation system produces an analysis, which is a compromise between the observation and the respective background value. Both pieces of information are weighted according to the error covariances, specified in the observation error covariance matrix $\mathbf{R}$ and the background error covariance matrix $\mathbf{B}$. If the background error variances were the only information provided to the system (i. e. $\mathbf{B}$ is a diagonal matrix), the analysed quantity would have a singular peak at the location of the observation, as shown in Figure 3.1a. Such an atmospheric state is usually not prohibited by physical or chemical laws, but it is known to be highly improbable. Atmospheric trace gases are subject to mixing with ambient air, for example, due to turbulent diffusion. Thus, sharp peaks in concentration are not expected to be found. In other words, the errors of the background field are spatially correlated, such that $\mathbf{B}$ is a matrix with non-diagonal entries. If an observation indicates that the background differs from the observed atmospheric state, a correction must be applied to neighboring locations as well (Figure 3.1b).

Note that any error covariance matrix $\mathbf{P}$ can be decomposed into a correlation matrix

$$\mathbf{C} = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,n} \\ \rho_{1,2} & 1 & \cdots & \rho_{2,n} \\ \vdots & \vdots & & \vdots \\ \rho_{1,n} & \rho_{2,n} & \cdots & 1 \end{pmatrix},$$

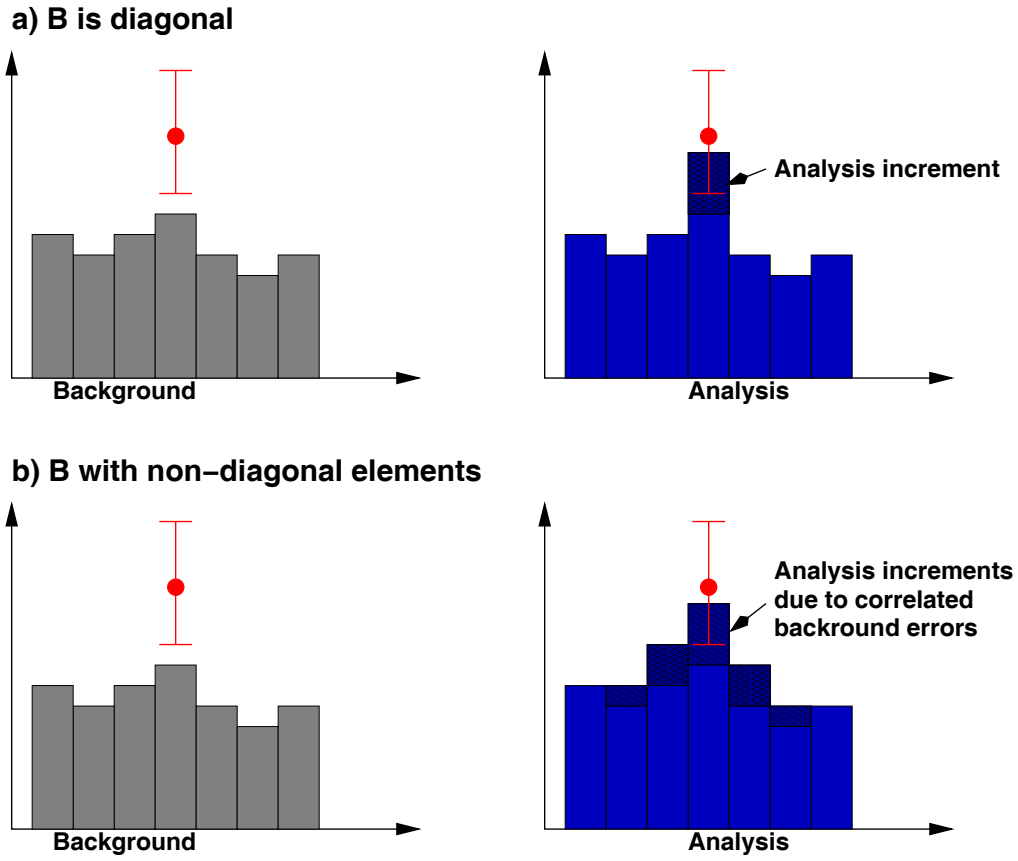

Figure 3.1: *Effect of background error covariances on the analysis. a) If there are only error variances encoded in* $\mathbf{B}$*, the analysis will show singular peaks at the location of observations. b) Known or estimated error covariances of the background field encoded in* $\mathbf{B}$ *lead to a more probable atmospheric analysis state.*

containing the error correlations $\rho_{i,j} = E\left[\,(\epsilon_i - E[\epsilon_i])(\epsilon_j - E[\epsilon_j])\,\right]/\sigma_i\sigma_j$, and diagonal matrices $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ containing the standard deviations:

$$\mathbf{P} = \mathbf{\Sigma}\,\mathbf{C}\,\mathbf{\Sigma}. \tag{3.1}$$

Consequently, the problem of constructing an error covariance matrix can be split into two separate tasks: The estimation of the variances σ for each control variable and the specification of error correlations between them.

Covariances restricted to a single model variable, for example ozone volume mixing ratios at different locations, form the *univariate* part of $\mathbf{B}$, which has a block diagonal structure. The remaining elements, the *multivariate* component of $\mathbf{B}$, represent cross-covariances between different model variables. For atmospheric trace gases however, such multivariate error covariances are

even more difficult to estimate than the univariate components, and until now no approach towards a multivariate BECM has been undertaken in this field. Furthermore, in contrast to meteorological data assimilation, where imbalances between different control variables can cause severe problems, for example, the generation of unrealistic strong gravity waves, the multivariate formulation of $\mathbf{B}$ can be assumed to be of less importance for a chemical 4D-var assimilation system. Here, the chemical equilibrium is attained after a short period of time and the analysis renders a chemically consistent atmospheric state. Thus, multivariate background error correlations are not considered in this work and $\mathbf{B}$ reduces to a block-diagonal matrix, a single block representing the background error covariances between one particular model variable.

In practice of atmospheric data assimilation, the straightforward implementation of the background error covariance matrix is hampered by two facts:

1. The number of control variables in atmospheric models is large. Even if cross-correlations between different model variables are neglected, the number of matrix elements is too large to be handled. In the SACADA assimilation system with its $32 \times 10\,242$ grid points, the resulting dimension for one model variable is $N \approx 3 \times 10^5$. The corresponding univariate BECM would be of size $(N^2 + N)/2 \approx 5 \times 10^{10}$ ($\mathbf{B}$ is symmetric). Storing a matrix of this size for each of the 41 model species would require approximately 8 Terrabytes of memory, which is still not feasible with today's available computational resources.

2. The statistical information to be encoded in the BECM is difficult to obtain. In principle, a sufficiently large set of realisations of the background field must be compared to the corresponding true atmospheric states to derive the background error statistics. Certainly, the true state is unknown and the best that may be available is an estimate thereof.

Consequently, as the background error covariances can neither be calculated exactly nor stored explicitly, it is necessary to develop a suitable parameterisation of $\mathbf{B}$, which captures the relevant features of the backround error covariances. Often it is assumed that the spatial background error correlations can be expressed by a correlation function $f(r)$ depending on the separation r between grid points only (isotropy), and that the same dependence is valid all over the globe (homogeneity). Examples for various variants of correlation functions used in data assimilation can be found in *Daley* [1991]. All of them, however, require treatment of the explicitly stored matrix $\mathbf{B}$ and do not allow for inhomogeneous, that is, location dependent, formulation.

The approach adopted for the BECM parameterisation within the SACADA assimilation system, was proposed by *Weaver and Courtier* [2001]. They express the cost function in terms of an incremental variable $\boldsymbol{v}_0$, which circumvents the need to compute the inverse $\mathbf{B}^{-1}$, as described in Section 3.1, and that provides an efficient preconditioning for the minimisation procedure. $\mathbf{B}$ is modelled using a generalised diffusion operator and a square root decomposition $\mathbf{B}^{1/2}$ and $\mathbf{B}^{T/2}$ of the BECM can be easily obtained (Section 3.2). The realisation of inhomogeneous and anisotropic correlations using a set of local coordinate stretching factors is outlined in Section 3.3.

3.1 An incremental formulation of the cost function

Since $\mathbf{B}$ is non-singular, one can introduce a new vector of control variables

$$\boldsymbol{v}_0 := \mathbf{B}^{-1/2}\,\delta\boldsymbol{x}_0 \quad \text{with} \quad \delta\boldsymbol{x}_0 := \boldsymbol{x}_0 - \boldsymbol{x}^{\mathrm{b}}, \tag{3.2}$$

where the square root of the background error covariance matrix is taken to be any square root decomposition of $\mathbf{B}$ such that $\mathbf{B} = \mathbf{B}^{1/2}\,\mathbf{B}^{T/2}$ holds (*Weaver and Courtier* [2001]). The relation of $\boldsymbol{x}_0$ to the new incremental variable is given by

$$\boldsymbol{x}_0 = \mathbf{B}^{1/2}\,\boldsymbol{v}_0 + \boldsymbol{x}^{\mathrm{b}}. \tag{3.3}$$

The cost function remains invariant under this transformation, i. e. $J(\boldsymbol{x}_0) = J(\boldsymbol{v}_0)$, while the gradient of the background cost portion simplifies to

$$\nabla_{\boldsymbol{v}_0} J^{\mathrm{b}} = \nabla_{\boldsymbol{v}_0}\left[\frac{1}{2}\,\boldsymbol{v}_0^T\,\boldsymbol{v}_0\right] = \boldsymbol{v}_0\,. \tag{3.4}$$

Calculating the gradient of the observational part of the cost function with respect to the new vector of control variable is accomplished in the same way as shown in Section 2.2, taking into account that, by definition, $\boldsymbol{v}_0$ represents a small variation around zero, with $\boldsymbol{v}_0 = 0$ corresponding to $\boldsymbol{x}_0 = \boldsymbol{x}^{\mathrm{b}}$:

$$\begin{aligned}\delta J^{\mathrm{o}} &= \langle \nabla_{\boldsymbol{x}_0} J^{\mathrm{o}}\,,\,\delta\boldsymbol{x}_0\rangle = \left\langle \nabla_{\boldsymbol{x}_0} J^{\mathrm{o}}\,,\,\mathbf{B}^{1/2}\boldsymbol{v}_0\right\rangle \\ &= \left\langle \mathbf{B}^{T/2}\nabla_{\boldsymbol{x}_0} J^{\mathrm{o}}\,,\,\boldsymbol{v}_0\right\rangle\,.\end{aligned}$$

Therefore, the gradient of J^{o} with respect to $\boldsymbol{v}_0$ is obtained by multiplying the gradient of J^{o} with respect to $\boldsymbol{x}_0$ from the left by the transposed square root of $\mathbf{B}$:

$$\nabla_{\boldsymbol{v}_0} J^{\mathrm{o}} = \mathbf{B}^{T/2}\nabla_{\boldsymbol{x}_0} J^{\mathrm{o}}\,. \tag{3.5}$$

Note that the quantity $\boldsymbol{v}_0$ is the vector of control variables and the gradient of $J^{\mathrm{b}}(\boldsymbol{v}_0)$ at the same time. Consequently, there is no need to recalculate $\nabla_{\boldsymbol{v}_0} J^{\mathrm{b}}$ by means of (3.2) for each iteration, because $\boldsymbol{v}_0$ is the output of the minimisation algorithm at the preceeding iteration. Moreover, a real world assimilation system will start the iterative process of minimising J by choosing $\boldsymbol{x}_0 = \boldsymbol{x}^{\mathrm{b}}$ as a first guess atmospheric state, resulting in $\boldsymbol{v}_0 = 0$ at the first iteration. Consequently, the transformation (3.2), and hence $\mathbf{B}^{-1/2}$, is never needed explicitly. The initial values for the next forward run are obtained from

$$\boldsymbol{x}_0 = \mathbf{B}^{1/2}\boldsymbol{v}_0 + \boldsymbol{x}^{\mathrm{b}} \tag{3.6}$$

using the improved $\boldsymbol{v}_0$ as computed by the minimisation algorithm.

3.2 Correlation modelling using a diffusion approach

As shown in the previous section, a 4D-var assimilation system using the incremental control variable $\boldsymbol{v}_0$ requires a square root decomposition

$$\mathbf{B} = \mathbf{B}^{1/2}\,\mathbf{B}^{T/2} = \boldsymbol{\Sigma}\,\mathbf{C}^{1/2}\,\mathbf{C}^{T/2}\,\boldsymbol{\Sigma}$$

instead of the inverse BECM. As it is not feasible to implement a correlation matrix $\mathbf{C}$ directly, a correlation *operator* has to be devised, which replaces the matrix-vector multiplications needed for the computation of the the cost function and its gradient $\nabla_{\boldsymbol{v}_0} J$. Furthermore, it is desirable that a square root decomposition of this operator can be obtained easily. *Weaver and Courtier* [2001] point out that the solution $\psi(z,t)$ of the one-dimensional diffusion equation

$$\frac{\partial\psi}{\partial t} - \kappa\,\frac{\partial^2\psi}{\partial z^2} = 0 \tag{3.7}$$

is given by the convolution of $\psi(z,0)$ with a Gaussian function:

$$\psi(z,t) = \frac{1}{\sqrt{4\pi\kappa t}}\int \exp\left(-\frac{(z-z')^2}{4\kappa t}\right)\psi(z',0)\,dz'.$$

It is well known that a Gaussian defines a valid correlation function, if it is normalised to the value one at the origin. Therefore, a correlation operator may be obtained by multiplying $\psi(z,t)$ by the constant factor

$$\lambda := \sqrt{4\pi\kappa t}\,. \tag{3.8}$$

The product $L^2 := 2\kappa t$ can be interpreted as the square of a correlation length-scale, meaning that the correlation between two locations z and z' decreases to a value of $\exp(-1/2)$ at a separation of L.

The same considerations apply for the solution of the two-dimensional diffusion equation

$$\frac{\partial \psi}{\partial t} - \kappa \nabla^2 \psi = 0 \tag{3.9}$$

on the sphere $S^2 := \{(x, y, z) : \sqrt{x^2 + y^2 + z^2} = a\}$ with radius a. As in the one-dimensional case, the solution $\psi(\lambda, \phi, t)$ at time t is given by a convolution of $\psi(\lambda, \phi, 0)$ with a *quasi-Gaussian* function $f(\theta, \kappa t)$, which depends on the angular separation θ between two points (λ, ϕ) and (λ', ϕ') on the sphere (see *Weaver and Courtier* [2001] for a comprehensive discussion). The term *quasi-Gaussian* refers to the fact that $f(\theta, \kappa t)$ closely approximates a Gaussian function provided that the involved length-scale is small compared to the radius of the sphere ($L \ll a$), which is the case for scales relevant in atmospheric modelling. Under this condition the correlation length-scale can, again in analogy with the one dimensional case, be approximated by

$$L^2 \approx 2\kappa t\,. \tag{3.10}$$

The essence of what has been outlined so far is that a Gaussian or quasi-Gaussian covariance operator is readily available in atmospheric models in form of a diffusion operator $\mathbf{L}$, which solves (3.7) and (3.9) numerically. *Weaver and Courtier* [2001] extend their approach to a larger class of isotropic correlation functions on the sphere, derived from the solution of a generalised diffusion equation (involving higher order spatial derivatives), which allow for negative correlations at particular angular separations θ. Backround errors of meteorological variables derived from climatologies often display such negative correlations at distances of about a few thousand kilometers (see e. g. *Daley* [1991]). However, this is not further considered in the work presented here, because very little is known yet about the spatial background error correlations within the new SACADA assimilation system and thus, the simple quasi-Gaussian assumption is a good point to start from.

The Laplacian operator, showing up in Equations 3.7 and 3.9, has the important property of being, under certain conditions, self-adjoint with respect to the scalar product $< f(\boldsymbol{x}), g(\boldsymbol{x}) >= \int fg\, d\boldsymbol{x}$ for functions $f, g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, that is,

$$< \nabla^2 f, g >=< f, \nabla^2 g > . \tag{3.11}$$

In the one dimensional case this can be seen as follows:

$$\int_a^b \left(f\frac{d^2 g}{dx^2} - g\frac{d^2 f}{dx^2} \right) dx = \left[f\frac{dg}{dx} \right]_a^b - \left[g\frac{df}{dx} \right]_a^b . \tag{3.12}$$

If –among other possibilities– the derivatives of f and g vanish at the interval boundaries a and b, the integral (3.12) equals zero and (3.11) holds. In general, the property (3.11) is a consequence of *Green's second theorem*, again under the condition that the partial derivatives of f and g normal to the boundary of D vanish.

3.2.1 Discrete formulation

In atmospheric models the one-dimensional diffusion equation (3.7) is solved along each column of grid points to represent vertical diffusion processes, while horizontal diffusion is accounted for by the solution of the two-dimensional equation (3.9) along the sphere. These horizontal and vertical diffusion operators can be utilised to devise a horizontal and a vertical correlation operator, respectively. An appropriate combination of both leads to a three-dimensional correlation model, as will be shown in the following section.

The discrete solution of the diffusion equations for a vector $\boldsymbol{\psi}$ of model variables may be obtained by

$$\boldsymbol{\psi}(t_n) = \boldsymbol{\psi}(t_{n-1}) + \kappa\Delta t\, \mathbf{D}\, \boldsymbol{\psi}(t_{n-1})\,, \tag{3.13a}$$

or alternatively, in implicit form[1]

$$\boldsymbol{\psi}(t_n) = \boldsymbol{\psi}(t_{n-1}) + \kappa\Delta t\, \mathbf{D}\, \boldsymbol{\psi}(t_n)\,. \tag{3.13b}$$

Here, $\mathbf{D}$ denotes a discrete representation of either the horizontal or vertical Laplacian operator. The corresponding discrete diffusion operator $\mathbf{L}$ which calculates the diffusion of $\boldsymbol{\psi}$ over M timesteps is given by

$$\boldsymbol{\psi}(t_M) = \mathbf{L}\, \boldsymbol{\psi}(t_0) = \{\mathbf{I} + \kappa\Delta t\, \mathbf{D}\}^M\, \boldsymbol{\psi}(t_0) \quad \text{and} \tag{3.14a}$$

$$\boldsymbol{\psi}(t_M) = \mathbf{L}\, \boldsymbol{\psi}(t_0) = \{\mathbf{I} - \kappa\Delta t\, \mathbf{D}\}^{-M}\, \boldsymbol{\psi}(t_0)\,, \tag{3.14b}$$

respectively. Note that $\mathbf{D}$, and hence $\mathbf{L}$, can be thought of as matrices allthough their realisation in terms of computer code has a different form. The implicit discretisation (3.14b) is not well suited for the horizontal diffusion operator, because the inverse of the expression $\{\mathbf{I} - \kappa\Delta t\, \mathbf{D}\}$ would be too difficult to obtain. However, the implementation for the vertical diffusion

[1]For convenience, allthough the implicit and the explicit formulation are *different* approximations to the true solution, no notational difference will be made between them in the following text. Hence, $\boldsymbol{\psi}(t_n)$ refers to any approximate solution of the diffusion equation at time t_n, which was obtained by either an implicit or an explicit scheme.

is easy (see Appendix B) and –apart from the superior numerical stability– has the appealing advantage, that the transpose diffusion operator can be simply constructed by transposing the resulting system of linear equations.

The Laplacian's property (3.11) of being self adjoint can be translated into discrete form:

$$\langle \mathbf{W}\mathbf{D}\boldsymbol{\psi}_1, \boldsymbol{\psi}_2 \rangle = \langle \boldsymbol{\psi}_1, \mathbf{W}\mathbf{D}\boldsymbol{\psi}_2 \rangle \,,$$

where $\mathbf{W} := \mathrm{diag}(\Delta s_1, \cdots, \Delta s_n)$ is a diagonal matrix containing the area (or line) elements of the corresponding grid cells, and $< \cdot, \cdot >$ now denotes the canonical scalar product. It can be concluded that

$$\mathbf{W}\mathbf{D} = \mathbf{D}^T\mathbf{W} \quad \Leftrightarrow \quad \mathbf{D} = \mathbf{W}^{-1}\mathbf{D}^T\mathbf{W} \quad \Leftrightarrow \quad \mathbf{D}^{-1} = \mathbf{W}^{-1}\mathbf{D}^{-T}\mathbf{W}, \tag{3.15}$$

and this property of $\mathbf{D}$ opens the possibility to find a square root decomposition of the discrete diffusion operator $\mathbf{L}$. Without loss of generality, it can be assumed that M denotes an even number of timesteps. Hence, $\mathbf{L}$ can be factored as follows (*Weaver and Courtier* [2001]):

$$\begin{aligned}
\mathbf{L} &= \mathbf{L}^{1/2}\mathbf{L}^{1/2} = \left\{\mathbf{I} + \kappa\Delta t\mathbf{D}\right\}^{M/2}\left\{\mathbf{I} + \kappa\Delta t\mathbf{D}\right\}^{M/2} \\
&= \left\{\mathbf{I} + \kappa\Delta t\mathbf{W}^{-1}\mathbf{D}^T\mathbf{W}\right\}^{M/2}\left\{\mathbf{I} + \kappa\Delta t\mathbf{D}\right\}^{M/2} \\
&= \mathbf{W}^{-1}\left\{\mathbf{I} + \kappa\Delta t\mathbf{D}^T\right\}^{M/2}\mathbf{W}\left\{\mathbf{I} + \kappa\Delta t\mathbf{D}\right\}^{M/2} \\
&= \mathbf{W}^{-1}\mathbf{L}^{T/2}\mathbf{W}\mathbf{L}^{1/2} \\
&= \mathbf{L}^{1/2}\mathbf{W}^{-1}\mathbf{L}^{T/2}\mathbf{W}\,.
\end{aligned} \tag{3.16}$$

Obviously, the same factorisation applies for the implicit definition (3.14b) of the diffusion operator $\mathbf{L}$.

As $\mathbf{L}$ realises the diffusion of a physical quantity averaged over the area of a particular grid cell, it will in general not be symmetric (cf. Equation (3.16)). A valid correlation operator has the fundamental property of being symmetric, reflecting the fact that $\rho_{i,j} = \rho_{j,i}$. By inspection of (3.16) it can be seen that a symmetric operator can be obtained by multiplying $\mathbf{L}$ either to the left by $\mathbf{W}$ or to the right by $\mathbf{W}^{-1}$. Here, we choose the first possibility

$$\mathbf{L}\mathbf{W}^{-1} = \mathbf{L}^{1/2}\mathbf{W}^{-1}\mathbf{L}^{T/2} = \mathbf{W}^{-1}\mathbf{L}^T, \tag{3.17}$$

in other words, the quantity to be diffused is weighted by the inverse area of the corresponding grid cells *before* applying the diffusion operator.

Another fundamental property of a correlation matrix is that the autocorrelations $\rho_{i,i}$ equal one, a property that, in analogy to (3.8), can be introduced by applying a diagonal normalisation matrix and defining

$$\begin{aligned}
\mathbf{C} &:= \boldsymbol{\Lambda}\mathbf{L}^{1/2}\mathbf{W}^{-1}\mathbf{L}^{T/2}\boldsymbol{\Lambda} = (\boldsymbol{\Lambda}\mathbf{L}^{1/2}\mathbf{W}^{-1/2})(\mathbf{W}^{-1/2}\mathbf{L}^{T/2}\boldsymbol{\Lambda}) \\
&= \mathbf{C}^{1/2}\mathbf{C}^{T/2}\,.
\end{aligned} \tag{3.18}$$

The action of the normalisation matrix can be understood by applying the correlation operator $\mathbf{C}$ to the unit vector $\boldsymbol{e}_i := (0, \cdots, 0, 1, 0, \cdots, 0)^T$ having a unit value at grid point i and zero values elsewhere. The result of the operation $\mathbf{C}\,\boldsymbol{e}_i$ is the discrete correlation function for grid point i. Without normalisation the value $\rho_{i,i}$ would clearly be less than unity because the diffusion operator $\mathbf{L}$ conserves the total 'mass' of the input field, and accordingly all other values $\rho_{i,j}$ would be too small as well. The normalisation factor (3.8) derived from the analytical solution can be applied yielding $\mathbf{\Lambda} = \mathrm{diag}(\sqrt{2\pi L^2})$. If anisotropic and/or inhomogeneous correlations are considered, as discussed below, computationally more expensive methods to calculate the normalisation matrix are needed.

3.2.2 Three-dimensional correlations

In atmospheric data assimilation it is often assumed that the background error correlations are *separable* in that they are a product of a horizontal and a vertical correlation function (see e. g. *Daley* [1991]). It is possible to construct a separable 3-dimensional diffusion operator and a square root decomposition, by combining the vertical and horizontal diffusion operator $\mathbf{L}_h\mathbf{W}_h^{-1}$ and $\mathbf{L}_v\mathbf{W}_v^{-1}$, respectively. The resulting operator remains symmetric by construction, however, a few restrictions arise, namely:

(a) κ_h must not depend on z and κ_v must not depend on λ or ϕ.

(b) The elements of $\mathbf{W}_h$ have to be constant for all levels at a given location (λ, ϕ) and the elements $\mathbf{W}_v$ must not vary horizontally.

Given the condition (a) the operators $\mathbf{L}_h$ and $\mathbf{L}_v$ commute because the horizontal and vertical Laplacian operators do ($\partial^2/\partial z^2\,\nabla_h^2 = \nabla_h^2\,\partial^2/\partial z^2$)[2]:

$$\begin{aligned}
&\{\mathbf{I} + \kappa_h \Delta t_h \mathbf{D}_h\}\ \{\mathbf{I} + \kappa_v \Delta t_v \mathbf{D}_v\}\ \boldsymbol{\psi} \\
&= \{\mathbf{I} + \kappa_v \Delta t_v \mathbf{D}_v + \kappa_h \Delta t_h \mathbf{D}_h + \kappa_h \Delta t_h \mathbf{D}_h \kappa_v \Delta t_v \mathbf{D}_v\}\ \boldsymbol{\psi} \\
&= \{\mathbf{I} + \kappa_v \Delta t_v \mathbf{D}_v + \kappa_h \Delta t_h \mathbf{D}_h + \kappa_v \Delta t_v \mathbf{D}_v \kappa_h \Delta t_h \mathbf{D}_h\}\ \boldsymbol{\psi} \\
&= \{\mathbf{I} + \kappa_v \Delta t_v \mathbf{D}_v\}\ \{\mathbf{I} + \kappa_h \Delta t_h \mathbf{D}_h\}\ \boldsymbol{\psi}\,.
\end{aligned} \tag{3.19a}$$

By repeated application of (3.19a) it follows that $\mathbf{L}_v\mathbf{L}_h = \mathbf{L}_h\mathbf{L}_v$. Obviously the same is true for $\mathbf{L}_v^{1/2}$ and $\mathbf{L}_h^{1/2}$ as well as for the transpose operators.

[2] This is only true because the dependence of ∇_h on the vertical coordinate is usually neglected in atmospheric models (see for example *Kalnay* [2003]).

Condition (b) ensures that the matrices $\mathbf{W}_v$ and $\mathbf{W}_h$ commute with the horizontal and vertical Laplacian operators, respectively:

$$\mathbf{D}_h\mathbf{W}_v^{-1} = \mathbf{W}_v^{-1}\mathbf{D}_h \quad \text{and} \quad \mathbf{D}_v\mathbf{W}_h^{-1} = \mathbf{W}_h^{-1}\mathbf{D}_v\,. \tag{3.19b}$$

Note that the first part of condition (b) is usually fulfilled in practise, as the area elements of grid cells do not depend on the vertical coordinate. If σ or hybrid coordinate systems are used, the layer thickness (in terms of pressure or distance) depends on the meteorological situation and orography. Hence, $\mathbf{W}_v$ is different from one column of grid points to the next, thereby violating condition (b) to some extent. Now, given (3.19), the symmetry of $\mathbf{L}_h\mathbf{W}_h^{-1}$ and $\mathbf{L}_v\mathbf{W}_v^{-1}$ can be inherited to a separable 3d diffusion operator

$$\mathbf{L}\mathbf{W}^{-1} := \mathbf{L}_v\mathbf{L}_h\mathbf{W}_v^{-1}\mathbf{W}_h^{-1} = \mathbf{W}_h^{-1}\mathbf{W}_v^{-1}\,\mathbf{L}_h^T\mathbf{L}_v^T = \mathbf{W}^{-1}\mathbf{L}^T\,,$$

and a square root decomposition of this operator can be obtained by

$$\begin{aligned}\mathbf{L}\mathbf{W}^{-1} &= \mathbf{L}_v\mathbf{W}_v^{-1}\,\mathbf{L}_h\mathbf{W}_h^{-1} = \mathbf{L}_v^{1/2}\mathbf{W}_v^{-1}\mathbf{L}_v^{T/2}\,\mathbf{L}_h^{1/2}\mathbf{W}_h^{-1}\mathbf{L}_h^{T/2}\\ &= \mathbf{L}_v^{1/2}\mathbf{L}_h^{1/2}\mathbf{W}^{-1}\mathbf{L}_h^{T/2}\mathbf{L}_v^{T/2}\\ &= \mathbf{L}^{1/2}\mathbf{W}^{-1}\mathbf{L}^{T/2}\,.\end{aligned} \tag{3.20}$$

Finally, the full three-dimensional background error covariance matrix in its operator form reads:

$$\mathbf{B} = (\mathbf{\Sigma}\mathbf{\Lambda}\mathbf{L}^{1/2}\mathbf{W}^{-1/2})\,(\mathbf{W}^{-1/2}\mathbf{L}^{T/2}\mathbf{\Lambda}\mathbf{\Sigma}) = \mathbf{B}^{1/2}\mathbf{B}^{T/2}. \tag{3.21}$$

The free parameters of this covariance model are the background error variances for each grid cell, which are encoded in the diagonal matrix $\mathbf{\Sigma}$, as well as a vertical and a horizontal correlation length scale L_v and L_h. The diffusion coefficients κ_h and κ_v and the diffusion times t_h and t_v are calculated according to (3.10). Then, the number of timesteps M_h, M_v together with Δt_h and Δt_v are chosen such that stability criteria for the respective schemes are fulfilled.

The conditions that allow for the application of the separable three-dimensional covariance operator are rather restrictive and only simple covariance structures can be considered. If a more elaborate estimate of background error covariances is available, a self adjoint, non separable three-dimensional operator can be constructed by applying an operator splitting approach (see Chapter 4):

$$\mathbf{L} := \{\,\{\mathbf{I} + \kappa_v\Delta t\mathbf{D}_v\}\{\mathbf{I} + \kappa_h\Delta t\mathbf{D}_h\}\,\}^M \tag{3.22}$$

Now, the more unstable of the two schemes determines the steplength Δt. Special care has to be taken, because, in contrast to the separable case, the

operator $\mathbf{LW}^{-1}$ is not symmetric by construction. This is a further limitation of the step length, as M has to be chosen greater than some minimum number of steps to ensure an approximate symmetry of the resulting covariance operator. There is no simple strategy to calculate the necessary minimum number of steps, and property (3.15) has to be tested numerically. The square root decomposition of $\mathbf{LW}^{-1}$ can then be derived according to (3.16).

3.3 Flow dependent inhomogeneous and anisotropic correlations

Allthough it is a good point to start from in practise, there is no reason to assume that the background error correlations are homogeneous and isotropic. Advanced data assimilation methods use a short-range forecast initialised with the previous analysis as a background. Consequently, the backround error covariance structure is given by the analysis error covariance of the previous analysis, advanced in time according to the model dynamics. Additionally, errors in the model also contribute to the background error covariances. Unlike Kalman filters, which, at least theoretically, take full account of error evolution, variational methods usually neither estimate the analysis error covariances nor propagate them explicitly. Hence, an alternative solution had to be devised for the SACADA system to account for this deficiency. Due to the linkage between model dynamics and backround error covariances, it is reasonable to assume, that background errors show a stronger correlation between air parcels belonging to the same air mass (*Riishøjgaard* [1998]). A way to distinguish between different air masses is due to *Ertel's potential vorticity* (PV), which is defined by

$$P = (\nabla \times \boldsymbol{v} + 2\boldsymbol{\Omega}) \cdot \nabla\theta = \boldsymbol{\zeta} \cdot \nabla\theta \,, \tag{3.23}$$

where θ is the potential temperature and $\boldsymbol{\zeta}$ the vector of absolute vorticity (see e. g. *Pichler* [1997]). Note that the potential vorticity can be approximated by

$$P \approx \boldsymbol{k} \cdot \boldsymbol{\zeta} \frac{\partial\theta}{\partial p} \,, \tag{3.24}$$

that is, the vertical component of the absolute vorticity multiplied by the vertical component of the potential temperature gradient. This is justified because for stratospheric conditions $\nabla\theta$ is nearly aligned with the vertical direction. In the lower stratosphere larger deviation may occur, but *Elbern*

et al. [1998] found, even in this region, approximation (3.24) to be accurate within 15%. It can be shown (see e. g. *Pichler* [1997]) that the potential vorticity is conserved for adiabatic and frictionless motion. Hence, it can be concluded that the chemical composition of individual air parcels is correlated along lines of constant potential vorticity, as long as there is negligible energy exchange with the environment. An example is shown in Figure 3.2, where the potential vorticity for 28 October 2003 at the 100 hPa pressure level (top panel) is shown together with an ozone analysis for the same date. To a large extend, the distribution of ozone is governed by the PV field and it is clearly desirable to devise a *flow dependent* background error covariance parameterisation. Obviously, the correlations will neither be homogeneous nor isotropic in this case.

The diffusion approach described above can be generalised to account for anisotropic and inhomogeneous background error correlations. Following *Weaver and Courtier* [2001], this is accomplished by introducing a symmetric coordinate stretching tensor $\mathbf{S}$ in the horizontal two-dimensional diffusion equation:

$$\frac{\partial \psi}{\partial t} - \kappa_h \mathrm{div}\,(\mathbf{S}\,\mathrm{grad}_h \psi) = 0\,. \tag{3.25}$$

Tensor $\mathbf{S}$ is composed of a diagonal tensor $\tilde{\mathbf{S}}$ containing stretching factors $\tilde{s}_1$ and $\tilde{s}_2$ and a rotation tensor which rotates the local coordinate system such that the stretching can be applied along the two coordinate axes of the rotated system:

$$\mathbf{S} = \mathbf{T}(\alpha)\tilde{\mathbf{S}}\,\mathbf{T}^T(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \tilde{s}_1 & 0 \\ 0 & \tilde{s}_2 \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\,. \tag{3.26}$$

The discrete three-dimensional diffusion operator takes the same form as the isotropic scheme (3.22) with the exception that the horizontal Laplacian operator $\mathbf{D}_h$ is replaced by a discrete representation of div $(\mathbf{S}\,\mathrm{grad}_h)$. The implementation of diffusion schemes for the SACADA system is described in Chapter 4 and in Appendix B.

3.3.1 The normalisation matrix

If the anisotropic diffusion operator is to be employed, no simple analytic expression for the elements λ_i of the normalisation matrix $\mathbf{\Lambda}$ can be derived. Consequently, a numerical method for the evaluation of λ_i must be applied. The straightforward method would be to directly evaluate the diagonal elements of the diffusion operator $l_i := \boldsymbol{e}_i^T \mathbf{L}\mathbf{W}^{-1}\boldsymbol{e}_i$ and setting $\lambda_i = 1/\sqrt{l_i}$. This

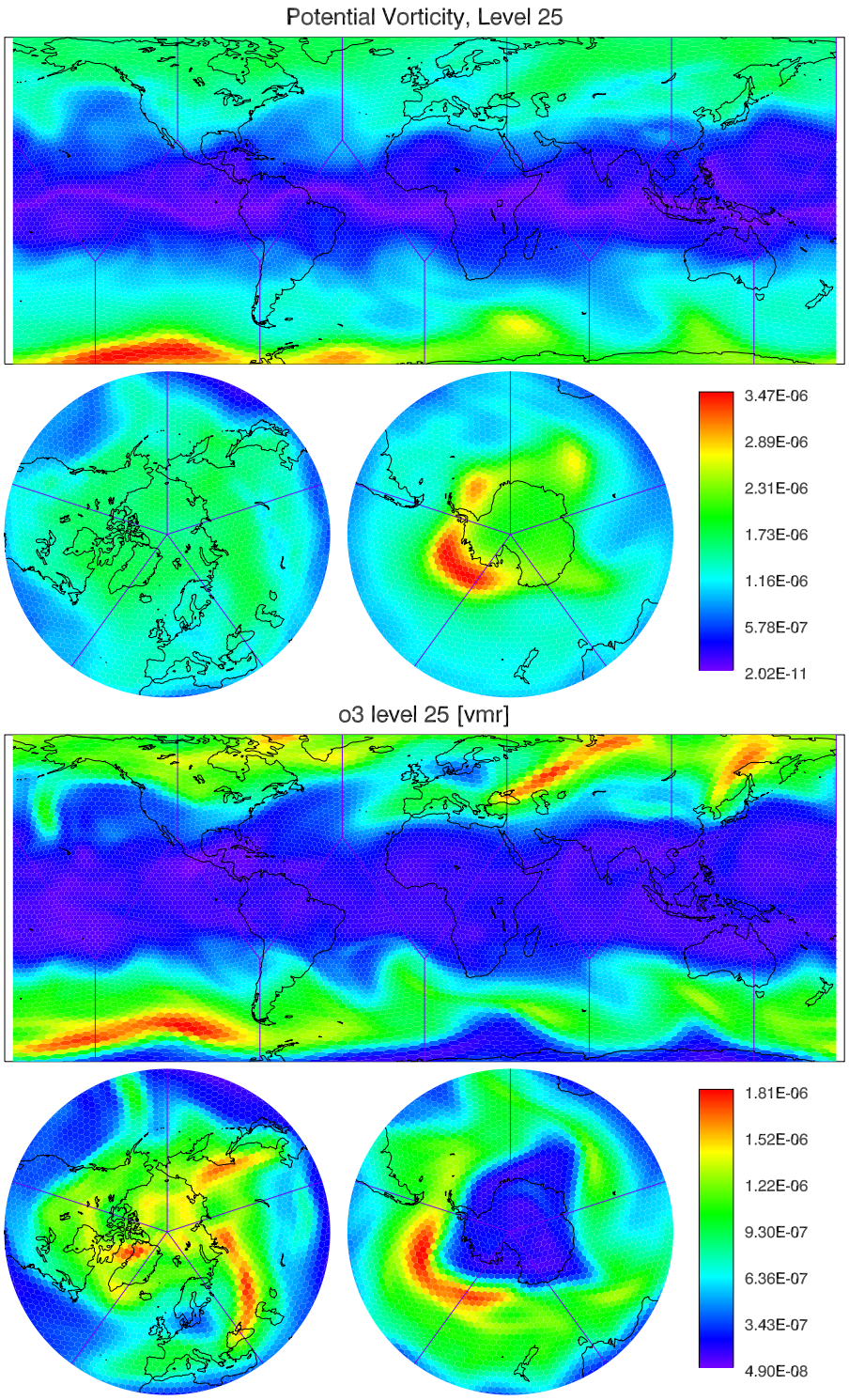

Figure 3.2: *Potential vorticity (absolute values in* $\frac{K\ m^2}{s\ kg}$*, top) and analysed ozone volume mixing ratio (bottom) fields for 28 October 2003 at the 100 hPa pressure level.*

approach, however, is computationally far too expensive in practise, as the full diffusion operator has to be applied for each grid point of the model. An alternative algorithm, which calculates an approximation of the normalisation factors, has been proposed by *Weaver and Courtier* [2001]: An ensemble of Q random vectors $\boldsymbol{v}_q$ having zero mean and unit variance is generated, and the transformed ensemble

$$\tilde{\boldsymbol{v}}_q := \mathbf{L}^{1/2}\mathbf{W}^{-1/2}\boldsymbol{v}_q$$

is calculated. Since $\mathrm{E}[\boldsymbol{v}] = 0$ and $\mathrm{E}[\boldsymbol{v}\boldsymbol{v}^T] = \mathbf{I}$, the diagonal elements $\tilde{l}_i$ may be estimated from

$$\mathbf{L}\mathbf{W}^{-1} \approx \frac{1}{Q}\sum_{q=1}^{Q}\tilde{\boldsymbol{v}}_q\tilde{\boldsymbol{v}}_q^T = \frac{1}{Q}\sum_{q=1}^{Q}\left(\mathbf{L}^{1/2}\mathbf{W}^{1/2}\boldsymbol{v}_q\right)\left(\mathbf{L}^{1/2}\mathbf{W}^{1/2}\boldsymbol{v}_q\right)^T . \qquad (3.27)$$

The standard deviation of the $\tilde{l}_i$ derived on the basis of (3.27) can be shown to be $1/\sqrt{2Q}$ if $\boldsymbol{v}$ is a Gaussian random vector (see *Weaver and Courtier* [2001] and references therein). Hence, an efficient way to approximate the normalisation matrix is given by setting $\mathbf{\Lambda} = \mathrm{diag}(1/\sqrt{\tilde{l}_1}, \cdots, 1/\sqrt{\tilde{l}_n})$.

CHAPTER 4

Description of the SACADA assimilation system

This chapter gives an overview of the main components of the SACADA system, particularly of the computer code that realises the model operator M in the cost function (Section 4.1) and the corresponding adjoint model (Section 4.2). The implementation of the background error covariance operator is described in Section 4.3. Finally, the assimilation system set-up is summarised in Section 4.4. As outlined in the introduction, a main focus was placed on a sophisticated representation of the background error covariance matrix within the assimilation system, as this can drastically improve the quality of the analyses, particularly if data are sparse. The SACADA system is designed to analyse operationally the chemical state of the stratosphere and lower mesosphere region in *near real time*. The near real time requirement implies that the system must be able to keep pace with the incoming observational data, in other words, observations that have been collected during the past 24 hours have to be processed within a timespan shorter than one day. Efficiency therefore was an important criterion taken into account during the development.

4.1 The SACADA model

A model M that is employed within a four-dimensional variational data assimilation system has to satisfy different conflicting requirements. On the

one hand the model has to be fast and efficient to allow the computationally expensive method to be applied in real time, on the other hand the model error should be kept as small as possible over the time span of the assimilation window, as M is assumed to be perfect, that is, to reproduce the time evolution of a given initial state with negligible error. As traditional Chemistry Transport Models (CTMs) are usually driven offline by external meteorological analyses, enforcing temporal and spatial interpolation, the meteorological parameters, particularly the vertical wind fields, are often poorly represented. It has therefore been decided to keep the meteorological driver *online* in the assimilation system. This approach uses a state of the art meteorological forecast model to compute the meteorological fields at the same temporal and spatial locations where they are needed by the chemistry module. Note that the SACADA model is the first model designed for 4D-var data assimilation of stratospheric constituents that combines a state of the art meteorological model and a comprehensive stratospheric chemistry module. However, the possibility to include the feedback of changing constituent concentrations on the meteorological forecast was not implemented for the work presented here. The term SACADA-CTM is used throughout this text to denote the chemistry-transport module only, which is the part of the SACADA model that computes the transport and chemical transformations of atmospheric trace gases.

A key issue concerning the efficiency of atmospheric models is the spatial discretisation used for the solution of the governing equations. Traditional grid structures of global atmospheric models are aligned to lines of equal latitude and longitude. These grids suffer from the poleward convergence of grid points, introducing unwanted high zonal resolution. Spectral methods do not show this so called *pole problem* (*Kalnay* [2003]), but the necessary transformations between spectral space and grid point space make these methods less efficient for increasing resolution. This is especially true, if costly chemistry calculations are involved, as these have to be accomplished in the grid point domain, where the poleward convergence is still present. As an alternative, the *icosahedral grid* preserves an almost isotropic grid point distribution all over the globe while, at the same time, requiring approximately 25% less grid cells than a latitude/longitude grid of comparable resolution.

The requirement of producing analyses of stratospheric trace gases in near real time is difficult to satisfy, because state of the art atmospheric models tend to absorb all available computational power, leaving few resources for the actual assimilation procedure. The use of parallel computing environments can ease this conflict. Especially models including complex chemistry encounter most computational burden at each grid point, due to the solu-

tion of the ordinary differential equations of the chemical reaction mechanism. As this solution can be obtained independently from neighboring grid points, these models can draw above-average advantage of parallel compute platforms.

4.1.1 Theoretical background

The atmospheric state, which is to be forecasted by a General Circulation Model, is determined by the wind field $\boldsymbol{v}$, temperature T, pressure p, density ρ, the mixing ratio of water in its various phases $\boldsymbol{q}$ and the mixing ratio $\boldsymbol{c}$ of trace gases. The partial differential equations (PDEs) that govern the time evolution of these quantities are given by (see e. g. *Kalnay* [2003]):[1]

$$\frac{d\boldsymbol{v}}{dt} = -\alpha\nabla p - \nabla\phi + \boldsymbol{F} - 2\boldsymbol{\Omega}\times\boldsymbol{v} \tag{4.1a}$$

$$\frac{\partial\rho}{\partial t} = -\nabla(\boldsymbol{v}\rho) \tag{4.1b}$$

$$\frac{\partial\boldsymbol{q}}{\partial t} = -\boldsymbol{v}\nabla\boldsymbol{q} + (\boldsymbol{E} - \boldsymbol{C}) \tag{4.1c}$$

$$\frac{\partial\boldsymbol{c}}{\partial t} = -\boldsymbol{v}\nabla\boldsymbol{c} + (\boldsymbol{P} - \boldsymbol{L}) \tag{4.1d}$$

$$Q = C_p\frac{dT}{dt} - \alpha\frac{dp}{dt} \tag{4.1e}$$

Here, $\alpha = 1/\rho$ is the specific volume of air, ϕ the geopotential, $\boldsymbol{F}$ represents frictional forces, and $\boldsymbol{\Omega}$ is the angular velocity of the earth. Equation (4.1c) is the continuity equation for atmospheric water with $\boldsymbol{E}$ and $\boldsymbol{C}$ being the rate of change due to phase transitions between liquid, solid and gas phase. For atmospheric trace gases the continuity equation (4.1d) applies with $\boldsymbol{P}$ and $\boldsymbol{L}$ denoting production and loss due to chemical reactions. The coupling between (4.1c) and (4.1d) and the other governing equations is via the thermodynamic energy equation (4.1e), as the rate of heating Q is strongly influenced by cloud and radiation processes which in turn are determined by the amount of liquid and solid water and the mixing ratio of absorbing trace gases. Note that for the SACADA assimilation system in its current version equations (4.1d) and (4.1e) are decoupled, because the radiation scheme uses a climatological ozone distribution. Thus equations (4.1a) - (4.1c) and (4.1e)

[1] Troughout this text no notational difference is made between the continuous atmospheric quantities, which are functions of time and three spatial coordinates and their approximate discrete counterparts. The correct interpretation will always be clear from the respective context.

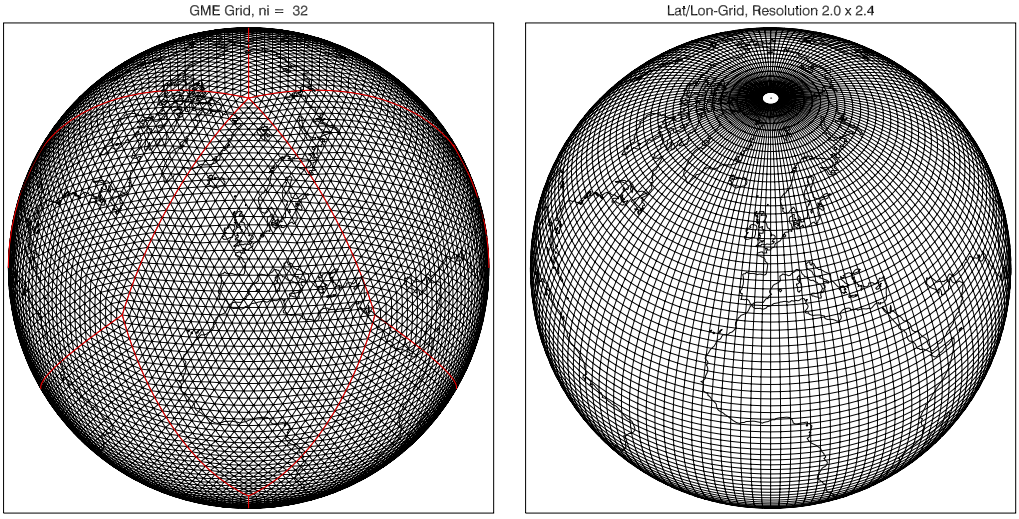

Figure 4.1: *Icosahedral grid with a resolution of $n_i = 32$ (left hand side). The distance between neighbouring grid points is about 240 km. The boundaries of diamonds (see text) are marked by red lines. A conventional latitude-longitude grid with a resolution of $2.0^o \times 2.4^o$ is shown on the right hand side for comparison.*

can independently be solved by the meteorological module followed by the integration of (4.1d) using the precomputed meteorological fields.

4.1.2 Meteorological module, icosahedral grid and parallelisation

By courtesy of the German Weather Service, the global forecast model GME, which is the only operational meteorological model utilising an icosahedral grid structure, was made available (model version 1.22) and has been implemented as the meteorological driver module of the SACADA assimilation system. Care has been taken, to introduce as little modifications as possible to the GME software, to make a later update feasible without laborious revision of the whole assimilation system. Since the new parts of the SACADA model, which have been developed in the framework of this study, adopt some of the numerical concepts, it is necessary to introduce the principal features of GME in this section. The reader may refer to *Majewski et al.* [2001] for a more detailed description of the GME model.

The icosahedral grid is constructed as follows: An icosahedron, that is, the highest Platonic body with 20 equilateral triangles, is placed into a sphere with the 12 vertices connected to theirs neighbours along the sphere. As shown in Figure 4.1, the resulting sections of great circles are equally sub-

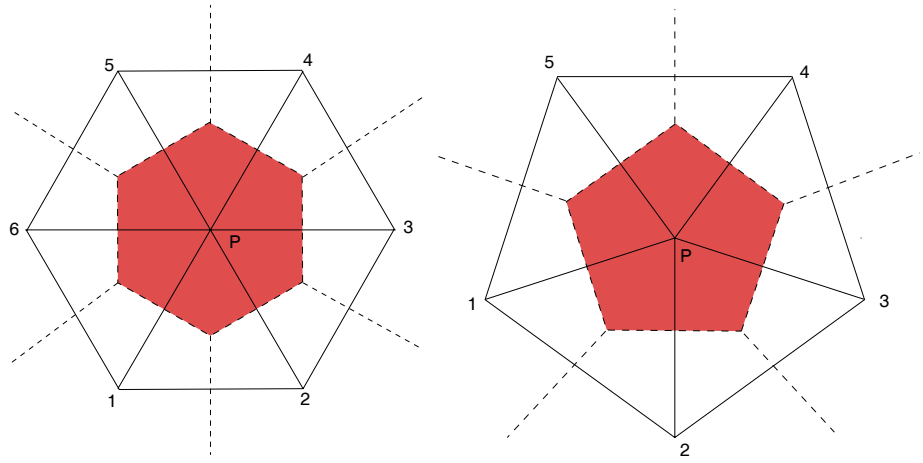

Figure 4.2: *Geometry of grid cells in the icosahedral grid. The area of representativeness assigned to each grid cell is a hexagon or a pentagon at the twelve special points, respectively.*

divided into a number of n_i intervals each, to form an almost isotropic grid. Each grid point has six nearest neighbours with the exception of the twelve points located at the vertices of the original icosahedron (called *special points* hereafter), which have only five direct neighbours. The area of representativeness for a grid cell is a hexagon and pentagon at the twelve special points, respectively (see Figure 4.2). This approach results in a mesh with virtually constant mesh size all over the globe, with the minimum and maximum separation between neighbouring grid points $\Delta_{\min}$ and $\Delta_{\max}$ varying about 20% only. For the SACADA system it has been decided that $n_i = 32$ gives a sufficient resolution, resulting in 10 242 grid points per level. The minimum and maximum distances are $\Delta_{\min} = 220$ km and $\Delta_{\max} = 263$ km, respectively, and the average area of representativeness is about 50 000 km^2. A comparable traditional latitude-longitude grid, that is, a grid with the same area represented by one grid cell at the equator (where the resolution is coarsest), requires a grid spacing of $2.0^{\circ} \times 2.4^{\circ}$ resulting in 13 500 grid points per level which is about 30% more than the icosahedral grid. To obtain a rectangular data structure, two adjacent spherical triangles are combined to form a *diamond*, partitioning the grid into ten logically rectangular sub-grid domains as marked by the thick red lines in Figure 4.1.

An orthonormal system $(\boldsymbol{x}_0, \boldsymbol{e}_\lambda, \boldsymbol{e}_\phi)$ is attached to each grid point of the GME mesh, where $\boldsymbol{x}_0$ is the position vector of the grid point on the unit sphere and $\boldsymbol{e}_\lambda$, $\boldsymbol{e}_\phi$ are unit vectors aligned with the east and north direction, respectively. Local spherical coordinates (η, χ) are used to describe the position of an arbitrary unit vector $\boldsymbol{x}$ relative to $(\boldsymbol{x}_0, \boldsymbol{e}_\lambda, \boldsymbol{e}_\phi)$, as illustrated in Figure 4.3.

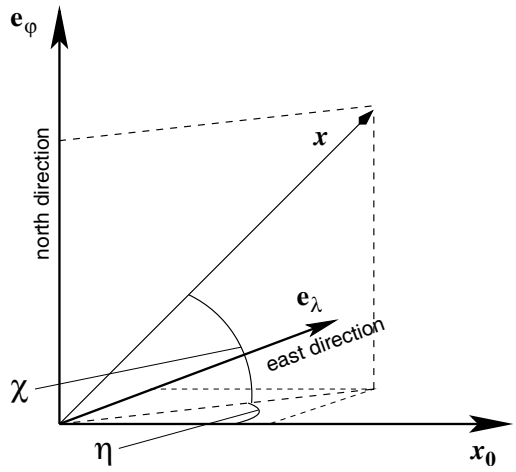

Figure 4.3: *Definition of local coordinates (η, χ) relative to the orthonormal unit vectors $\boldsymbol{x}_0, \boldsymbol{e}_\lambda, \boldsymbol{e}_\phi$ attached to each grid point of the GME mesh.*

Hence, grid points are located at the intersection of the 'equator' ($\chi = 0$) and the 'zero meridian' ($\eta = 0$) of their own spherical system. Thereby, the singularities, which are present at the poles of spherical coordinate systems, are removed from the neighbourhood of each grid point. The zonal wind component u and the meridional wind component v are given by

$$u = R_E \cos\chi \frac{d\eta}{dt} \qquad\qquad v = R_E \frac{d\chi}{dt}\,, \tag{4.2}$$

where R_E is the radius of earth. The GME grid is horizontally unstaggered, i. e. all variables are defined at the centre of each grid cell (Arakawa A-grid). The meteorological equations are formulated and solved in the (η, χ) local coordinate system. To this end the first and second order spatial derivatives with respect to the local coordinates, which are needed for the gradient and Laplacian operators, must be constructed. A discrete representation of these derivatives for an arbitrary scalar function ψ, which takes into account the six (five) nearest neighbours of a grid point, is given by

$$\begin{aligned}
\left.\frac{\partial\psi}{\partial\eta}\right|_{\psi_0} &\approx \sum_{i=1}^{6} G_{\eta,i}(\psi_i - \psi_0) \qquad & \left.\frac{\partial\psi}{\partial\chi}\right|_{\psi_0} &\approx \sum_{i=1}^{6} G_{\chi,i}(\psi_i - \psi_0) \\
\left.\frac{\partial^2\psi}{\partial\eta^2}\right|_{\psi_0} &\approx \sum_{i=1}^{6} L_{\eta,i}(\psi_i - \psi_0) \qquad & \left.\frac{\partial^2\psi}{\partial\chi^2}\right|_{\psi_0} &\approx \sum_{i=1}^{6} L_{\chi,i}(\psi_i - \psi_0) \\
\left.\frac{\partial^2\psi}{\partial\eta\partial\chi}\right|_{\psi_0} &\approx \sum_{i=1}^{6} M_i(\psi_i - \psi_0)\,.
\end{aligned} \tag{4.3}$$

The coefficients $G_{\eta,i}$, $G_{\chi,i}$, $L_{\eta,i}$ $L_{\chi,i}$ and M_i depend on the relative position of grid points only and are computed during the grid generation process of

the GME software as described by *Majewski et al.* [2001]. Note that for the twelve special points the coefficients for $i = 6$ are set to zero. Components of vector valued quantities like the zonal or meridional wind u and v have to be rotated into the local coordinate system of the centre node before the spatial derivatives (4.3) are formed. Time derivatives of prognostic fields are discretised using a three time level leapfrog scheme, which reads

$$\frac{\psi(t_{n+1}) - \bar{\psi}(t_{n-1})}{2\Delta t} = F(\psi(t_n)) \tag{4.4}$$

for an arbitrary prognostic variable ψ. To avoid separation of time levels, a phenomenon typically occurring when using leapfrog time stepping schemes (*Kalnay* [2003]), an Asselin time filter is applied after each time step:

$$\bar{\psi}(t_n) = \psi(t_n) + \alpha \left[\psi(t_{n+1}) - 2\psi(t_n) + \bar{\psi}(t_{n-1})\right] . \tag{4.5}$$

The GME uses two different algorithms for horizontal advection. Cloud water and cloud ice are advected by a semi-Lagrangeian scheme (*Staniforth and Côté* [1991], *Kalnay* [2003]) as monotonicity and positive definiteness is particularly easy to obtain with a scheme of this kind, while advection of all other prognostic variables is done by an Eulerian scheme (*Majewski et al.* [2001]). The departure points of the semi-Lagrange trajectories are found by a two step iterative procedure. After temporal discretisation (4.2) gives

$$\Delta\eta_1 = -\frac{u(t_n)}{R_E}\Delta t \qquad \Delta\chi_1 = -\frac{v(t_n)}{R_E}\Delta t\,,$$

taking into account $\cos\chi = 1$, because the trajectory starts at the grid cell centre with $(\eta, \chi) = 0$. The trajectory midpoint is computed at a second and final iteration

$$\Delta\eta_2 = -\frac{I_l(\boldsymbol{u}(t_n))}{R_E \cos\Delta\chi_1}\Delta t \qquad \Delta\chi_2 = -\frac{I_l(\boldsymbol{v}(t_n))}{R_E}\Delta t\,, \tag{4.6}$$

where I is an operator that interpolates the wind fields to the position $(\Delta\eta_1, \Delta\chi_1)$. The departure point of the trajectory is then approximated by $(2\Delta\eta_2, 2\Delta\chi_2)$. Note that the wind field at time level n is used to compute the semi-Lagrangian trajectory connecting the grid point at time $n+1$ with the departure point at time level $n-1$. Once the departure point has been found, the new value of the advected quantity is found by interpolating the values corresponding to time level $n-1$ from surrounding grid points to the departure point. Two different interpolation operators are available: A linear operator I_l, taking into account the three nearest neighbours of the

departure point, and an operator I_q performing quadratic interpolation using the values at twelve neighbouring grid points (see *Majewski et al.* [2001] for details). The interpolation of wind fields in (4.6) is performed using the linear operator I_l.

In the vertical direction the GME utilises a hybrid σ-pressure coordinate system and an advection scheme that was proposed by *Simmons and Burridge* [1981]. The pressure for each level is given by a set of two parameters

$$\begin{aligned} p_{f,k} &= a_{f,k} + b_{f,k}\, p_s \quad \text{for} \quad k = l_s, \ldots, l_e \\ p_k &= a_k \quad + b_k\, p_s \quad \text{for} \quad k = l_s, \ldots, l_e + 1\,, \end{aligned} \tag{4.7}$$

where p_f is the pressure at the centre of each layer, p the pressure at the layer boundaries and p_s is the surface pressure. The coefficients a_k, b_k and $a_{f,k}$, $b_{f,k}$ determine the vertical structure of the grid. A staggered (Charney-Phillips) grid is used with the geopotential and the vertical wind specified at the boundaries of layers. Following the focus being placed on the stratosphere, the number of model layers has been increased from 32 to 42 for the SACADA assimilation system compared to the operational GME (version 1.22), and the top level pressure has been reduced from 10.0 hPa to 0.1 hPa. The coefficients for the vertical grid are listedt in Appendix C, Table C.1; the resulting pressure values together with the corresponding heights are shown in Figure 4.4.

To facilitate the use of the model on parallel computers a diamond-wise domain decomposition is performed as shown in Figure 4.5. Each processor works on one portion of each diamond. This is a simple yet effective strategy to achieve a good load balancing between processors. Each domain has a *halo* of two rows and columns of grid points that have to be exchanged among processors. To this end the MPICH library (*Gropp et al.* [1996]), which implements the **M**essage **P**assing **I**nterface (MPI) standard is used. On a parallel computer system representing the state of technology at the time of writing, the assimilation of observational data collected during 24 hours can be accomplished in less than four hours, provided twelve or more processors are available (see Figure 4.5).

4.1.3 Chemistry transport module

The SACADA chemistry transport module solves equation (4.1d), which reads, in the local coordinate system of the icosahedral grid

$$\frac{\partial \boldsymbol{c}}{\partial t} = -\frac{u}{R_E}\frac{\partial \boldsymbol{c}}{\partial \eta} - \frac{v}{R_E}\frac{\partial \boldsymbol{c}}{\partial \chi} - w\frac{\partial \boldsymbol{c}}{\partial p} + (\boldsymbol{P} - \boldsymbol{L})\,. \tag{4.8}$$

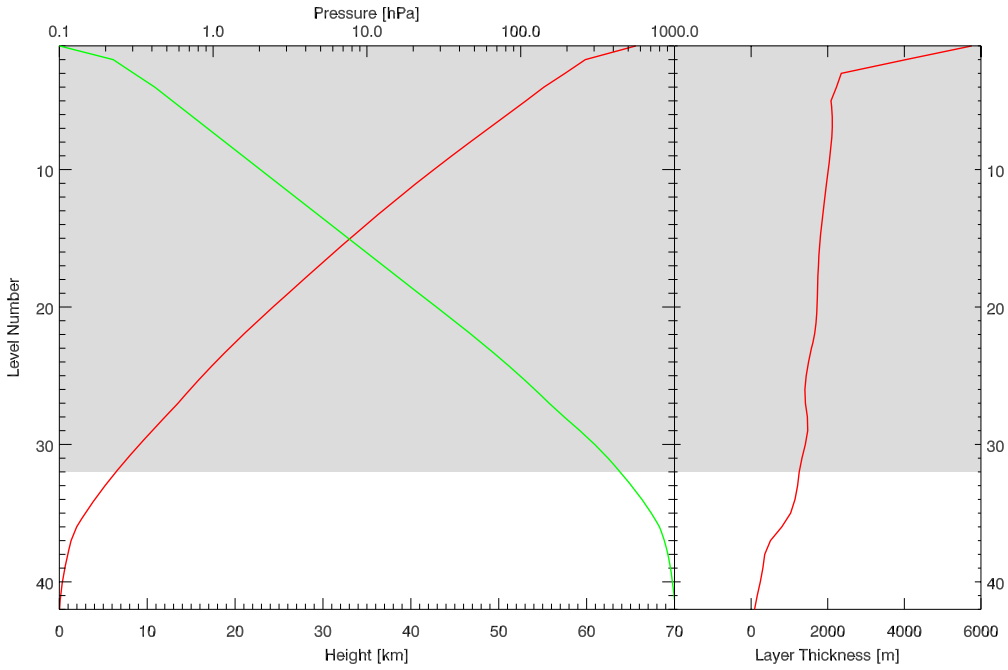

Figure 4.4: *Pressure (top scale, green line) and height (bottom scale, red line) of SACADA model layers. On the right hand side the resulting layer thickness is shown. Values are valid for standard sea level surface pressure (1013.25 hPa) and heights have been computed from pressure according to the U.S. Standard Atmosphere from 1976. Chemistry computations and consequently the assimilation of trace gas observations is restricted to the grey shaded region in the work presented here.*

The discrete approximate solution of this partial differential equation is obtained using an operator splitting approach (*McRae et al.* [1982]), which is common practice in atmospheric modelling. The PDE (4.8) is split into subproblems which describe the rate of change in volume mixing ratio due to horizontal advection, vertical advection and chemical production or loss, denoted by the superscripts h, v and c respectively:

$$\left(\frac{\partial \boldsymbol{c}}{\partial t}\right)^{h} = -\frac{u}{R_E}\frac{\partial \boldsymbol{c}}{\partial \eta} - \frac{v}{R_E}\frac{\partial \boldsymbol{c}}{\partial \chi} \tag{4.9a}$$

$$\left(\frac{\partial \boldsymbol{c}}{\partial t}\right)^{v} = -w\frac{\partial \boldsymbol{c}}{\partial p} \tag{4.9b}$$

$$\left(\frac{\partial \boldsymbol{c}}{\partial t}\right)^{c} = (\boldsymbol{P} - \boldsymbol{L})\,. \tag{4.9c}$$

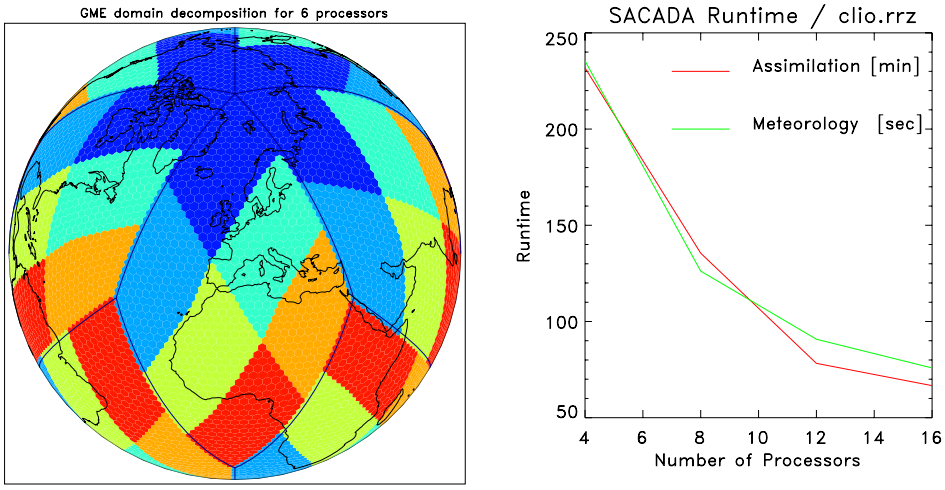

Figure 4.5: Left hand side: *Domain decomposition for six processors. Each colour indicates a region that is assigned to one processor.* Right hand side: *Run time for the meteorological module (GME) and the data assimilation part of the SACADA software (four iterations + analysis run) on a SUN PC-Cluster with AMD-Opteron processors and INFINIBAND network. Note that the GME run time is in seconds.*

These subproblems are sequentially solved by dedicated algorithms. Each scheme may be explicit or implicit and a different number of intermediate time steps may be chosen. This is of particular importance for the operator that treats the chemical reactions, because the differential equations describing this process are know to be stiff and an elaborate, preferably implicit scheme with an adaptive step size control should be employed. Let M^h, M^v and M^c be the generally non-linear discrete operators that solve (4.9a) – (4.9c). Then, an approximate solution of (4.8) may be obtained by

$$\boldsymbol{c}(t_{n+1}) = \left[M^c(t_n) \circ M^v(t_n) \circ M^h(t_n)\right] \boldsymbol{c}(t_{n-1})\,. \tag{4.10}$$

Here, n is the time step index as used by the meteorological module, and the meteorological data at time t_n is used to compute the new volume mixing ratio $\boldsymbol{c}(t_{n+1})$ from $\boldsymbol{c}(t_{n-1})$ using the same leapfrog time stepping as the meteorological module (see equation (4.4)). Note however, that in the case of the SACADA-CTM all operators at the right hand side of equation (4.10) are two time level schemes as they do *not* depend on $\boldsymbol{c}(t_n)$. Consequently, there is no need to compute this quantity. Hence, equation (4.10) describes an ordinary forward time stepping using twice the meteorological time step and the meteorological data at the centre of the time interval. This is advantageous as the outer time step of the chemistry-transport module is effectively

doubled in comparison to the meteorological time step and only two time levels of the field of chemical constituents have to be stored.

Note that a symmetric operator splitting $M^{h/2} \circ M^{v/2} \circ M^c \circ M^{v/2} \circ M^{h/2}$ would reduce the splitting error at only little additional computational cost. The storage and recomputation strategy for the adjoint model (see Section 4.2), however, becomes significantly more demanding in this case. For the SACADA assimilation system the simpler and more efficient scheme (4.10) is considered to be sufficient.

Horizontal and vertical advection

Semi-Lagrangian algorithms are very efficient for models with complex chemistry, where about 50 different species are to be transported, as the departure points have to be computed only once. For this reason, the semi-Lagrangian scheme from GME has been adopted as the horizontal transport scheme for the CTM. Departure points are computed according to (4.6) and saved together with the meteorological data during the GME run. Consequently, the operator M^h is remarkably simple:

$$\boldsymbol{c}(t_{n+1}) = M^h\, \boldsymbol{c}(t_{n-1}) = I_q(\boldsymbol{c}(t_{n-1}))\,. \tag{4.11}$$

The interpolation operator I_q performs a quadratic interpolation using the twelve nearest grid points surrounding the departure point. Optionally the operator is available in a positive definite and a monotonic form. To suppress negative volume mixing ratios, the SACADA-CTM employs the operator in its positive definite form. It should be noted that the semi-Lagrangian scheme is not conservative, as no measures are taken to ensure that the total mass remains constant before and after one advection step. While this would be a severe deficiency for long term integrations, as necessary for climate modelling, this is fully acceptable in the context of 4D-var data assimilation. Since the total mass of constituents is continuously updated from day to day by the assimilation procedure, a comparatively small gain or loss due to the transport scheme does not degrade the quality of analyses.

The vertical transport equation is solved by means of a simple and efficient implicit upwind algorithm:

$$\boldsymbol{c}(t_{n+1}) = M^v\, \boldsymbol{c}(t_{n-1}) = \left[\mathbf{I} + 2\Delta t \mathbf{A}\right]^{-1}\, \boldsymbol{c}(t_{n-1})\,. \tag{4.12}$$

Matrix $\mathbf{A}$ depends on the vertical wind given in the hybrid vertical coordinate system and the pressure differences between adjacent layers. As $\mathbf{A}$ is tridiagonal, the scheme can be implemented such that a tridiagonal linear system has to be solved for each column of grid points. Details can be

found in Appendix A. The scheme is unconditionally stable because of its implicit time stepping, thus no constraints regarding the magnitude of Δt are imposed. Furthermore the scheme is *linear*, which has the advantage that the adjoint operator $\mathbf{M}^{v*}$ does not depend on $\boldsymbol{c}(t_{n-1})$ and, consequently, this quantity has neither to be saved during the forward model integration nor to be recomputed during the course of adjoint computations. The drawback, however, is that the scheme shows a relatively strong numerical diffusion, a fact that can be problematic in regions where sharp vertical gradients of volume mixing ratios occur. In the tropopause region, where the mixing ratio of H_2O drops by several orders of magnitude, the scheme may show artificial 'transport' of water vapour into the lower stratosphere.

The vertical wind at the surface and at the top of the model domain are zero. Therefore, no boundary conditions for equation (4.9b) have to be provided. However, the SACADA assimilation system offers the flexibility to restrict the vertical domain of the chemistry-transport module. As the focus of the system is the stratosphere and, consequently, the chemistry-scheme has been developed to represent stratospheric processes, computational effort may be saved by excluding the middle and lower troposphere from the chemistry-transport module. In this case the question of how to choose the boundary conditions for the vertical advection scheme arises. Dirichlet conditions could be used by specifying fixed (e. g. climatological) volume mixing ratios at the boundary, but these climatological values are difficult to obtain for at least the most important among the 41 constituents that are treated within the chemistry scheme. In the case of a 'free' model run (a model integration without data assimilation) over an extended period of time, this would be the mandatory solution, because the boundary values would have a considerable influence on the whole model domain. This also implies that the model skill would depend partly on the quality of the chosen boundary values. If the model, however, is assimilating observational data, which is sufficiently dense in space and time, the impact of the boundary values becomes negligible. Hence, the need for climatological boundary values can be circumvented in this case by using von Neumann boundary conditions and prescribing that the spatial derivative of $\boldsymbol{c}$ normal to the boundary vanishes. The latter concept has been adopted for the SACADA assimilation system and details on the implementation can be found in Appendix A.

Chemistry scheme and solver

The set of reactions that are included in the SACADA chemistry-transport module comprises 148 gas phase and 7 heterogeneous reactions on surfaces of

Polar Stratospheric Cloud (PSC) particles and in sulphate aerosol droplets. The treatment of these solid and liquid stratospheric particles is discussed below. For a detailed description of the reaction mechanism, the reader may refer to *Hendricks et al.* [2001]. The reaction equations together with their rate constants are listed in Tables C.2 and C.3. An evaluation of this scheme in comparison with other state of the art mechanisms has been presented by *Krämer et al.* [2003]. The current implementation uses updated values for gas phase reaction rates taken from *Sander et al.* [2003].

A second order Rosenbrock method is applied to solve the chemistry subproblem (4.9c). It is a two stage linear-implicit scheme, which reads for an arbitrary autonomous differential equation $d\boldsymbol{x}/dt = f(\boldsymbol{x})$ with $f : \mathbb{R}^m \to \mathbb{R}^m$:

$$\begin{aligned} \boldsymbol{x}(t+\tau) &= \boldsymbol{x}(t) + \frac{3}{2}\tau\boldsymbol{k}_1 + \frac{1}{2}\tau\boldsymbol{k}_2 \\ (\mathbf{I} - \gamma\tau\mathbf{J})\boldsymbol{k}_1 &= f(\boldsymbol{x}(t)) \\ (\mathbf{I} - \gamma\tau\mathbf{J})\boldsymbol{k}_2 &= f(\boldsymbol{x}(t) + \tau\boldsymbol{k}_1) - 2\boldsymbol{k}_1\,, \end{aligned} \tag{4.13}$$

where $\gamma = 1 + 1/\sqrt{2}$ and τ is the step length. Details on this solver and its properties can be found in *Verwer et al.* [1997]. In contrast to fully implicit schemes, which require the solution of a non-linear system of equations at each stage, the Rosenbrock stages $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ can be computed by solving a system of linear equations involving the Jacobian $\mathbf{J} = \partial f/\partial\boldsymbol{x}$ at $\boldsymbol{x}(t)$. If appropriate measures are taken and the sparsity of $\mathbf{J}$ is exploited, Rosenbrock methods outperform other types of implicit solvers (*Sandu et al.* [1997a], *Sandu et al.* [1997b]). The KPP chemical solver tool (*Sandu et al.* [2003]) was employed to implement (4.13) for $f := \boldsymbol{P} - \boldsymbol{L}$ including a sparse representation of $\mathbf{J}$. Generally, a number of adaptive intermediate time steps has to be taken when propagating the volume mixing ratio $\boldsymbol{c}(t_{n-1})$ over the time span $2\Delta t$ to yield $\boldsymbol{c}(t_{n+1})$. As the meteorological parameters (pressure and temperature) as well as the photolysis rates are kept fixed at their values for t_n, the production and loss rates $\boldsymbol{P}$ and $\boldsymbol{L}$ are not explicitly time-dependent and the autonomous scheme (4.13) can be used. Observe that $\tilde{\boldsymbol{x}}(t+\tau) := \boldsymbol{x}(t) + \tau\boldsymbol{k}_1$ is a first order consistent approximation of the true solution and can be employed to provide a cheap estimate of the local truncation error for the adaptive step size control (see e. g. *Stoer and Bulirsch* [1990]):

$$\tau_{\text{next}} = \alpha\,\tau\,\left(\frac{1}{m}\sum_{l=1}^{m}\left[\frac{\tilde{x}_l(t+\tau) - x_l(t+\tau)}{\epsilon_a + \epsilon_b\,|x_l(t+\tau)|}\right]^2\right)^{-1/2} \tag{4.14}$$

The factor $\alpha = 0.9$ is introduced to avoid overoptimistic estimates of the new stepsize and ϵ_a, ϵ_b are some prescribed absolute and relative error tolerances.

If the expression in brackets on the right hand side of (4.14) is greater than one, the current step is rejected and repeated with the reduced stepsize τ_{next}.

Stratospheric particles

Three kinds of stratospheric particles are included in the SACADA chemistry-transport module, namely sulfate aerosol droplets, solid nitric acid trihydrate particles (NAT, type Ia PSC) and water-ice (type II PSC). In contrast to sulfate aerosol particles, which are present throughout the lower and middle stratosphere, occurrence of PSC particles is restricted to the polar winter stratosphere, where temperature can reach extremely low values, especially over the Antarctic region. Heterogeneous reactions take place at the surfaces of or, in the case of liquid droplets, within these particles. The reaction rates are determined by

$$k = \frac{1}{4}\gamma(r, T, c_{\mathrm{H_2O}}, c_{\mathrm{HCl}})\,\bar{v}NS\,, \tag{4.15}$$

where N is the number density and $\bar{v}$ the mean molecular velocity of the reacting species in the gas phase. S is the surface area density of particles in units of m^2/m^3. The factor γ is called *uptake coefficient* and can be interpreted as the fraction of molecule-particle collisions, that lead to an irreversible uptake of molecules from the gas phase into the solid/liquid phase. The uptake coefficient generally is a complicated function of particle radius, temperature and, for certain reactions, the gas phase concentrations of H_2O and HCl. The set of heterogenous reaction considered within the SACADA CTM is listed together with the respective uptake coefficients in Table C.4. A comprehensive discussion of the treatment of heterogenous reactions can be found in *Hendricks* [1997], *Hendricks et al.* [1999] and references therein. As a consequence of (4.15), the surface area density S and some representative radius r of stratospheric particles must be known in order to calculate the rates of heterogeneous reactions.

Since circa 1997 up to the present time, the stratospheric aerosol layer can be considered as undisturbed by the influence of major volcanic eruptions[2] (*Thomason et al.* [1997], *Deshler et al.* [2003]). Hence, it was decided that a single reference profile of aerosol surface area and median radius was sufficient for uptake coefficient calculations. This reference profile, as shown in Figure 4.6, was calculated using data from 20 balloon-borne measurements made between 1997 and 2002 at Laramie, Wyoming (*Deshler et al.* [2003]).

[2]The eruption of Mt Pinatubo in 1991 was the last one with considerable impact on stratospheric aerosol.

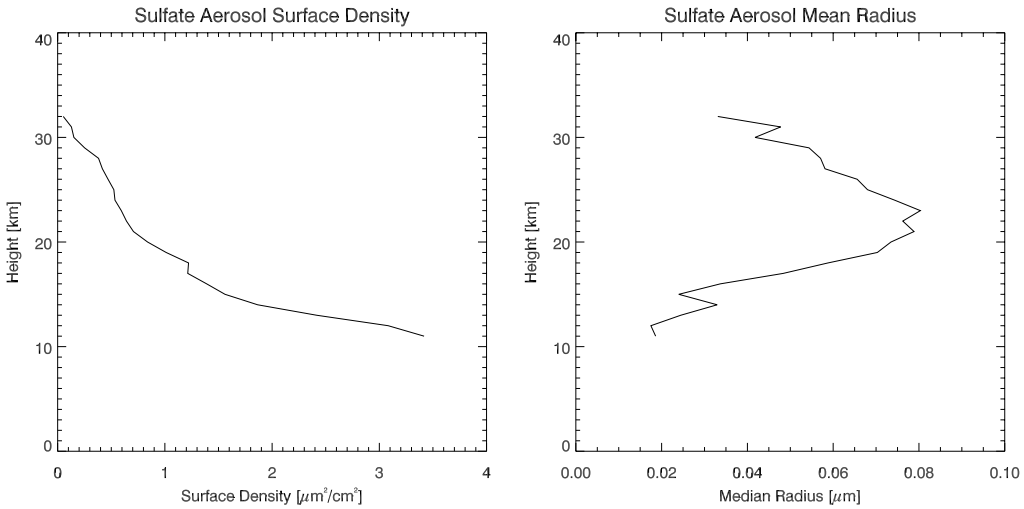

Figure 4.6: *Reference profiles of sulfate aerosol surface area density (left) and median particle radius (right), derived for undisturbed stratospheric conditions from balloon-borne measurements over Laramie, Wyoming between 1997 and 2002.*

Surface area densities and radii for PSC particles are calculated using thermodynamic equilibrium constraints for H_2O and HNO_3 over ice and NAT surfaces, as specified by *Marti and Mauersberger* [1993] and *Hanson and Mauersberger* [1988]. Following *Hendricks et al.* [2001], a constant number density $N_{\mathrm{ice}} = 1\ \mathrm{cm}^{-3}$ and $N_{\mathrm{NAT}} = 0.01\ \mathrm{cm}^{-3}$ for ice and NAT particles is assumed. If the number of H_2O or HNO_3 molecules is larger/smaller than predetermined by the saturation vapour pressure, the appropriate number of molecules is condensed/evaporated. Surface area and radius are inferred using N_{ice} and N_{NAT}. It should be noted that the formation of PSC particles is still not very well understood (see *World Meteorological Organization* [1999] and references therein). Thermodynamic considerations allow for ice and NAT particles below a threshold temperature of $T_{\mathrm{ice}} \approx 189$ K and $T_{\mathrm{NAT}} \approx 195$ K under stratospheric conditions. However, observations indicate that a substantial supercooling may occur before the onset of PSC formation. The SACADA PSC-module therefore uses reduced temperature thresholds $T_{\mathrm{ice,s}} = 187$ K and $T_{\mathrm{NAT,s}} = 191$ K. Furthermore a time constant $\tau = 5$ hours was introduced, such that only a fraction $\Delta t/\tau$ of the theoretically possible number of molecules may condense or evaporate per time interval Δt.

4.2 The adjoint model

The adjoint model operator $\mathbf{M}^*$, which is needed for the computation of the gradient $\nabla_{\boldsymbol{c}_0} J$ of the cost function with respect to the initial volume mixing ratios $\boldsymbol{c}_0$ (compare Equation 2.14), has been constructed starting from the forward code, detailed above. The transpose of the tangent-linear (or Jacobian) of the operators M^h and M^c is obtained by forming the tangent-linear of each individual line of code and transposing it, as described in Section 2.2. In the case of the linear operator M^v, the adjoint is particularly easy to obtain, as the Jacobian

$$\frac{\partial M^v}{\partial \boldsymbol{c}} = [\mathbf{I} + 2\Delta t \mathbf{A}]^{-1}$$

is the operator itself and hence,

$$\mathbf{M}^{v*} = [\mathbf{I} + 2\Delta t \mathbf{A}]^{-T} \ .$$

Note that, in contrast to $\mathbf{M}^{v*}$, $\mathbf{M}^{c*}$ and $\mathbf{M}^{h*}$ involve the recomputation of required variables starting from the volume mixing ratios values *before* the respective forward operator was applied. These values are saved to disk at each time step as shown in Figure 4.7. During the course of adjoint model integration, the gradient due to observations $\boldsymbol{y}_{n+1}$, which are available within the time interval $[t_n, t_{n+2}]$, is added to the adjoint variable $\tilde{\boldsymbol{c}}^*(t_{n+1})$, which then is propagated backward in time by means of the adjoint model:

$$\begin{aligned} \boldsymbol{c}^*(t_{n+1}) &= \tilde{\boldsymbol{c}}^*(t_{n+1}) + \mathbf{H}^T \mathbf{R}^{-1} [\boldsymbol{y}_{n+1} - \mathbf{H}\boldsymbol{c}(t_{n+1})] \\ \boldsymbol{c}^*(t_{n-1}) &= [\mathbf{M}^{h*} \circ \mathbf{M}^{v*} \circ \mathbf{M}^{c*}]\, \boldsymbol{c}^*(t_{n+1}) \end{aligned} \tag{4.16}$$

Finally, at $n - 1 = 0$, the gradient of J^o with respect to the initial volume mixing ratios $\boldsymbol{c}(t_0)$ has been obtained:

$$\nabla_{\boldsymbol{c}_0} J^o = \boldsymbol{c}^*(t_0) \, .$$

4.3 BECM implementation

Two background error covariance parameterisations, an isotropic and a generalised scheme based on (3.25) have been developed for the SACADA system. The core of both schemes is the diffusion operator

$$\mathbf{L} = \left\{ \{\mathbf{I} - \kappa_v \Delta t \mathbf{D}_v\}^{-1} \{\mathbf{I} + \kappa_h \Delta t \mathbf{D}_h\} \right\}^M \tag{4.17}$$

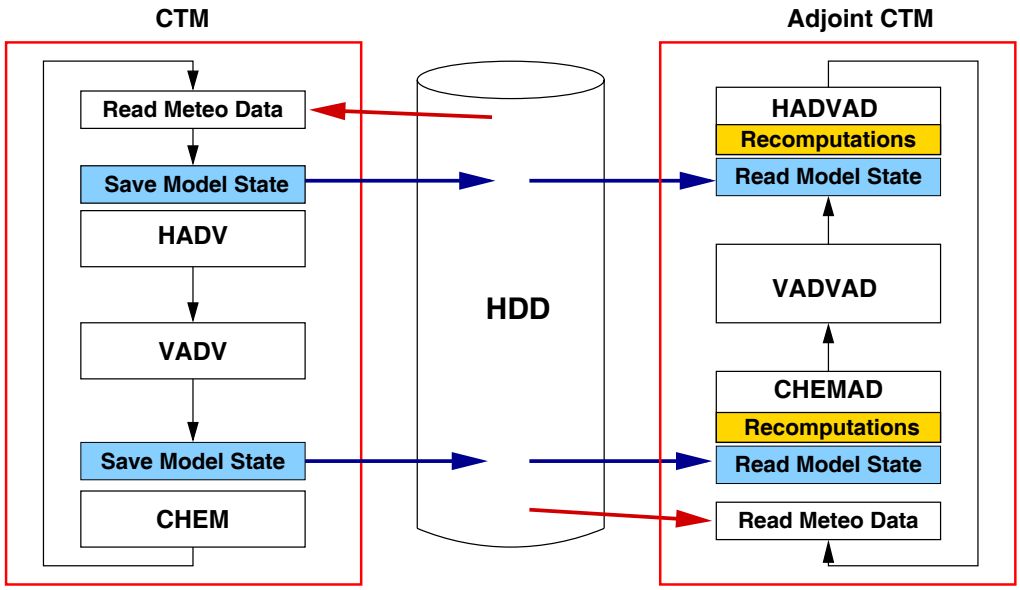

Figure 4.7: *Storage strategy of the SACADA assimilation system. Volume mixing ratios are saved to disk before the horizontal advection and chemistry routines are called. These model states serve as a starting point for the recomputation of required variables during the course of adjoint integration. Note that the vertical advection scheme is linear and does not require any recomputations.*

utilising an implicit vertical diffusion algorithm. The only difference between the two schemes is the formulation of the horizontal Laplacian operator $\mathbf{D}_h$. The practical implementation, including the derivation of stability criteria for the horizontal operators, is detailed in Appendix B. Normalisation factors for the anisotropic scheme are computed according to the random method (3.27). The distribution of relative differences between the random normalisation factors and the exact solution is shown in Figure 4.8 for different numbers of ensemble members Q. Standard deviations of the three tested realisations of random normalisation factors are in good accordance with the theoretical value $1/\sqrt{2Q}$ (see Chapter 3). For $Q = 5000$, which was the ensemble size selected for subsequent tests, the resulting standard deviation is 1% and hence, the probability is less than 1% that a single normalisation factor differs more than 3% from the exact value. The computational effort that has to be spend on the generation of an ensemble of this size depends on the diffusion length scale L_h and the stretching factors $\tilde{s}_1$, $\tilde{s}_2$, but is generally less than the time consumed by one iteration of the assimilation algorithm.

The rotation angle α is calculated at each grid point as the angle between the gradient of the potential vorticity and the north-direction of the local coordinate system. Hence, according to (3.26), $\mathbf{T}^T \nabla_h \psi$ is the gradient of an

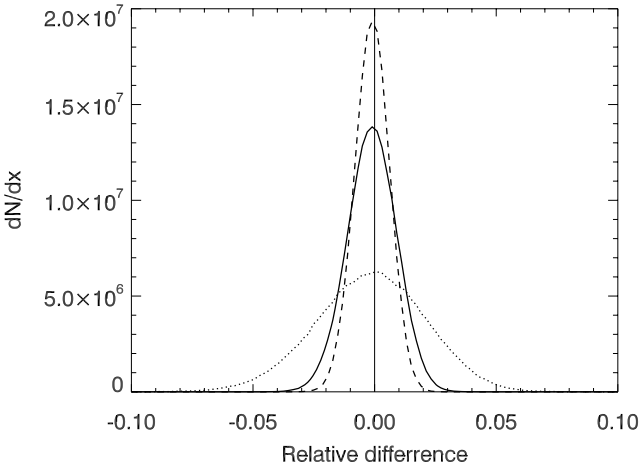

Figure 4.8: *Distribution of relative differences between random and exact normalisation factors for* 1000 *(dotted),* 5000 *(solid) and* 10 000 *(dashed) ensemble members* Q*. The standard deviations of the three distributions are* 2.25 %*,* 1.0 % *and* 0.71 %*, respectively.*

arbitrary scalar field ψ transformed into a $(\tilde{\eta}, \tilde{\chi})$ coordinate system, where the $\tilde{\chi}$-axis is aligned to the direction of the PV-gradient. Consequently, the stretching factors $\tilde{s}_1$ and $\tilde{s}_2$ specify the stretching or shrinking of coordinates $(\tilde{\eta}, \tilde{\chi})$ in the direction perpendicular and parallel to the PV-gradient, respectively. In the current version of the SACADA system, the full stretching is applied for PV-gradients greater than $0.2(\nabla P)_{\max}$, where $(\nabla P)_{\max}$ is the maximum PV-gradient at the corresponding model level. Below this value there is a linear decrease of stretching factors, and finally at locations where the PV-gradient equals zero, no coordinate stretching is applied.

As an example, the horizontal correlations generated by the two schemes are shown in Figure 4.9. Isotropic correlations with a horizontal length scale of 500 km are displayed at the top, and the outcome of the anisotropic scheme using the settings $L_h = 500$ km, $\tilde{s}_1 = 2$ and $\tilde{s}_2 = 0.25$ is shown at the bottom. The meteorological situation from 28 October 2003 was taken to compute the anisotropic correlations; the corresponding PV-field is shown in Figure 3.2. The larger correlation between grid cells along the edge of the polar vortex is clearly visible.

4.4 Assimilation system set-up

The main components of the SACADA assimilation system are the meteorological module, the forward model M, its adjoint version $\mathbf{M}^*$ and the BECM module as described above. A limited memory BFGS (L-BFGS, **l**imited-memory **B**royden **F**letcher **G**oldfarb **S**hanno) algorithm, which was devised by *Nocedal* [1980] and *Liu and Nocedal* [1989], is employed for the

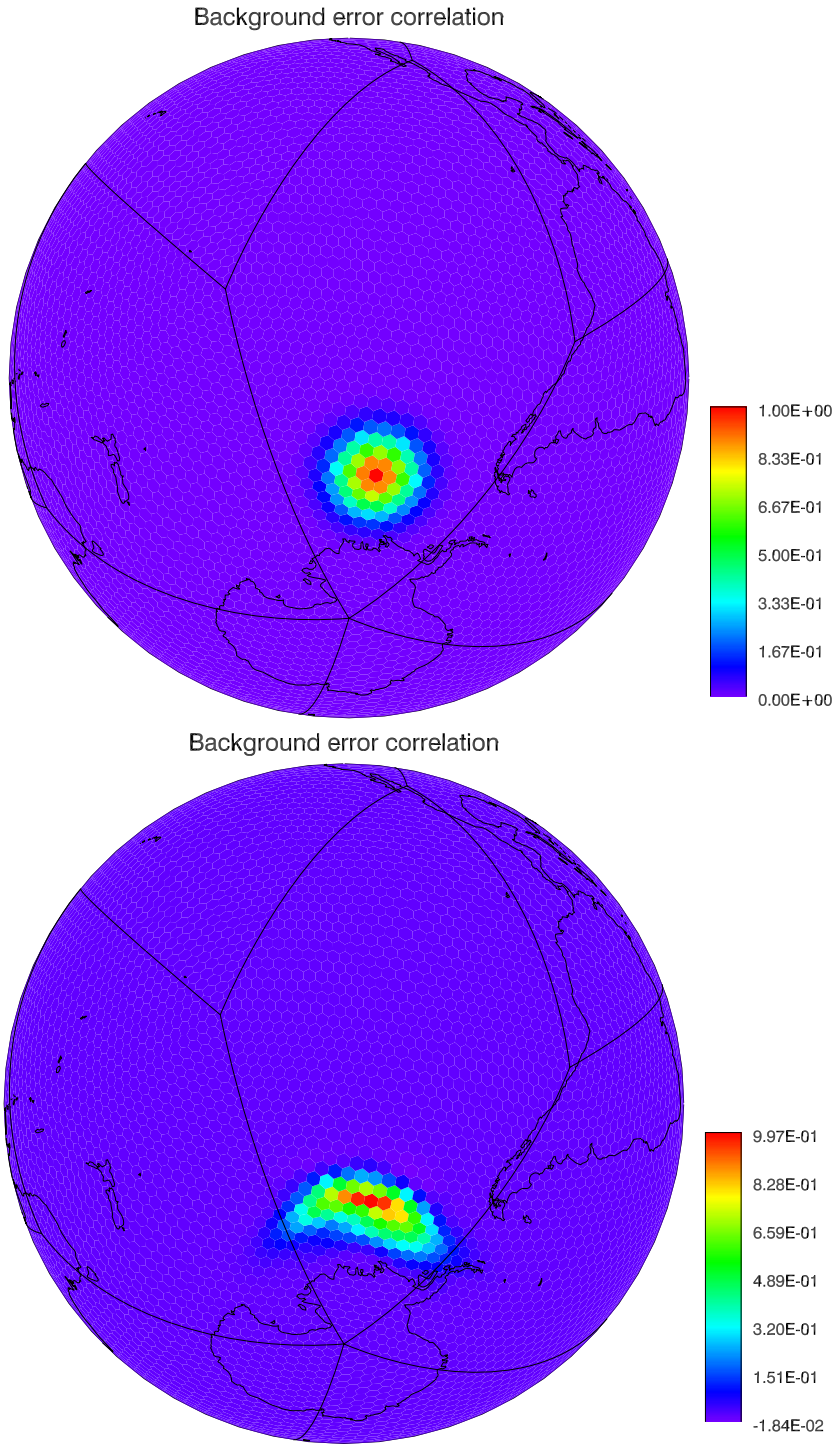

Figure 4.9: *Example of a discrete correlation function at the 100 hPa level for a grid point located at 72°S and 130°W, as calculated by the isotropic scheme (top) and the anisotropic scheme (bottom, PV field as shown in Figure 3.2).*

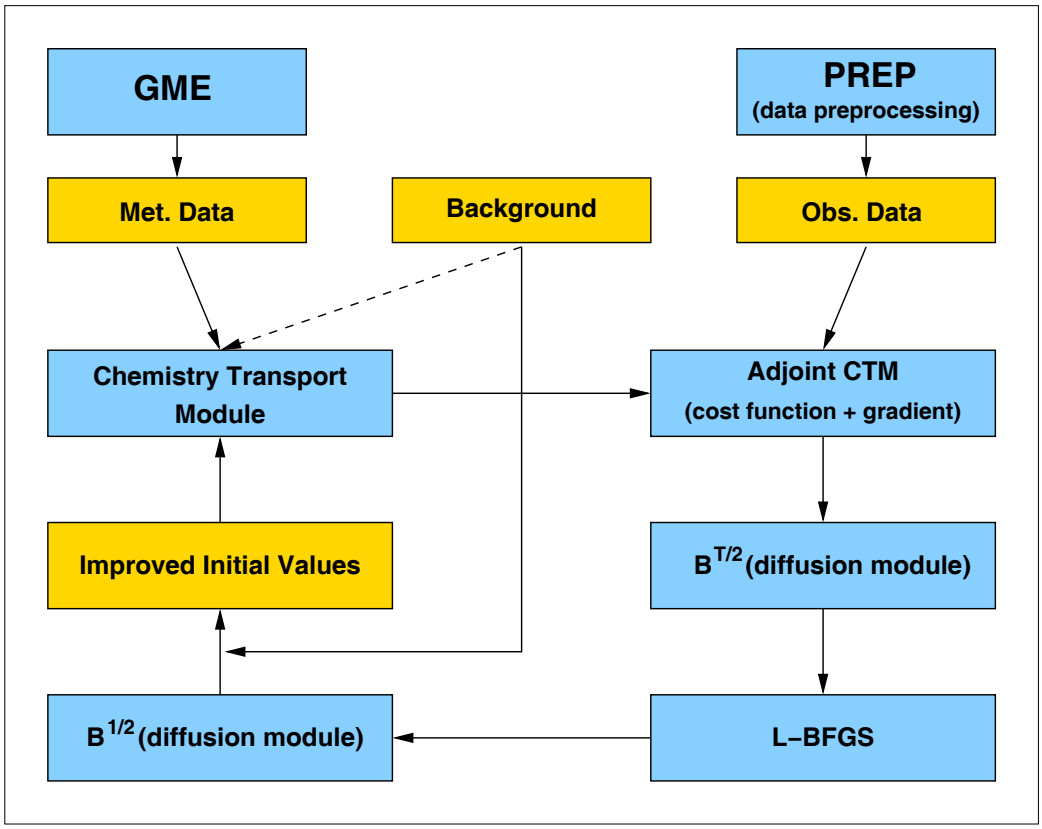

Figure 4.10: *SACADA system set-up.*

minimisation of the cost function. The interplay of all components within the SACADA system is shown in Figure 4.10. The first link in the process chain is the generation of meteorological data by GME. Observational data are collected by a preprocessing tool PREP, which provides data from different sources in a unified format. If necessary, i. e. if data quality is poor, data may be filtered at this step. The first iteration of the minimisation process is started with the background field $\boldsymbol{c}^{\mathrm{b}}$, which is the analysis result from the previous day, as a first guess. The chemistry-transport module performs a forward integration, thereby storing the chemical state twice at each timestep, as described in Section 4.2, for subsequent use by the adjoint CTM. Outcome of the adjoint calculations are the cost function and its gradient with respect to the initial volume mixing ratios. The resulting gradient is transformed according to (3.5), as the minimisation is performed with respect to the incremental variable $\boldsymbol{v}_0$. Improved values for $\boldsymbol{v}_0$ are delivered by the L-BFGS routine, which are transformed back by multiplying with $\mathbf{B}^{1/2}$ and adding the background volume mixing ratios according to (3.6). It was

found, that twelve iterations of this procedure suffice to approximate the minimum of $J(\boldsymbol{c}_0)$ adequately well.

CHAPTER 5

Observational basis

Space borne remote sounding instruments measure electromagnetic radiation that is emitted, scattered or transmitted by the earth's atmosphere. Information on atmospheric particles, temperature, pressure and the amount of various trace gases can be derived from the spectra recorded by these instruments. This process, which involves the inversion of a radiative transfer model, is called *retrieval*. Instruments differ in their viewing geometry and the spectral range of radiation that they are able to analyse. *Nadir viewing* instruments have a relatively small field of view, which is generally directed downwards, but may be scanned slightly away from the nadir direction. The horizontal resolution can be refined down to approximately 1×1 km for recent instruments with special high resolution observation modes. However, the information, that can be infered on vertical atmospheric structures remains limited. In contrast, the *limb viewing* geometry, as depicted in Figure 5.1, offers an improved vertical resolution at the cost of a limited information content along the line of sight. Instruments that analyse emitted radiation can operate on both, day- and night-side of the earth, while instruments recording scattered sunlight can take their measurements at the sunlit side of their orbit only. Some instruments use *occultation* techniques, recording spectra of sunlight that is transmitted through the atmosphere during sunrise or sunset. Occultation measurements are also possible with the moon or bright stars as a source of radiation.

From the data assimilation point of view several implications arise from the

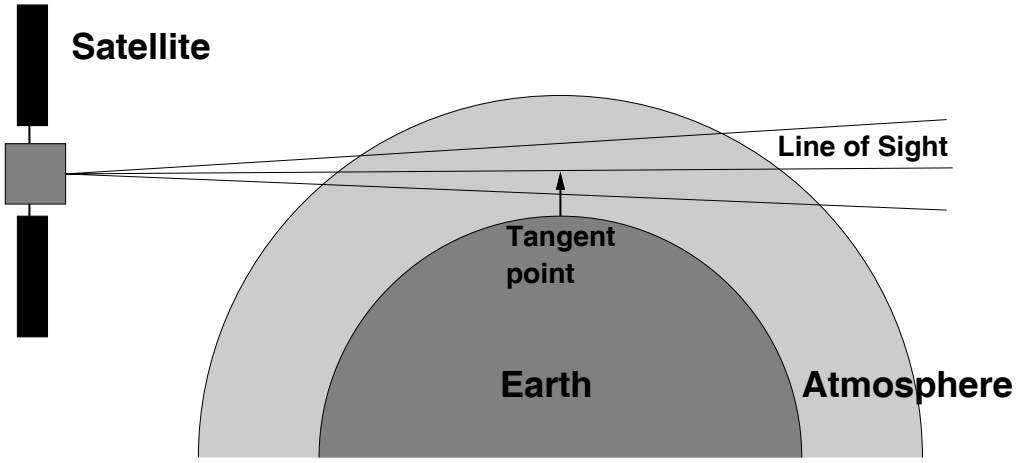

Figure 5.1: *A limb viewing space borne remote sounding instrument is recording a sequence of spectra scanning through the atmosphere as seen from the orbit. The information contained in these spectra mainly comes from the region around the tangent point, where the atmospheric density reaches its maximum along the line of sight.*

fact that retrieved profiles are *not* direct measurements of the atmospheric state, but rather a derivation through a complex inversion process is required. Therefore, the basic facts about retrieval methods for atmospheric sounding will be summarised in Section 5.1. A detailed discussion of this topic can be found in the textbook of *Rodgers* [2000]. In the work presented here, the SACADA assimilation system has been tested and evaluated by means of three case studies. Data from four different limb viewing instruments, namely MIPAS, SCIAMACHY, HALOE, and SAGE II, have been used for this purpose. A short description of each instrument and the respective data products will be given in Section 5.2.

5.1 Retrieval methods

From a theoretical point of view, the retrieval process is in no way different from the data assimilation problem. A radiative transfer model H calculates the radiance $\boldsymbol{y}$ that a certain instrument would measure for a given atmospheric state $\boldsymbol{x}$:

$$\boldsymbol{y} = H(\boldsymbol{x}) + \boldsymbol{\epsilon}_r + \boldsymbol{\epsilon}_s \,.$$

The measurement error $\boldsymbol{\epsilon}_r$ is usually regarded as a Gaussian distributed random quantity. Thus, it is called *random error* or *instrument noise*. In practice, additional errors $\boldsymbol{\epsilon}_s$ arise from the fact that a numerical model has to

be employed for the forward mapping to measurement space. Such error components originate from sources like uncertain spectral data, instrument calibration parameters or simplifying assumptions in the model H relative to the true physics. As these errors lead to differences between the retrieval and the true state, which are correlated in space and time, they are traditionally called *systematic errors* in the remote sounding literature. For the derivation of an optimal inverse method, systematic errors are neglected, which is the *perfect model assumption* already encountered in Chapter 2. Due to the relative smallness of the profile retrieval problem, compared to the enormous number of control variables that is encountered in atmospheric data assimilation, it is possible to a posteriori calculate the sensitivity of the retrieved profile to variations in uncertain model parameters. By this measure an estimate of the systematic error component $\boldsymbol{\epsilon}_s$ can be derived as outlined in *Rodgers* [2000]. However, by their nature, the systematic errors are *not* Gaussian distributed and therefore, the inclusion of these components in the observation error covariance matrix of the assimilation system violates the underlying statistical asumptions to some extent. On the other hand, the random component of the total retrieval error is often comparatively small and hence, the retrieved profiles recieve an unrealistic large weight, if the systematic error is neglected.

Most practical retrieval algorithms consider a one dimensional atmosphere, which composition varies with height only. The retrieval method solves the problem of finding an atmospheric profile $\boldsymbol{x}^{\mathrm{a}}$, which combines optimally the information content of the measurements and some background state $\boldsymbol{x}^{\mathrm{b}}$. The Bayesian approach, as described in Chapter 2, may be applied and under the common Gaussian assumption the maximum of the posteriori PDF can be found by minimising the cost function (2.3). The Gauss-Newton iteration (2.6) may be used to find the minimum of J and the error covariance matrix of the retrieval is given by (2.7) (*Rodgers* [2000]). As systematic error sources have been neglected, $\mathbf{P}^{\mathrm{a}}$ does *not* characterise the total retrieval error, but the contributions due to background error and instrument noise only. Note, that even if the measurement errors were uncorrelated, i. e. $\mathbf{R}$ is diagonal, the contribution of measurement noise to the analysis error covariance is correlated due to the retrieval.

Most inverse methods for atmospheric sounding are based on (2.6) or variants thereof. Sometimes, a Levenberg-Marquardt method is used, which employs slightly modified weights

$$\tilde{\mathbf{W}}_i = \left[\mathbf{B}^{-1} + \mathbf{H}_i'^T \mathbf{R}^{-1} \mathbf{H}_i' + \gamma_i \mathbf{I}\right]^{-1} \mathbf{H}_i'^T \mathbf{R}^{-1}\,, \tag{5.1}$$

a measure that stabilises the matrix inversion and effectively restricts the

magnitude of the update added at each iteration. Several strategies of choosing γ_i for each iteration step are discussed by *Rodgers* [2000]. Another common modification is to introduce some constraints on the *smoothness* of the retrieved profile in addition to or instead of the a priori information. This is done by including a term

$$\left[\boldsymbol{x} - \boldsymbol{x}^b\right]^T \mathbf{H}_r \left[\boldsymbol{x} - \boldsymbol{x}^b\right]$$

in the cost function (2.3). $\mathbf{H}_r$ is called *regularisation matrix*, which may encode different kinds of constraints. For example, the matrix $\mathbf{H}_r := \mathbf{S}^T\mathbf{S}$ with

$$\mathbf{S} = \begin{pmatrix} 0 & 0 & \ldots\ldots\ldots & 0 \\ -1 & 1 & 0 & \ldots\ldots & 0 \\ \vdots & & & & \vdots \\ 0 & \ldots. & 0 & -1 & 1 \end{pmatrix}$$

constraints the variability of the retrieved profile. In practical applications the regularisation constraint must be properly weighted against the observational part of the cost function and more sophisticated constraints can be applied (*Rodgers* [2000]). Note, that a regularisation matrix cannot be interpreted, in general, as the inverse of some covariance matrix, as $\mathbf{H}_r$ is usually singular.

5.1.1 Averaging kernels

An important characteristic of any inverse method is its sensitivity to variations in the state vector $\boldsymbol{x}$. Let R be the retrieval method, which produces an analysis from a given set of observations:

$$\boldsymbol{x}^{\mathrm{a}} = R(\boldsymbol{y})\,.$$

Now consider a retrieval that is carried out with a simulated vector of observations $\boldsymbol{y} = H(\boldsymbol{x})$. This will yield

$$\boldsymbol{x}^{\mathrm{a}} = R(H(\boldsymbol{x}))$$

and the sensitivity of the retrieval $\boldsymbol{x}^{\mathrm{a}}$ with respect to a small variation about the atmospheric state $\boldsymbol{x}_l$ can be computed as

$$\begin{aligned} &\boldsymbol{x}^{\mathrm{a}} - \boldsymbol{x}_l^{\mathrm{a}} \approx \mathbf{A}(\boldsymbol{x} - \boldsymbol{x}_l) \\ &\text{with} \quad \mathbf{A} := \left.\frac{\partial \boldsymbol{x}^{\mathrm{a}}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_l} = \left.\frac{\partial R}{\partial \boldsymbol{y}}\right|_{H(\boldsymbol{x}_l)} \left.\frac{\partial H}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_l} . \end{aligned} \tag{5.2}$$

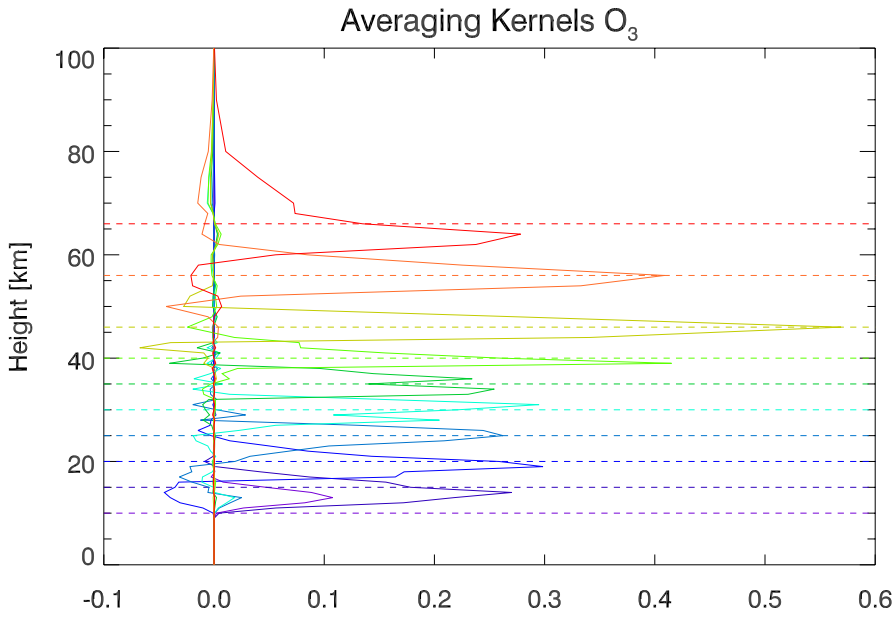

Figure 5.2: *Averaging kernels for an O_3 retrieval based on MIPAS spectra as calculated by Institut für Meteorologie und Klimaforschung (IMK), Forschungszentrum Karlsruhe. Note that for clarity reasons only a subset of kernels has been plotted. The IMK retrieval grid has a spacing between 1 and 10 kilometres and the horizontal dashed lines mark the heights for which kernels have been plotted.*

$\mathbf{A}$ is called *averaging kernel matrix* or *sensitivity matrix*. Considering the rows $\boldsymbol{a}_i$ of $\mathbf{A}$ as discrete functions of height, $\boldsymbol{a}_i(h_j)$ quantifys the response of the retrieval at level i to a perturbation $\Delta x_j = x_j - x_{j,l}$ at height h_j, as shown in Figure 5.2. For a perfect retrieval method, $\mathbf{A}$ would be the unit matrix, but in reality the functions $\boldsymbol{a}_i$ are peaked around the level i. The half-width of these peaks is often taken as a measure of the spatial resolution of the observing system. The area of the averaging kernel rows indicates how much information in the retrieved profile at a certain height originates from the observations. If the area of $\boldsymbol{a}_i$ is small, the actual atmospheric state has only little influence on the retrieved profile. Consequently, the analysis x_i^{a} is mainly determined by the a priori profile or smoothing constraints.

5.1.2 Assimilation of retrievals vs. assimilation of radiance

The assimilation of retrived profiles into an atmospheric model has a few drawbacks from the theoretical point of view. The vector $\boldsymbol{x}^{\mathrm{a}}$ which is the

outcome of a complex retrieval algorithm, as outlined above, becomes a part of the vector of 'observations' $\boldsymbol{y}$ in the assimilation system. An independent a priori profile has been used within the retrieval process to constrain the solution, and often the retrieval grid spacing is much finer than the actual resolution of the observing system. Hence, taking the retrieved profile as a direct observation of the quantity, which is to be assimilated, will overestimate the information content of $\boldsymbol{y}$.

For this reason weather prediction centres use observed radiances directly for the assimilation process. In this case the observation operator H involves a radiative transfer model, which calculates the radiance that a certain instrument would observe, given the atmospheric state computed by the assimilating model. The background field of the data assimilation system is superior to the climatological profiles, which usually represent the background in the retrieval, and it is possible to take into account a full three dimensional state of the atmosphere in the radiative transfer model. However, the effort to be spend on the development and implementation of a radiative transfer model (and its adjoint) suitable for the use in an atmospheric data assimilation system is considerable, and each different instrument and each trace gas requires its own specific treatment. Recently, a fast radiative transfer model intended for the assimilation of MIPAS limb radiances has been presented by *Bormann et al.* [2005].

A promising strategy to escape this dilemma is the use of averaging kernel information in the H-operator of an atmospheric data assimilation system instead of the full radiative transfer model. The observation increment due to a single retrieved trace gas profile[1] $\boldsymbol{y}^r$ can be computed according to (5.2) as

$$\boldsymbol{y}^r - H(\boldsymbol{x}^m) \approx \boldsymbol{y}^r - [\boldsymbol{y}_l^r + \mathbf{A}(G(\boldsymbol{x}^m) - \boldsymbol{x}_l^r)] \; . \tag{5.3}$$

The superscripts m and r have been introduced to distinguish quantities that are defined on different grids, namely the grid of the assimilating model (m) and the retrival grid (r). Operator G calculates the model equivalent on the retrieval grid, and $\boldsymbol{y}_l^r$ is the retrieval at the linearisation point $\boldsymbol{x}_l^r$, for which the averaging kernel matrix has been calculated. Provided the model state $\boldsymbol{x}^m$ is close enough to $\boldsymbol{x}_l^r$, such that the linear assumption underlying (5.2) is valid, it is possible to compute what would have been retrieved under the atmospheric conditions $\boldsymbol{x}^m$ given by the assimilating model. *Segers et al.* [2005] have successfully applied (5.3) in their Kalman filter based assimilation

[1]The similarity of retrieval algorithms and data assimilation methods has been emphasised by denoting retrieved profiles with $\boldsymbol{x}^{\mathrm{a}}$ until here. To maintain notational consistency, the vector of observations within the data assimilation system, which may consist of retrieved profiles, is refered to as $\boldsymbol{y}$ again, as introduced in Chapter 2.

of GOME profile data. Unfortunately, only few data products exist at the time of writing that provide full averaging kernel information. Hence, in the work presented here the approach outlined above could not be used because no or only incomplete averaging kernel information was available from the respective data provider.

5.2 Instrument and data product description

On 1 March 2002 MIPAS and SCIAMACHY were launched on board of ESA's Environmental Satellite (EnviSat) into a sun-synchronous polar orbit with an inclination of 98.55°. The orbit is almost circular at about 800 km altitude resulting in a total of 14.3 orbits that are performed each day. The descending node of the orbit (crossing of the equatorial plane from north to south) is located at 10 a. m. local time. MIPAS has been the main data source for assimilation case studies presented in this work. Data from SCIAMACHY, SAGE II and HALOE have been kept back as independent (not assimilated) control data sets. A cross validation of the MIPAS instrument with data from the latter sensors can be performed, as outlined in Chapter 6. The most important characteristics of the different MIPAS and SCIAMACHY data products are summarised in Table 5.1.

5.2.1 MIPAS

The Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) is a Fourier-transform spectrometer measuring high resolution emission spectra in the mid-infrared from 4.1 to 14.7 μm wavelength in a limb viewing mode. The instrument's field of view is about 30 km in the horizontal and 3 km along the vertical direction at the tangent point. A single limb-scan covers an altitude range from approximately 6 to 68 km by 17 steps. A detailed description of the instrument design can be found in *Fischer and Oelhaf* [1996]. Two different data products derived from MIPAS spectra have been assimilated, namely the trace gas profiles that are operationally delivered by ESA as described by *Ridolfi et al.* [2000], called MPE herafter, and retrievals that have been produced by the Institut für Meteorologie und Klimaforschung (IMK), Forschungszentrum Karlsruhe, referred to as MPI in this text. More detailed information about the latter data product can be found in *von Clarmann et al.* [2003], *Glatthor et al.* [2004], *Höpfner et al.* [2004], *Funke et al.* [2005] and *Mengistu Tsidu et al.* [2005]. Profiles of O_3, NO_2, CH_4, HNO_3, H_2O and N_2O are contained in both of these data products. The set of retrieved

species available from IMK additionally comprises HNO_4, $ClONO_2$, N_2O_5, ClO, NO, CFC-11 and CFC-12. The IMK and ESA retrieval algorithms are based on the Levenberg-Marquardt formula (5.1), however, without using an explicit a priori profile but a smoothing constraint instead.

The retrieval grid of the MPI data product is fixed with a 1–2 km spacing in the region where the retrieval is sensitive to the respective target species. A detailed error characterisation of the profile data comprising the full retrieval noise error covariance matrix, a total error estimate, and the averaging kernel matrix is provided, however, only for two of the three case study periods. Note, that the linearisation profile $\boldsymbol{x}_l$ for the averaging kernel matrix is not regularly distributed with the data. Hence, the approach (5.3) cannot be used.

Retrievals that are delivered by ESA are sampled on a grid, the spacing of which varies from one profile to the next, as the volume mixing ratios are given at the tangent location of the respective scan. Thus, the retrieval grid has approximately the same spatial resolution as the observing system. Retrieval noise error covariance matrices (but no systematic error estimates) as well as averaging kernels are part of the data product. Unfortunately, the averaging kernel matrices are provided for nominal (fixed) tangent altitudes, which makes the practical application of (5.3) complicated, if feasible at all (compare *Migliorini et al.* [2004]).

5.2.2 SCIAMACHY

The Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) measures sunlight transmitted, reflected and scattered by the earth's atmosphere or surface. The spectral range of the instrument covers the ultraviolet, visible and near infrared wavelength region between 240 and 2380 nm. SCIAMACHY can be operated in three different measurement modes, a nadir viewing, a limb viewing and a solar/lunar occultation mode. A description of the instrument and viewing modes can be found in *Bovensmann et al.* [1999]. Limb and solar occultation profiles retrieved by Institut für Umweltphysik, Universität Bremen have been used for the case study based evaluation of the SACADA assimilation system.

The solar occultation technique and the retrieval aspects are described in full detail by *Meyer* [2004]. As a solar occultation takes place only once per orbit, a maximum of 14 profiles per day can be obtained. Due to EnviSat's special orbit geometry the location of these profiles is confined to a relatively small latitude band between approximately 50°N and 70°N. The main advantages

of the solar occultation method arise from the facts that the sun is a strong source of radiation leading to a favourable signal to noise ratio, and that reference spectra of the sun well above the atmosphere can be recorded. Hence, this technique is self-calibrating. For these reasons, profiles retrieved from solar occultation measurements are considered to be more precise than profiles derived using other viewing geometries. Because of the sparse spatial coverage on the one hand, and the superior quality of these data on the other hand, profiles from solar occultation measurements have not been assimilated but have been retained to serve as independent data for the assessment of the quality of the analysed atmospheric fields.

The Bremen solar occultation and limb data products (denoted SCO and SCL, respectively) contain profiles of O_3 and NO_2 while the limb data additionally provides profiles of BrO. Due to the nature of the measurement process, SCIAMACHY limb profiles can only be retrieved at the sunlit side of the orbit. Consequently, spatial coverage is less dense compared to MIPAS data. Global estimates of the absolute retrieval error are provided for both data products. Averaging kernels are not distributed, but the retrieval is reported to have a good sensitivity between 15 and 40 km altitude (*Meyer* [2004]).

5.2.3 SAGE II and HALOE

The Stratospheric Aerosol and Gas Experiment II (SAGE II) and the Halogen Occultation Experiment (HALOE) provide a long term record of important stratospheric constituents. Their data has been compared to and validated against a wealth of direct and indirect measurements from other instruments (see *Wang et al.* [2002], *Nazaryan et al.* [2005] and references therein). The retrieval algorithms and data quality have been improved continuously over the years. Trace gas profiles from SAGE II and HALOE can therefore be considered to be the best validated stratospheric remote sounding data available at the time of writing.

The SAGE II instrument was launched in October 1984 on board the ERBS spacecraft. The instrument operated in the UV, visible and near-infrared (385-1020 nm) wavelength region employing the solar occultation technique. In contrast to SCIAMACHY, SAGE II was able to take measurements during sunrise *and* sunset, leading to a maximum of 30 profiles per day. After technical problems in July 2000, the instrument was operated at a 50% duty cycle and about 15 retrievals per day are available. Due to the orbit geometry, tangent point locations vary slowly from 60°N to 60°S within approximately one month. Retrieved profiles of aerosol extinction, O_3, NO_2 and H_2O are

	MPE	MPI	SCO	SCL
Data product version	4.61	V1 / V2 / V3	1.0	O_3: V1.62 NO_2: V2.0
Available species	O_3, H_2O, CH_4, N_2O, HNO_3 and NO_2	as MPE + N_2O_5, HNO_4, $ClONO_2$, ClO, NO, CFC-11 and CFC-12	O_3 and NO_2	O_3, NO_2 and BrO
Sensitive altitude range	stratospheric, depending on species	stratospheric, depending on species	15-40 km	15-40 km
Retrieval grid	variable, ~3 km spacing	fixed, 1-2 km spacing	fixed, 1 km spacing	variable, 2-3 km spacing
Vertical resolution	3-5 km depending on species and altitude	3-5 km depending on species and altitude	3-5 km	3-5 km
Error estimates	random error covariance matrix	random error covariance matrix + systematic error estimate	global absolute error estimate	global absolute error estimate
Averaging kernel	available (see text)	available	not available	not available
A priori information	no explicit background, regularisation	no explicit background, regularisation	background + regularisation	background + regularisation

Table 5.1: *Summary of MIPAS and SCIAMACHY data product characteristics. MPE and MPI trace gas profiles are derived from MIPAS spectra by ESA and by Institut für Meteorologie und Klimaforschung, Forschungszentrum Karlsruhe, respectively. SCO and SCL are the SCIAMCHY solar occultaion and limb data product, which have been provided by Institut für Umweltphysik, Universität Bremen.*

provided by the NASA Langley Radiation and Aerosol Branch. Recently, a new data product version 6.2 (referred to as SG2 in the following text) was released with improvements mainly concerning the water vapour retrieval (*Thomason et al.* [2004]). SAGE II was switched off in August 2005 after almost 21 years of nearly continuous operation.

HALOE was in operation from October 1991 to November 2005. On board the Upper Atmosphere Research Satellite (UARS) it performed solar occultation measurements at sunrise and sunset in the infrared wavelength region. Profiles of O_3, HCl, CH_4, H_2O, NO, NO_2, HF, temperature and aerosol extinction have been derived routinely. During nominal operation, about 30 occultation events per day have been recorded. As in the case of SAGE II, the latitude of tangent point location changes slowly from day to day, covering a range of 80°N to 80°S within approximately one month. The latest data

product release (version 19, abbreviated HLO hereafter) has been obtained from the NASA Langley Research Center in Hampton, Virginia.

CHAPTER 6

Case studies

Three periods of time in 2002 and 2003 each spanning about one and a half month have been selected to test and evaluate the new SACADA assimilation system. There have been dedicated assimilation runs for each of the data products MPE and MPI, while the SCL, SCO, SG2 and HLO data served as independent control data sets (see Chapter 5 for the definition of data product acronyms). The three case study periods are:

CS1 1 September 2002 – 15 October 2002

CS2 1 July 2003 – 15 August 2003

CS3 21 October 2003 – 30 November 2003.

Two noteworthy events took place during the periods **CS1** and **CS3**. At the end of September 2002, a major stratospheric warming occurred at southern latitudes, which caused the antarctic polar vortex to split into two fragments around 25 September. It was for the first time since the beginning of observations in 1958 that a major stratospheric warming event occurred in the southern hemisphere. This event is clearly visible in the assimilated trace gas fields as will be shown in Section 6.2. Furthermore, a series of exceptionally strong solar proton events was observed during late October and early November 2003. Energetic particles penetrated the atmosphere at high geomagnetic latitudes and caused an enhancement of NO_x together with depletion of ozone and various other effects in the polar upper stratosphere and

lower mesosphere (see *López-Puertas et al.* [2005*a*] and *López-Puertas et al.* [2005*b*] and references therein). The occurrence of the latter events is not optimal for the purpose of testing and evaluating the new assimilation system, since the energetic particles act as sources of chemical constituents and these processes are not accounted for in the SACADA chemistry module and its adjoint. However, it was found that the performance of the assimilation system was generally not degraded as discussed in Section 6.4.

An interesting application of 4D-var data assimilation is the possibility to *cross validate* retrieved profiles from different instruments. Usually this task is accomplished by comparing an ensemble of *collocated* profiles, which are separated by not more than some predefined maximum distance and a maximum time interval. This technique is common practise, but the selection of collocation criteria is somewhat arbitrary and problematic: If strict collocation criteria are chosen, only few coincident profiles will be found and the statistical basis for cross validation remains poor. If, however, the criteria are relaxed, the two instruments may have sounded different air masses, which makes a comparison worthless. The use of chemically consistent constituent fields obtained by 4D-var data assimilation offers a more satisfying approach to the cross validation challenge. Without applying any artificial criteria, a global analysis produced by processing observations of one instrument can be compared to profiles of another sensor, which have not been used within the assimilation procedure. Naturally, this approach requires the errors of the analysed field to be small enough in order to enable a useful comparison. Unfortunately, the 4D-var method does not provide for estimates of the analysis error, but it can be argued that if the analysis is based on sufficiently dense data sets –as it is the case for MIPAS observations– the analysis error will be small enough for cross validation purposes. A few results of cross validation between MIPAS and SCIAMACHY, SAGE II as well as HALOE ozone profiles are presented in Section 6.4. An in-depth discussion of this issue, however, is beyond the scope of this work.

Note, that not all case study periods are completely covered by all data sets. Data availability is summarised in Figure 6.1. For notational convenience, different configurations of the assimilation system will be referred to as following the scheme **CS*n*-*ins*-*v***, where n is the case study number ranging from 1 to 3, *ins* is the instrument/data product acronym of the assimilated data set and v is a version number distinguishing assimilation runs with different BECM parameter settings. For example, **CS2-MPE-1** denotes an assimilation run where MIPAS data processed by ESA have been assimilated with BECM parameter settings as detailed in the corresponding column of Table 6.1.

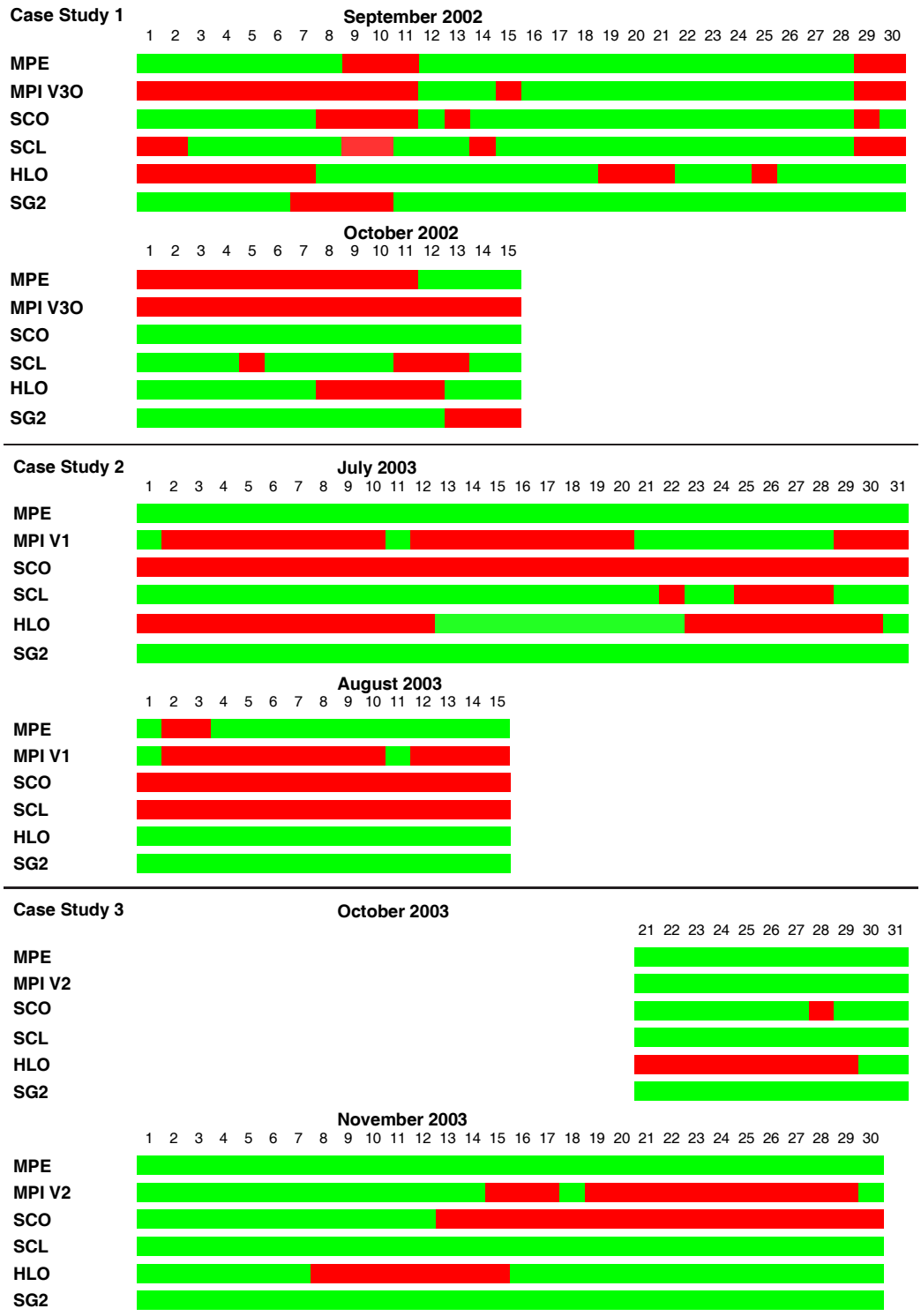

Figure 6.1: *Data availability for the three case study periods. Red colour indicates that no data from the respective instrument is available at the particular day.*

6.1 System set-up

Initial values for the first day of each of the case study periods have been derived from the two dimensional atmospheric model SOCRATES (*Brasseur et al.* [1995]). The output of this model contains latitude/height distributions of zonal mean volume mixing ratios of all stratospheric constituents included in the SACADA reaction mechanism, except for HNO_4, for which a constant volume mixing ratio of 1.0×10^{-3} ppb was assumed. After distributing these data to the icosahedral grid, a 48 hour model spin-up run was performed to relax the chemical constituents towards their equilibrium state. As the estimate of the atmospheric state represented by these initial values is relatively coarse, a *spin-up assimilation* covering the first week of each case study period has been performed. For the first six days, the PSC-scheme has been turned off, and a large relative background error ϵ_b as well as large correlation length scales L_h, L_v have been assumed. At day seven the PSC-scheme was switched on and reduced values for ϵ_b, L_h and L_v have been inserted, which were the settings for the rest of the case study periods. The definition of BECM parameter values for different configurations of the system can be found in Table 6.1. Note that during September/October 2002 and July/August 2003 only limited MPI data were available. Therefore, the configurations **CS1-MPI-*n*** and **CS2-MPI-*n*** have been run with a three day spin-up assimilation, which was initialised using the assimilation result from **CS1-MPE-1** and **CS2-MPE-1** for 12 September 2002 and 21 July 2003, respectively. The standard set of species, which comprises O_3, NO_2, CH_4, N_2O, HNO_3 and H_2O, has been assimilated in all but the **CS3-MPE-2/3** configurations presented here. Furthermore, all configurations involving the MPI data product, additionally included the assimilation of N_2O_5, $ClONO_2$, CFC-11 and CFC-12. In order to enable a comparison between the analyses attained with the assimilating model and a free model run, a *control model run* without data assimilation starting from the final analysis of the spin-up assimilation has been accomplished.

To investigate the benefits of the anisotropic background error covariance formulation, two pairs of assimilation experiments have been conducted. The first experiment with anisotropic correlations (**CS3-MPE-3**) was carried out with the domain of the chemistry-transport module restricted to model levels 10–25 (2–100 hPa) and the assimilation confined to ozone only. The local coordinate stretching factors have been set to $\tilde{s}_1 = 4$ and $\tilde{s}_2 = 0.25$, corresponding to halving the correlation length scale in the direction aligned with the local PV-gradient, and doubling it in the perpendicular direction. As the results have been very encouraging, case study **CS1-MPE-5** used

Configuration	**1**	**2**	**3**	**4**	**5**
L_h (spin-up)	600 km	1200 km	1200 km	800 km	800 km
L_h	300 km	800 km	800 km	500 km	500 km
L_v (spin-up)	2 km	2 km	2 km	2 km	2 km
L_v	1.5 km	1.5 km	1.5 km	1.5 km	1.5 km
ϵ_b (spin-up)	100%	100%	100%	80%	80%
ϵ_b	50%	50%	50%	40%	40%
$\tilde{s}_1$	1.0	1.0	4.0	1.0	4.0
$\tilde{s}_2$	1.0	1.0	0.25	1.0	0.25

Table 6.1: *BECM parameter settings for the different configurations of the assimilation system.*

the same settings for $\tilde{s}_1$ and $\tilde{s}_2$, but the standard set of species was assimilated within the full domain of the chemistry-transport module. Case studies **CS3-MPE-2** and **CS1-MPE-4**, for which assimilation was performed with isotropic background error covariance settings, serve as a reference. The results can be found in Section 6.3.

The observation error covariance matrix $\mathbf{R}$ has been assumed to be diagonal with the variances taken from the respective data products. If, however, only random (instrument noise) errors were available, the error margins had to be increased in order to account for the systematic error components (see discussion in Section 5.1). Furthermore, some of the MPI error estimates turned out to be unrealistic small, even allthough the systematic error component was included. As observations with unrealistic small errors receive too much weight within the data assimilation procedure, it has been decided to assume some minimum relative error for the MPE and MPI data products. Different values for the minimum relative error margin have been tested and the results presented here, have been obtained with the MPI and MPE error margins increased to 7% and 10%, respectively. The statistical evaluation below indicates that even the latter value is presumably too optimistic. Additionally, the operational ESA data MPE had to be filtered for extreme outliers before they could be used for assimilation.

6.2 General results

The evolution of the cost function during the course of the three case studies for configurations **CS3-MPI-1**, **CS1-MPE-1** and **CS2-MPE-1** is shown in Figures 6.2, 6.3 and 6.4, respectively. The cost function values are divided

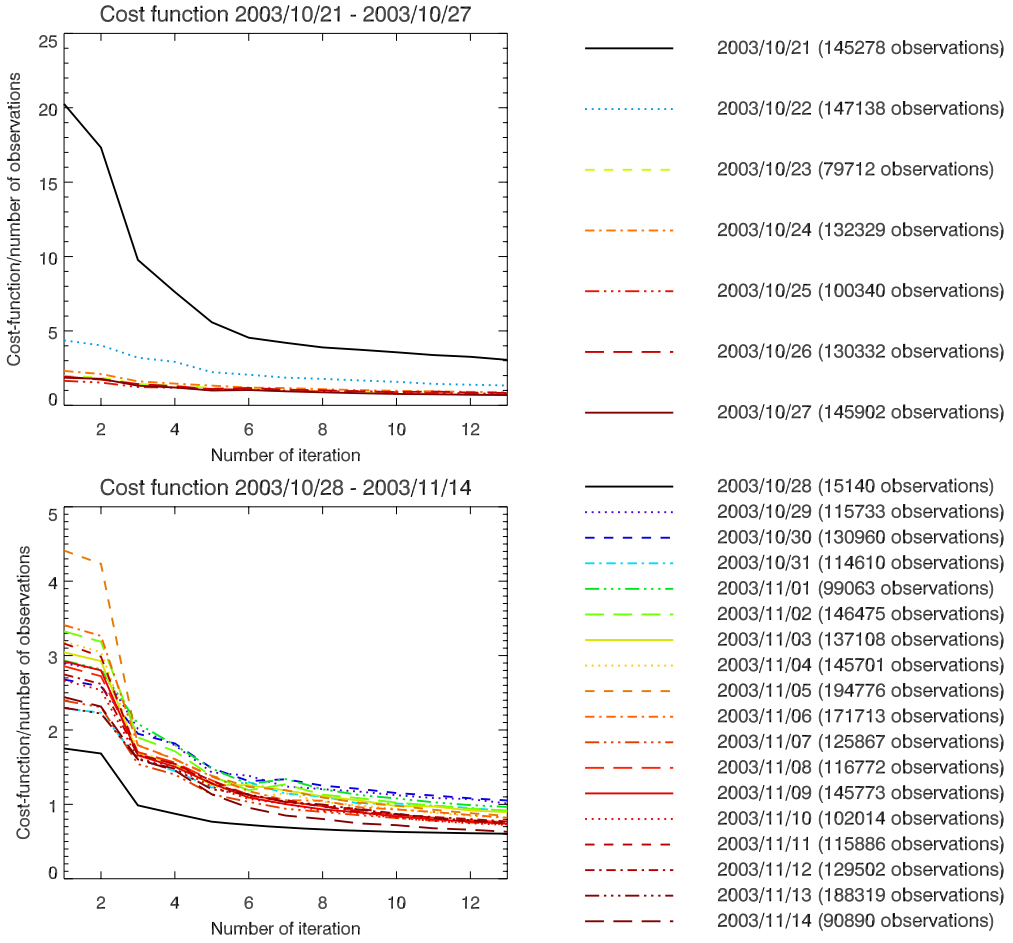

Figure 6.2: *Evolution of the normalised cost function (cost divided by the number of observations) for configuration* ***CS3-MPI-1****. The x-axis gives the iteration count, and the number of observations available at a particular day is displayed at the right hand side of each plot.*

by the number of available observations p to make the results comparable among different days. Further, it can be shown (*Talagrand* [1998]) that at the minimum of the cost function the expected value of $J_p^{\mathrm{a}} := J(\boldsymbol{x}^{\mathrm{a}})/p$ is 0.5, if the true covariance matrices $\mathbf{R}^{\mathrm{t}}$ and $\mathbf{B}^{\mathrm{t}}$ had been used. Consequently, a mean value of $J_p^{\mathrm{a}} = 0.5$ constitutes a necessary condition for $\mathbf{B}$ and $\mathbf{R}$ to approximate the true covariances adequately well. The analysis cost function value J_p^{a} stays slightly above the optimal value of 0.5. This may indicate, among other possible causes, that the observations error margins are still too small, as the background with its large error margins does not contribut significantly to the total cost. Results for the remaining configurations (not

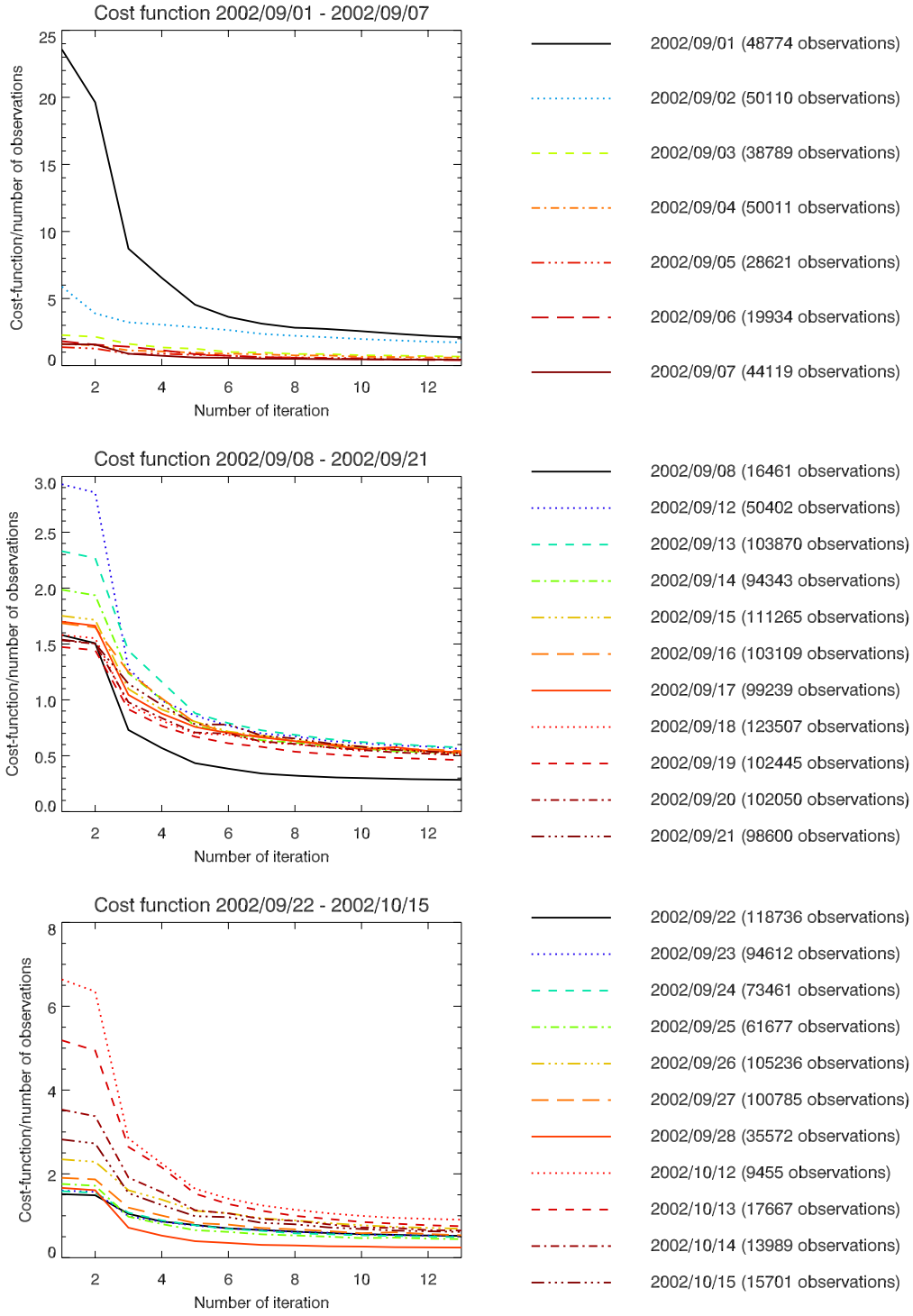

Figure 6.3: *As Figure 6.2 for configuration* ***CS1-MPE-1****.*

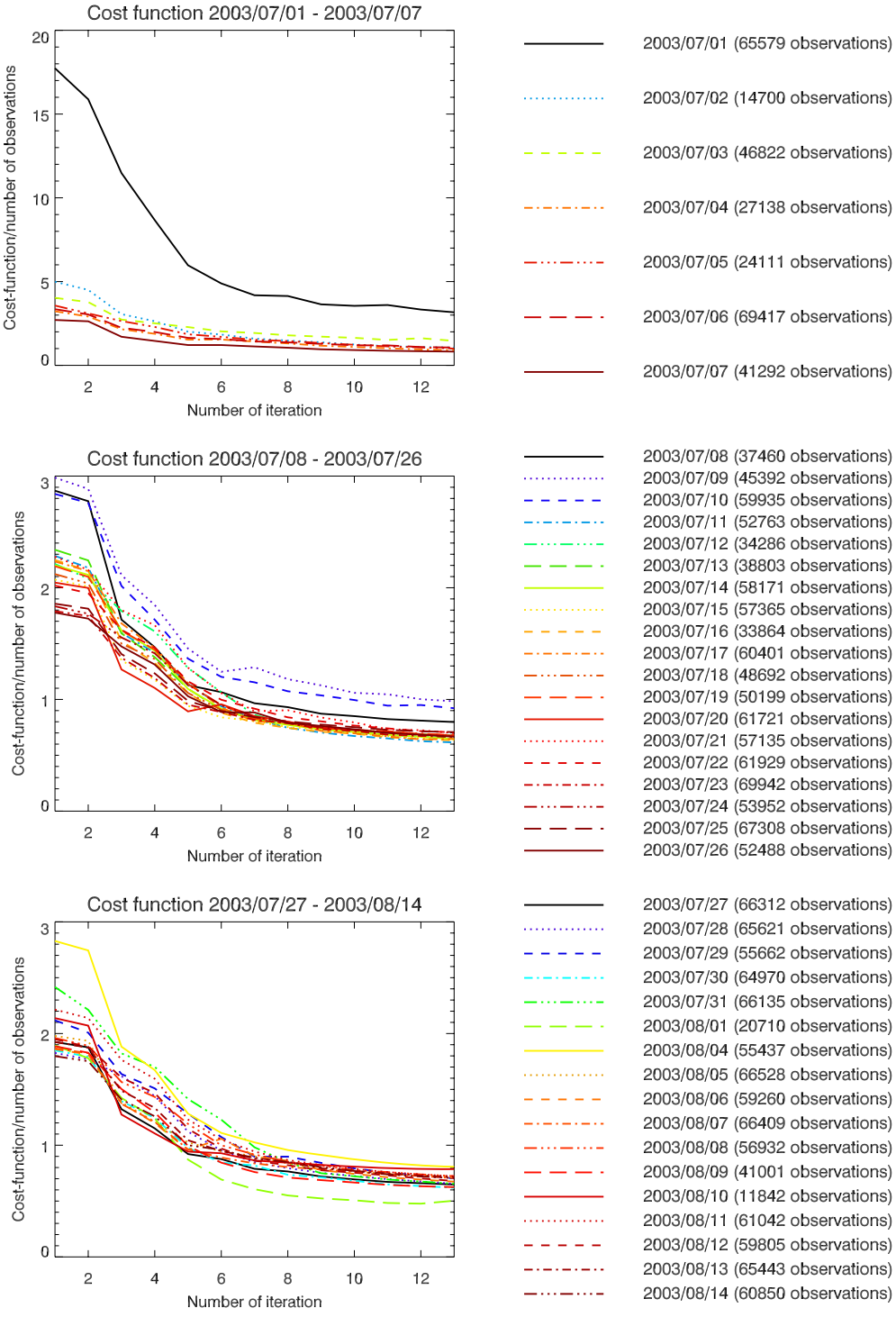

Figure 6.4: *As Figure 6.2 for configuration* ***CS2-MPE-1****.*

shown) are very similar.

During the spin-up assimilation of seven days (Figures 6.2–6.4, top panel), the normalised cost function value at iteration one

$$J_p^{\mathrm{f}} := J(\boldsymbol{x}^{\mathrm{b}})/p = \frac{1}{2p} \sum_{i=0}^{N} \left[H\left(M_i(\boldsymbol{x}^{\mathrm{b}})\right) - \boldsymbol{y}_i \right]^T \mathbf{R}^{-1} \left[H\left(M_i(\boldsymbol{x}^{\mathrm{b}})\right) - \boldsymbol{y}_i \right]$$

rapidly decreases within the first few days. During the rest of the period (middle and bottom panel) the cost function is reduced by a factor of approximately three with little day to day variation. Note that J_p^{f} gives the mean square difference between the first iteration forward run and observations, weighted by the observation error covariances. Hence, this quantity can be regarded as a measure of forecast skill: The first forward integration of the model, initialised with the analysis from the previous day, has not been corrected towards the current observations. Hence, it is a pure one-day forecast, the skill of which depends on both, the quality of the previous analysis, and the model skill, particularly if data is sparse in time or space.

Observe that during **CS1-MPE-1** there are two data gaps of 3 and 13 days. To bridge these gaps, the model has been operated in forecast mode starting from the last available analysis. Consequently, J_p^{f} at 12 September 2002 and 12 October 2002 (Figure 6.3) reflect the forecast skill of a four-day and a two-week forecast, respectively. The factor by which J_p^{f} is larger for the four-day forecast in comparison to the J_p^{f} value at other days is less than two, and even after bridging the data gap of 14 days this factor is not larger than four, which is a fully satisfying result.

Assimilated ozone fields for the vortex split case study **CS1-MPE-1**, together with the corresponding ozone distribution from the control run are shown in Figures 6.5–6.7 for model level 21, corresponding to a pressure of 36 hPa. At 20 September 2002 the polar vortex is still intact, but already distorted. The vortex split occurs around 25 September and the smaller of the two fragments rapidly mixes with mid-latitude air masses. A comprehensive discussion of the vortex split event can be found in *Allen et al.* [2003]. The control model run basically captures the relevant features of the ozone field during the September/October 2002 period. However, the ozone values, especially within the polar vortex and the two vortex fragments, are much too high. Another noticeable feature is, that the analysed ozone fields show a noisy structure in some regions. Since the parameters of the backround error covariance matrix, i. e. the relative backround error ϵ_b as well as the length scales L_h and L_v, have been chosen very conservatively, an overfitting of single observations may have occurred.

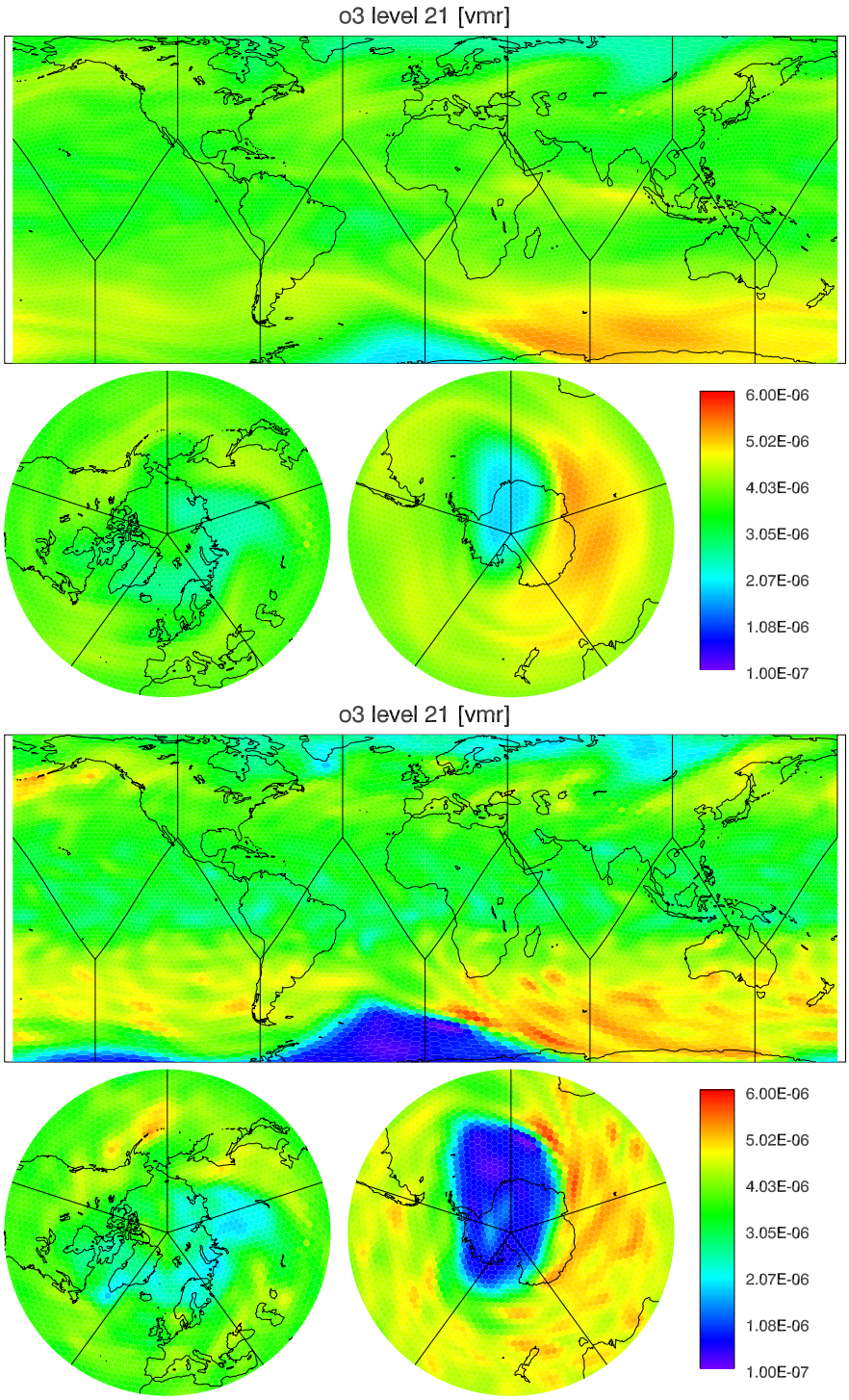

Figure 6.5: CS1-MPE-1 *ozone analysis for 20 September 2002 (bottom) and ozone control run distribution (top) at model level 21, corresponding to 36 hPa.*

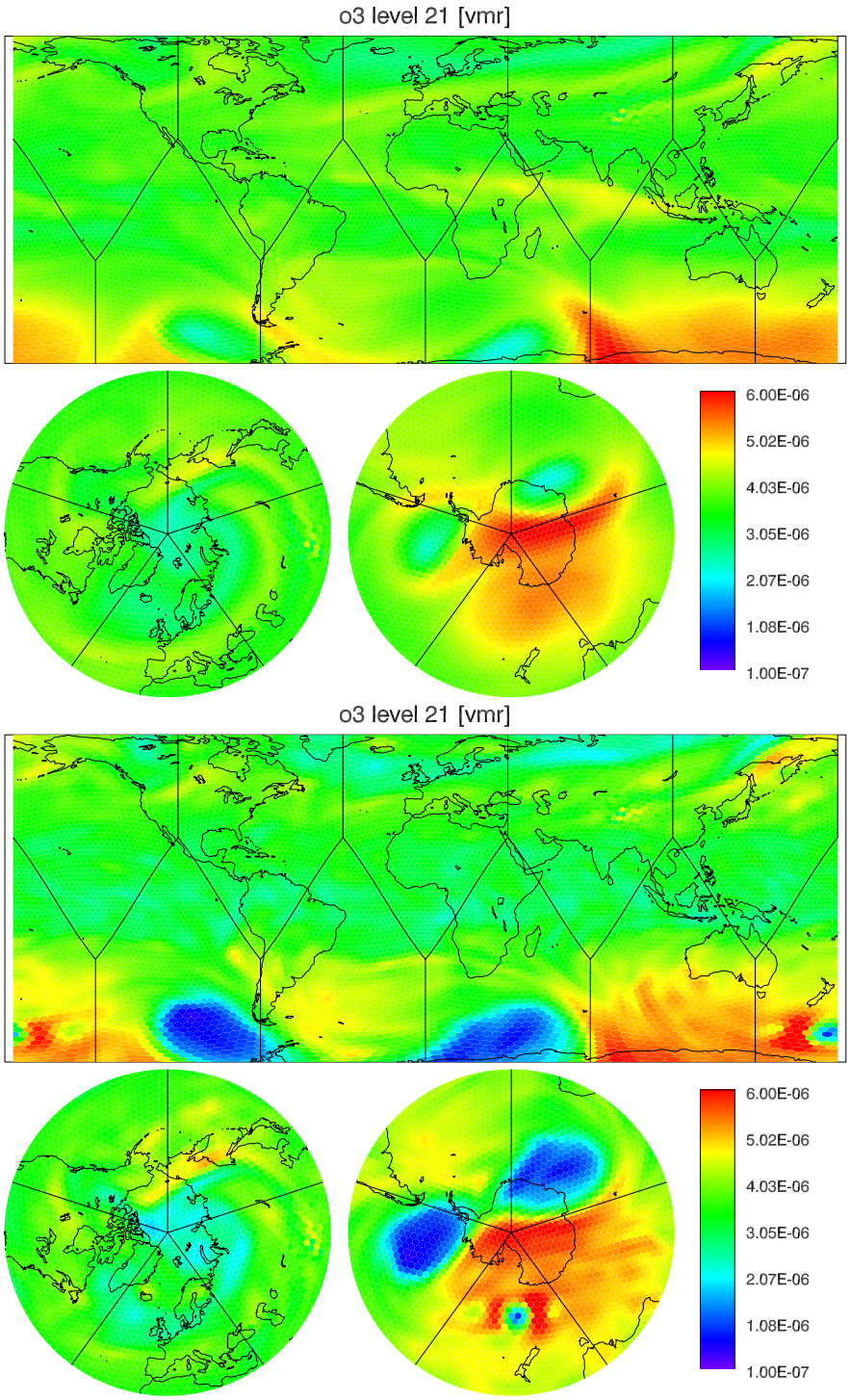

Figure 6.6: CS1-MPE-1 *ozone analysis for 25 September 2002 (bottom) and ozone control run distribution (top) at model level 21, corresponding to 36 hPa.*

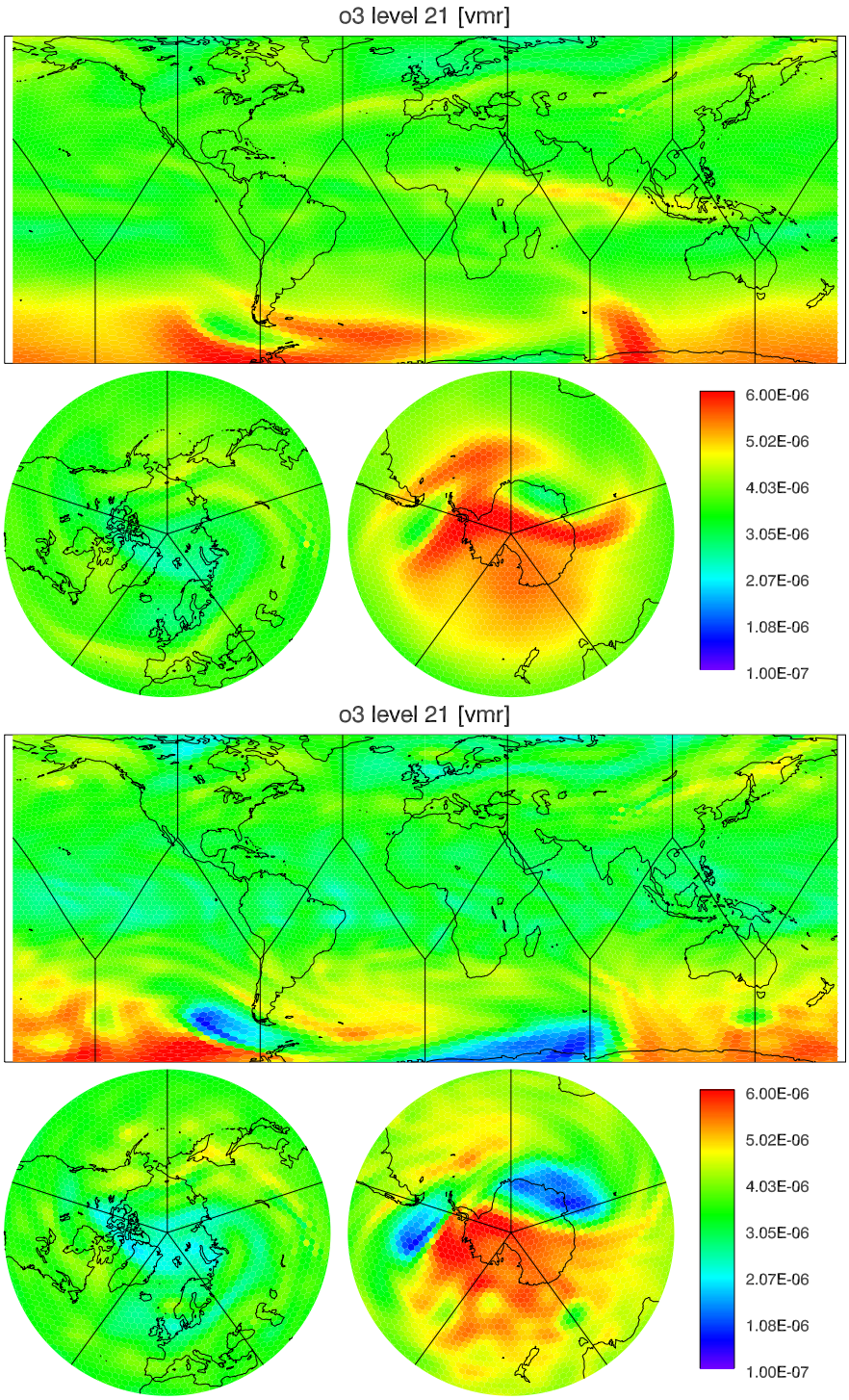

Figure 6.7: CS1-MPE-1 *ozone analysis for 28 September 2002 (bottom) and ozone control run distribution (top) at model level 21, corresponding to 36 hPa.*

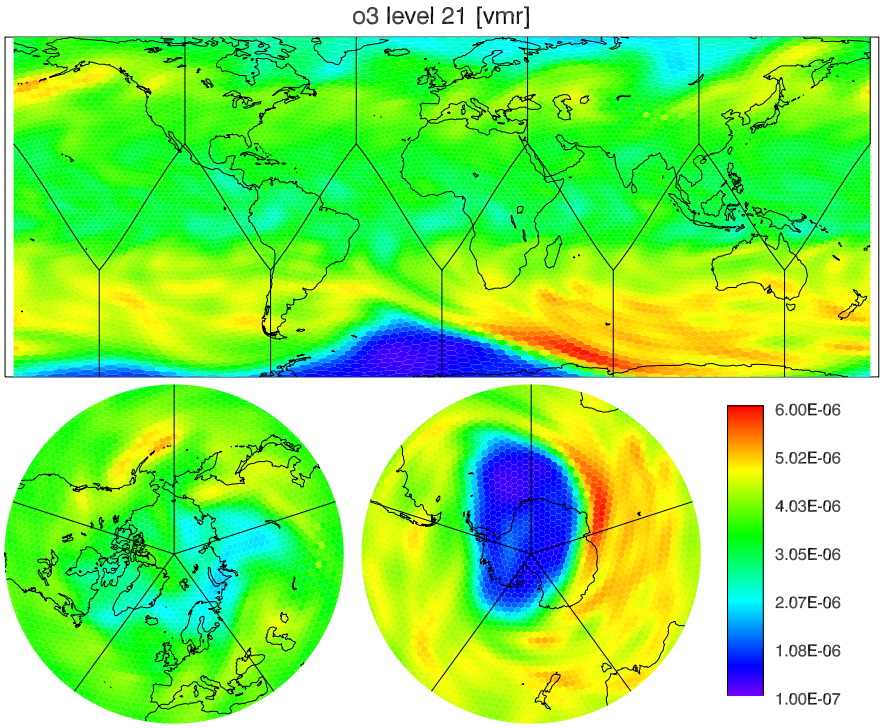

Figure 6.8: CS1-MPE-4 *ozone analysis for 20 September 2002 at model level 21, corresponding to 36 hPa. The effect of smaller background errors and larger correlation length scales is most pronounced in high southern latitudes (compare Figure 6.5).*

Consequently, due to the good performance of the assimilating model, it seemed to be justified to further reduce the value of ϵ_b and to assume larger horizontal correlation length scales for the configuration **CS1-MPE-4**. The resulting ozone analysis at 20 September 2002, as shown in Figure 6.8, is smoother and the signature of individual observations is less obvious (compare Figure 6.5). In order to decide if the latter analysis is objectively better, the cost function values J_p^{f} can serve as a measure of analysis quality. Thereby, as outlined above, analysis quality is quantified in terms of the weighted mean square difference between the subsequent forecast and observations. Figure 6.9 clearly indicates that the J_p^{f} values are generally lower for the **CS1-MPE-4** configuration. Even after the data gap of 13 days, J_p^{f} is smaller by 8% with the **CS1-MPE-4** BECM settings. This demonstrates that the choice of the covariance parameters has considerable influence on the analysis quality. As will be shown in the next section, further improvement can be achieved by employing the more sophisticated flow dependend background error covariance parameterisation.

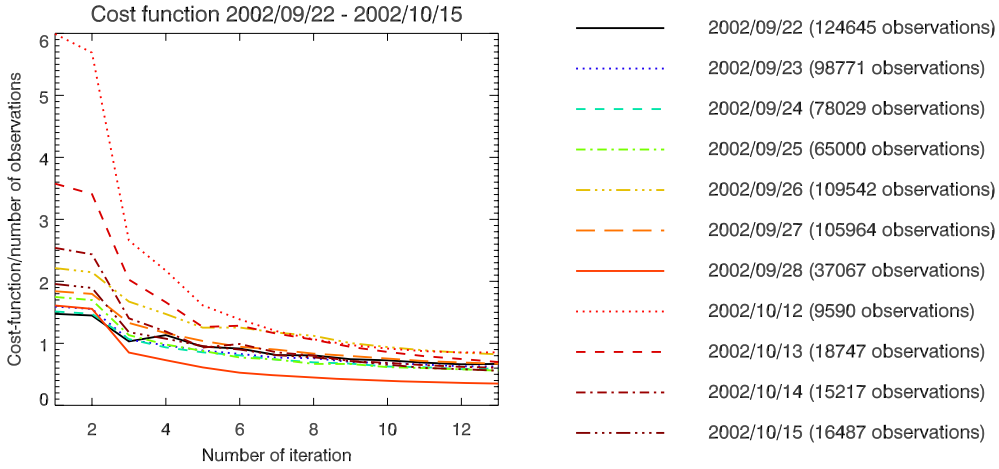

Figure 6.9: *Evolution of the normalised cost function (cost divided by the number of observations) for configuration* ***CS1-MPE-4****. The x-axis gives the iteration count, and the number of observations available at a particular day is displayed at the right hand side of each plot. Observe the lower J_p^f values compared to* ***CS1-MPE-1*** *shown in Figure 6.3 (bottom).*

6.3 Flow dependent BECM parameterisation

The benefit of the flow dependent, anisotropic and inhomogeneous background error covariance formulation is demonstrated by Figure 6.11. Assimilated ozone for 27 October 2003 at 28 hPa obtained with the isotropic BECM parameterisation (**CS3-MPE-2**, top panel) is shown together with the ozone analysis for **CS3-MPE-3** (bottom). The corresponding potential vorticity distribution can be found in Figure 6.10. Clearly, the gradient of the ozone field at the edge of the polar vortex remains steeper if the anisotropic BECM formulation is applied, and the collar of high ozone volume mixing ratios around the vortex does not show a noisy structure.

A second example taken from **CS1-MPE-4** and **CS1-MPE-5** is given in Figure 6.12. Again, an obvious feature of the analysis utilising the flow dependent background error covariances is the better representation of filament structures in the ozone field. Note that here, in contrast to **CS3-MPE-2/3**, the complete set of standard species has been assimilated on the full domain of the chemistry-transport module, which poses a greater challenge to the assimilation system. The relative difference of J_p^f values between the two configurations (Figure 6.14) reveals an overall positive impact of the anisotropic BECM scheme. The only notable exception is for 2 September (day 245) where J_p^f is about 10% larger for configuration **CS1-MPE-5**. This,

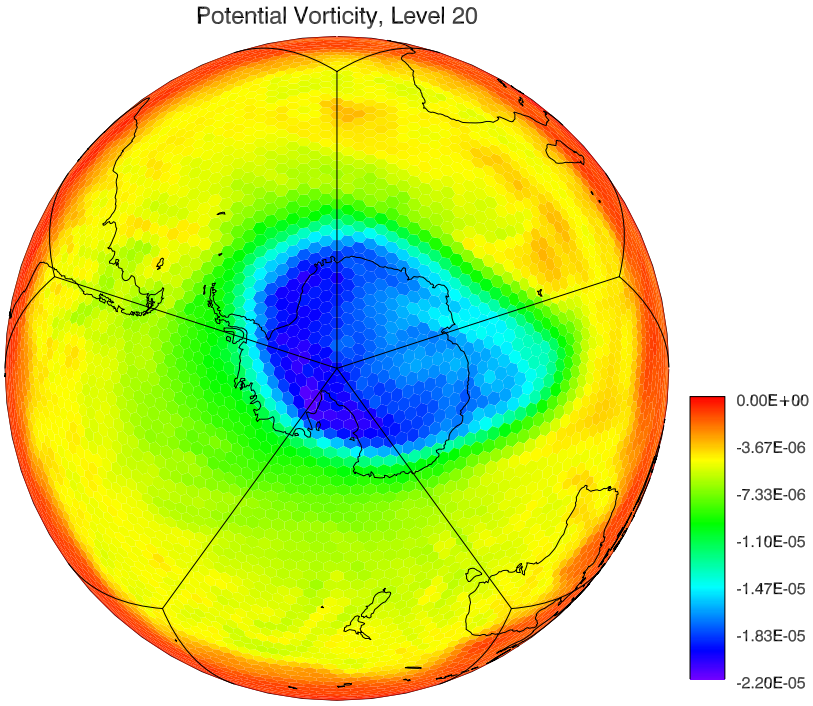

Figure 6.10: *Potential vorticity distribution* $[\frac{K\ m^2}{s\ kg}]$*, 27 October 2003 at 28 hPa.*

however, is not a deficiency of the anisotropic configuration, as at the second day of the spin-up assimilation a positive impact of the flow dependent BECM parameterisation cannot be expected. Here, the background field has still large errors with correlations, which are not governed by the model dynamics.

It should be noted that the estimate of parameters for the diffusion scheme is still relatively coarse: The basic correlation length scales L_h and L_v as well as the relative background error ϵ_b, are not varied within the model domain and the coordinate stretching factors $\tilde{s}_1$ and $\tilde{s}_2$ are calculated using a simple dependence on the gradient of potential vorticity. A closer inspection of J_p^{f} relative differences for individual species and different parts of the model domain shows that there is considerable variation. In Figure 6.15 the ozone J_p^{f} relative differences in the latitude bands $[-90^{\mathrm{o}}, -60^{\mathrm{o}}]$ and $[30^{\mathrm{o}}, 60^{\mathrm{o}}]$ are shown for different altitude regions. Between approximately 100 and 15 hPa (model levels 25–18) the current choice of parameters for the anisotropic diffusion scheme leads to a consistently positive impact at high southern latitudes. However, this is not the case for northern mid-latitudes where no significant changes are visible within this altitude region. Hence, further improvements of analysis quality can be expected, if the parameters L_h, L_v, ϵ_b, $\tilde{s}_1$ and $\tilde{s}_2$ are calculated location dependent and based on rigorous statistics.

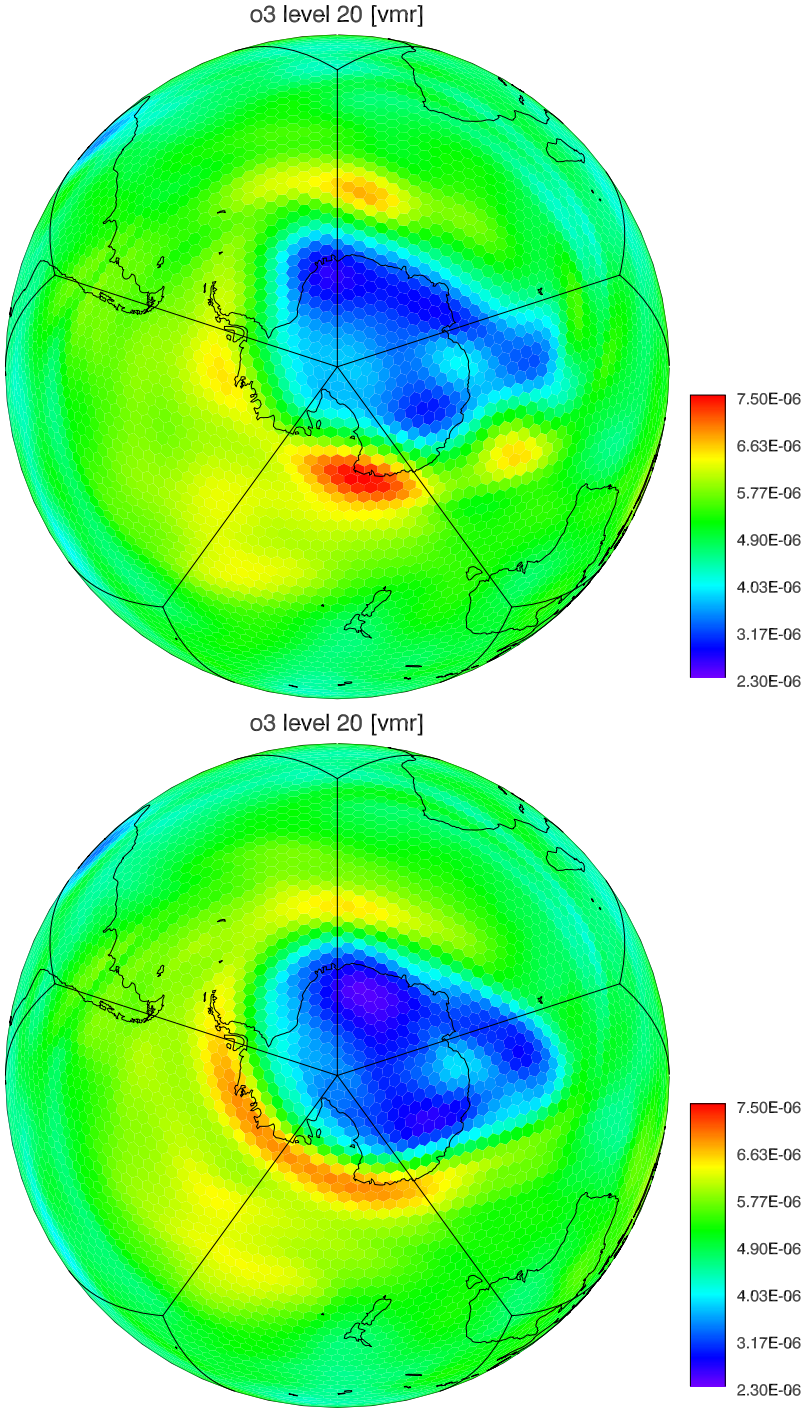

Figure 6.11: ***CS3-MPE-2*** *(top) and* ***CS3-MPE-3*** *(bottom) ozone analysis for 27 October 2003 at 28 hPa. The benefit of the anisotropic BECM formulation is clearly visible in the region around the polar vortex. The corresponding PV-distribution is shown in Figure 6.10.*

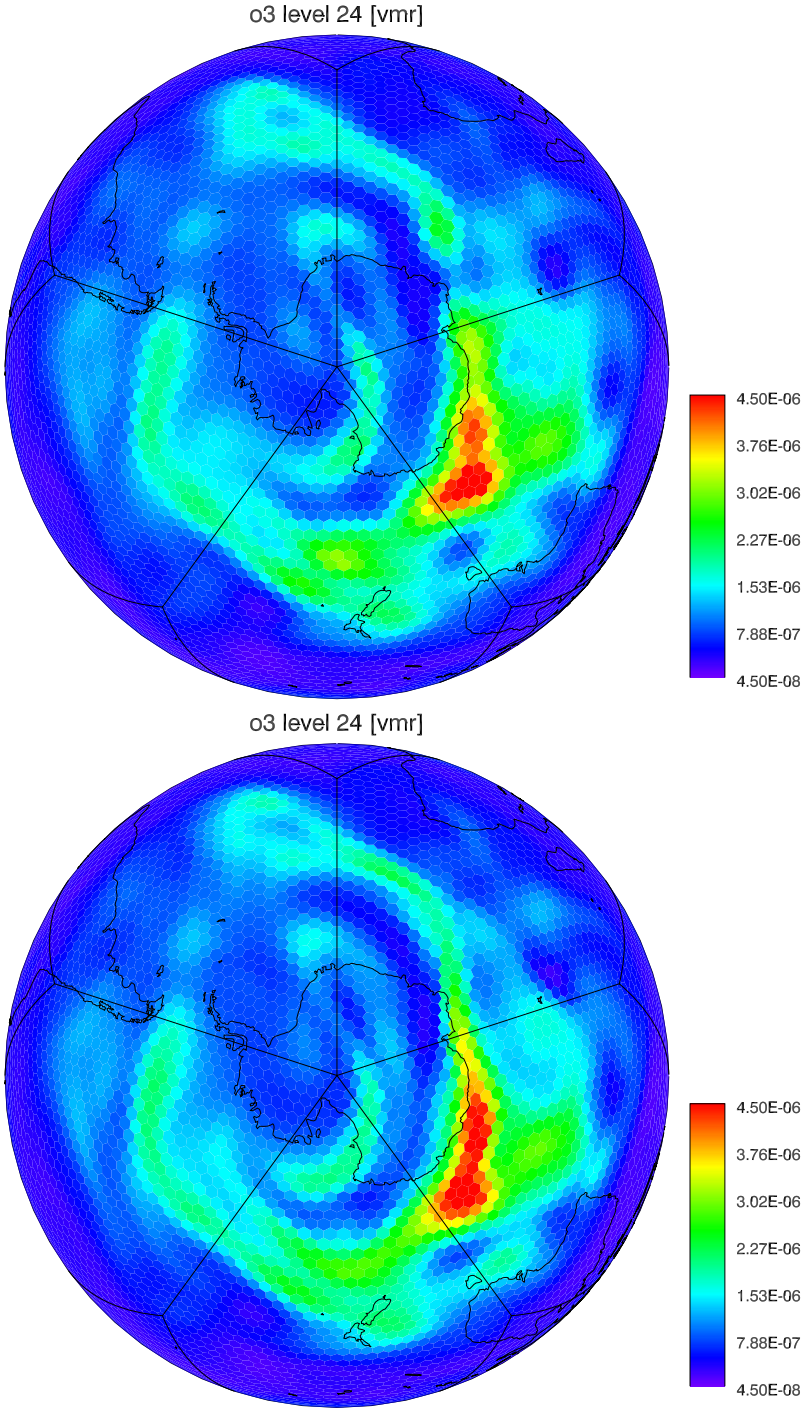

Figure 6.12: ***CS1-MPE-4*** *(top) and* ***CS1-MPE-5*** *(bottom) ozone analysis for 13 September 2002 at 80 hPa. The benefit of the anisotropic BECM formulation is clearly visible in the region around the polar vortex. The corresponding PV-distribution is shown in Figure 6.13.*

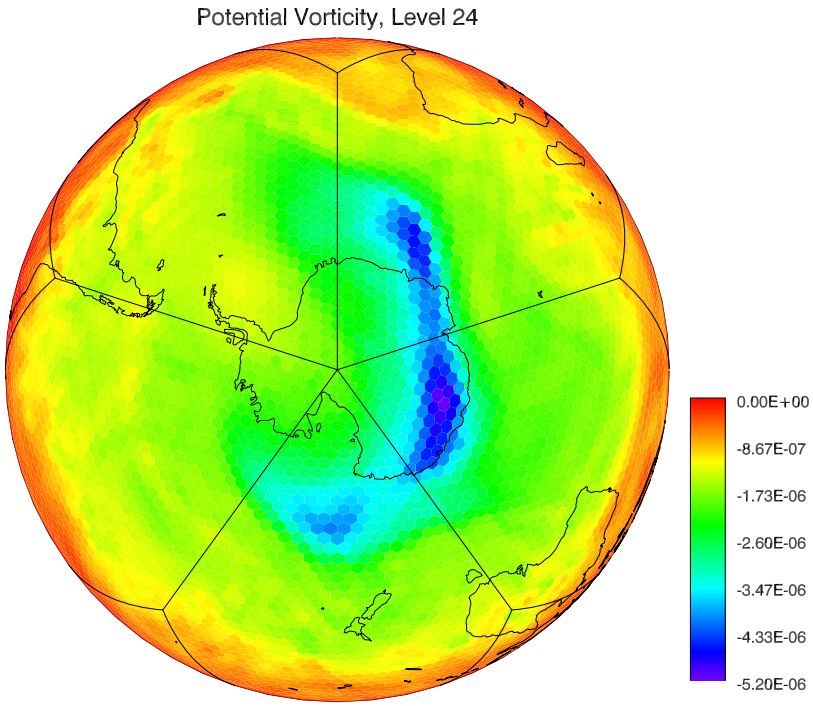

Figure 6.13: *Potential vorticity distribution* $[\frac{K\ m^2}{s\ kg}]$*, 13 September 2002 at 80 hPa.*

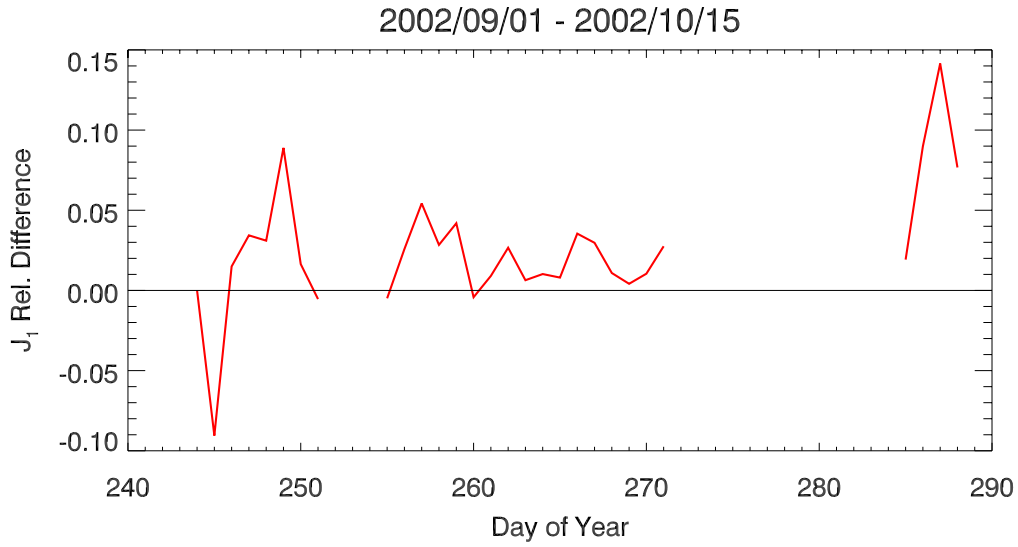

Figure 6.14: *Relative difference between the* J_p^{f} *value for* ***CS1-MPE-4*** *and* ***CS1-MPE-5****. A positive difference indicates that the* J_p^{f} *cost function value for the anisotropic scheme has been smaller than for the isotropic configuration.*

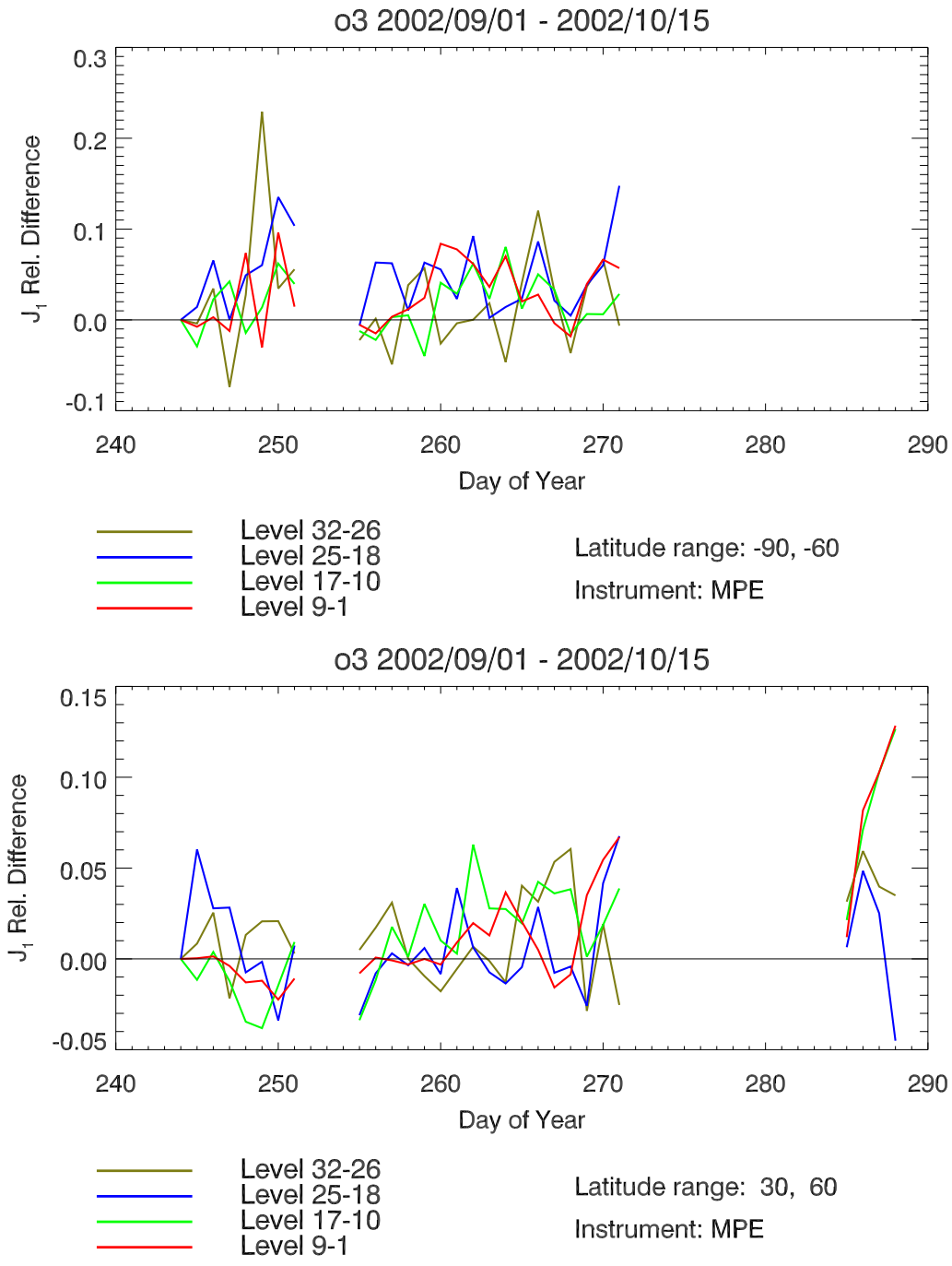

Figure 6.15: *Relative difference between the ozone J_p^{f} value for* ***CS1-MPE-4*** *and* ***CS1-MPE-5*** *in the latitude bands $[-90^o, -60^o]$ (top) and $[30^o, 60^o]$ (bottom). Positive differences indicates that the J_p^{f} cost function value for the anisotropic scheme has been smaller than for the isotropic configuration.*

6.4 Statistical evaluation and cross validation

Validation of assimilation results is presented in this setion. A detailed discussion is devoted to ozone, including the cross validation of ozone retrievals originating from different sensors. A statistical analysis of differences between background field and observations (O−B) is presented in order to assess the validity of the Gaussian assumption, which the derivation of the cost function (2.9) was based on.

6.4.1 Ozone

Ozone profiles, zonally averaged over five latitude bands are presented in Figures 6.16–6.21 for **CS1-MPI-1** and in Figures 6.22–6.26 for **CS2-MPE-1**. All profile data present within two adjacent model layers and the respective latitude range have been averaged to form the corresponding mean profile value. Note that the altitude intervals of 3–5 km, over which has been averaged, are comparable to the vertical resolution of the MIPAS instrument. The mean retrieved profiles are plotted together with the mean error margins, the corresponding analysis and the control run data. The grey shaded area marks the standard deviation of the ensemble of averaged observational profiles.

The mean analysis profiles show a perfect compliance with the MPE data up to approximately 3 hPa and a small bias towards lower volume mixing ratios above (Figure 6.22). The analysis produced with the MPI data product additionally shows a small positive bias in the region of the ozone volume mixing ratio maximum at tropical and mid latitudes, which is most pronounced between 30°S and 30°N (Figure 6.16). Note that the profile values are given in ppm. The ozone maximum in units of molecules/cm^3, usually referred to as the ozone layer, is located at approximately 30 hPa in the tropics and at about 90 hPa over the poles. The negative bias at the top of the model domain points to an inconsistency between the model and observational data. An evaluation of retrieved profiles with independent in-situ data is not possible at this altitude range and hence, it is not undoubtedly clear, wether the observed bias is due to systematic errors in the retrieved profiles or due to a bias in the assimilating model. However, as a matter of fact, the bias is not only present with respect to the assimilated data, but is also seen in comparison to the control data from HALOE (see Figures 6.18 and 6.24) and SAGE II (see Figures 6.19 and 6.25). An inspection of the time evolution of ozone analysis and control run biases, as shown in Figure 6.27, reveals similar results. Biases are completely removed by the assimilation procedure in the

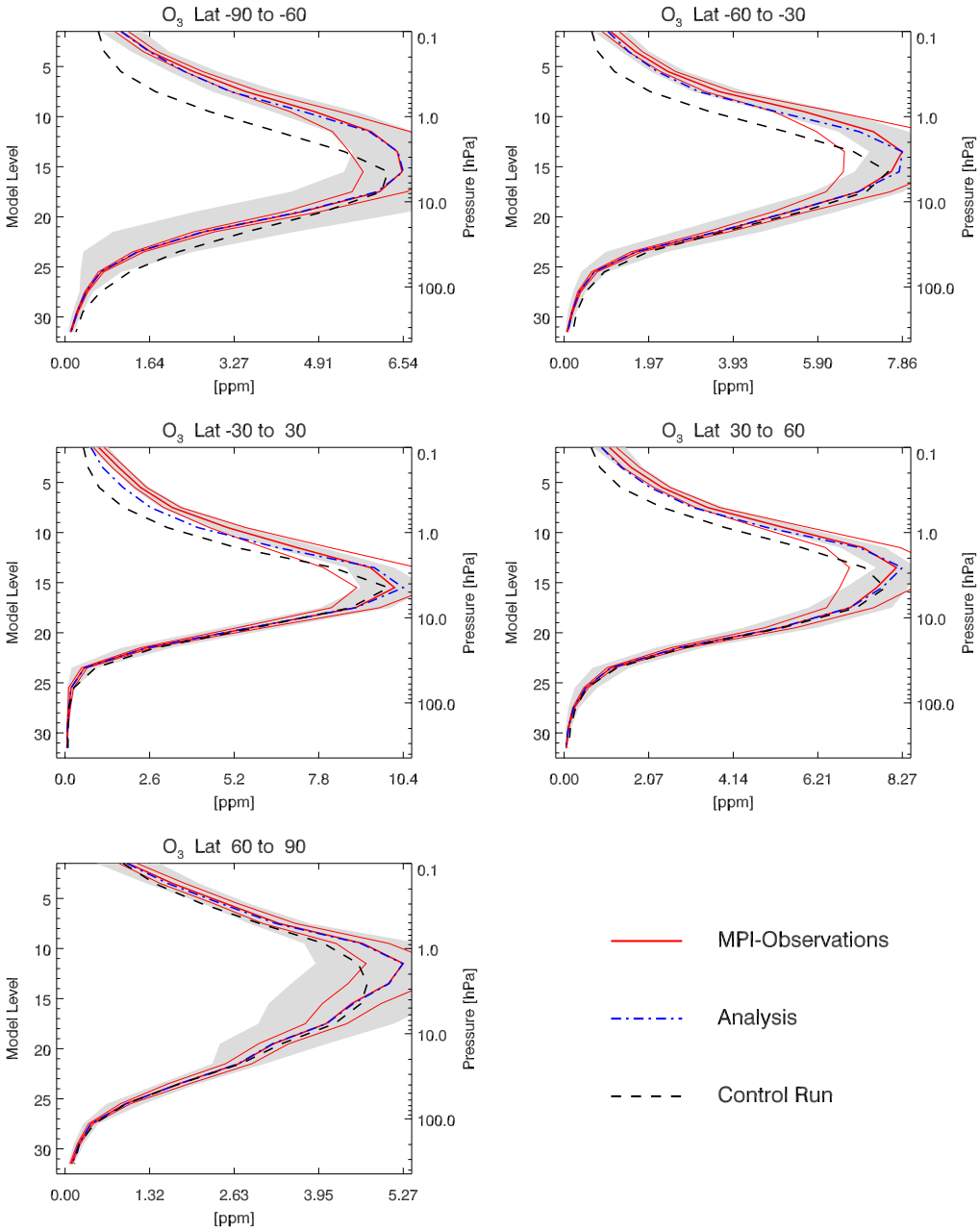

Figure 6.16: ***CS1-MPI-1*** *mean ozone profiles. MIPAS IMK retrieved profiles (assimilated) are given in red, the control run is the black dashed line, the analysis the dash-dotted blue line. The grey-shaded area indicates values within the standard deviation of observations.*

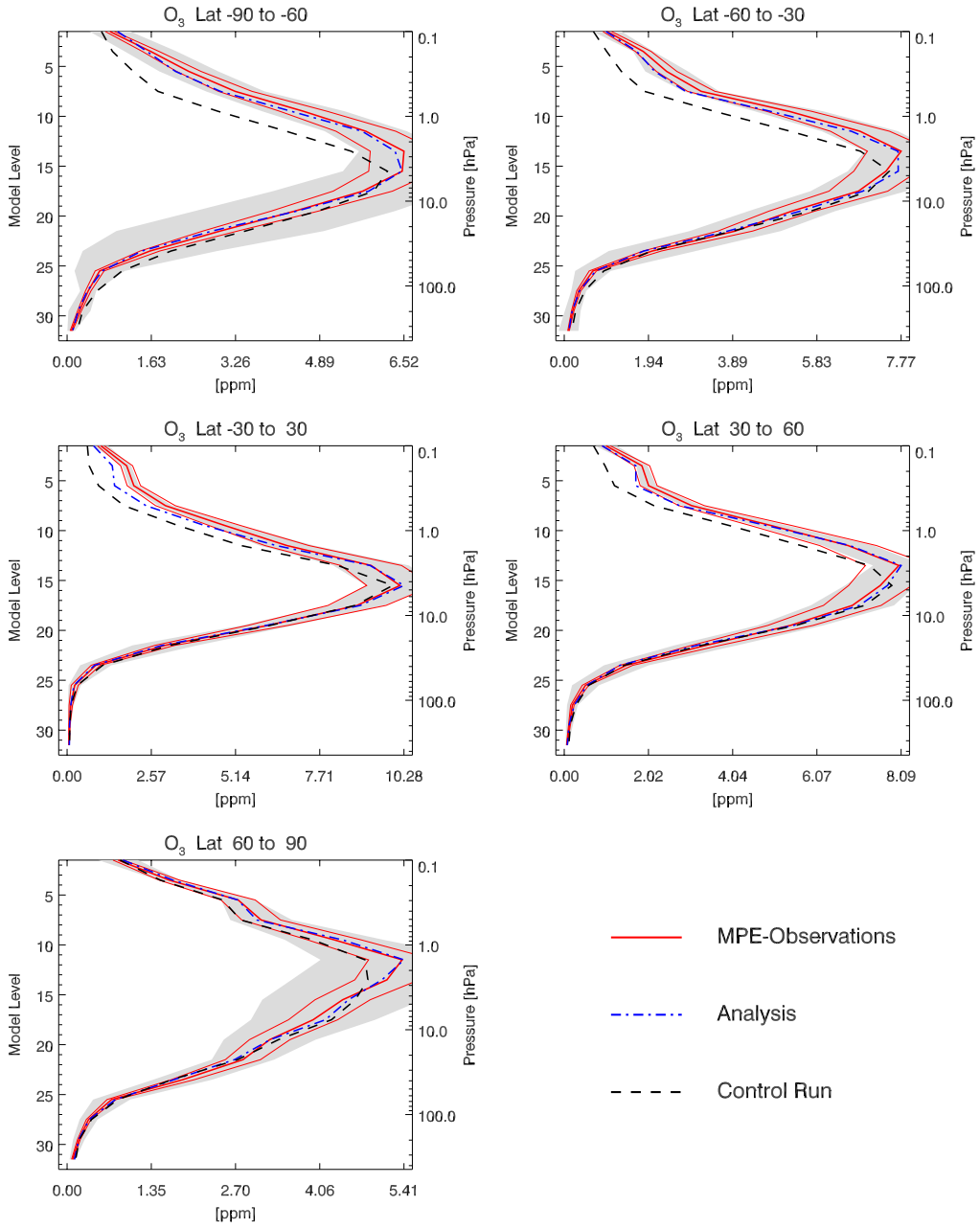

Figure 6.17: CS1-MPI-1 *mean MPE ozone profiles (not assimilated) compared to analysis and control run.*

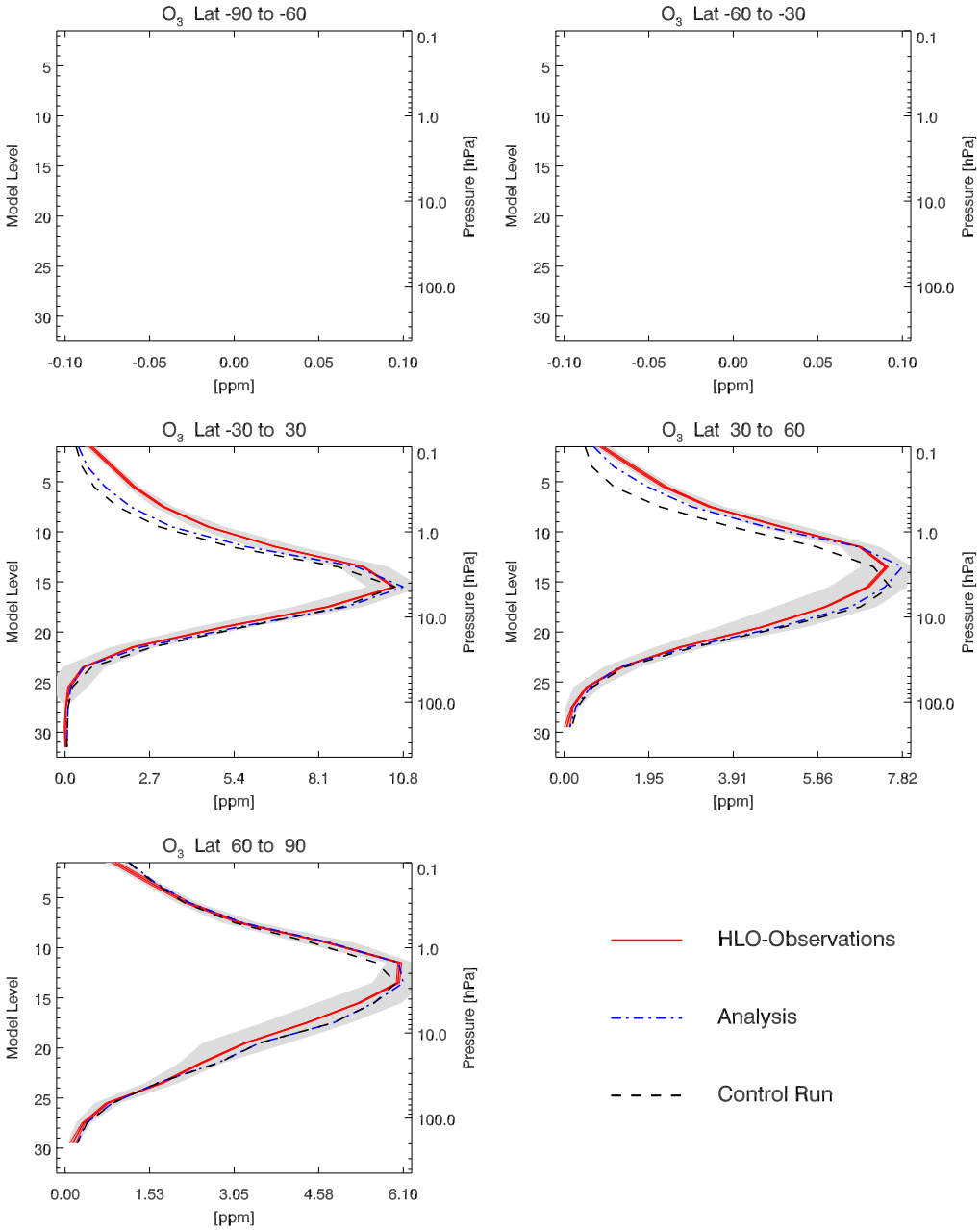

Figure 6.18: CS1-MPI-1 *mean HALOE ozone profiles (not assimilated) compared to analysis and control run. No HALOE observations took place in the latitude range* $[-90^o, -30^o]$ *during the* ***CS1*** *period.*

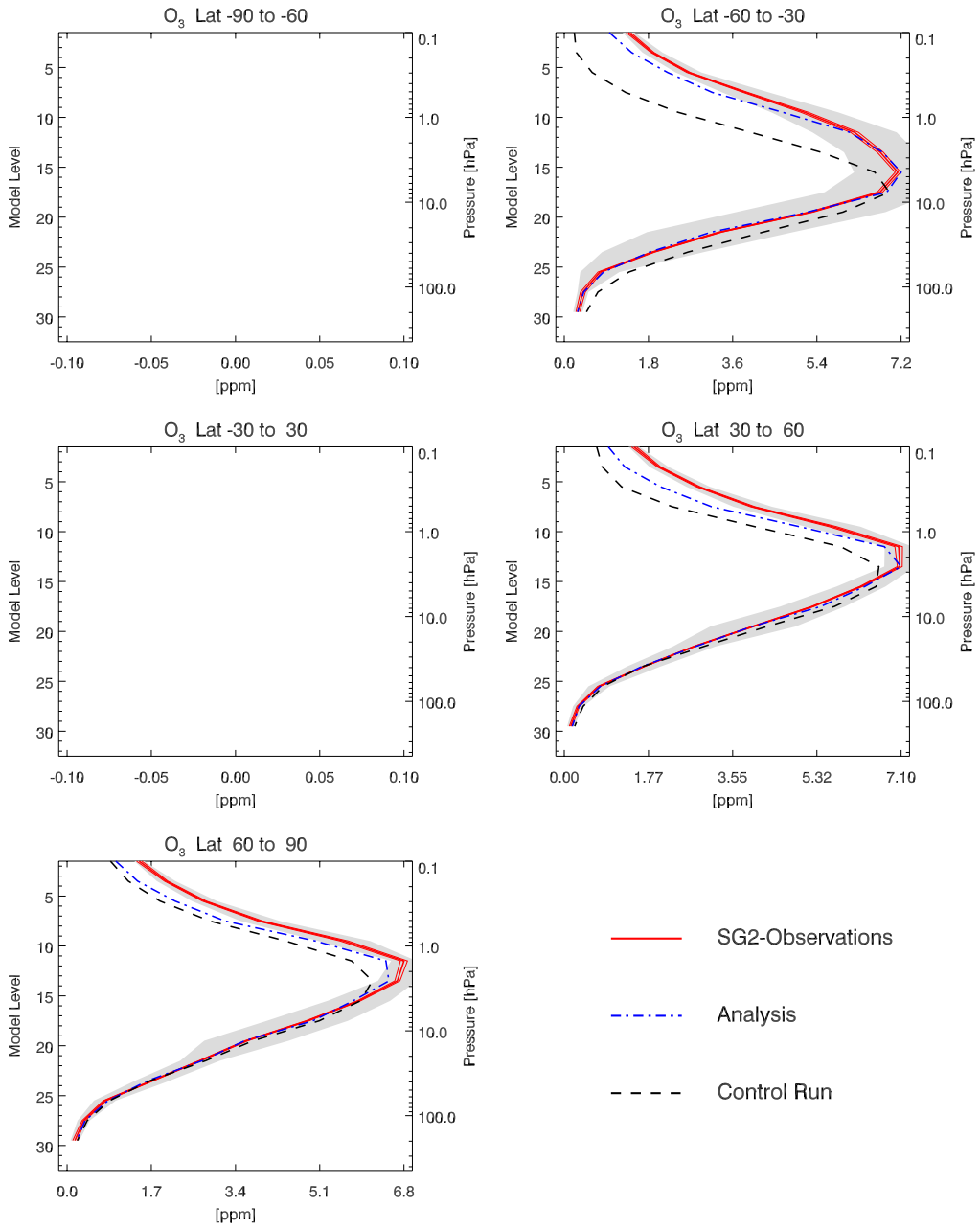

Figure 6.19: CS1-MPI-1 *mean SAGE II ozone profiles (not assimilated) compared to analysis and control run. No SAGE II observations took place in the latitude ranges* $[-90^o, -60^o]$ *and* $[-30^o, 30^o]$ *during the* ***CS1*** *period.*

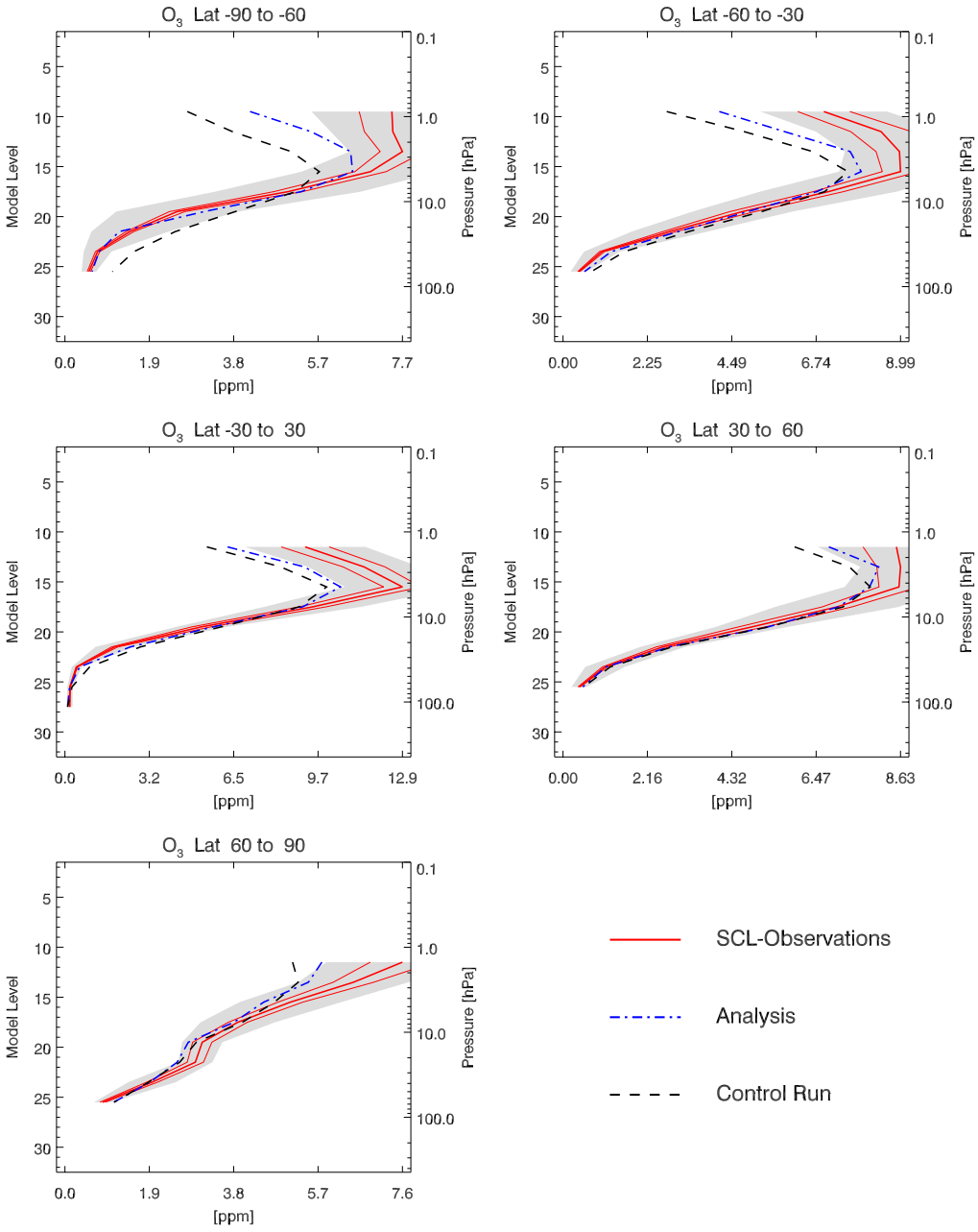

Figure 6.20: CS1-MPI-1 *mean Sciamachy limb ozone profiles (not assimilated) compared to analysis and control run.*

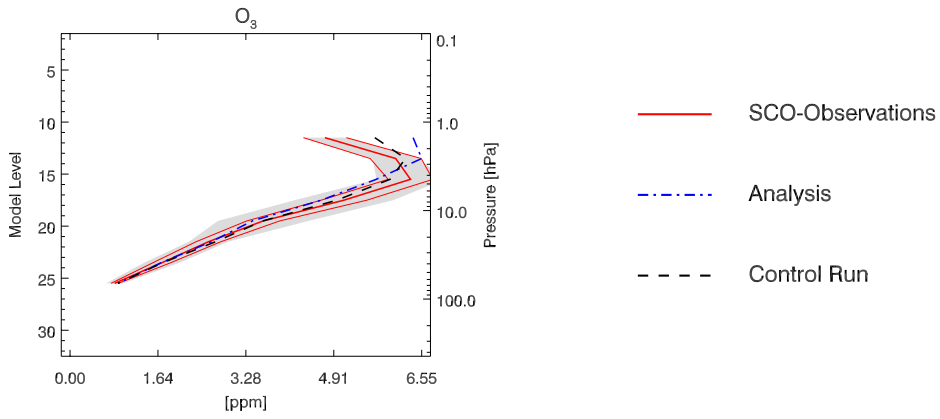

Figure 6.21: CS1-MPI-1 *mean SCO ozone profile (not assimilated) compared to analysis and control run. Note that SCIAMACHY solar occultaion measurements are confined to a latitude band between approximately* $50^o N$ *and* $70^o N$.

model domain covering the region of the ozone layer and clearly reduced in the altitude range above. However, as a bias remains in the *analysed* ozone field, a process, which proceeds on timescales shorter than one day –the length of the assimilation window used in this study– must be responsible. The photochemical lifetime of ozone decreases from about 3 years at 15 km altitude to about one day at 40 km in equatorial regions (*Seinfeld and Pandis* [1998]). The latter altitude corresponds to a pressure of approximately 3 hPa. Therefore, it can be suspected that the photochemical equilibrium is represented not entirely correct by the model above about 3 hPa.

Since the assimilation system shows a fully satisfying performance, it can be concluded that the assimilated ozone fields can be used to cross-validate different instruments at least up to about 3 hPa. Mean profiles of the HALOE, SAGE II and SCIAMACHY instruments have been compared to analyses obtained with MIPAS MPE and MPI data products. There is a reasonable agreement between the assimilated ozone fields and the HALOE profiles with the tendency of the MIPAS analyses to higher ozone volume mixing ratios (Figures 6.18 and 6.24). The maximum differences are found around the peak of the ozone volume mixing ratio, in good accordance with the results of *Wang et al.* [2005], who compared collocated HALOE and MPI profiles for the September/October 2002 period. They reported maximum differences of 0.4 ppm at the ozone volume mixing ratio maximum in tropical latitudes. An even better compliance can be observed between SAGE II mean ozone profiles and the corresponding MIPAS analyses based on either the MPI or the MPE data product up to an pressure of about 3 hPa (Figures 6.19 and 6.25).

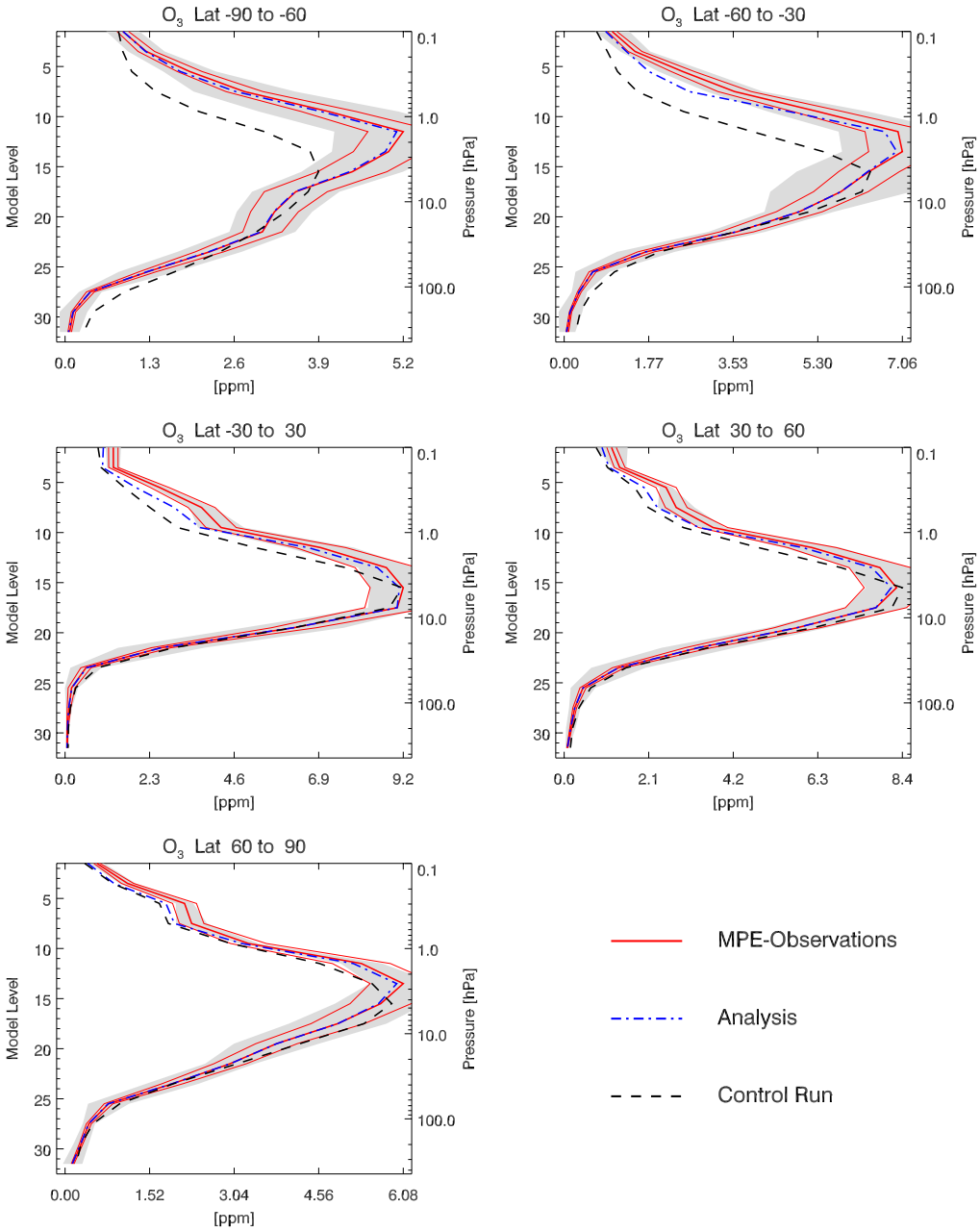

Figure 6.22: CS2-MPE-1 *mean ozone profiles. Mipas MPE retrieved profiles (assimilated) are given in red, the control run is the black dashed line, the analysis the dash-dotted blue line. The grey-shaded area indicates values within the standard deviation of observations.*

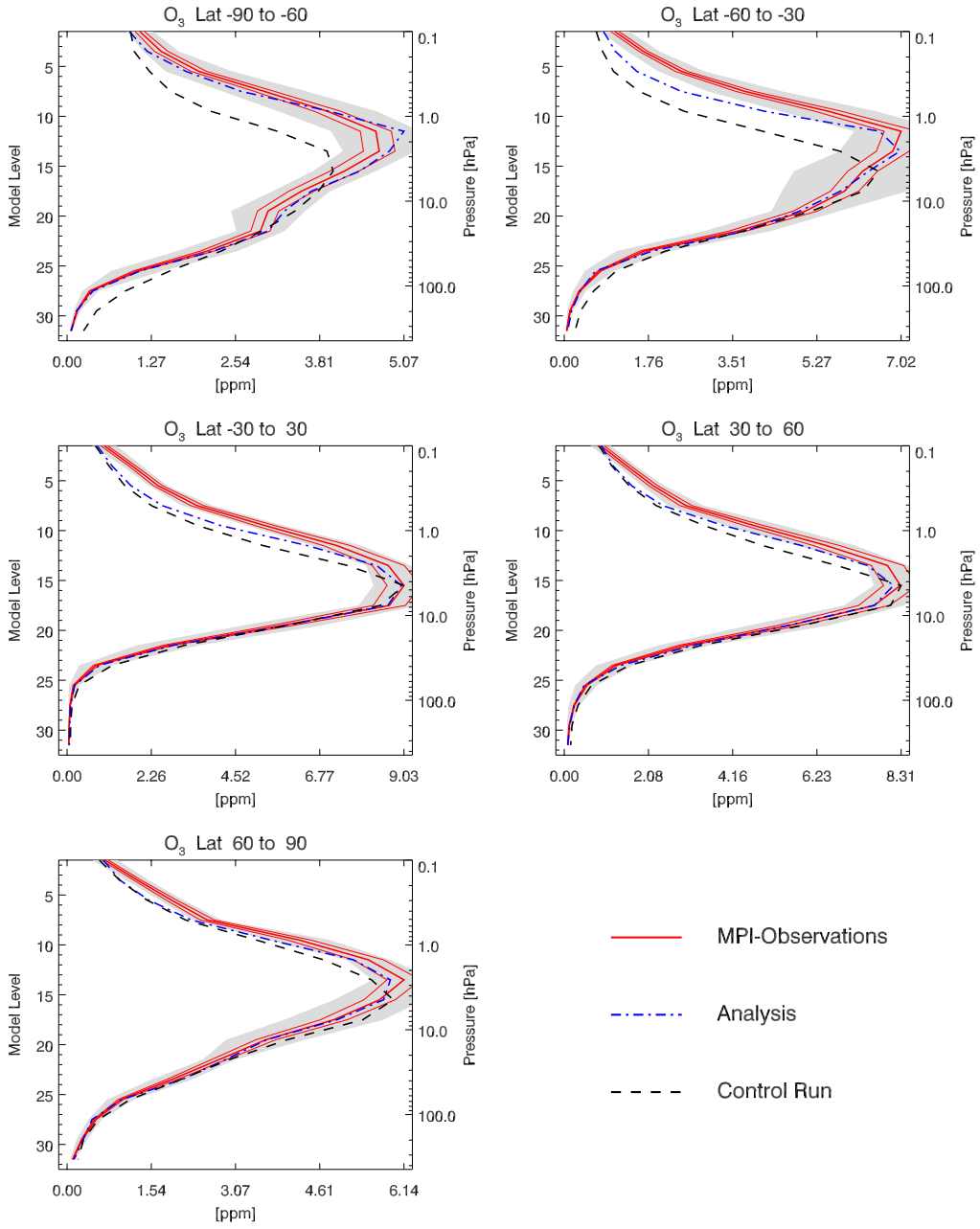

Figure 6.23: CS2-MPE-1 *mean MPI ozone profiles (not assimilated) compared to analysis and control run.*

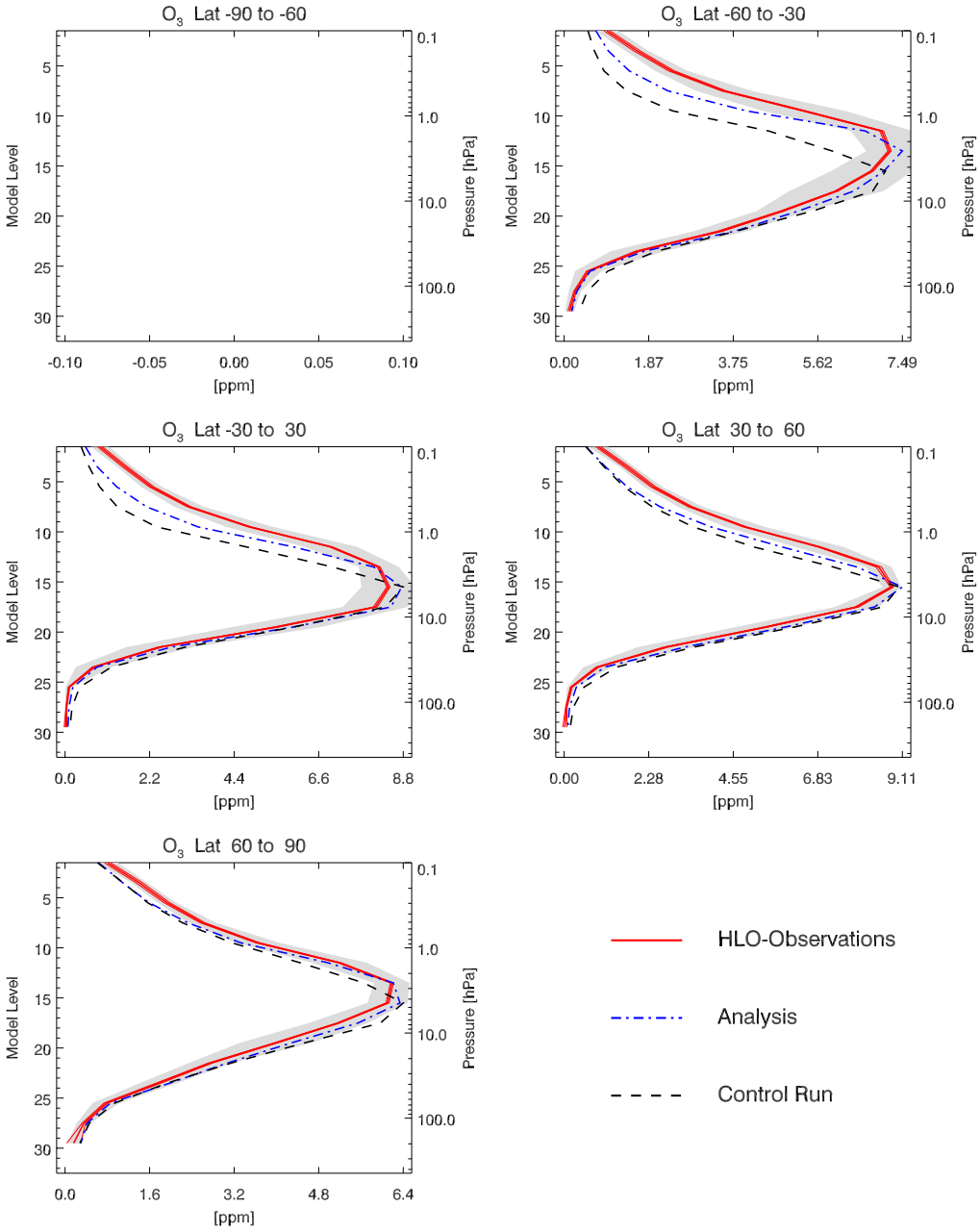

Figure 6.24: CS2-MPE-1 *mean HALOE ozone profiles (not assimilated) compared to analysis and control run. No HALOE observations took place in the latitude range* $[-90^o, -60^o]$ *during the* ***CS2*** *period.*

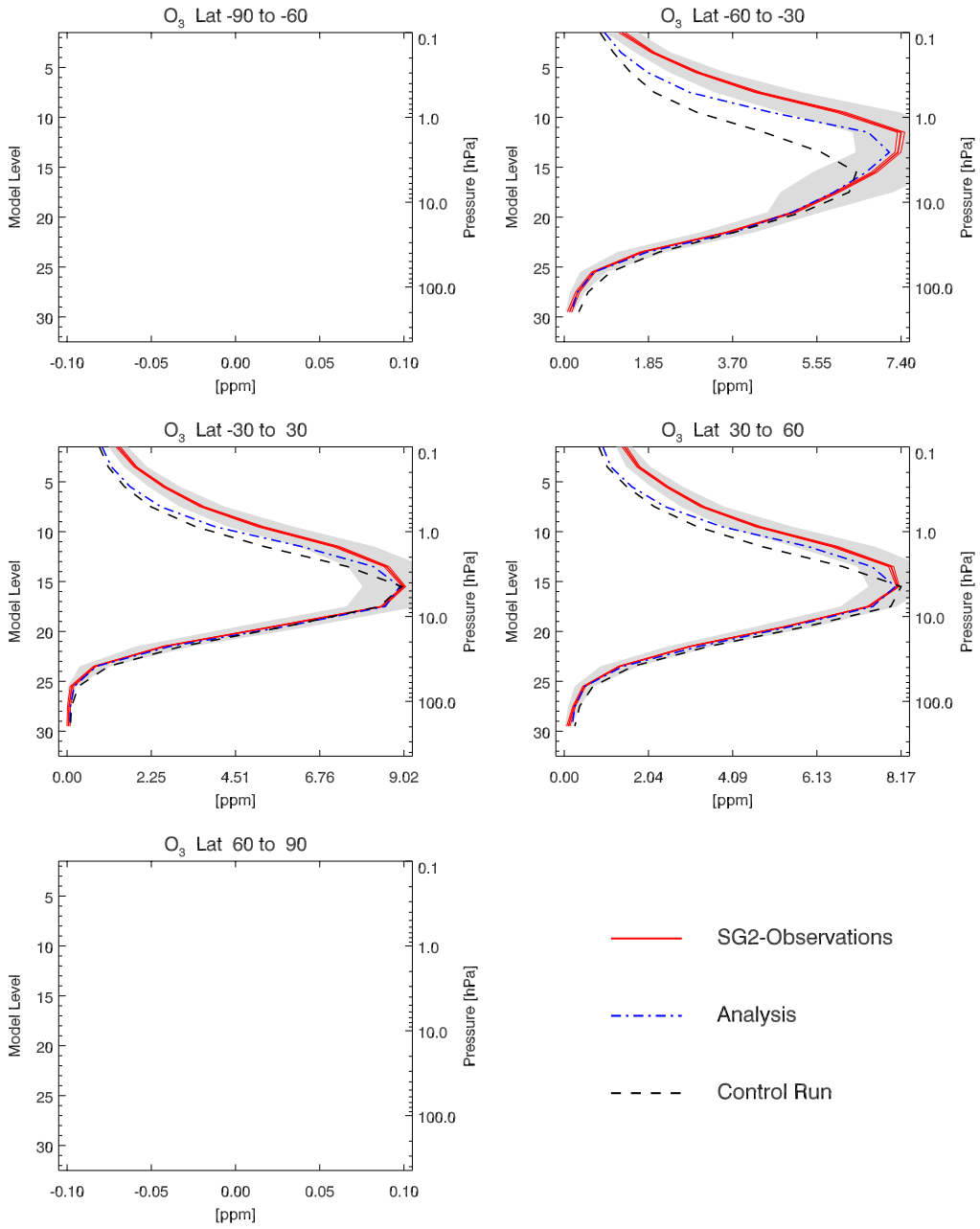

Figure 6.25: CS2-MPE-1 *mean SAGE II ozone profiles (not assimilated) compared to analysis and control run. No SAGE II observations took place in the latitude ranges* $[-90^o, -60^o]$ *and* $[60^o, 90^o]$ *during the* ***CS2*** *period.*

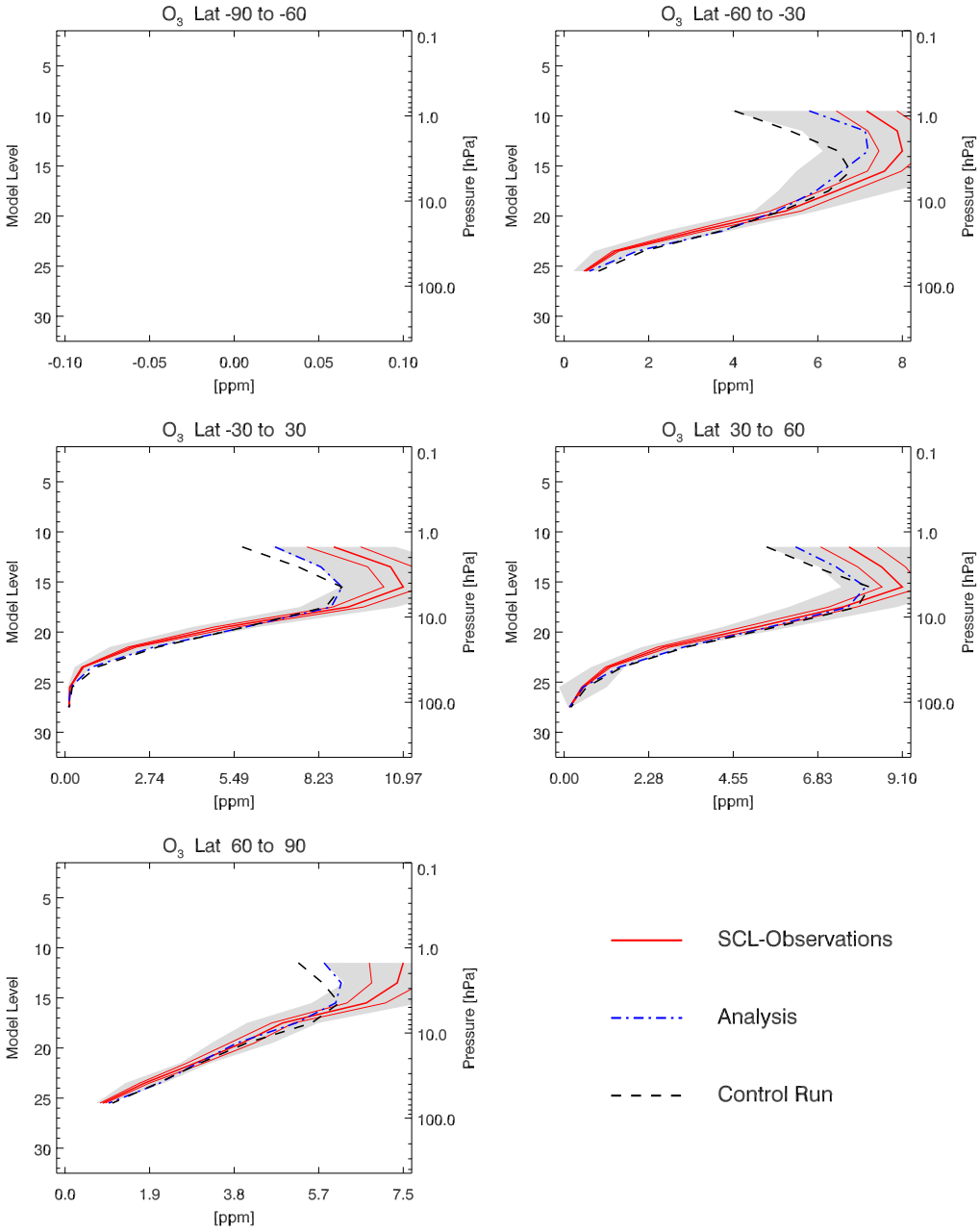

Figure 6.26: ***CS1-MPE-1*** *mean Sciamachy limb ozone profiles (not assimilated) compared to analysis and control run. No SCIAMACHY limb observations took place in the latitude range* $[-90^o, -60^o]$ *during the* ***CS2*** *period.*

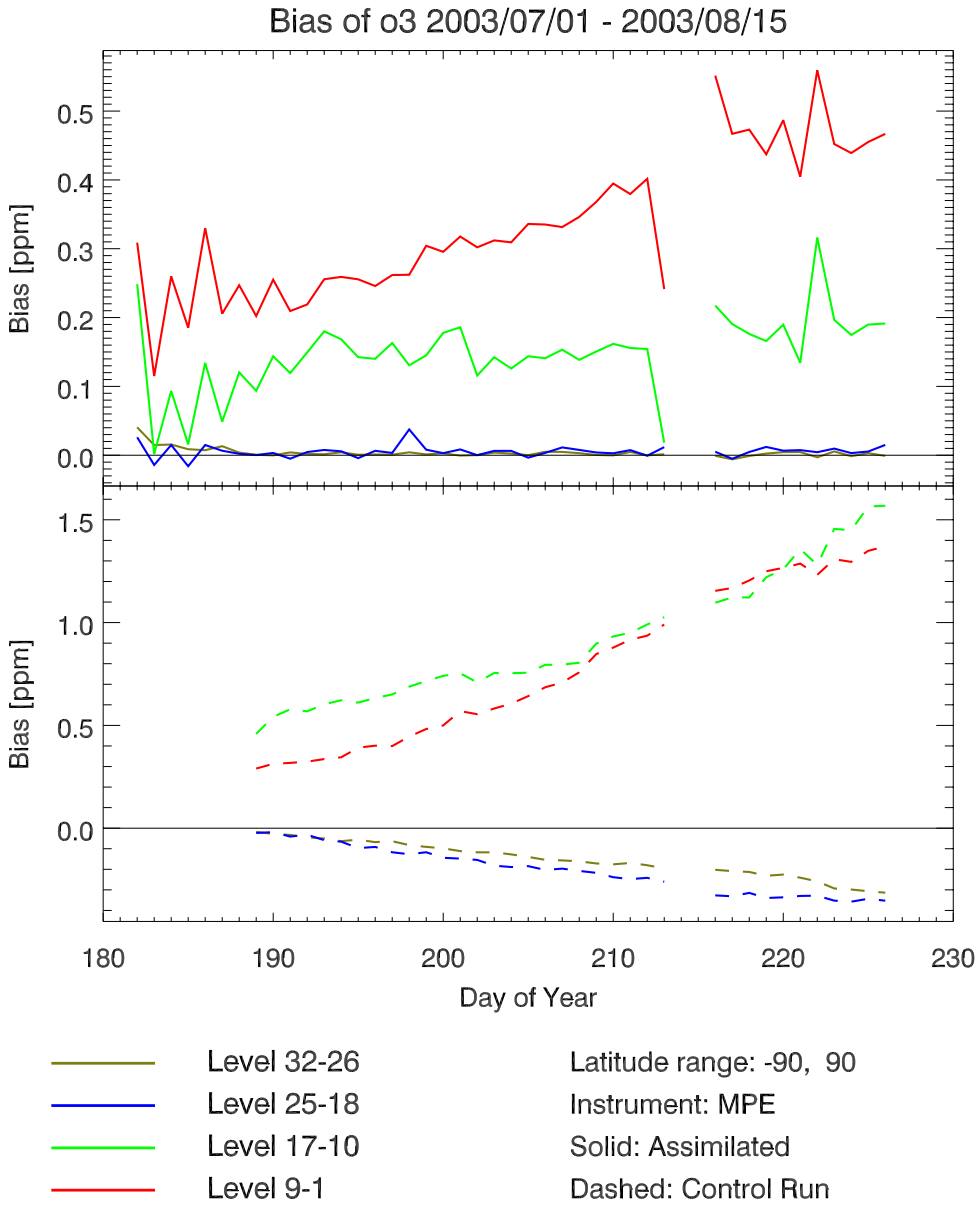

Figure 6.27: *Bias of ozone with respect to MPE observations for* ***CS2-MPE-1****. Solid lines (top panel) show the bias of O_3 analyses for each day of the* ***CS2*** *period against the assimilated MPE observations for different height ranges. Dashed lines (bottom panel) give the control run bias.*

Larger differences with respect to the analysed mean O_3 profiles are visible in the case of the Sciamachy limb retrival SCL, as shown in Figures 6.20 and 6.26, especially above approximately 10 hPa. Here, the retrived ozone mixing ratios are significantly higher than the analysed fields with differences up to 3 ppm in the region of the maximum ozone volume mixing ratio. The Sciamachy occultation data product shows an ozone peak of about the same magnitude as the analysed mean profiles, but shifted downwards by approximately 5 km (Figure 6.21).

6.4.2 CH_4, N_2O, NO_2, H_2O and HNO_3

Mean profiles of the remaining MIPAS standard species are shown in Figures 6.28–6.34 for the configuration **CS3-MPI-1**. Mean assimilated profiles of CH_4, N_2O and HNO_3 (see Figures 6.28, 6.29, and 6.30, respectively) perfectly match the observation mean profiles. HNO_3 analyses show a small negative bias relative to the observations in the region between 1 and 3 hPa at tropical and mid latitudes.

H_2O analyses, shown in Figure 6.31, are generally very good in the lower and middle stratosphere and show a slight positive bias against the MPI observations above approximately 5 hPa. In the tropopause region and below, water vapour has a very large natural variability and space borne remote sounding instruments have a weak sensitivity, leading to large error bars there. Consequently, the H_2O observational mean profiles show a large standard deviation in this region and the mean analysis profiles stay close to the control run. Note that the SACADA system in its current set-up has not been intended to analyse water vapour in the tropopause region. A higher horizontal and vertical model resolution has to be used for a fully satisfying H_2O-analysis in this altitude range.

Mean profiles of NO_2 (see Figure 6.32) have been additionally separated into nighttime and daytime profiles, as NO_2 shows a strong diurnal cycle with higher values occurring at night. Note that the sunset at the top of the model domain (65 km) occurs at zenith angles of about 100°. To exclude twilight observations, a nighttime profile was defined to have a zenith angle greater than 110° while a daytime observation was assumed at zenith angles smaller than 90°. The assimilation of highly reactive species is generally more challenging than the assimilation of more inert gases, as the linear approximation, which is implied by the use of the adjoint model, may become invalid, especially if an observation took place during daylight and the analysis at the beginning of the assimilation window is valid for local nighttime conditions, or vice versa. Nevertheless, the analyses of both, night- and daytime NO_2 show a good agreement with observations, as can be seen in Figure 6.33 and Figure 6.34, respectively. A notable exception is the nighttime profile at high northern latitudes, where the mean retrieved NO_2 profile shows substantially larger values than the analyses above 3 hPa. The reason for this discrepancy can be identified to be the solar proton event. An inspection of the time evolution of the NO_2 analysis bias, shown in Figure 6.35, reveals that the bias begins to rise at 27 October 2003, exactly at the time where the first solar proton event was recorded (see *López-Puertas et al.* [2005*a*]). It can be demonstrated (see Figure 6.36) that a corresponding bias is absent

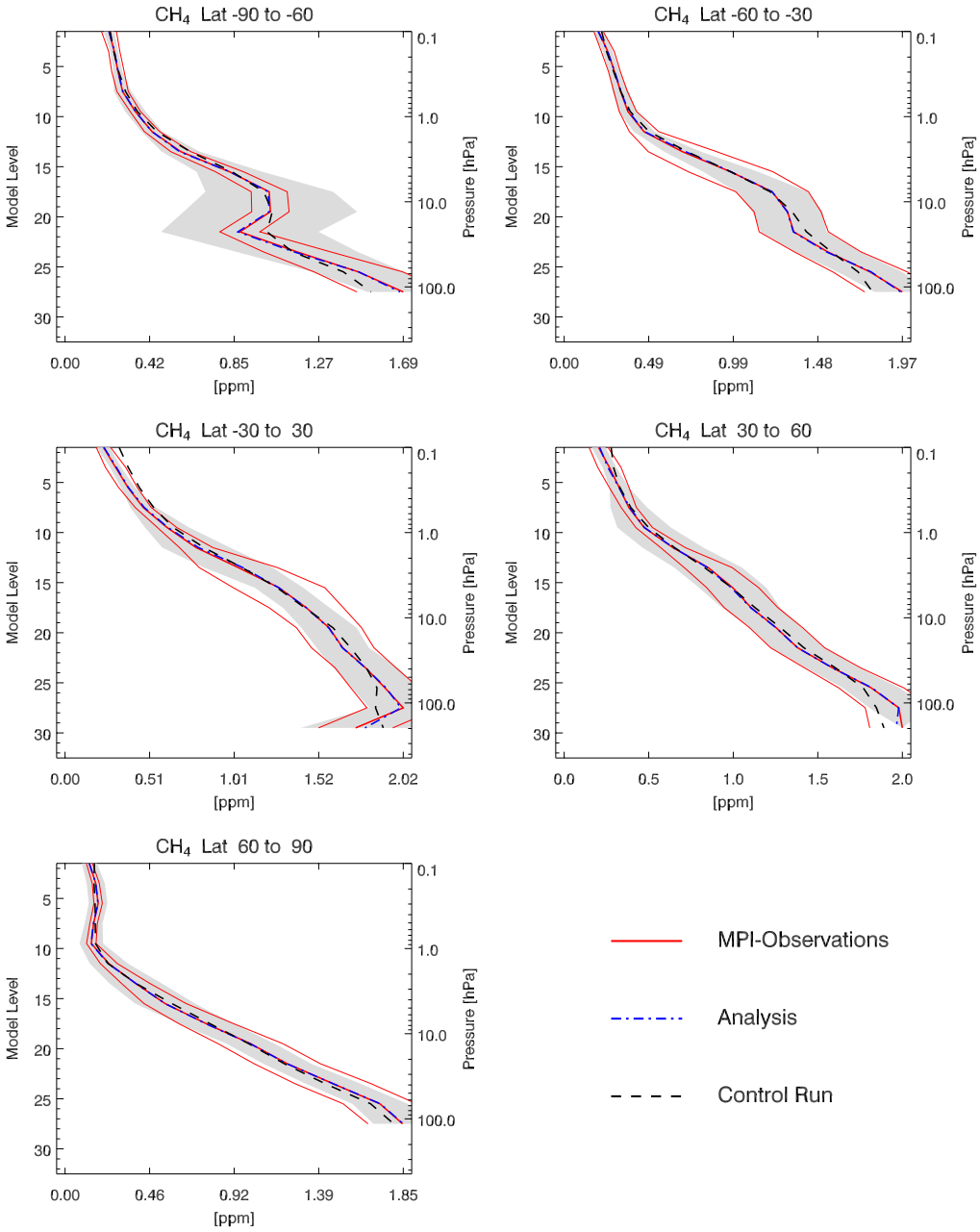

Figure 6.28: ***CS3-MPI-1*** *mean methane profiles. MIPAS IMK retrieved profiles (assimilated) are given in red, the control run is the black dashed line, the analysis the dash-dotted blue line. The grey-shaded area indicates values within the standard deviation of observations.*

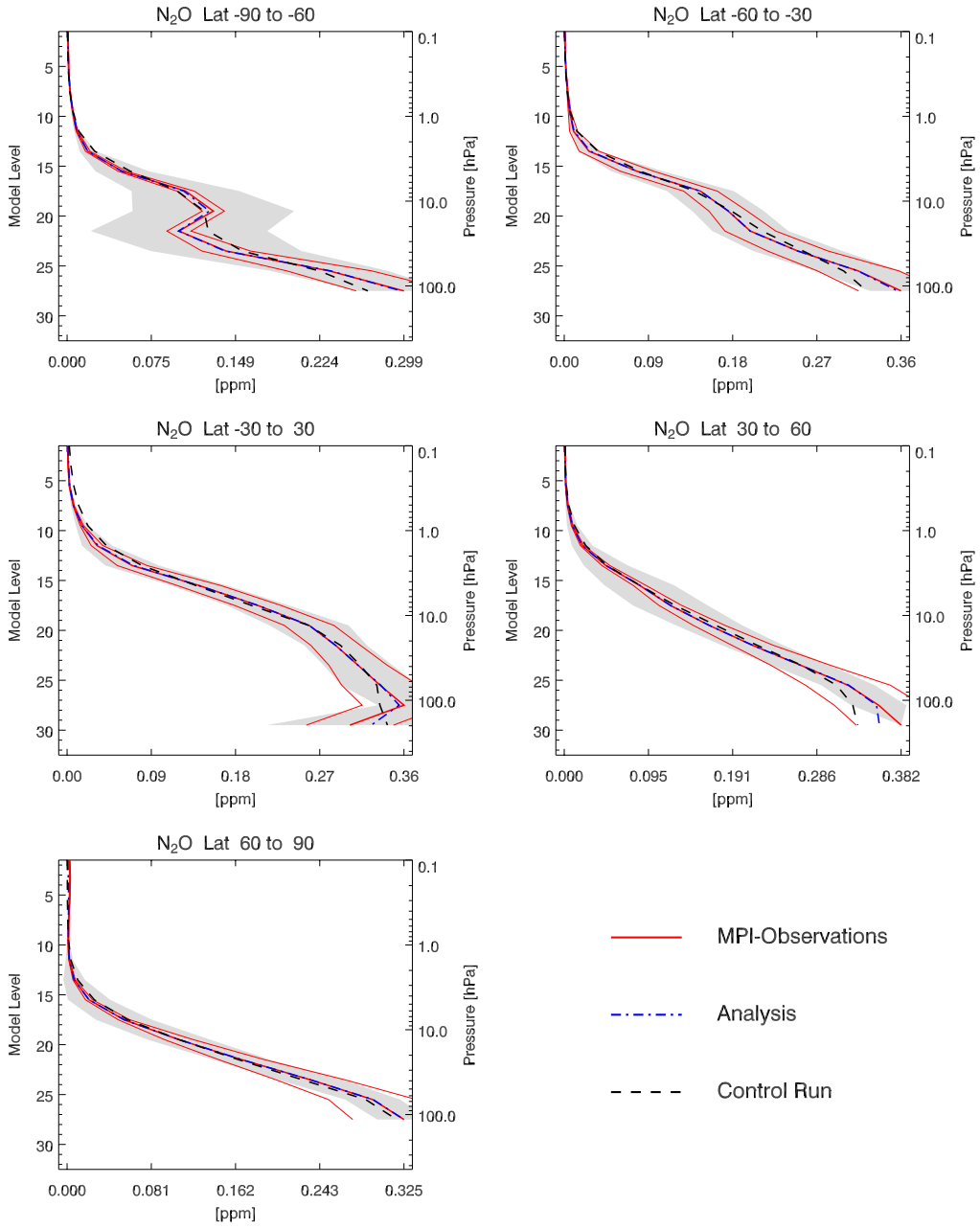

Figure 6.29: ***CS3-MPI-1*** *mean MPI N_2O profiles (assimilated) compared to analysis and control run.*

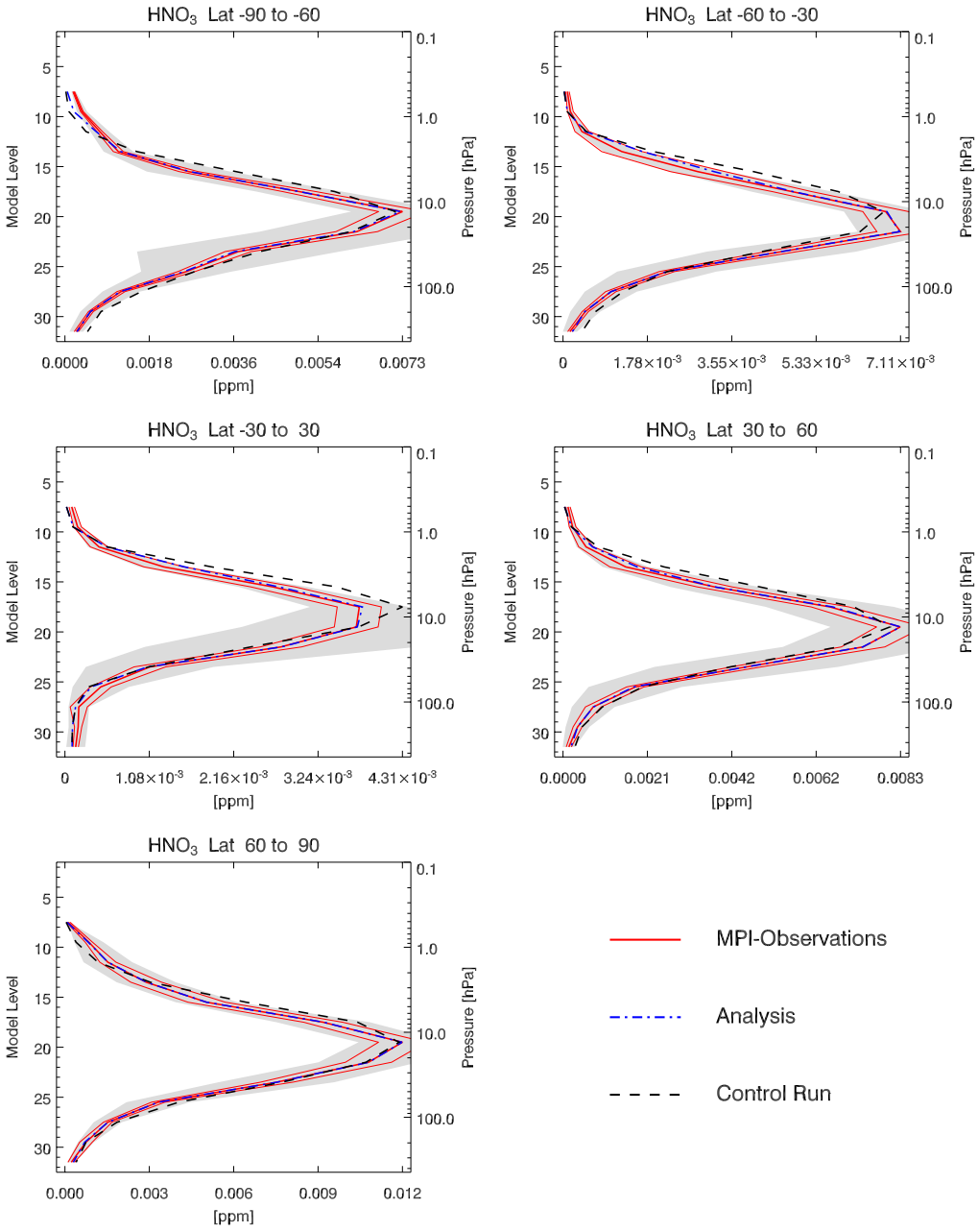

Figure 6.30: ***CS3-MPI-1*** *mean MPI HNO_3 profiles (assimilated) compared to analysis and control run.*

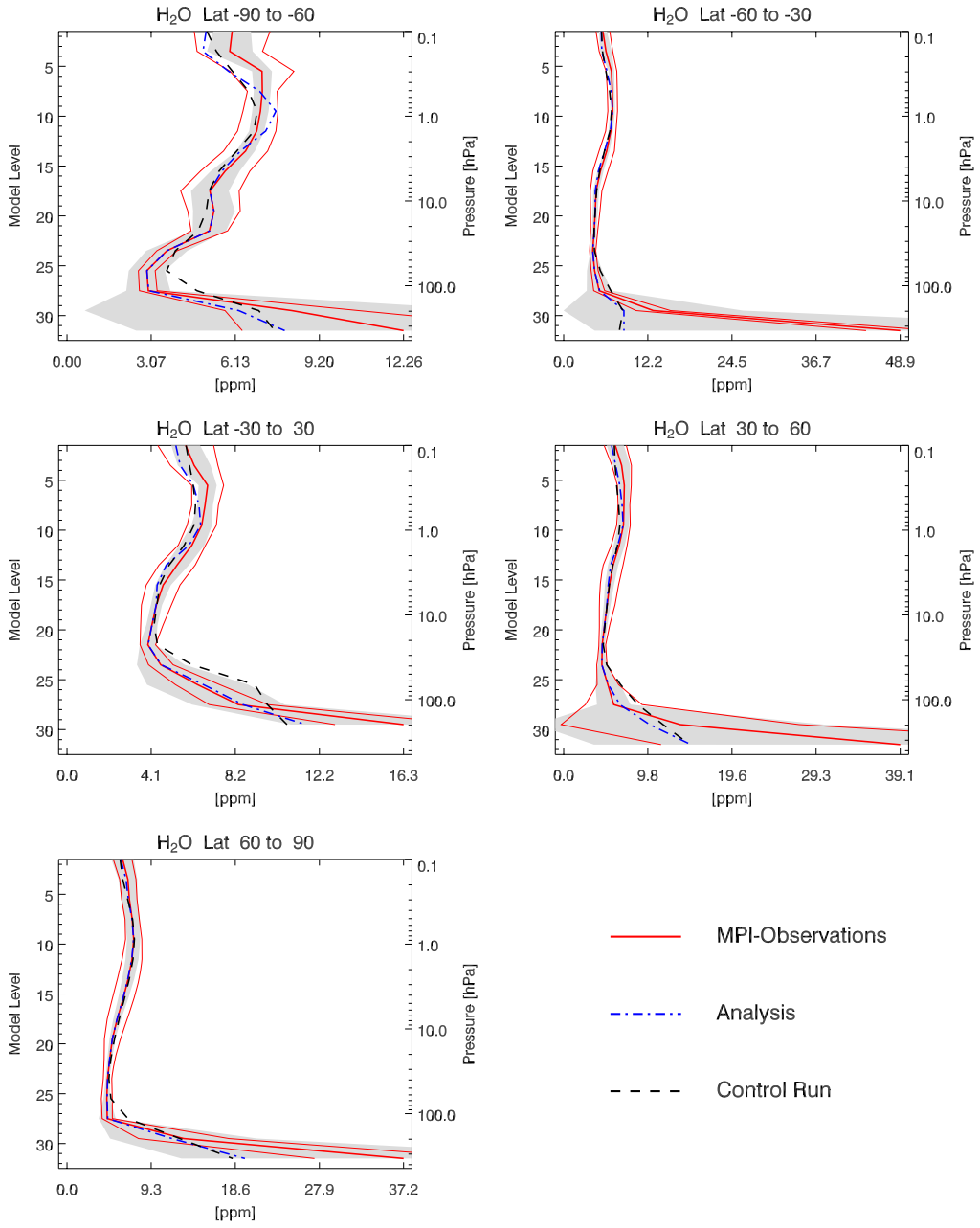

Figure 6.31: ***CS3-MPI-1*** *mean MPI H_2O profiles (assimilated) compared to analysis and control run.*

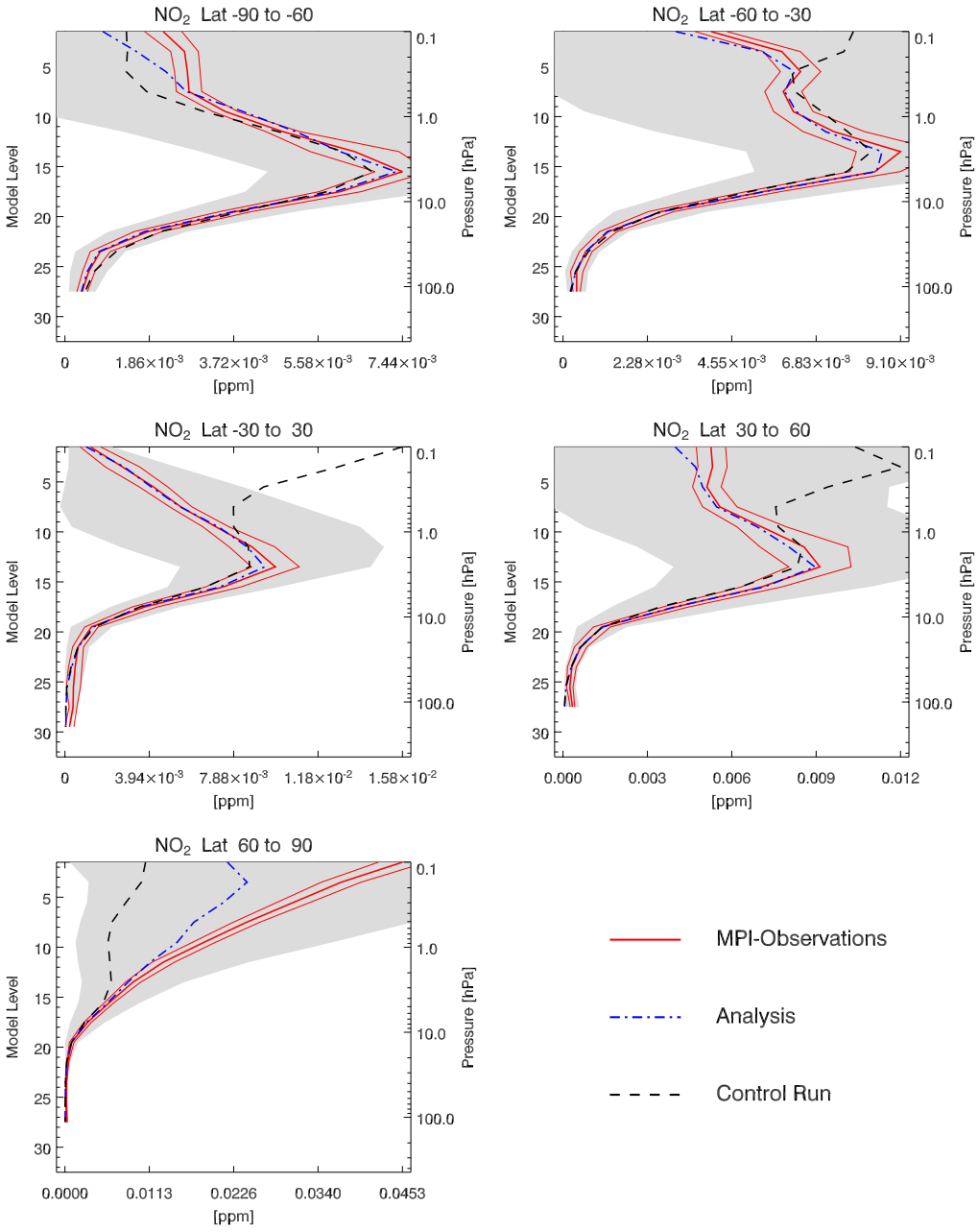

Figure 6.32: ***CS3-MPI-1*** *mean MPI NO_2 profiles (assimilated) compared to analysis and control run.*

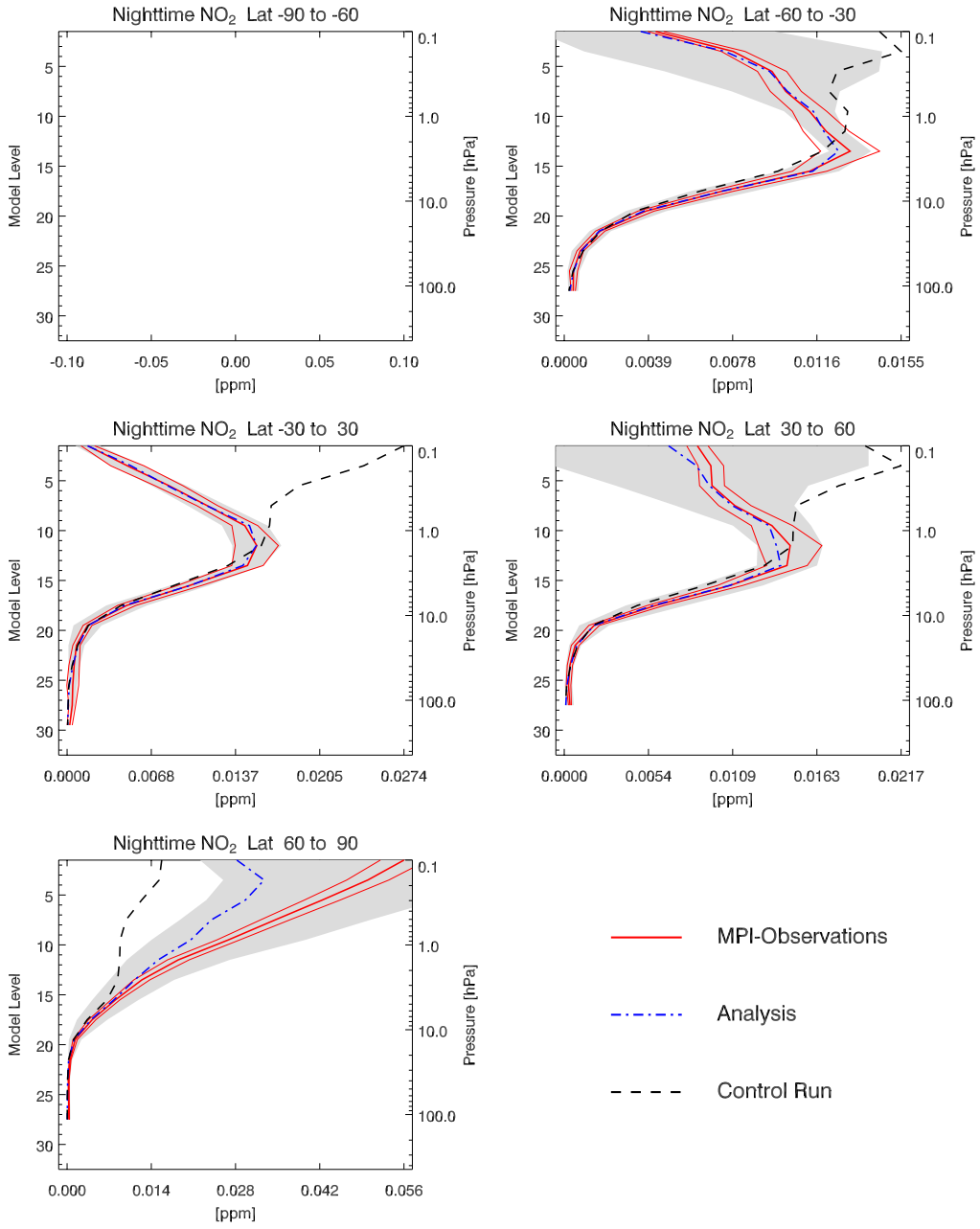

Figure 6.33: ***CS3-MPI-1*** *mean MPI nightime NO_2 profiles (assimilated) compared to analysis and control run.*

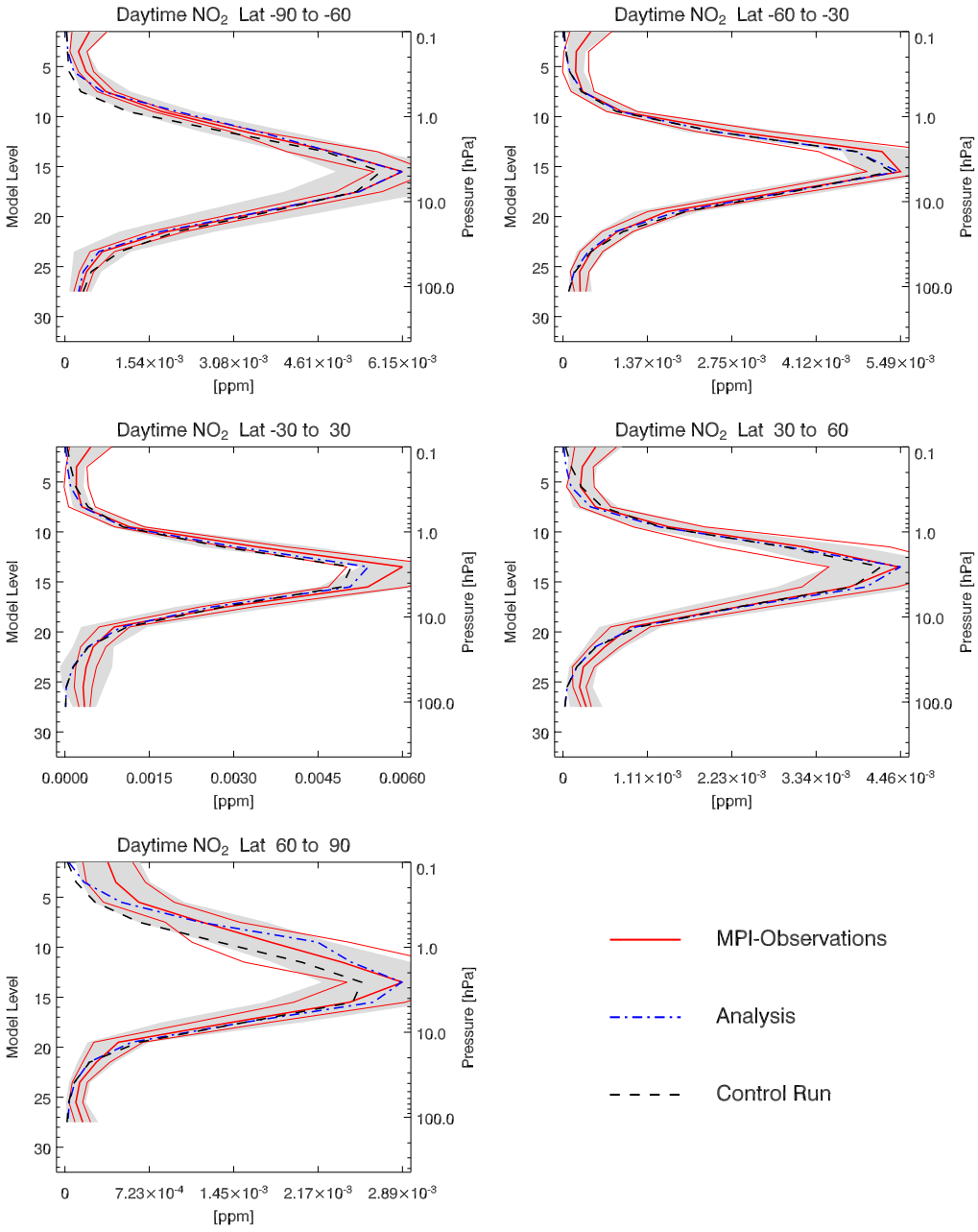

Figure 6.34: ***CS3-MPI-1*** *mean MPI daytime NO_2 profiles (assimilated) compared to analysis and control run.*

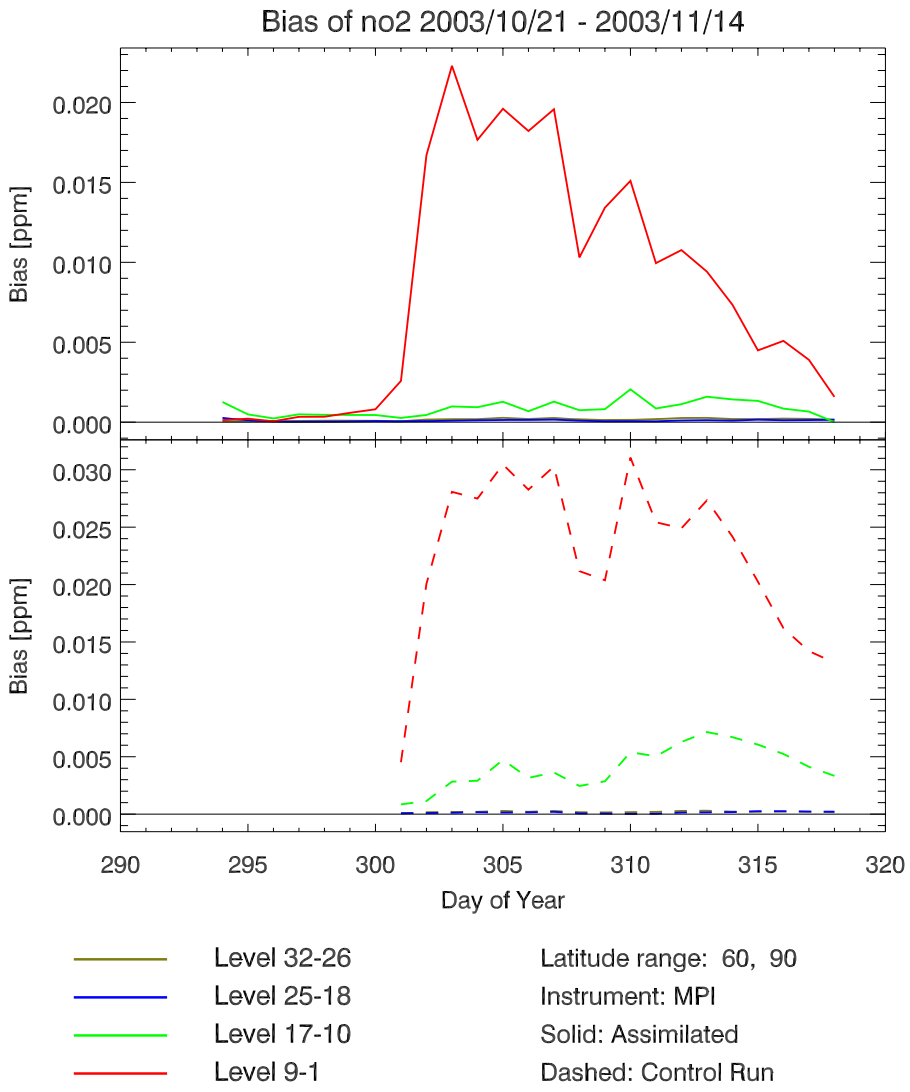

Figure 6.35: *Bias of analysed NO_2 against MPI observations (assimilated) for* ***CS3-MPI-1*** *at high northern latitudes. The sharp increase after 27 October 2003 at upper model layers is due to a series of exceptional strong solar proton events. This feature is absent at tropical latitudes (compare Figure 6.36, note the different scales).*

at tropical latitudes, where the NO_x budget was not influenced by the solar proton events.

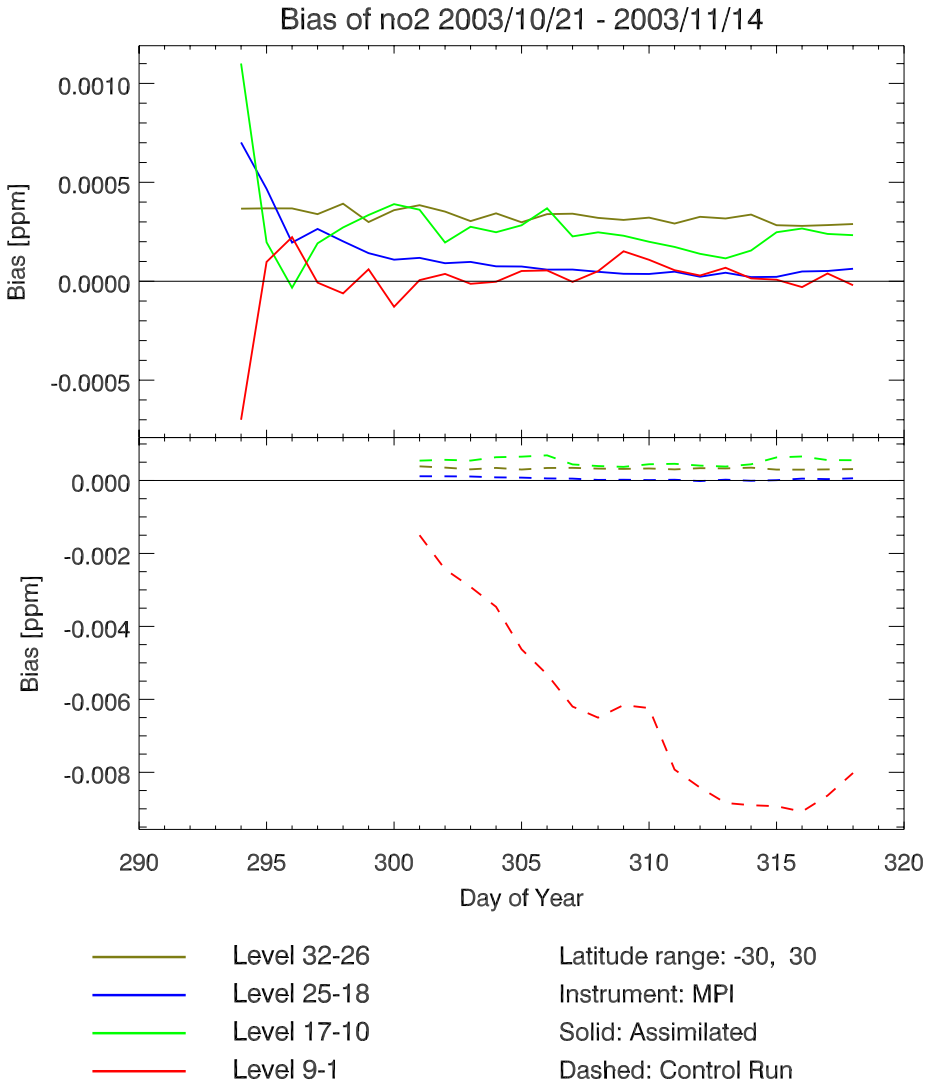

Figure 6.36: *Bias of analysed NO_2 against MPI observations (assimilated) for* ***CS3-MPI-1*** *at tropical latitudes. Here, the NO_x budget is not influenced by the solar proton events (compare Figure 6.35, note the different scales).*

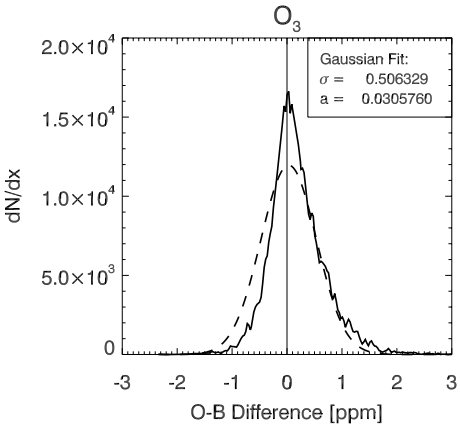

Figure 6.37: CS3-MPE-2 *O−B distribution for O_3 at 27 October 2003. The dashed curve is a Gaussian with the same variance as the data centred about the maximum data value. The standard deviation σ and the mean a are given in the upper right corner of the plot.*

6.4.3 Statistical evaluation

A means to test the consistency of the assumption that the probability density functions of the background field and observations are Gaussian, is given by inspecting the differences between observations and backround (O−B). The difference between two uncorrelated Gaussian distributed random quantities is likewise Gaussian distributed, and consequently a necessary condition for the correctness of the underlying statistical assumptions is that the O−B PDF is Gaussian. The distribution of this quantity for configurations **CS3-MPE-2** and **CS2-MPI-1** are shown in Figure 6.37 and Figures 6.38 – 6.39, respectively.

All distributions are peaked around zero and are approximately symmetric, whereas the most pronounced deviations from symmetry occur for O_3 and N_2O_5, which have a small tail of positive O−B differences. The reason for these deviations is the bias of modelled ozone against the observations, which is present in the upper stratosphere and lower mesosphere region, as discussed above. The fact that the distributions are generally more peaked than a Gaussian with the same variance, which is shown as a dashed line, is a well known phenomenon. It is caused partly by observational data with gross errors (see e. g. *Kalnay* [2003]) and to some extent by inconsistencies between data and the model physics. Observe that, as could be expected, the Gaussian fit matches much closer for configuration **CS3-MPE-2**, where ozone only has been assimilated within the restricted domain between 100 and 2 hPa.

It can be concluded, that the underlying Gaussian assumptions are at least not grossly violated. Improvements can be expected from a more rigorous quality control of observational data, and from future model updates, which

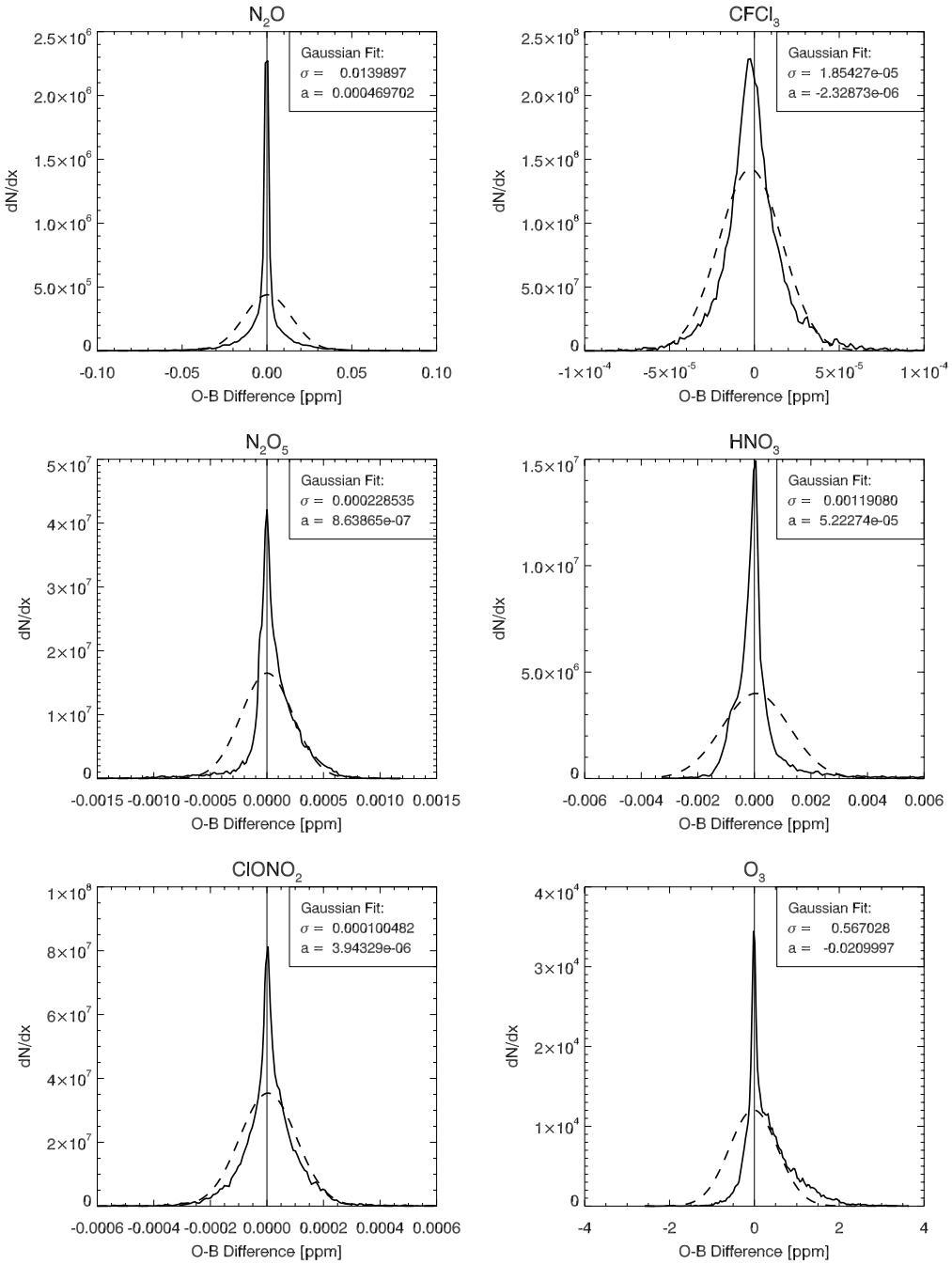

Figure 6.38: CS2-MPI-1 *$O-B$ distribution for N_2O, $CFCl_2$, N_2O_5, HNO_3, $ClONO_2$ and O_3 at 28 July 2003. The dashed curve is a Gaussian with the same variance as the data, centred about the maximum data value. The standard deviation σ and the mean a are given in the upper right corner of each plot.*

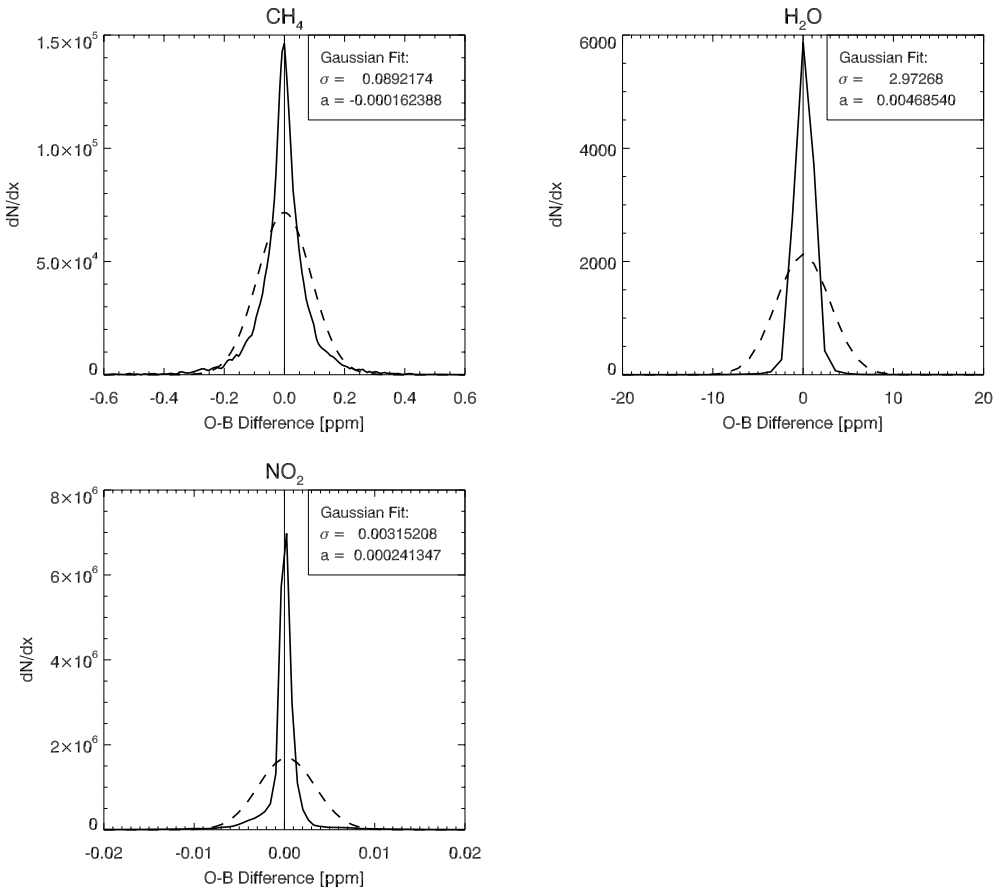

Figure 6.39: *As Figure 6.38 for CH_4, H_2O and NO_2.*

will have to aim for a removal of ozone and related biases in the upper model domain.

CHAPTER 7

Summary and Conclusions

A new 4D-var data assimilation system for stratospheric trace gas observations has been developed from scratch, aiming to address some of the shortcomings of presently existing assimilation systems.

Particular efforts were devoted to the development of a sophisticated parameterisation of the background error covariance matrix. A diffusion approach following *Weaver and Courtier* [2001] was implemented, as this solution is not only efficient, but also allows for anisotropic and inhomogeneous correlations by inserting a local coordinate stretching tensor into the diffusion operator. This feature was employed to devise a flow dependent BECM formulation, since the model dynamics and the evolution of background errors are coupled. This approach remedies a cardinal dilemma of data assimilation: While Kalman filter based methods, which account for evolving error covariances, cannot be implemented for the large problem of stratospheric trace gas assimilation unless the complexity of the problem is severely reduced, the efficient 4D-var method provides no means of estimating error covariances. The solution presented in this work assumes larger background error correlations along isopleths of potential vorticity, as proposed by *Riishøjgaard* [1998]. The feasibility of this idea in the context of 4D-var has been investigated and it was found that –in combination with the diffusion approach– very encouraging results can be achieved.

The kernel of the assimilation system is a novel stratospheric chemistry-transport model, which uses the global forecast model GME of German

Weather Service as an online meteorological driver. This concept improves the model skill, since a consistent representation of wind fields is available for the solution of the advection-reaction equation. The icosahedral grid, which provides a nearly uniform distribution of grid cells over the globe, has been adopted from GME to overcome the problem of poleward convergence and zonal crowding of grid cells. Thereby, the computational cost is reduced by about 25%, compared to CTMs employing a traditional latitude-longitude grid. Furthermore, a semi-Lagrange horizontal transport scheme together with a parallel implementation of the model lead to an excellent efficency of the new system. The chemical differential equations are solved by means of an accurate second order Rosenbrock solver without any simplifying assumptions. Heterogeneous reactions on PSC and aerosol surfaces are included in the reaction mechanism and its adjoint. The efficient system design enables the application of the computationally costly 4D-var technique in near real time. Actually, the new system exceeded all expectations in this respect: Given that twelve or more processors are available on a state of the technology parallel computer, an analysis can be calculated within less than four hours.

A comprehensive set of case studies based on EnviSat-MIPAS data products was accomplished and the assimilation results have been validated with independent (not assimilated) data from SCIAMACHY, SAGE II, and HALOE. Based on statistical a posteriori evaluation, the system proved to run stable producing fully satisfying results. Different BECM parameter settings have been tested and compared. Analysis quality has been objectively quantified on the basis of weighted mean square differences between a model forecast starting from analysed fields and observations. It was shown that the anisotropic formulation of background error covariances can have a considerable benefit, especially in regions where large gradients of potential vorticity prevail. It can be expected that a more sophisticated estimation of the BECM parameters used in the diffusion scheme will further improve the quality of analysed trace gas fields.

The 4D-var assimilation technique has been used to identify inconsistencies between the model and observational data. A bias of assimilated ozone fields against retrieved profiles in the upper stratosphere and lower mesosphere has been detected. This problem can most probably be attributed to a slight misrepresentation of the photochemical equilibrium in the upper model domain of the SACADA CTM. A cross validation of ozone profiles obtained from different remote sounding devices has been performed. The use of assimilated ozone fields for this task is a natural solution and circumvents the need to select colocated profiles for comparison. A drawback of the 4D-

var method in this context, is the fact that no analysis error estimates are provided. Hence, it is not possible to quantify a deviation from the analysis as 'large' or 'small'. Furthermore, a profound cross validation will require the use of averaging kernel information to account for different instrument characteristics. It is therefore highly desirable that suitable averaging kernel information becomes an integral part of each satellite data product. Despite these shortcomings, it has been demonstrated that data assimilation is a valuable tool for cross validation of different satellite data products.

Future applications, like the assimilation of water vapour in the tropopause region, will require a further refinement of spatial resolution, which is clearly possible due to the efficient system design. However, it will be advantageous –if not necessary– to reduce the computational cost and the amount of data to be stored for adjoint integrations. This may be achieved by employing approximate implementations of the adjoint operators, which circumvent the need of recomputations. A similar approach, which is already used in operational meteorological data assimilation, is the incremental variant of 4D-var. There, a lower resolution model is used for parts of the iterative process, as described by *Courtier et al.* [1994].

The SACADA system has been delivered to the German Remote Sounding Data Centre for operational application. The assimilation of total ozone columns from the Global Ozone Monitoring Experiment II (GOME II) is being prepared there.

APPENDIX A

Vertical advection scheme: Implementation

The scheme that is applied to solve the vertical advection subproblem (4.9b) is an implicit upwind scheme, which was developed by R. Botchorisvili (Fraunhofer-Institute for Algorithms and Scientific Computing) in the framework of the SACADA project, and implemented with slight modifications for the SACADA assimilation system. Remember that the vertical wind w in units of Pa/s is defined in the hybrid σ-pressure coordinate system at the boundary of layers. If the movement is directed upwards towards lower pressure values w is negative. Volume mixing ratios are defined at the centre of layers, which are separated by pressure differences $\Delta p_{i-1/2} := p_i - p_{i-1}$ and $\Delta p_{i+1/2} := p_{i+1} - p_i$. The scheme offers the flexibility to be applied on a subset of model layers $i = l_s, \cdots, l_e$ only. The numbering of layers is from top to bottom. This configuration is shown in Figure A.1. The characteristic property of an upwind scheme is that the

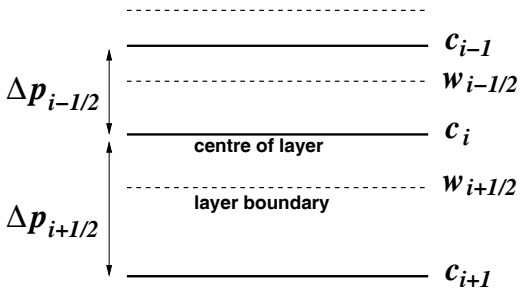

Figure A.1: *Definition of quantities for the vertical advection scheme*

spatial discretisation of the term $\partial \boldsymbol{c}/\partial p$ depends on the wind direction:

$$\frac{\partial \boldsymbol{c}}{\partial p} \approx \begin{cases} \frac{1}{2}\left(\frac{c_{i+1}-c_i}{\Delta p_{i+1/2}} + \frac{c_i - c_{i-1}}{\Delta p_{i-1/2}}\right) & \text{if } w_{i-1/2} > 0 \text{ and } w_{i+1/2} < 0 \\ \frac{c_i - c_{i-1}}{\Delta p_{i-1/2}} & \text{if } w_{i-1/2} > 0 \text{ and } w_{i+1/2} > 0 \\ \frac{c_{i+1}-c_i}{\Delta p_{i+1/2}} & \text{if } w_{i-1/2} < 0 \text{ and } w_{i+1/2} < 0 \end{cases}$$

$$\frac{\partial \boldsymbol{c}}{\partial p} = 0 \qquad \text{if } w_{i-1/2} < 0 \text{ and } w_{i+1/2} > 0$$

The last condition, which applies in the case of divergent vertical wind, may seem inconsistent at a first glance. However, it is just a natural consequence of the three dimensional continuity equation together with the fact that this discretisation is applied within an operator splitting environment. The divergent vertical wind implies a convergence in the horizontal wind, and any changes in volume mixing ratio at this point must be due to horizontal advection. Observe that the horizontal semi-Lagrangian transport scheme has the same property for a divergent horizontal wind field.

The time discretisation is chosen to be implicit. Let $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ denote volume mixing ratios at times t_n and $t_{n+1} := t_n + \Delta t$, respectively. A compact notation can be introduced by defining

$$w^+_{i-1/2} = \frac{1}{2}\left(w_{i-1/2} - |w_{i-1/2}|\right) \geq 0$$
$$w^-_{i+1/2} = \frac{1}{2}\left(w_{i+1/2} - |w_{i+1/2}|\right) \leq 0$$

$$\text{and} \quad \kappa = \begin{cases} 1/2 & \text{if } \operatorname{sign}(w^+)\operatorname{sign}(w^-) < 0 \\ 1 & \text{if } w^+ > 0 \text{ or } w^- < 0 \\ 0 & \text{otherwise} \end{cases} .$$

With these abbreviations the scheme reads

$$\frac{\hat{c}_i - c_i}{\Delta t} + \kappa_i \left(w^+_{i-1/2} \frac{\hat{c}_i - \hat{c}_{i-1}}{\Delta p_{i-1/2}} + w^-_{i+1/2} \frac{\hat{c}_{i+1} - \hat{c}_i}{\Delta p_{i+1/2}} \right) = 0 \,, \tag{A.1}$$

leading to a linear system of equations

$$\Delta t\, b_i \hat{c}_{i+1} + (1 + \Delta t\, d_i)\hat{c}_i + \Delta t\, a_i \hat{c}_{i-1} = c_i \,, \tag{A.2}$$

with the coefficients a_i, b_i and d_i given by

$$\begin{aligned} a_i &:= -\kappa_i \frac{w^+_{i-1/2}}{\Delta p_{i-1/2}} \\ b_i &:= \quad \kappa_i \frac{w^-_{i+1/2}}{\Delta p_{i+1/2}} \\ d_i &:= -(a_i + b_i)\,. \end{aligned} \tag{A.3}$$

At the boundary points $i = l_s$ and $i = l_e$ von Neumann boundary conditions $\partial c/\partial p = 0$ are introduced by setting

$$\begin{aligned} \frac{\hat{c}_{l_s} - c_{l_s}}{\Delta t} + \kappa_{l_s} w^-_{l_s+1/2} \frac{\hat{c}_{l_s+1} - \hat{c}_{l_s}}{\Delta p_{l_s+1/2}} &= 0 \\ \frac{\hat{c}_{l_e} - c_{l_e}}{\Delta t} + \kappa_{l_e} w^+_{l_e-1/2} \frac{\hat{c}_{l_e} - \hat{c}_{l_e-1}}{\Delta p_{l_e-1/2}} &= 0\,, \end{aligned} \tag{A.4}$$

leading to $a_{l_s} = 0$ and $b_{l_e} = 0$. The same does hold if $w^+_{l_s-1/2} = 0$ and $w^-_{l_e+1/2} = 0$ is assumed, so the numerical treatment of boundaries is the same for both cases. To calculate $\hat{\boldsymbol{c}}$ for a column of grid points a tridiagonal linear system $[\mathbf{I} + \Delta t \mathbf{A}]\ \hat{\boldsymbol{c}} = \boldsymbol{c}$ of equations has to be solved, where $\mathbf{A}$ is given by

$$\mathbf{A} := \begin{pmatrix} d_{l_s} & b_{l_s} & 0 & \dots & & 0 \\ a_{l_s+1} & d_{l_s+1} & b_{l_s+1} & 0 & \dots & 0 \\ . & & & & & . \\ . & & & & & . \\ . & & & & & 0 \\ 0 & \dots & & 0 & a_{nl_e-1} & d_{l_e-1} & b_{l_e-1} \\ 0 & \dots & & & 0 & a_{l_e} & d_{l_e} \end{pmatrix}.$$

The solution of this system can be easily obtained with numerically efficient recursive schemes (see *Press et al.* [1992]).

APPENDIX B

Implementation of diffusion schemes

B.1 Horizontal diffusion

As outlined in Chapter 3 both, an isotropic and an anisotropic diffusion scheme is available within the SACADA system, differing in the form of the horizontal Laplacian operator. In the isotropic case, the Laplacian of an arbitrary scalar field ψ defined on the icosahedral grid is given, according to (4.3), by

$$\nabla_h^2 \psi_0 = \sum_{1=1}^{6} \left[L_{\eta,i} + L_{\chi,i} \right] (\psi_i - \psi_0) = \sum_{1=1}^{6} L_i \psi_i - \sum_{i=1}^{6} L_i \psi_0 \tag{B.1}$$

with $L_i = L_{\eta,i} + L_{\chi,i}$. In analogy to the one dimensional case (see *Kalnay* [2003]), a stability criterion for the resulting diffusion scheme can be derived as follows: At each grid point the solution at time $n+1$

$$\psi_0^{n+1} = \psi_0^n + \kappa \Delta t \sum_{i=1}^{6} L_i \psi_i^n - \kappa \Delta t \sum_{i=1}^{6} L_i \psi_0^n$$

must be bounded, i. e. $|\psi_0^{n+1}| \leq |\psi_0^n|$, in oder to ensure stability. Taking $\psi_{\max}^n := \max \{\psi_i^n | i = 0, \ldots, 6\}$ and defining $\mu := \kappa \Delta t \sum_{i=1}^{6} L_i$, such that

$$|\psi_0^{n+1}| \leq (|1 - \mu| + |\mu|) \, |\psi_{\max}^n| \tag{B.2}$$

holds, it turns out that the necessary condition for stability of the horizontal diffusion scheme is $0 < \mu \leq 1$. A descriptive interpretation of this criterion can be given in the case that $\psi_i = 0$ for $i = 1, \ldots, 6$ and $\psi_0 \neq 0$. Then $0 < \mu \leq 1$ ensures that the total diffusive outflow is not greater than the current content of the respective grid cell.

The anisotropic Laplacian operator $\tilde{\nabla}_h^2 \psi = \mathrm{div}\,(\mathbf{S}\,\mathrm{grad}_h \psi)$ in the local coordinate system of the icosahedral grid reads

$$\tilde{\nabla}_h^2 \psi = \begin{pmatrix} \frac{\partial}{\partial \eta} \\ \frac{\partial}{\partial \chi} \end{pmatrix} \cdot \begin{pmatrix} s_1 & s_3 \\ s_3 & s_2 \end{pmatrix} \begin{pmatrix} \frac{\partial \psi}{\partial \eta} \\ \frac{\partial \psi}{\partial \chi} \end{pmatrix} , \tag{B.3}$$

where the coordinate stretching matrix $\mathbf{S}$ has been computed from stretching factors $\tilde{s}_1$, $\tilde{s}_2$ and a rotation angle α as described in Chapter 3. Using the expressions (4.3) for the spatial derivatives yields

$$\begin{aligned} \tilde{\nabla}_h^2 \psi_0 = \sum_{i=1}^{6} \Bigg[& \left(\frac{\partial s_1}{\partial \eta} + \frac{\partial s_3}{\partial \chi} \right) G_{\eta,i} + \left(\frac{\partial s_2}{\partial \chi} + \frac{\partial s_3}{\partial \eta} \right) G_{\chi,i} \\ & + s_1 L_{\eta,i} + s_2 L_{\chi,i} + 2 s_3 L_{\eta,\chi,i} \Bigg] (\psi_i - \psi_0) = \sum_{i=1}^{6} \tilde{L}_i \, (\psi_i - \psi_0) \; . \end{aligned} \tag{B.4}$$

The spatial derivatives of the stretching tensor components are evaluated according to (4.3). However, because the elements of $\mathbf{S}$ are valid in the local coordinate system of each grid point only, the stretching tensor has to be expressed in terms of the local coordinates of the centre node as

$$\hat{\mathbf{S}} = \mathbf{T}(\beta_i) \mathbf{S} \mathbf{T}^T(\beta_i) \, ,$$

before the differences $\hat{s}_i - s_0$ can be computed. Here, β_i is the rotation angle between the local (η, χ) system at the centre node and the local coordinate system at the neighbouring grid point i. These angles are calculated during the grid generation process of the GME software.

A stability criterion can be derived in exactly the same way as for the isotropic case, yielding $0 < \tilde{\mu} \leq 1$ as a necessary condition for stability, with $\tilde{\mu} := \kappa \Delta t \sum_{i=1}^{6} \tilde{L}_i$.

B.1.1 Transpose horizontal diffusion

The transpose operators $\mathbf{D}^T$ and $\tilde{\mathbf{D}}^T$ have been build by reversing and transposing the code line-by-line as described in Chapter 2. Let $d_h := \nabla_h \psi_0$ be the

Laplacian of ψ_0 according to (B.1). Then, the transpose of this instruction is given by the mapping

$$\mathbf{D}^T : \quad \begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_6 \end{pmatrix} \longrightarrow \begin{pmatrix} \psi_0 - d_h \sum_{i=1}^{6} L_i \\ \psi_1 + d_h L_1 \\ \vdots \\ \psi_6 + d_h L_6 \end{pmatrix}, \tag{B.5}$$

and the same is valid for the anisotropic scheme, replacing L_i with $\tilde{L}_i$.

B.2 Vertical diffusion

For the vertical diffusion scheme, a discrete solution of the one-dimensional diffusion-equation

$$\frac{\partial \psi(z,t)}{\partial t} - \kappa \frac{\partial^2 \psi(z,t)}{\partial z^2} = 0 \qquad z \in [a,b],\ t \in [0,T] \tag{B.6}$$

is to be found, given the boundary conditions

$$\left.\frac{\partial \psi(z,t)}{\partial z}\right|_{z=a} = \left.\frac{\partial \psi(z,t)}{\partial z}\right|_{z=b} = 0 \tag{B.7}$$

on a vertical grid, which is not equally spaced. Equation (B.6) is valid if the diffusion coefficient κ is assumed to be constant. As described in Chapter 4, the vertical grid of GME is a hybrid σ-pressure grid and most quantities are defined at the centre of pressure levels, which (in the lower part of the grid) depend on surface pressure and thus on the current meteorological situation. The diffusion module may be applied to a subset of model layers in the same manner as the vertical advection algorithm described in Appendix A. Geometric heights z_i in meters are calculated for $i = l_s, \ldots, l_e$, using the geopotential as computed by GME. Define $\Delta z_i := z_i - z_{i+1}$ and $\Delta z_i^c := \frac{1}{2}(\Delta z_{i-1} + \Delta z_i)$. Note that numbering of levels in GME is from top to bottom leading to $z_i > z_{i+1}$. As the vertical diffusion is a purely one-dimensional problem, it is easy to implement a discretisation that is implicit with respect to the time derivative. Let ψ_i and $\hat{\psi}_i$ be the discrete function values at time t_n and $t_{n+1} := t_n + \Delta t$, respectively. At model layer i the one-dimensional diffusion equation (B.6) can be discretised as follows:

$$\frac{\hat{\psi}_i - \psi_i}{\Delta t} - \frac{\kappa}{\Delta z_i^c} \left(\frac{\hat{\psi}_{i-1} - \hat{\psi}_i}{\Delta z_{i-1}} - \frac{\hat{\psi}_i - \hat{\psi}_{i+1}}{\Delta z_i} \right) = 0 \tag{B.8}$$

The second term is a symmetric discretisation of the Laplacian, using the difference of the left handed derivative and the right handed derivative as an discrete approximation for the curvature of $\psi(z)$. The boundary conditions given in (B.7) lead to $\hat{\psi}_{l_s-1} = \hat{\psi}_{l_s}$ and $\hat{\psi}_{l_e+1} = \hat{\psi}_{l_e}$. The resulting system of equations reads

$$[\mathbf{I} - \kappa \Delta t \mathbf{D}]\, \hat{\boldsymbol{\psi}} = \boldsymbol{\psi}\,, \tag{B.9}$$

where $\mathbf{D}$ is the discretised Laplacian whose matrix representation is given by

$$\mathbf{D} := \begin{pmatrix} b_{l_s} & c_{l_s} & 0 & \dots & & 0 \\ a_{l_s+1} & b_{l_s+1} & c_{l_s+1} & 0 & \dots & 0 \\ . & & & & & . \\ . & & & & & . \\ . & & & & & 0 \\ 0 & \dots & 0 & a_{nl_e-1} & b_{l_e-1} & c_{l_e-1} \\ 0 & \dots & & 0 & a_{l_e} & b_{l_e} \end{pmatrix} \tag{B.10}$$

with

$$a_i := \frac{1}{\Delta z_i^c \Delta z_{i-1}}\,, \quad c_i := \frac{1}{\Delta z_i^c \Delta z_i} \quad \text{and} \quad b_i := -(a_i + c_i)\,.$$

As $\mathbf{D}$ is tridiagonal matrix, the system (B.10) can easily be solved with numerically efficient recursive schemes (see *Press et al.* [1992]).

APPENDIX C

Tables

C.1 Vertical grid parameters

Table C.1: *Coefficients defining the vertical grid. k is the layer index, $a_{f,k}$ and $b_{f,k}$ are the coefficients defining the pressure at the centre of layer k according to (4.7). The coeffients a_k, b_k define the pressure at the upper boundary and a_{k+1}, b_{k+1} the pressure at the bottom of layer k. The last column gives the height of the centre of layer, computed from pressure according to the U. S. Standard Atmosphere from 1976.*

k	a_k [Pa]	b_k [-]	$a_{f,k}$ [Pa]	$b_{f,k}$ [-]	Height [km]
1	0.0	0.0	0.1	0.0	65.620
2	0.2	0.0	0.2240	0.0	59.850
3	0.2479	0.0	0.3053	0.0	57.540
4	0.3626	0.0	0.4173	0.0	55.150
5	0.4719	0.0	0.5430	0.0	53.080
6	0.6141	0.0	0.7066	0.0	50.980
7	0.7991	0.0	0.9195	0.0	48.860
8	1.0399	0.0	1.1966	0.0	46.740
9	1.3532	0.0	1.5571	0.0	44.660
10	1.7610	0.0	2.0263	0.0	42.620

k	a_k [Pa]	b_k [-]	$a_{f,k}$ [Pa]	$b_{f,k}$ [-]	Height [km]
11	2.2916	0.0	2.6368	0.0	40.630
12	2.9820	0.0	3.4313	0.0	38.690
13	3.8805	0.0	4.4651	0.0	36.780
14	5.0498	0.0	5.8105	0.0	34.920
15	6.5713	0.0	7.5613	0.0	33.100
16	8.5513	0.0	9.8395	0.0	31.320
17	11.1278	0.0	12.8042	0.0	29.550
18	14.4807	0.0	16.6622	0.0	27.800
19	18.8438	0.0	21.6827	0.0	26.060
20	24.5216	0.0	28.2159	0.0	24.340
21	31.9102	0.0	36.7175	0.0	22.630
22	41.5249	0.0	47.7808	0.0	20.930
23	54.0367	0.0	61.5368	0.0	19.320
24	69.0369	0.0	78.2029	0.00005025	17.780
25	87.3690	0.00010049	97.8525	0.00061173	16.320
26	108.3360	0.00112296	119.1935	0.00363116	14.910
27	130.0510	0.00613936	139.5775	0.01323663	13.510
28	149.1040	0.02033390	155.6260	0.03568065	12.070
29	162.1480	0.05102740	163.0295	0.07848870	10.570
30	163.9110	0.10595000	159.6420	0.14116749	9.120
31	155.3730	0.17638499	146.2010	0.22071898	7.740
32	137.0290	0.26505297	124.3030	0.31453148	6.460
33	111.5770	0.36400998	97.4859	0.42082497	5.230
34	83.3948	0.47763997	70.6652	0.53588000	4.040
35	57.9356	0.59412003	46.9460	0.65086651	2.940
36	35.9564	0.70761299	27.7032	0.75972000	1.970
37	19.4500	0.81182700	12.9157	0.83839142	1.340
38	6.3815	0.86495584	4.5489	0.88633585	0.965
39	2.7163	0.90771586	1.7185	0.92596453	0.629
40	0.7206	0.94421321	0.3603	0.95859921	0.352
41	0.0000	0.97298521	0.0000	0.98263335	0.148
42	0.0000	0.99228150	0.0000	0.99614075	0.033
43	0.0000	1.0	-	-	-

C.2 Reaction equations and rate constants

Table C.2: *Photolysis reactions included in the model. The term* products *represents constituents that are not considered in the present work.*

Reaction	
(J1)	$O_2 + h\nu \rightarrow 2O(^3P)$
(J2)	$O_3 + h\nu \rightarrow O(^3P) + O_2$
(J3)	$O_3 + h\nu \rightarrow O(^1D) + O_2$
(J4)	$HO_2 + h\nu \rightarrow O(^3P) + OH$
(J5)	$H_2O + h\nu \rightarrow H + OH$
(J6)	$H_2O_2 + h\nu \rightarrow 2OH$
(J7)	$NO_2 + h\nu \rightarrow O(^3P) + NO$
(J8)	$NO_3 + h\nu \rightarrow NO_2 + O(^3P)$
(J9)	$NO_3 + h\nu \rightarrow NO + O_2$
(J10)	$N_2O + h\nu \rightarrow N_2 + O(^1D)$
(J11)	$N_2O_5 + h\nu \rightarrow NO_2 + NO_3$
(J12)	$HNO_3 + h\nu \rightarrow OH + NO_2$
(J13)	$HNO_4 + h\nu \rightarrow HO_2 + NO_2$
(J14)	$Cl_2O_2 + h\nu\ (+ M) \rightarrow 2Cl + O_2\ (+ M)$
(J15)	$Cl_2 + h\nu \rightarrow 2Cl$
(J16)	$OClO + h\nu \rightarrow O(^3P) + ClO$
(J17)	$HCl + h\nu \rightarrow Cl + H$
(J18)	$HOCl + h\nu \rightarrow Cl + OH$
(J19)	$ClONO_2 + h\nu \rightarrow Cl + NO_3$
(J20)	$CH_3Cl + h\nu \rightarrow Cl +$ products
(J21)	$CCl_4 + h\nu \rightarrow 4Cl +$ products
(J22)	$CFCl_3 + h\nu \rightarrow 3Cl +$ products
(J23)	$CF_2Cl_2 + h\nu \rightarrow 2Cl +$ products
(J24)	$CHF_2Cl + h\nu \rightarrow Cl +$ products
(J25)	$CF_2ClCFCl_2 + h\nu \rightarrow 3Cl +$ products
(J26)	$CH_3CCl_3 + h\nu \rightarrow 3Cl +$ products
(J27)	$BrO + h\nu \rightarrow Br + O(^3P)$
(J28)	$BrCl + h\nu \rightarrow Br + Cl$
(J29)	$HBr + h\nu \rightarrow Br + H$
(J30)	$HOBr + h\nu \rightarrow Br + OH$
(J31)	$BrONO_2 + h\nu \rightarrow Br + NO_3$
(J32)	$CH_3Br + h\nu \rightarrow Br +$ products
(J33)	$CF_2ClBr + h\nu \rightarrow Cl + Br +$ products
(J34)	$CF_3Br + h\nu \rightarrow Br +$ products

Table C.3: *Gas phase reactions that are included in the model. Rate constants for first- and second-order reactions are given in units of* s^{-1} *and* $molecules^{-1}cm^3s^{-1}$*, respectively. Rate constants for third-order reactions are given as effective second-order rate constants in units of* $molecules^{-1}cm^3s^{-1}$*.* $M \in \{N_2,\ O_2\}$*. For third-order reactions,* $f(k_0, k_\infty)$ *has to be evaluated according to Sander et al. [2003]:*

$$f(k_0, k_\infty) = (k_0[\mathrm{M}]/(1 + k_0[\mathrm{M}]/k_\infty)) \times 0.6^{(1+(\log_{10}(k_0[\mathrm{M}]/k_\infty))^2)^{-1}}.$$

The term products *represents constituents which are not considered in the reaction scheme.*

Reaction		Rate constant
(R1)	$O(^1D) + N_2 \rightarrow O(^3P) + N_2$	$1.8\times10^{-11}\exp(110/T)$
(R2)	$O(^1D) + O_2 \rightarrow O(^3P) + O_2$	$3.2\times10^{-11}\exp(70/T)$
(R3)	$O(^1D) + O_3 \rightarrow 2O_2$	1.2×10^{-10}
(R4)	$O(^3P) + O(^3P) + M \rightarrow O_2 + M$	$4.7\times10^{-33}[M](T/300.)^{-2.0}$
(R5)	$O(^3P) + O_2 + M \rightarrow O_3 + M$	$6.0\times10^{-34}[M](T/300.)^{-2.4}$
(R6)	$O(^3P) + O_3 \rightarrow 2O_2$	$8.0\times10^{-12}\exp(-2060/T)$
(R7)	$O(^1D) + H_2O \rightarrow 2OH$	2.2×10^{-10}
(R8)	$O(^1D) + H_2 \rightarrow OH + H$	1.1×10^{-10}
(R9)	$H + O_3 \rightarrow OH + O_2$	$1.4\times10^{-10}\exp(-470/T)$
(R10)	$H + O_2 + M \rightarrow HO_2 + M$	$f(k_0, k_\infty)$, $k_0=5.7\times10^{-32}(T/300)^{-1.6}$, $k_\infty=7.5\times10^{-11}$
(R11)	$H + HO_2 \rightarrow 2OH$	7.3×10^{-11}
(R12)	$OH + O(^3P) \rightarrow H + O_2$	$2.2\times10^{-11}\exp(120/T)$
(R13)	$OH + OH \rightarrow H_2O + O(^3P)$	$4.2\times10^{-12}\exp(-240/T)$
(R14)	$OH + OH + M \rightarrow H_2O_2 + M$	$f(k_0, k_\infty)$, $k_0=6.2\times10^{-31}(T/300)^{-1.0}$, $k_\infty=2.6\times10^{-11}$
(R15)	$OH + H_2O_2 \rightarrow H_2O + HO_2$	$2.9\times10^{-12}\exp(-160/T)$
(R16)	$OH + H_2 \rightarrow H_2O + H$	$5.5\times10^{-12}\exp(-2000/T)$
(R17)	$OH + O_3 \rightarrow HO_2 + O_2$	$1.7\times10^{-12}\exp(-940/T)$
(R18)	$OH + HO_2 \rightarrow H_2O + O_2$	$4.8\times10^{-11}\exp(250/T)$
(R19)	$HO_2 + O(^3P) \rightarrow OH + O_2$	$3.0\times10^{-11}\exp(200/T)$
(R20)	$HO_2 + O_3 \rightarrow OH + 2O_2$	$1.0\times10^{-14}\exp(-490/T)$
(R21)	$HO_2 + HO_2 \rightarrow H_2O_2 + O_2$	$2.3\times10^{-13}\exp(600/T)$
(R22)	$HO_2 + HO_2 + M \rightarrow H_2O_2 + O_2 + M$	$1.7\times10^{-33}[M]\exp(1000./T)$
(R23)	$H_2O_2 + O(^3P) \rightarrow HO_2 + OH$	$1.4\times10^{-12}\exp(-2000/T)$
(R24)	$CH_4 + OH \rightarrow H_2O +$ products	$2.45\times10^{-12}\exp(-1775/T)$
(R25)	$CH_4 + O(^1D) \rightarrow OH +$ products	1.125×10^{-10}
(R26)	$N_2O + O(^1D) \rightarrow N_2 + O_2$	4.9×10^{-11}
(R27)	$N_2O + O(^1D) \rightarrow 2NO$	6.7×10^{-11}
(R28)	$N_2 + O(^1D) + M \rightarrow N_2O + M$	$3.5\times10^{-37}[M](T/300.)^{-0.6}$
(R29)	$NO + O_3 \rightarrow NO_2 + O_2$	$3.0\times10^{-12}\exp(-1500/T)$
(R30)	$NO + O(^3P) + M \rightarrow NO_2 + M$	$f(k_0, k_\infty)$, $k_0=9.0\times10^{-32}(T/300)^{-1.5}$, $k_\infty=3.0\times10^{-11}$

Table C.3: *Gas phase reaction equations and rate constants (continued)*

Reaction		Rate constant
(R31)	$NO + HO_2 \rightarrow OH + NO_2$	$3.5\times10^{-12}\exp(250/T)$
(R32)	$NO + NO_3 \rightarrow 2NO_2$	$1.5\times10^{-11}\exp(170/T)$
(R33)	$NO_2 + O(^3P) \rightarrow NO + O_2$	$5.6\times10^{-12}\exp(180/T)$
(R34)	$NO_2 + O(^3P) + M \rightarrow NO_3 + M$	$f(k_0, k_\infty)$, $k_0=9.0\times10^{-32}(T/300)^{-2.0}$, $k_\infty=2.2\times10^{-11}$
(R35)	$NO_2 + O_3 \rightarrow NO_3 + O_2$	$1.2\times10^{-13}\exp(-2450/T)$
(R36)	$NO_2 + H \rightarrow OH + NO$	$4.0\times10^{-10}\exp(-340/T)$
(R37)	$NO_2 + OH + M \rightarrow HNO_3 + M$	$f(k_0, k_\infty)$, $k_0=2.5\times10^{-30}(T/300)^{-4.4}$, $k_\infty=1.6\times10^{-11}(T/300)^{-1.7}$
(R38)	$NO_2 + HO_2 + M \rightarrow HNO_4 + M$	$f(k_0, k_\infty)$, $k_0=1.8\times10^{-31}(T/300)^{-3.2}$, $k_\infty=4.7\times10^{-12}(T/300)^{-1.4}$
(R39)	$HNO_4 + M \rightarrow HO_2 + NO_2 + M$	$k_{R38}/(2.1\times10^{-27}\exp(10900/T))$
(R40)	$NO_2 + NO_3 + M \rightarrow N_2O_5 + M$	$f(k_0, k_\infty)$, $k_0=2.0\times10^{-30}(T/300)^{-4.4}$, $k_\infty=1.4\times10^{-12}(T/300)^{-0.7}$
(R41)	$N_2O_5 + M \rightarrow NO_3 + NO_2 + M$	$k_{R40}/(3.0\times10^{-27}\exp(10990/T))$
(R42)	$NO_3 + O(^3P) \rightarrow NO_2 + O_2$	1.0×10^{-11}
(R43)	$NO_3 + OH \rightarrow NO_2 + HO_2$	2.2×10^{-11}
(R44)	$NO_3 + HO_2 \rightarrow HNO_3 + O_2$	1.5×10^{-12}
(R45)	$HNO_3 + OH \rightarrow NO_3 + H_2O$	$k_0+k_3/(1+k_3/k_2)$, $k_0=2.4\times10^{-14}\exp(460/T)$, $k_2=2.7\times10^{-17}[M]\exp(2199/T)$, $k_3=6.5\times10^{-34}\exp(1335/T)$
(R46)	$HNO_4 + OH \rightarrow NO_2 + H_2O + O_2$	$1.3\times10^{-12}\exp(380/T)$
(R47)	$Cl + O_3 \rightarrow ClO + O_2$	$2.3\times10^{-11}\exp(-200/T)$
(R48)	$Cl + OH \rightarrow HCl + O(^3P)$	$9.8\times10^{-12}\exp(-2860/T)$
(R49)	$Cl + HO_2 \rightarrow HCl + O_2$	$1.8\times10^{-11}\exp(170/T)$
(R50)	$Cl + HO_2 \rightarrow OH + ClO$	$4.1\times10^{-11}\exp(-450/T)$
(R51)	$Cl + H_2 \rightarrow HCl + H$	$3.7\times10^{-11}\exp(-2300/T)$
(R52)	$Cl + CH_4 \rightarrow HCl + \text{products}$	$9.6\times10^{-12}\exp(-1360/T)$
(R53)	$Cl + H_2O_2 \rightarrow HCl + HO_2$	$1.1\times10^{-11}\exp(-980/T)$
(R54)	$Cl + H_2O \rightarrow HCl + OH$	$2.79\times10^{-11}\exp(-8670/T)$
(R55)	$Cl + NO_3 \rightarrow ClO + NO_2$	2.4×10^{-11}
(R56)	$Cl + Cl + M \rightarrow Cl_2 + M$	$6.14\times10^{-34}[M]\exp(906/T)$
(R57)	$Cl + OClO \rightarrow 2ClO$	$3.4\times10^{-11}\exp(160/T)$
(R58)	$Cl + HOCl \rightarrow Cl_2 + OH$	$6.0\times10^{-13}\exp(-130/T)$
(R59)	$Cl + Cl_2O_2 \rightarrow Cl + Cl_2 + O_2$	1.0×10^{-10}
(R60)	$Cl + ClONO_2 \rightarrow Cl_2 + NO_3$	$6.5\times10^{-12}\exp(135/T)$
(R61)	$ClO + O(^3P) \rightarrow Cl + O_2$	$3.0\times10^{-11}\exp(70/T)$
(R62)	$ClO + OH \rightarrow Cl + HO_2$	$7.4\times10^{-12}\exp(270/T)$
(R63)	$ClO + HO_2 \rightarrow HOCl + O_2$	$2.7\times10^{-12}\exp(220/T)$
(R64)	$ClO + NO \rightarrow Cl + NO_2$	$6.4\times10^{-12}\exp(290/T)$

Table C.3: *Gas phase reaction equations and rate constants (continued)*

Reaction		Rate constant
(R65)	$ClO + NO_2 + M \rightarrow ClONO_2 + M$	$f(k_0, k_\infty)$, k_0=1.8×10^{-31}(T/300)$^{-3.4}$, k_∞=1.5×10^{-11}(T/300)$^{-1.9}$
(R66)	$ClO + NO_3 \rightarrow Cl + NO_2 + O_2$	4.7×10^{-13}
(R67)	$ClO + ClO + M \rightarrow Cl_2O_2 + M$	$f(k_0, k_\infty)$, k_0=1.6×10^{-32}(T/300)$^{-4.5}$, k_∞=2.0×10^{-12}(T/300)$^{-2.4}$
(R68)	$Cl_2O_2 + M \rightarrow 2ClO + M$	$k_{R67}/(1.27\times10^{-27}\exp(8744/T))$
(R69)	$OClO + O(^3P) \rightarrow ClO + O_2$	2.4×10^{-12}exp(-960/T)
(R70)	$OClO + OH \rightarrow HOCl + O_2$	4.5×10^{-13}exp(800/T)
(R71)	$OClO + NO \rightarrow ClO + NO_2$	2.5×10^{-12}exp(-600/T)
(R72)	$Cl_2 + M \rightarrow 2Cl + M$	exp(ln(3.85×10^{-11}[M]) - 23630/T)
(R73)	$Cl_2 + O(^1D) \rightarrow ClO + Cl$	2.8×10^{-10} (25% quenching)
(R74)	$Cl_2 + OH \rightarrow HOCl + Cl$	1.4×10^{-12}exp(-900/T)
(R75)	$HCl + O(^1D) \rightarrow OH + Cl$	1.0×10^{-10}
(R76)	$HCl + O(^1D) \rightarrow H + ClO$	3.6×10^{-11}
(R77)	$HCl + O(^3P) \rightarrow Cl + OH$	1.0×10^{-11}exp(-3300/T)
(R78)	$HCl + OH \rightarrow Cl + H_2O$	2.6×10^{-12}exp(-350/T)
(R79)	$HOCl + O(^3P) \rightarrow ClO + OH$	1.7×10^{-13}
(R80)	$HOCl + OH \rightarrow ClO + H_2O$	3.0×10^{-12}exp(-500/T)
(R81)	$ClONO_2 + O(^3P) \rightarrow ClO + NO_3$	2.9×10^{-12}exp(-800/T)
(R82)	$ClONO_2 + OH \rightarrow HOCl + NO_3$	1.2×10^{-12}exp(-330/T)
(R83)	$CH_3Cl + O(^1D) \rightarrow ClO +$ products	4.0×10^{-10}
(R84)	$CH_3Cl + OH \rightarrow H_2O +$ products	2.4×10^{-12}exp(-1250/T)
(R85)	$CCl_4 + O(^1D) \rightarrow ClO +$ products	3.3×10^{-10} (14% quenching)
(R86)	$CFCl_3 + O(^1D) \rightarrow ClO +$ products	2.3×10^{-10} (40% quenching)
(R87)	$CF_2Cl_2 + O(^1D) \rightarrow ClO +$ products	1.4×10^{-10} (14% quenching)
(R88)	$CHF_2Cl + O(^1D) \rightarrow ClO +$ products	1.0×10^{-10} (28% quenching)
(R89)	$CHF_2Cl + OH \rightarrow H_2O +$ products	1.05×10^{-12}exp(-1600/T)
(R90)	$CF_2ClCFCl_2 + O(^1D) \rightarrow ClO +$ products	2.0×10^{-10}
(R91)	$CH_3CCl_3 + O(^1D) \rightarrow ClO +$ products	4.0×10^{-10}
(R92)	$CH_3CCl_3 + OH \rightarrow H_2O +$ products	1.6×10^{-12}exp(-1520/T)
(R93)	$Br + O_3 \rightarrow BrO + O_2$	1.7×10^{-11}exp(-800/T)
(R94)	$Br + HO_2 \rightarrow HBr + O_2$	1.5×10^{-11}exp(-600/T)
(R95)	$Br + OClO \rightarrow BrO + ClO$	2.6×10^{-11}exp(-1300/T)
(R96)	$BrO + O(^3P) \rightarrow Br + O_2$	1.9×10^{-11}exp(230/T)
(R97)	$BrO + OH \rightarrow Br + HO_2$	7.5×10^{-11}
(R98)	$BrO + HO_2 \rightarrow HOBr + O_2$	3.4×10^{-12}exp(540/T)
(R99)	$BrO + NO \rightarrow Br + NO_2$	8.8×10^{-12}exp(260/T)
(R100)	$BrO + NO_2 + M \rightarrow BrONO_2 + M$	$f(k_0, k_\infty)$, k_0=5.2×10^{-31}(T/300)$^{-3.2}$, k_∞=6.9×10^{-12}(T/300)$^{-2.9}$
(R101)	$BrO + ClO \rightarrow Br + OClO$	9.5×10^{-13}exp(550/T)
(R102)	$BrO + ClO$ (+M) $\rightarrow$	

Table C.3: *Gas phase reaction equations and rate constants (continued)*

Reaction		Rate constant
	$Br + Cl + O_2$ (+M)	$2.3\times10^{-12}\exp(260/T)$
(R103)	$BrO + ClO \rightarrow BrCl + O_2$	$4.1\times10^{-13}\exp(290/T)$
(R104)	$BrO + BrO \rightarrow 2Br + O_2$	$1.5\times10^{-12}\exp(230/T)$
(R105)	$HBr + O(^1D) \rightarrow OH + Br$	1.5×10^{-10} (20% quenching)
(R106)	$HBr + O(^3P) \rightarrow Br + OH$	$5.8\times10^{-12}\exp(-1500/T)$
(R107)	$HBr + OH \rightarrow Br + H_2O$	1.1×10^{-11}
(R108)	$HOBr + O(^3P) \rightarrow OH + BrO$	$1.2\times10^{-10}\exp(-430/T)$
(R109)	$HOBr + OH \rightarrow H_2O + BrO$	1.1×10^{-12}
(R110)	$HOBr + Cl \rightarrow HCl + BrO$	1.1×10^{-10}
(R111)	$CH_3Br + O(^1D) \rightarrow BrO +$ products	1.8×10^{-10}
(R112)	$CH_3Br + OH \rightarrow H_2O +$ products	$2.35\times10^{-12}\exp(-1300/T)$
(R113)	$CF_2ClBr + O(^1D) \rightarrow BrO +$ products	1.5×10^{-10} (36% quenching)
(R114)	$CF_3Br + O(^1D) \rightarrow BrO +$ products	1.0×10^{-10} (59% quenching)

Table C.4: *Heterogeneous reactions included in the model*

Reaction		Uptake coefficient		
		liquid	NAT	ice
(H1)	$N_2O_5 + H_2O \rightarrow 2HNO_3$	0.1^a	0.0003^d	0.01^d
(H2)	$ClONO_2 + H_2O \rightarrow HOCl + HNO_3$	$f(T, p_{H_2O}, p_{HCl})^{b,c}$	$f(T, p_{H_2O})^{e,f}$	0.3^i
(H3)	$ClONO_2 + HCl \rightarrow Cl_2 + HNO_3$	$f(T, p_{H_2O}, p_{HCl})^{b,c}$	$f(T, p_{H_2O}, p_{HCl})^{e,f}$	$f(T, p_{HCl})^j$
(H4)	$HOCl + HCl \rightarrow Cl_2 + H_2O$	$f(T, p_{H_2O}, p_{HCl})^{b,c}$	$f(T, p_{H_2O}, p_{HCl})^{g,f}$	$f(T, p_{HCl})^k$
(H5)	$BrONO_2 + H_2O \rightarrow HOBr + HNO_3$	$f(T, p_{H_2O})^b$	-	-
(H6)	$HOBr + HCl \rightarrow BrCl + H_2O$	$f(T, p_{H_2O}, p_{HCl})^{b,c}$	-	-
(H7)	$N_2O_5 + HCl \rightarrow + ClNO_2 + HNO_3$	-	$f(T, p_{HCl})^j$	$f(T, p_{HCl})^j$

[a] *DeMore et al.* [1997]; [b] as described by *Hendricks et al.* [1999]; [c] H_{HCl} for ternary solutions from *Luo et al.* [1995]; [d] *Hanson and Ravishankara* [1993]; [e] *Hanson and Ravishankara* [1993a]; [f] *Carslaw et al.* [1997] ("scheme 1"); [g] *Abbatt and Molina* [1992]; [h] *Zhang et al.* [1994]; [i] *Hanson and Ravishankara* [1991]; [j] as described by *Drdla et al.* [1993], data from *Leu* [1988] and *Hanson and Mauersberger* [1990]; [k] γ=0.3×γ(H3)/γ_{max}(H3) (γ as recommended by *DeMore et al.* [1997], but with (T, p_{HCl})-dependence as γ(H3))

Bibliography

Abbatt, J. D. P. and M. J. Molina, The Heterogeneous Reaction of HOCl + HCl → Cl_2 + H_2O on Ice and Nitric Acid Trihydrate: Reaction Probabilities and Stratospheric Implications, *Geophys. Res. Lett.*, *19*, (5), 461–464, 1992.

Allen, D. R., R. M. Bevilacqua, G. E. Nedoluha, C. E. Randall and G. L. Manney, Unusual stratospheric transport and mixing during the 2002 Antarctic winter, *Geophys. Res. Lett.*, *30*, (12), 2003.

Bormann, N., M. Matricardi and S. Healy, A fast radiative-transfer model for the assimilation of infrared limb radiances from MIPAS, *Q. J. R. Meteorol. Soc.*, *131*, 1631–1653, 2005.

Bovensmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, S. Noël and V. V. Rozanov, SCIAMACHY: Mission objectives and measurement modes, *J. Atmosph. Sci.*, *56*, (2), 127–150, 1999.

Brasseur, G., M. H. Hitchman, S. Walters, M. Dymek, E. Falise and M. Pirre, An interactive chemical dynamical radiative two-dimensional model of the middle atmosphere, *J. Geophys. Res.*, *95*, 5639–5655, 1995.

Carslaw, K. S., T. Peter and R. Müller, Uncertainties in reactive uptake coefficients for solid stratospheric particles — 2. Effect on ozone depletion, *Geophys. Res. Lett.*, *24*, (14), 1747–1750, 1997.

Chipperfield, M. P., B. V. Khattatov and D. J. Lary, Sequential assimilation of stratospheric chemical observations in a three-dimensional model, *J. Geophys. Res.*, *107*, (D21), 2002.

Cohn, S. E., An introduction to estimation theory, *J. Meteor. Soc. Japan*, *75*, 257–288, 1997.

Courtier, P., J. N. Thépaut and A. Hollingsworth, A strategy for operational implementation of 4D-Var, using an incremental approach, *Q. J. R. Meteorol. Soc.*, *120*, (519), 1367–1387, 1994.

Daley, R., *Atmospheric Data Analysis*, Cambridge Univ. Press, 1991.

DeMore, W. B. et al., Chemical kinetics and photochemical data for use in stratospheric modeling, JPL Publication 97-4, Jet Propulsion Laboratory, Januar 1997.

Derber, J. C. and W.-S. Wu, The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system, *Month. Weath. Rev.*, *126*, (8), 2287–2299, 1998.

Deshler, T., M. E. Hervig, D. J. Hofmann, J. M. Rosen and J. B. Liley, Thirty years of in situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41°N), using balloon-borne instruments, *J. Geophys. Res.*, *108*, (D5), 2003.

Drdla, K., R. P. Turco and S. Elliott, Heterogeneous chemistry on antarctic polar stratospheric clouds: A microphysical estimate of the extent of chemical processing, *J. Geophys. Res.*, *98*, (D5), 8965–8981, 1993.

Elbern, H., H. Schmidt and A. Ebel, Variational data assimilation for tropospheric chemistry modeling, *J. Geophys. Res.*, *102*, (D13), 15,967–15,985, 1997.

Elbern, H., J. Hendricks and A. Ebel, A climatology of tropopause folds by global analyses, *Theor. Appl. Climatol.*, *59*, 181–200, 1998.

El Serafy, G., R. van der A, H. Eskes and H. Kelder, Assimilation of 3D ozone field in global chemistry-transport models using Kalman filter, *Adv. Space Res.*, *30*, (11), 2473–2478, 2002.

Errera, Q. and D. Fonteyn, Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, *J. Geophys. Res.*, 2001.

Eskes, H., P. F. J. van Velthoven and H. M. Kelder, Global ozone forecasting based on ERS-2 GOME observations, *Atmos. Chem. Phys.*, *2*, 271–278, 2002.

Eskes, H., A. Segers and P. van Velthoven, Ozone forecasts of the stratospheric polar vortex-splitting event in September 2002, *J. Atmosph. Sci.*, *62*, 812–821, 2004.

Faure, C. and Y. Papegay, Odyssée Version 1.7 Users Guide, Technical Report 0224, Institut National de Recherche en Informatique et en Automatique, September 1998.

Fierli, F., A. Hauchecorne, S. Bekki, B. Théodore and O. F. d'Andon, Data assimilation of stratospheric ozone using a high-resolution transport model, *Geophys. Res. Lett.*, *29*, (10), 2002.

Fischer, H. and H. Oelhaf, Remote sensing of vertical profiles of atmospheric trace constituents with MIPAS limb-emission spectrometers, *Appl. Opt.*, *35*, (16), 2787–2796, 1996.

Fisher, M. and D. Lary, Lagrangian four-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, *121*, (527), 1681–1704, 1995.

Funke, B. et al., Retrieval of stratospheric NO_x from 5.3 and 6.2 μm nonlocal thermodynamic equilibrium emissions measured by Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) on Envisat, *J. Geophys. Res.*, *110*, (D09302), 2005.

Geer, A. J., C. Peubey, R. N. Bannister, R. Brugge, D. R. Jackson, W. A. Lahoz, S. Migliorini, A. O'Neill and R. Swinbank, Assimilation of stratospheric ozone from MIPAS into a global general-circulation model: The September 2002 vortex split, *Q. J. R. Meteorol. Soc.*, *132*, 231–257, 2006.

Giering, R., *Tangent linear and Adjoint Model Compiler, Users manual TAMC Version 5.2*, September 1999.

Giering, R. and T. Kaminski, Recipes for adjoint code construction, *ACM Trans. Mathematical Software*, *24*, (4), 437–474, 1998.

Glatthor, N. et al., Spaceborne ClO observations by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) before and during the Antarctic major warming in September/October 2002, *J. Geophys. Res.*, *109*, (D11307), 2004.

Gropp, W., E. Lusk, N. Doss and A. Skjellum, A high-performance, portable implementation of the MPI message passing interface standard, *Parallel Computing*, *22*, (6), 789–828, 1996.

Hanson, D. and K. Mauersberger, Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere, *Geophys. Res. Lett.*, *15*, (8), 855–858, 1988.

Hanson, D. R. and K. Mauersberger, HCl/H_2O solid-phase vapor pressures and HCl solubility in ice, *J. Phys. Chem.*, *94*, 4700–4705, 1990.

Hanson, D. R. and A. R. Ravishankara, The reaction probabilities of $ClONO_2$ and N_2O_5 on polar stratospheric cloud materials, *J. Geophys. Res.*, *96*, (D3), 5081–5090, 1991.

Hanson, D. R. and A. R. Ravishankara, Response to "Comment on Pororities of Ice Films Used to Simulate Stratospheric Cloud Surfaces", *J. Phys. Chem.*, *97*, 2802–2803, 1993.

Hanson, D. R. and A. R. Ravishankara, Reaction of $ClONO_2$ with HCl on NAT, NAD and frozen sulfuric acid and hydrolysis of N_2O_5 and $ClONO_2$ on frozen sulfuric acid, *J. Geophys. Res.*, *98*, 22931–22936, 1993a.

Hendricks, J., *Modellstudien zur Bedeutung heterogener Reaktionen auf und in Sulfataerosolen für die Photochemie der Tropopausenregion mittlerer Breiten*, Ph.D. Thesis, Institut für Geophysik und Meteorologie der Universität zu Köln, 1997.

Hendricks, J., E. Lippert, H. Petry and A. Ebel, Heterogeneous reactions on and in sulfate aerosols: Implications for the chemistry of the midlatitude tropopause region, *J. Geophys. Res.*, *104*, (D5), 5531–5550, 1999.

Hendricks, J., F. Baier, G. Günter, B. C. Krüger and A. Ebel, Stratospheric ozone depletion during the 1995-1996 Artic winter: 3-D simulations on the potential role of different PSC types, *Ann. Geophysicae*, *19*, 163–1181, 2001.

Hoffmann, L. and M. Riese, Quantitative transport studies based on trace gas assimilation, *Adv. Space Res.*, *33*, 1068–1072, 2004.

Höpfner, M. et al., Spaceborne ClO observations by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) before and during the Antarctic major warming in September/October 2002, *J. Geophys. Res.*, *109*, (D11308), 2004.

Intergovernmental Panel on Climate Change, Climate change 2001, technical report, IPCC, 2001.

Kalman, R. E., A new approach to linear filtering and prediction problems, *Trans. ASME, J. Basic Eng.*, *82D*, 35–45, 1960.

Kalnay, E., *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge Univ. Press, 2003.

Khattatov, B. V., J.-F. Lamarque, L. V. Lyjak, R. Menard, P. Levelt, X. Tie, G. P. Brasseur and J. C. Gille, Assimilaton of satellite observations of long-lived chemical species in global chemistry transport models, *J. Geophys. Res.*, *105*, (D23), 29,135–29,144, 2000.

Krämer, M. et al., Intercomparison of stratospheric chemistry models under polar vortex conditions, *J. Atmos. Chem.*, *45*, 51–77, 2003.

Lahoz, W. A., Atmospheric modelling, in *Data Assimilation for the Earth System*, edited by R. Swinbank, V. Shutyaev, and W. A. Lahoz, vol 26 of *NATO Science Series*, Kluwer Academic, 2003.

Lary, D. J., B. Khattatov and H. Y. Mussa, Chemical data assimilation: A case study of solar occultation data from the ATLAS 1 mission of the Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS), *J. Geophys. Res.*, *108*, (D15), 2003.

Leu, M.-T., Laboratory studies of sticking coefficients and heterogeneous reactions important in the antarctic stratosphere, *Geophys. Res. Lett.*, *15*, 17–20, 1988.

Levelt, P. F., B. V. Khattatov, J. C. Gille, G. P. Brasseur, X. X. Tie and J. Waters, Assimilation of MLS ozone measurements in the global three-dimensional chemistry-transport model ROSE, *Geophys. Res. Lett.*, 1998.

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

López-Puertas, M., B. Funke, S. Gil-López, T. von Clarmann, G. P. Stiller, M. Höpfner, S. Kellmann, H. Fischer and C. H. Jackmann, Observation of NO_x enhancement and ozone depletion in the Northern and Southern Hemispheres after the October–November 2003 solar proton events, *J. Geophys. Res.*, *110*, 2005*a*.

López-Puertas, M., B. Funke, S. Gil-López, T. von Clarmann, G. P. Stiller, M. Höpfner, S. Kellmann, G. Mengistu Tsidu, H. Fischer and C. H. Jackmann, HNO_3, N_2O_5 and $ClONO_2$ enhancements after the October–November 2003 solar proton events, *J. Geophys. Res.*, *110*, 2005*b*.

Luo, B. P., K. S. Carslaw, T. Peter and S. L. Clegg, Vapor pressures of $H_2SO_4/HNO_3/HCl/HBr/H_2O$ solutions to low stratospheric temperatures, *Geophys. Res. Lett.*, *22*, 247–250, 1995.

Lyster, P. M., S. E. Cohn, R. Ménard, L.-P. Chang, S.-J. Lin and R. G. Olsen, Parallel implementation of a Kalman filter for constituent data assimilation, *Month. Weath. Rev.*, *125*, 1674–1686, 1997.

Majewski, D., D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen and J. Baumgardner, The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests, *Month. Weath. Rev.*, *130*, 319–338, 2001.

Marti, J. and K. Mauersberger, A survey and new measurements of ice vapor pressure at temperatures between 170 and 250K, *Geophys. Res. Lett.*, *20*, (5), 363–366, 1993.

McRae, G. J., W. R. Goodin and J. H. Seinfeld, Numerical solution of the atmospheric diffusion equation for chemically reacting flows, *J. Comp. Phys.*, *45*, 1–42, 1982.

Ménard, R., S. E. Cohn, L.-P. Chang and P. M. Lyster, Assimilation of stratospheric chemical tracer observations using a Kalman Filter, Part I: Formulation, *Month. Weath. Rev.*, *128*, (8), 2654–2671, 2000*a*.

Ménard, R., S. E. Cohn, L.-P. Chang and P. M. Lyster, Assimilation of stratospheric chemical tracer observations using a Kalman Filter, Part II: χ^2-validated results and analysis of variance and correlation dynamics, *Month. Weath. Rev.*, *128*, (8), 2672–2686, 2000*b*.

Mengistu Tsidu, G. et al., NO_y from Michelson Interferometer for Passive Atmospheric Sounding on Environmental Satellite in September/October 2002, *J. Geophys. Res.*, *110*, (D11301), 2005.

Meyer, J., *Solar Occultation Measurements with SCIAMACHY in the UV-Visible-IR Wavelength Region*, Ph.D. Thesis, Universität Bremen, 2004.

Migliorini, S., C. Piccolo and C. D. Rodgers, Intercomparison of direct and indirect measurements: Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) versus sonde ozone profiles, *J. Geophys. Res.*, *109*, (D19316), 2004.

Nazaryan, H., M. P. McCormick and J. M. Russel III, New studies of SAGE II and HALOE ozone profile and long-term change comparisons, *J. Geophys. Res.*, *110*, (D09305), 2005.

Nocedal, J., Updating quasi-newton matrices with limited storage, *Math. Comput.*, *35*, (151), 773–782, 1980.

Pichler, H., *Dynamik der Atmosphäre*, Spektrum, 3. Auflage, 1997.

Polavarapu, S., S. Ren, Y. Rochon, D. Sankey, N. Ek, J. Koshyk and D. Tarasick, Data assimilation with the Canadian Middle Atmosphere Model, *Atmos.-Ocean*, *43*, (1), 77–100, 2005.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipies*, Cambridge Univ. Press, 1992.

Ridolfi, M. et al., Optimized forward model and retrieval scheme for MIPAS near-real-time processing, *Appl. Opt.*, *39*, (8), 1323–1340, 2000.

Riishøjgaard, L. P., A direct way of specifying flow-dependent background error correlations for meteorological analysis systems, *Tellus*, *50A*, 42–57, 1998.

Rodgers, C. D., *Inverse Methods for Atmospheric Sounding*, World Scientific, 2000.

Sander, S. P. et al., Chemical kinetics and photochemical data for use in atmospheric studies, JPL Publication 02-25, Jet Propulsion Labaratory, Februar 2003.

Sandu, A., J. G. Verwer, M. V. Loon, G. R. Carmichael, F. A. Potra, D. Dabdub and J. H. Seinfeld, Benchmarking stiff ODE solvers for atmospheric chemistry problems — I. Implicit vs explicit, *Atmos. Environ.*, *31*, (19), 3151–3166, 1997a.

Sandu, A., J. G. Verwer, J. G. Blom, E. J. Spee, G. R. Carmichael and F. A. Potra, Benchmarking stiff ODE solvers for atmospheric chemistry problems II: Rosenbrock solvers, *Atmos. Environ.*, *31*, (20), 3459–3472, 1997b.

Sandu, A., D. N. Daescu and G. R. Carmichael, Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I – Theory and software tools, *Atmos. Environ.*, *37*, 5083–5096, 2003.

Segers, A. J., H. J. Eskes, R. J. van der A, R. F. van Oss and P. F. J. van Velthoven, Assimilation of GOME ozone profiles and a global chemistry-transport model using a Kalman filter with Anisotropic covariance, *Q. J. R. Meteorol. Soc.*, *131*, 477–502, 2005.

Seinfeld, J. H. and S. N. Pandis, *Atmospheric chemistry and physics*, Wiley-Interscience, 1998.

Simmons, A. J. and D. M. Burridge, An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates, *Month. Weath. Rev.*, *109*, 758–766, 1981.

Stajner, I., N. Winslow, R. B. Rood and S. Pawson, Monitoring of observation errors in the assimilation of satellite ozone data, *J. Geophys. Res.*, *109*, (D06309), 2004.

Staniforth, A. and J. Côté, Semi-lagrangian integration schemes for atmospheric models — a rewiev, *Month. Weath. Rev.*, *119*, 2206–2223, 1991.

Stoer, J. and R. Bulirsch, *Numerische Mathematik 2*, Springer, 3. Auflage, 1990.

Struthers, H., R. Brugge, W. A. Lahoz, A. O'Neill and R. Swinbank, Assimilation of ozone profiles and total column measurements into a general circulation model, *J. Geophys. Res.*, *107*, (D20), 2002.

Talagrand, O., Assimilation of observations, an introduction, *J. Meteor. Soc. Japan*, *75*, (1B), 191–209, 1997.

Talagrand, O., A posteriori evaluation and verification of analysis and assimilation algorithms, in *Workshop on Diagnosis of Data Assimilation Systems*, European Centre for Medium-range Weather Forecasts Reading, England, 2-4 November, 1998.

Talagrand, O. and P. Courtier, Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. R. Meteorol. Soc.*, *113*, 1311–1328, 1987.

Thomason, L. W., L. R. Poole and T. Deshler, A global climatology of stratospheric aerosol surface area densety deduced from Stratospheric Aerosol and Gas Experiment II measurements: 1984–1994, *J. Geophys. Res.*, *102*, (D7), 8967–8976, 1997.

Thomason, L. W., S. P. Burton, N. Iyer, J. M. Zawodny and J. Anderson, A revised water vapor product for the Stratospheric Aerosol and Gas Experiment (SAGE) II version 6.2 data set, *J. Geophys. Res.*, *109*, (D06312), 2004.

van Leeuwen, P. J., A variance-minimizing filter for large-scale applications, *Month. Weath. Rev.*, *131*, 2071–2084, 2003.

van Leeuwen, P. J. and G. Evensen, Data assimilation and inverse methods in terms of a probabilistic formulation, *Month. Weath. Rev.*, *124*, 2898–2913, 1996.

Verwer, J. G., E. J. Spee, J. G. Blom and W. Hundsdorfer, A second order rosenbrock method applied to photochemical dispersion problems, Technical Report MAS R9717, Centrum voor Wiskunde en Informatica, 1997.

von Clarmann, T. et al., Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), *J. Geophys. Res.*, *108*, (D23), 2003.

Wang, H. J., D. M. Cunnold, L. W. Thomason, J. M. Zawodny and G. E. Bodeker, Assessment of SAGE II version 6.1 ozone data quality, *J. Geophys. Res.*, *107*, (D23), 2002.

Wang, D. Y. et al., Comparisons of MIPAS/ENVISAT ozone profiles with SMR/ODIN and HALOE/UARS observations, *Adv. Space Res.*, *36*, (5), 927–931, 2005.

Weaver, A. and P. Courtier, Correlation modelling on the sphere using a generalized diffusion equation, *Q. J. R. Meteorol. Soc.*, *127*, 1815–1846, 2001.

World Meteorological Organization, Scientific assessment of ozone depletion: 1998, technical report, WMO, 1999.

Zhang, R., J. T. Jayne and M. J. Molina, Heterogeneous interaction of $ClONO_2$ and HCl with sulfuric acid tetrahydrate: Implications for the stratosphere, *J. Phys. Chem.*, *98*, 867–874, 1994.

Acknowledgements

In the first place I want to thank Prof. Dr. Adolf Ebel, who gave me the opportunity to learn and work at RIU. I am thankful for his support and his interest in my work.

I am especially grateful to Dr. Hendrik Elbern for a lot of things: For sparking my interest in data assimilation, for his ideas and suggestions, for being always available for questions and discussions, for his advice and support during the whole time of this work, for managing the SACADA project, and finally for reading and considerably improving this manuscript.

Further, I want to thank Prof. Dr. Michael Kerschgens, who kindly accepted to act as a referee for this thesis.

Considerable computational resources have been used to accomplish this work and I want to thank those, who provided and maintained them. Georg Piekorz kept the computational infrastructure at RIU running and solved all problems fast and thoroughly whenever they occurred. A large part of the case studies has been computed at the ZAIK/RRZK cluster Clio. Special thanks to Lars Packschies and Daniel Tiggemann, who helped me a lot.

This work would not have been possible without support from a number of groups, persons and institutions. I am especially grateful to the German Weather Service for providing the GME software for the SACADA project. Prof. R. Botchorishvili contributed a lot to the development of the SACADA system. His ideas and suggestions on numerical issues improved the model considerably. Dr. Johannes Hendricks gave me helpful advice with the stratospheric reaction scheme. MIPAS retrievals were provided by M. Kiefer, G. Stiller, T. von Clarmann and the AME-group at Forschungszentrum Karls-

ruhe. H. Bovensmann, J. Meyer, A. Rozanov and C. von Savigny made the SCIAMACHY limb and solar occultation data products available. The preparation of meteorological initial data was accomplished by S. Pott. T. Erbertseder provided preprocessed MPE observational data. SAGE II profiles have been obtained from the NASA Langley Research Center / NASA Langley Radiation and Aerosol Branch. The HALOE trace gas profiles have been supplied by Hampton University, Virginia and the NASA Langley Research Center. I grateful acknowledge the funding of the SACADA project through the atmospheric research programe AFO-2000, FZK 07ATF48, by the German Federal Ministry of Education and Research.

I want to thank Achim Strunk. It would have been tedious without him. Thanks for the many discussions about –and beyond– data assimilation and the diffusion approach. Many thanks to Lars Nieradzik for proof-reading this work and for participating in our 'Early Bird Contest'. Thank you to Nadine Gorris, Lars Nieradzik, and Elmar Friese for putting me into good humour almost every day. In the same way, everyone else at RIU contributed to the positive atmosphere. Thanks to everyone being "halb eins in der Mensa".

Finally, I want to thank my parents Christa and Eberhardt Schwinger for their support during all the years.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit –einschließlich Tabellen, Karten und Abbildungen–, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie –abgesehen von unten angegebenen Teilpublikationen– noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. A. Ebel betreut worden.

Teilpublikationen:

Elbern, H., F. Baier, M. Bittner, R. Bochorishvili, H. Bovensmann, L. Hoffmann, W. Joppich, J. Meyer, M. Riese, J. Schwinger, G. Stiller, Th. von Clarmann, SACADA, a new chemical data assimilation system for earth observation, in *Results of the German atmospheric research programme AFO2000 final publication*, Editor GSF, 2005.

Schwinger, J., H. Elbern, und R. Botchorishvili, A new 4-Dimensional variational Assimilation System applied to ENVISAT MIPAS observations, *Proceedings of the Envisat & ERS Symposium*, Salzburg, 2004.