

# Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration

H. Elbern,<sup>1,2</sup> J. Schwinger,<sup>1</sup> and R. Botchorishvili<sup>3,4</sup>

Received 21 February 2009; revised 12 October 2009; accepted 29 October 2009; published 20 March 2010.

[1] A novel stratospheric chemical data assimilation system has been developed and applied to Environmental Satellite Michelson Interferometer for Passive Atmospheric Sounding (ENVISAT/MIPAS) data, aiming to combine the sophistication of the four-dimensional variational (4D-var) technique with flow-dependent covariance modeling and also to improve numerical performance. The system is tailored for operational stratospheric chemistry state monitoring. The atmospheric model of the assimilation system includes a state-of-the-art stratospheric chemistry transport module along with its adjoint and the German weather service's global meteorological forecast model, providing meteorological parameters. Both models share the same grid and same advection time step, to ensure dynamic consistency without spatial and temporal interpolation errors. A notable numerical efficiency gain is obtained through an icosahedral grid. As a novel feature in stratospheric variational data assimilation a special focus was placed on an optimal spatial exploitation of satellite data by dynamic formulation of the forecast error covariance matrix, providing potential vorticity controlled anisotropic and inhomogeneous influence radii. In this first part of the study the design and numerical features of the data assimilation system is presented, along with analyses of two case studies and a posteriori validation. Assimilated data include retrievals of O<sub>3</sub>, CH<sub>4</sub>, N<sub>2</sub>O, NO<sub>2</sub>, HNO<sub>3</sub>, and water vapor. The analyses are compared with independent observations provided by Stratospheric Aerosol and Gas Experiment II (SAGE II) and Halogen Occultation Experiment (HALOE) retrievals. It was found that there are marked improvements for both analyses and assimilation based forecasts when compared with control model runs without any data ingestion.

**Citation:** Elbern, H., J. Schwinger, and R. Botchorishvili (2010), Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: System configuration, *J. Geophys. Res.*, 115, D06302, doi:10.1029/2009JD011953.

## 1. Introduction

[2] In March 2002 the European research satellite ENVISAT was launched into a Sun-synchronous orbit, carrying the sensors Michelson Interferometer for Passive Atmospheric Sounding (MIPAS), the Global Ozone Monitoring by Occultation of Stars (GOMOS), and the Scanning Imaging Absorption Spectrometer for Atmospheric Cartography (SCIAMACHY) aboard, which are delivering an unprecedented wealth of observations of stratospheric trace gases with global coverage.

[3] It is the objective of data assimilation to provide an estimate of the state of the atmosphere from heterogeneous, irregularly distributed observational data of differencing accuracies, fused with a numerical model of the atmosphere. This is achieved mostly by the use of estimation methods adapted to large-scale problems. For the sake of mathematical rigor, an objective optimality criterion must be invoked. In most cases a Best Linear Unbiased Estimation (BLUE) is applied, which implies a least square optimum [see Kalnay, 2003].

[4] Sun-synchronous satellite observations are limited to measurements at a single local time, which is clearly a limitation for atmospheric chemistry. The use of spatial-temporal DA algorithms (such as 4D-Var) can effectively propagate the observation information to other times, and thus provide a complete temporal estimation of the chemical state of the atmosphere. A thorough overview of chemical data assimilation systems is provided by Lahoz *et al.* [2007] and Geer *et al.* [2006].

[5] In the realm of advanced spatiotemporal data assimilation algorithms resting on Gaussian error characteristics

<sup>1</sup>Rhenish Institute for Environmental Research, University of Cologne, Cologne, Germany.

<sup>2</sup>Also at Institute for Chemistry and Dynamics of the Geosphere, Institute 2: Troposphere, Research Centre Jülich, Jülich, Germany.

<sup>3</sup>Scientific Computing and Algorithms Institute, Fraunhofer Society, Saint Augustin, Germany.

<sup>4</sup>Now at Department of Exact and Natural Sciences, Tbilisi State University, Tbilisi, Georgia.

providing for a BLUE, there are only two families of techniques, namely the Kalman filter [e.g., Kalman, 1960; Cohn, 1997] and the 4D-var method [e.g., Talagrand and Courtier, 1987]. The former is a sequential method; that is, the model state is corrected at times when observations are encountered. The Kalman filter possesses the theoretical advantage that the background error covariance matrix (BECM) is evolved by a prognostic equation, and the analysis error covariances are provided by a diagnostic equation. However, with two model integrations per dimension of model-phase space, the implementation of the full Kalman filter algorithm is not feasible for atmospheric applications and complexity reduced Kalman filter algorithms must be applied [Hanea et al., 2004]. To the knowledge of the authors it was for regional scale chemical data assimilation that Kalman filter implementations, which provide analysis error covariance matrices, were studied first. In the Netherlands two chemistry transport models (CTMs) were furnished with sophisticated implementations of complexity reduced Kalman filters. These include the reduced rank square root Kalman filter of the Long Term Ozone Simulation (LOTOS) model [van Loon et al., 2000] and the EUROS model [Hanea et al., 2004]. The reduced rank square-root approach was selected to factorize covariance matrices by a few principal components [Verlaan and Heemink, 1995].

[6] 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 minimized in a root-mean-square sense. The 4D-var method is sufficiently efficient to be implemented without serious simplifications. However, it lacks means to update the BECM, which must be prescribed in some way instead. While in most cases this is implemented in a static way, any dynamical evolution of the BECM must be constructed by additional information. Given typical state space vectors with dimensions of  $\mathcal{O}(10^6\text{--}10^7)$ , the data volume of the squared dimension of the BECM is explicitly intractable for comprehensive three-dimensional chemistry models. As the background field is given by a short-range forecast in present assimilation systems, it is the error statistics of this forecast that is to be approximated. Further, in 4D-var there is no direct strategy to derive an analysis error estimate. Addressing the latter problem, approaches like those proposed by Fisher and Courtier [1995] can be applied.

[7] A first application of the 4D-var technique with a heavily reduced 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. A state of the art tropospheric chemistry mechanism was introduced to variational assimilation by Elbern et al. [1997]. First extensions to the full 4D-var are given by Elbern and Schmidt [1999, 2001]. The 4D-var method proved flexible enough for generalization to emission rate inversion with reactive chemistry, as shown by Elbern et al. [2007].

[8] Errera and Fonteyn [2001] presented the first application of the 4D-var method in the stratosphere with a comprehensive stratospheric CTM, now the Belgian Assimilation System of Chemical Observations from ENVISAT (BASCOE; see also Errera et al. [2008] for further developments).

[9] Since an optimal analysis requires a realistic representation of error statistics, the treatment of the BECM constitutes a core task in designing an assimilation system. A properly specified BECM does not only balance the forecast or background error with respect to observation errors, but also guides the spreading of measurement information given a statistically well estimated influence or decorrelation length. By using known correlations, the BECM can therefore constitute a key instrument to exploit the information contents of a retrieval or observation as thoroughly as possible. In practice, the treatment of the BECM involves two independent general problems: First, an algorithmic formulation must be found to model an extremely high dimensional matrix that is generally too big to be represented explicitly. Secondly, statistically useful entries must be inferred in some suitable way.

[10] In practice, either only variances are considered (diagonal BECM [Errera et al., 2008]), or the specification of covariances in chemical data assimilation mainly rests on assumptions of homogeneity (constant horizontal correlation lengths all over the globe) and isotropy (constant horizontal correlation lengths in each direction). Despite the necessary simplicity, a skillful parameterization should be capable of representing the relevant structures of the background error covariances. This includes the possibility of modeling inhomogeneous and anisotropic correlations length scales.

[11] In the realm of meteorological data assimilation, the formulation of the BECM has attracted much research (see Bannister [2008a, 2008b] for a comprehensive survey). An early attempt to move to anisotropic and inhomogeneous background errors is presented by Thiebaut [1976] by a suitably defined autoregressive scheme. Thépaut et al. [1996] and Otte et al. [2001] further demonstrated the need to relax the constraint of isotropy. A latitudinally dependent correlation function was defined by Wu et al. [2002] through recursive filtering. Purser et al. [2003] generalized the approach to variable anisotropy and inhomogeneity adaptive to geographic location. Weaver and Courtier [2001] applied a diffusion method, providing the same statistical properties as the recursive filtering. While for the former authors claim higher numerical efficiency, the latter show promise to better account for abrupt correlation changes, like at fronts and other air mass boundaries, due to the local control by diffusion coefficients.

[12] A straightforward approach to obtain flow dependency by anisotropy and inhomogeneity has been proposed by Riishøjgaard [1998], where the BECM is controlled by a function of variability of concentration levels, but also mentioned the possibility to use potential vorticity (PV) fields. While this is a direct and easy to implement method without a need to form an ensemble of model fields, the validity of the method rests on the assumption that similar field values imply similar origin, separated only by distorting flows.

[13] In the challenging domain of tropospheric chemical data assimilation, Hölzemann et al. [2001] introduced inhomogeneous BECMs, to account for urban-to-rural chemical regimes changes of the boundary layer. On the global scale, Segers et al. [2005] applied a complexity reduced global stratospheric Kalman filter system with anisotropic covariance formulation by a parameterization of correlations. Using MIMOSA (Modèle Isentrope de transport MésO échelle de Ozone Stratosphérique par Advection), a horizontally high-

resolution transport model with 16 isentropic levels, it was possible to preserve fine-scale structures in the analyzed ozone field [Fierli *et al.*, 2002]. Background error correlations were flow dependent and anisotropic, specified in terms of distance and by PV field.

[14] The general objective of the first part of this study is to introduce a data assimilation system, which combines the full sophistication of a dynamically controlled BECM formulation with a complex state of the art reactive chemistry model using the 4D-var technique.

[15] The specific objectives of this paper is to validate an efficient flow-dependent formulation of the spatial background error covariances, while maintaining ability to (1) make best use of all available (satellite) data, by algorithmic capability to extend observation results spatially, while preserving the BLUE property, (2) ensure chemical and dynamical consistency by application of a state of the art chemistry mechanism, and (3) to provide numerical efficiency (grid design, parallelization), to allow for near real time operation. The system SACADA (Synoptic Analysis of Chemical Constituents by Advanced Data Assimilation), presented in this study, has been designed efficiently enough to provide for daily routine operations.

[16] In order to comply with the first item, the study is designed to implement and test the diffusion approach proposed by Weaver and Courtier [2001]. The resulting background error covariance operator will be shown to be well suited for the application with large models and allows for anisotropic and inhomogeneous background error correlations, a feature that was utilized to devise a flow-dependent formulation of the BECM.

[17] This work lays the foundation for a follow-up study addressing a more sophisticated a posteriori validation of the assimilation results in observation space [Desroziers *et al.*, 2005]. This paper is organized as follows. Section 2 describes the theoretical background of the assimilation approach used. In section 3 the meteorological driver model and its geodetic grid, together with the chemistry transport model are presented. Further, the data assimilation setup is given in section 4, with emphasis placed on background error covariance modeling. The satellite data involved in this study are introduced in section 5. Results and statistical evaluations of the assimilation runs are presented in section 6, followed by conclusions made in section 7.

## 2. Theoretical Basics of 4D-var

[18] For the sake of consistency a brief outline of the variational formalism is given, which is applied in this paper. More comprehensive expositions are given by Talagrand and Courtier [1987], Elbern *et al.* [1997] and Wang *et al.* [2001].

[19] Let  $\mathcal{M}_i$  denote the nonlinear model operator that propagates the atmospheric state  $\mathbf{x}_0$  at time  $t_0$  to time  $t_i$ . Further, it is assumed that the model and its controlling parameters, except initial values, are of sufficient quality for a forecast of the time span of interest, with statistics of forecast errors known in terms of its error covariance matrix. In addition, given a set of observations or satellite retrievals over a time interval or assimilation window  $[t_0, t_N]$ , we seek for a BLUE of the model's  $n$  initial values of the state variable  $\mathbf{x}_0 = \mathbf{x}(t_0)$ . It is assumed that background and observation errors are normally distributed and uncorrelated.

[20] The four-dimensional variational method minimizes the cost function:

$$J(\mathbf{x}_0) = J^b + J^o$$

$$= \frac{1}{2} [\mathbf{x}_0 - \mathbf{x}^b]^T \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^b]$$

$$+ \frac{1}{2} \sum_{i=0}^N [H(\mathcal{M}_i(\mathbf{x}_0)) - \mathbf{y}_i]^T \mathbf{R}^{-1} [H(\mathcal{M}_i(\mathbf{x}_0)) - \mathbf{y}_i]. \quad (1)$$

Adopting the Ide *et al.* [1997] notation as far as possible,  $\mathbf{y}_i$  is the vector of observations available within time step  $i$ ,  $\mathbf{x}^b$  is the first guess or background state at initial time, typically obtained from an earlier model forecast. The initial chemical state  $\mathbf{x}_0$  of the model at time step  $i = 0$ , at the beginning of the assimilation window, is the optimization parameter, where the most probable state is to be identified. The background error and observation error covariance matrices are denoted by  $\mathbf{B}$  and  $\mathbf{R}$ , respectively. The observation operator  $H$  maps the model state onto the observation state. Superscript  $T$  denotes transposition. As is obvious from (1),  $J^b$  are the partial costs resulting from the deviation of initial state  $\mathbf{x}_0$  from the background state  $\mathbf{x}^b$ , while  $J^o$  gives the partial costs arising from the deviation of the observations  $\mathbf{y}_i$  from the model equivalents  $H(\mathcal{M}_i(\mathbf{x}_0))$ . Equation (1) implements the model  $\mathcal{M}$  as a strong constraint.

[21] Minimization algorithms, which are suitable for high-dimensional problems, include quasi-Newton or Conjugate-Gradient methods and require gradients of the cost function with respect to the control variables  $\mathbf{x}_0$ . The calculation of  $\nabla_{\mathbf{x}_0} J^o$  is the computationally most demanding task of 4D-var data assimilation, given the number of control variables in atmospheric models is on the order  $10^6$ – $10^7$ . The feasible strategy to accomplish the calculation of the gradient makes use of the adjoint model operator. The gradient of  $J$  with respect to the initial model values  $\mathbf{x}_0$  is given by [Talagrand, 1997]

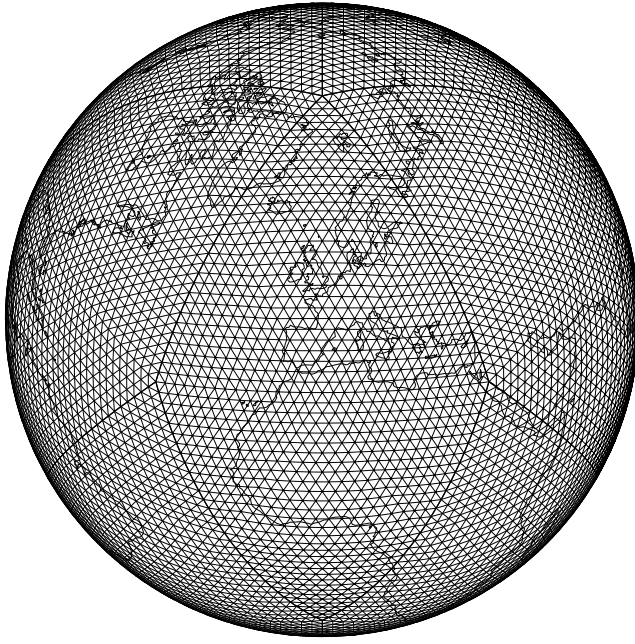
$$\nabla_{\mathbf{x}_0} J = \mathbf{B}^{-1} [\mathbf{x}_0 - \mathbf{x}^b] + \sum_{i=0}^N \mathbf{M}_i^* \mathbf{H}^T \mathbf{R}^{-1} [H(\mathbf{x}_i) - \mathbf{y}_i]. \quad (2)$$

[22] Here,  $\mathbf{M}_i^*$  is the adjoint model operator, linearized backward in time from time step  $N$  until time step  $i$ , while  $\mathbf{H}$  is the linearized observation operator  $H$ .

## 3. System Description

### 3.1. Meteorological Driver Model

[23] The SACADA chemistry 4D-var system has been assembled from scratch to allow for a couple of novel features. In 1999 the operational icosahedral meteorological forecast model of the German weather service replaced the earlier global model (GM) as well as the regional model (EM), and was named GME [Majewski *et al.*, 2001]. Its version 1.22 is adopted for SACADA and applied to compute the meteorological fields at the same temporal and spatial locations where they are needed by the chemistry module. Storing the meteorological states for each time step and recovery for forward and adjoint integration of the chemistry model avoids temporal and spatial interpolation errors of meteorological



**Figure 1.** Icosahedral grid with a resolution of  $n_i = 32$ . The distance between neighboring grid points is about 240 km.

parameters. For realization of this approach GME is integrated in the chemical 4D-var model code as a subroutine.

[24] GME has been selected because of its icosahedral grid design, which provides a nearly homogeneous distribution of grid points over the globe, avoiding singularities at the poles and crowding of grid points due to meridional convergence at high latitudes. Given the fact that most computational burden is associated with the solution of the stiff ordinary differential equations of chemistry at each grid point, a high numerical efficiency can be expected by the quasi-isotropic grid design. Since the new chemistry parts of the SACADA model, which have been developed in the framework of this study, adopt some of the numerical concepts, the relevant grid features with related discretized operators of GME are given below.

[25] In three-dimensional space an icosahedron is the highest Platonic body, shaped by 20 equilateral triangles as faces. To approximate the globe it is placed into a sphere of the same diameter, onto which the 12 icosahedral vertices and connection lines are projected. As shown in Figure 1, the resulting sections of great circles are equally subdivided into a number of  $n_i$  intervals each, to form an almost isotropic grid. Each grid point has six direct neighbors except the twelve points located at the vertices of the icosahedron (termed “special points” hereafter), which have only five direct neighbors. The area of representativeness for a grid cell is a hexagon and pentagon at the twelve special points. This approach results in a mesh with approximately common mesh size all over the globe, with the minimum and maximum separation between neighboring grid points  $\Delta_{\min}$  and  $\Delta_{\max}$  varying about 20% only. For this first version of the SACADA system it has been decided that  $n_i = 32$  intersections give 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 grid cell, denoting the horizontal representativeness in data assimilation, is about 50,000 km<sup>2</sup>. A traditional latitude-longitude grid of similar resolution, that is, a grid with the same area represented by one grid cell at the equator requires a grid spacing of, say,  $2.0^\circ \times 2.4^\circ$  resulting in 13,500 grid points per level which is about 30% more than the icosahedral grid.

[26] An orthonormal coordinate system  $(\mathbf{x}_0, \mathbf{e}_\lambda, \mathbf{e}_\phi)$  is attached to each grid point of the triangular grid, where  $\mathbf{x}_0$  is the position vector of the grid point on the unit sphere pointing in the vertical direction, and  $\mathbf{e}_\lambda$  and  $\mathbf{e}_\phi$  are affine unit vectors aligned from  $\mathbf{x}_0$  to the east and north directions, respectively. Local spherical coordinates  $(\eta, \chi)$  are used to describe the position of an arbitrary unit vector  $\mathbf{x}$  relative to  $(\mathbf{x}_0, \mathbf{e}_\lambda, \mathbf{e}_\phi)$  (see Majewski *et al.* [2001] for details). With this approach, polar singularities are avoided. The zonal wind component  $u$  and the meridional wind component  $v$  are given by

$$u = R_E \cos \chi \frac{d\eta}{dt} \quad v = R_E \frac{d\chi}{dt}, \quad (3)$$

respectively, where  $R_E$  is the radius of Earth. The GME grid is horizontally unstaggered, (Arakawa A grid). The meteorological equations are formulated and solved in the  $(\eta, \chi)$  local coordinate system. See Majewski *et al.* [2001] for details.

[27] GME uses two different algorithms for horizontal advection. Cloud water and cloud ice are advected by a semi-Lagrangian scheme [Staniforth and Côté, 1991] with necessary adaptations to the icosahedral grid, while other prognostic variables are advected by an Eulerian scheme [Majewski *et al.*, 2001]. The departure points of the trajectories are found by a two step iterative procedure. After temporal discretization, (3) gives

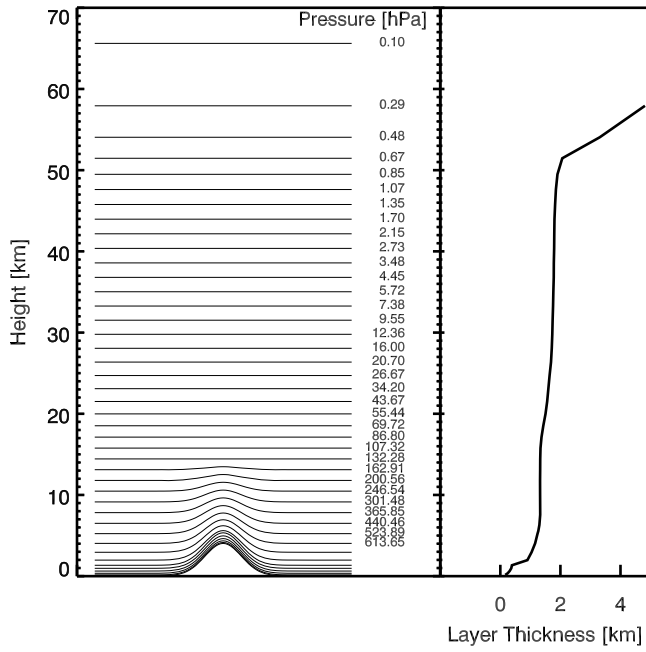
$$\Delta\eta_1 = -\frac{u(t_n)}{R_E} \Delta t \quad \Delta\chi_1 = -\frac{v(t_n)}{R_E} \Delta t, \quad (4)$$

as  $\cos \chi = 1$  at the first step, starting at the grid point with  $\chi = 0$ . The trajectory midpoint is computed at a second step

$$\Delta\eta_2 = -\frac{I_l(\mathbf{u}(t_n))}{R_E \cos \Delta\chi_1} \Delta t \quad \Delta\chi_2 = -\frac{I_l(\mathbf{v}(t_n))}{R_E} \Delta t, \quad (5)$$

where  $I$  is an operator that interpolates the local wind field at the position  $(\Delta\eta_1, \Delta\chi_1)$ . The departure point of the trajectory is then approximated by the relative position  $(2\Delta\eta_2, 2\Delta\chi_2)$ . Note that the wind field at time level  $n$  is used to compute the trajectory connecting the grid point at time  $n + 1$  with the departure point at time level  $n - 1$ . Two different interpolation operators are available: A linear operator  $I_l$ , taking into account the three nearest neighbors of the departure point, and an operator  $I_q$  performing quadratic interpolation using the values at twelve neighboring grid points (see Majewski *et al.* [2001] for details). Wind fields in (5) are calculated by linear interpolation  $I_l$ .

[28] In the vertical direction GME utilizes a hybrid  $\sigma$ -pressure coordinate system and an advection scheme that was proposed by Simmons and Burridge [1981]. A staggered (Charney-Phillips) grid is used with the geopotential and the vertical wind specified at the boundaries of layers. As the



**Figure 2.** SACADA hybrid model levels with extensions to the lower mesosphere and refinements for the (left) middle atmosphere and (right) layer thickness profile.

focus is 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, giving about 65 km height. The resulting pressure values together with the corresponding heights are shown in Figure 2.

### 3.2. Chemistry Transport Module

[29] The SACADA chemistry transport module solves the tendency equation in the local coordinate system of the icosahedral grid for a vector  $\mathbf{c}$  of volume mixing ratios,

$$\frac{\partial \mathbf{c}}{\partial t} = -\frac{u}{R_E} \frac{\partial \mathbf{c}}{\partial \eta} - \frac{v}{R_E} \frac{\partial \mathbf{c}}{\partial \chi} - w \frac{\partial \mathbf{c}}{\partial p} + (\mathbf{P} - \mathbf{L}), \quad (6)$$

with chemical production  $\mathbf{P}$  and loss  $\mathbf{L}$ . Making use of an operator split approach [McRae *et al.*, 1982; Yanenko, 1971], the partial differential equation (6) is separated 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,

$$\begin{aligned} \left(\frac{\partial \mathbf{c}}{\partial t}\right)^h &= -\frac{u}{R_E} \frac{\partial \mathbf{c}}{\partial \eta} - \frac{v}{R_E} \frac{\partial \mathbf{c}}{\partial \chi} \\ \left(\frac{\partial \mathbf{c}}{\partial t}\right)^v &= -w \frac{\partial \mathbf{c}}{\partial p} \\ \left(\frac{\partial \mathbf{c}}{\partial t}\right)^c &= (\mathbf{P} - \mathbf{L}). \end{aligned} \quad (7)$$

Let  $M^h(t_n)$ ,  $M^v(t_n)$  and  $M^c(t_n)$  be the generally nonlinear discretized integration operators, which solve (7) using meteorological parameters at time step  $(t_n)$ . Then, an approximate solution of (6) for time step  $n + 1$  may be obtained by

$$\mathbf{c}(t_{n+1}) = [M^c(t_n) \circ M^v(t_n) \circ M^h(t_n)] \mathbf{c}(t_{n-1}). \quad (8)$$

[30] In data assimilation algorithms, care for mass conservation is typically not taken, given the fact that global trace gas masses are not known with sufficient accuracy. Therefore, for efficiency considerations, the semi-Lagrangian scheme from GME has been selected as the horizontal transport algorithm for short-range forecasts. Departure points are computed according to (5) and saved together with the meteorological data during the GME run. Consequently, the operator  $M^h = I_q$  is remarkably simple. The interpolation operator  $I_q$  performs a quadratic interpolation using the twelve nearest grid points surrounding the departure point. To suppress negative volume mixing ratios, the SACADA-CTM employs the operator in its positive definite form, which means here that interpolated volume mixing ratios smaller than zero are set to zero.

[31] In the vertical direction the transport equation is solved by a linear and efficient implicit upwind algorithm  $M^v = [\mathbf{I} + 2\Delta t \mathbf{A}]^{-1}$ , where the tridiagonal matrix  $\mathbf{A}$  depends on the vertical wind given in the hybrid vertical coordinate system and the pressure differences between adjacent layers. The scheme is unconditionally stable and linear, and consequently needs no chemical states saved for adjoint calculation.

[32] Forty-eight atmospheric trace constituents are treated in the SACADA chemistry transport module, all of which are advected. The reaction scheme comprises 167 gas-phase reactions (see Tables 1 and 2) and 10 heterogeneous reactions on surfaces of Polar Stratospheric Cloud (PSC) particles and in sulphate aerosol droplets (see Table 3). The mechanism follows partly Hendricks *et al.* [2001], with extensions for the upper stratosphere and lower mesosphere. Reaction velocities follow recommendations from Sander *et al.* [2006]. A second-order Rosenbrock method is applied to solve the gas-phase chemistry (operator  $M^c$ ). Details on this solver and its properties are given by Verwer *et al.* [1997].

[33] Stratospheric particles in the SACADA chemistry transport module include sulfate aerosol droplets, solid nitric acid trihydrate particles (NAT, type Ia PSC), supercooled ternary  $\text{HNO}_3/\text{H}_2\text{O}/\text{H}_2\text{SO}_4$  solutions (STS, type Ib PSC), and water ice (type II PSC).

[34] Heterogeneous reactions rates are determined by [e.g., Seinfeld and Pandis, 1998]

$$k = \frac{1}{4} \gamma(r, T, \mathbf{c}) \bar{v} S, \quad (9)$$

where  $\bar{v}$  is the mean molecular velocity of the reacting species in the gas phase,  $S$  is the surface area density of particles in units of  $\text{m}^2/\text{m}^3$  and  $\gamma$  is the reactive uptake coefficient. The hydrolysis of  $\text{ClONO}_2$  and the reactions of  $\text{ClONO}_2$  and  $\text{HOCl}$  with  $\text{HCl}$  in sulfate aerosol/STS particles is treated following Shi *et al.* [2001], while all other  $\gamma$  are taken from the recommendations of Sander *et al.* [2006]. A comprehensive

**Table 1.** Photolysis Reactions Included in the SACADA Gas-Phase Mechanism

Reaction Number	Reaction <sup>a</sup>
(R1)	$O_2 + h\nu \rightarrow O(^3P) + O(^3P)$
(R2)	$O_3 + h\nu \rightarrow O(^3P) + O_2$
(R3)	$O_3 + h\nu \rightarrow O(^1D) + O_2$
(R4)	$H_2O + h\nu \rightarrow H + OH$
(R5)	$H_2O_2 + h\nu \rightarrow OH + OH$
(R6)	$NO_2 + h\nu \rightarrow O(^3P) + NO$
(R7)	$NO_3 + h\nu \rightarrow NO_2 + O(^3P)$
(R8)	$NO_3 + h\nu \rightarrow NO + O_2$
(R9)	$N_2O + h\nu \rightarrow N_2 + O(^1D)$
(R10)	$N_2O_5 + h\nu \rightarrow NO_2 + NO_3$
(R11)	$HNO_3 + h\nu \rightarrow OH + NO_2$
(R12)	$HNO_4 + h\nu \rightarrow HO_2 + NO_2$
(R13)	$Cl_2O_2 + h\nu \rightarrow Cl + Cl + O_2$
(R14)	$Cl_2 + h\nu \rightarrow Cl + Cl$
(R15)	$OCIO + h\nu \rightarrow O(^3P) + ClO$
(R16)	$HCl + h\nu \rightarrow Cl + H$
(R17)	$HOCl + h\nu \rightarrow Cl + OH$
(R18)	$ClONO_2 + h\nu \rightarrow Cl + NO_3$
(R19)	$CH_3Cl + h\nu \rightarrow Cl$
(R20)	$CCl_4 + h\nu \rightarrow 4 Cl + \text{products}$
(R21)	$CFCl_3 + h\nu \rightarrow 3 Cl + \text{products}$
(R22)	$CF_2Cl_2 + h\nu \rightarrow 2 Cl + \text{products}$
(R23)	$CHF_2Cl + h\nu \rightarrow Cl + \text{products}$
(R24)	$CF_2ClCFCl_2 + h\nu \rightarrow 3 Cl + \text{products}$
(R25)	$CH_3CCl_3 + h\nu \rightarrow 3 Cl + \text{products}$
(R26)	$BrO + h\nu \rightarrow Br + O(^3P)$
(R27)	$BrCl + h\nu \rightarrow Br + Cl$
(R28)	$HOBr + h\nu \rightarrow Br + OH$
(R29)	$BrONO_2 + h\nu \rightarrow Br + NO_3$
(R30)	$CH_3Br + h\nu \rightarrow Br + \text{products}$
(R31)	$CF_2ClBr + h\nu \rightarrow Cl + Br + \text{products}$
(R32)	$CF_3Br + h\nu \rightarrow Br + \text{products}$
(R33)	$HNO_4 + h\nu \rightarrow OH + NO_3$
(R34)	$ClONO_2 + h\nu \rightarrow ClO + NO_2$
(R35)	$N_2O_5 + h\nu \rightarrow NO_3 + NO + O(^3P)$
(R36)	$CH_2O + h\nu \rightarrow H + HCO$
(R37)	$CH_3O + h\nu \rightarrow H_2 + CO$

<sup>a</sup>The term “products” represents constituents that are not considered in this model.

discussion of the treatment of heterogeneous reactions is given by *Hanson et al.* [1994], *Carslaw et al.* [1997], *Hendricks et al.* [1999] and references therein. As a consequence of (9), 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.

[35] Given the fact that the stratosphere has hardly been affected by volcanic eruptions over the last decade [*Thomason et al.*, 1997; *Deshler et al.*, 2003], it was decided, that for the present assimilation system version, a single reference profile of aerosol surface area and median radius was sufficient for uptake coefficient calculations. This reference profile was calculated using data from 20 balloon-borne measurements made between 1997 and 2002 at Laramie, Wyoming [*Deshler et al.*, 2003]. A monodisperse size distribution and fixed number density is assumed for all particle classes.  $N_{\text{sulf}} = 5 \text{ cm}^{-3}$  was taken as a representative number density for stratospheric aerosols [*Deshler et al.*, 2003]. Following *Hendricks et al.* [2001], PSC particle number densities are prescribed by  $N_{\text{NAT}} = 1 \text{ cm}^{-3}$  and  $N_{\text{ice}} = 0.01 \text{ cm}^{-3}$ .

[36] Aerosol composition as well as the uptake of  $\text{HNO}_3$  into sulfate aerosol particles at low temperatures (and its

removal from the gas phase) is parameterized according to *Carslaw et al.* [1995]. Further, the solubility and the corresponding gas-phase removal of  $\text{HCl}$  and  $\text{HBr}$  in sulfate aerosol is accounted for as devised by *Luo et al.* [1995]. Radius and surface area are then calculated from known composition using  $N_{\text{sulf}}$ .

[37] Surface area densities and radii for NAT and ice particles are calculated using thermodynamic equilibrium constraints for  $\text{H}_2\text{O}$  and  $\text{HNO}_3$  over ice and NAT surfaces, as specified by *Marti and Mauersberger* [1993] and *Hanson and Mauersberger* [1988]. If the number of  $\text{H}_2\text{O}$  or  $\text{HNO}_3$  molecules is larger or smaller than predetermined by the saturation vapor pressure, the appropriate number of molecules is condensed or evaporated. Surface area and radius are inferred using  $N_{\text{ice}}$  and  $N_{\text{NAT}}$ . If ambient conditions favor ice formation, it is assumed that all other types of particles are incorporated into the ice, and the amount of  $\text{HNO}_3$  within ice is determined by its equilibrium pressure over ice surfaces [*Hanson and Mauersberger*, 1988]. As observations indicate that a supercooling of 2–4 K usually occurs before ice particles emerge [*World Meteorological Organization*, 1999, and references therein], ice formation is initiated if temperature falls 3 K below the ice frost point, or if water vapor reaches a 50% supersaturation with respect to ice surfaces. Further, it is well known [see, e.g., *World Meteorological Organization*, 1999] that homogenous nucleation of  $\text{HNO}_3$  and  $\text{H}_2\text{O}$  vapor to NAT particles is not favored energetically. Therefore, NAT formation is initiated only at rather high supersaturation of  $\text{HNO}_3$  with respect to NAT. In the presence of preexisting NAT or ice particles, no phase transition barrier is assumed. The composition of STS particles is calculated after gas-phase removal of  $\text{HNO}_3$  due to NAT formation.

## 4. Variational Assimilation System Components

### 4.1. Adjoint Chemistry Transport Module

[38] The development of the adjoint model operator  $\mathbf{M}^*$ , which is needed for the computation of the gradient  $\nabla_{\mathbf{c}_0} J$  of the cost function with respect to the initial volume mixing ratios  $\mathbf{c}_0$  (1) is constructed module by module, by forming the tangent linear of each individual line of code and its transpose. Adjoint operators  $\mathbf{M}^{c*}$  and  $\mathbf{M}^{h*}$  are constructed by application of KPP [*Sandu et al.*, 1997] and TAMC [*Giering*, 1999], respectively. In contrast to  $\mathbf{M}^v$ , operators  $\mathbf{M}^{c*}$  and  $\mathbf{M}^{h*}$  involve the recomputation of required variables starting from the volume mixing ratios values before the respective forward operator was applied [*Giering*, 1999]. For each iteration, these values are saved to disk at each forward time step prior to the horizontal advection step  $\mathbf{M}^h$  and prior to the call of the chemical solver  $\mathbf{M}^c$ . In total, two chemical model states and one meteorological state are stored per advection time step, the latter once and before the chemical data assimilation procedure commences. Consequently, during the adjoint/backward integration, the chemical states are retrieved from disc in reverse order. The overall computational costs of the adjoint integration is about a factor 2.4 larger than the forward integration, including the time of writing and reading the model states during the forward and backward integration. The nonlinear forward model integration is updated after each adjoint integration. Typical

**Table 2.** Gas-Phase Reactions Included in the SACADA Gas-Phase Mechanism

Reaction Number	Reaction <sup>a</sup>	Reaction Signature From Sander <i>et al.</i> [2006]
(R38)	$O(^3P) + O_3 \rightarrow O_2 + O_2$	A1
(R39)	$O(^1D) + O_2 \rightarrow O(^3P) + O_2$	A2, A3
(R40)	$O(^1D) + O_3 \rightarrow O_2 + O_2$	A2, A4
(R41)	$O(^1D) + O_3 \rightarrow O(^3P) + O(^3P) + O_2$	A2, A4
(R42)	$O(^1D) + H_2 \rightarrow OH + H$	A2, A5
(R43)	$O(^1D) + H_2O \rightarrow OH + OH$	A2, A6
(R44)	$O(^1D) + N_2 \rightarrow O(^3P) + N_2$	A2,A7
(R45)	$N_2O + O(^1D) \rightarrow N_2 + O_2$	A2,A8
(R46)	$N_2O + O(^1D) \rightarrow NO + NO$	A2,A8
(R47)	$CH_4 + O(^1D) \rightarrow OH + CH_3$	A2,A11
(R48)	$CH_4 + O(^1D) \rightarrow CH_2O + H_2$	A2,A11
(R49)	$HCl + O(^1D) \rightarrow OH + Cl$	A2,A12
(R50)	$HCl + O(^1D) \rightarrow H + ClO$	A2,A12
(R51)	$HBr + O(^1D) \rightarrow OH + Br$	A2,A15
(R52)	$Cl_2 + O(^1D) \rightarrow ClO + Cl$	A2,A16
(R53)	$CCl_4 + O(^1D) \rightarrow ClO + \text{products}$	A19
(R54)	$CFCl_3 + O(^1D) \rightarrow ClO + \text{products}$	A19
(R55)	$CF_2Cl_2 + O(^1D) \rightarrow ClO + \text{products}$	A19
(R56)	$CH_3Br + O(^1D) \rightarrow BrO + CH_3$	A19,A20
(R57)	$CHF_2Cl + O(^1D) \rightarrow ClO + \text{products}$	A19,A27
(R58)	$CF_2ClCFCl_2 + O(^1D) \rightarrow ClO + \text{products}$	A19,A46
(R59)	$CF_2ClBr + O(^1D) \rightarrow BrO + \text{products}$	A19,A30
(R60)	$CF_3Br + O(^1D) \rightarrow BrO + \text{products}$	A19,A32
(R61)	$OH + O(^3P) \rightarrow H + O_2$	B1
(R62)	$HO_2 + O(^3P) \rightarrow OH + O_2$	B2
(R63)	$H_2O_2 + O(^3P) \rightarrow HO_2 + OH$	B3
(R64)	$H + O_3 \rightarrow OH + O_2$	B4
(R65)	$H + HO_2 \rightarrow OH + OH$	B5
(R66)	$H + HO_2 \rightarrow H_2 + O_2$	B5
(R67)	$H + HO_2 \rightarrow H_2O + O(^3P)$	B5
(R68)	$OH + O_3 \rightarrow HO_2 + O_2$	B6
(R69)	$OH + H_2 \rightarrow H_2O + H$	B7
(R70)	$OH + OH \rightarrow H_2O + O(^3P)$	B9
(R71)	$OH + HO_2 \rightarrow H_2O + O_2$	B10
(R72)	$OH + H_2O_2 \rightarrow H_2O + HO_2$	B11
(R73)	$HO_2 + O_3 \rightarrow OH + O_2 + O_2$	B12
(R74)	$HO_2 + HO_2 \rightarrow H_2O_2 + O_2$	B13
(R75)	$NO_2 + O(^3P) \rightarrow NO + O_2$	C1
(R76)	$NO_3 + O(^3P) \rightarrow NO_2 + O_2$	C2
(R77)	$NO_2 + H \rightarrow OH + NO$	C6
(R78)	$NO_3 + OH \rightarrow NO_2 + HO_2$	C7
(R79)	$HNO_3 + OH \rightarrow NO_3 + H_2O$	C9
(R80)	$HNO_4 + OH \rightarrow NO_2 + H_2O + O_2$	C10
(R81)	$NO + HO_2 \rightarrow OH + NO_2$	C12
(R82)	$NO_3 + HO_2 \rightarrow HNO_3 + O_2$	C14
(R83)	$HO_2 + NO_3 \rightarrow OH + NO_2 + O_2$	C14
(R84)	$NO + O_3 \rightarrow NO_2 + O_2$	C20
(R85)	$NO + NO_3 \rightarrow NO_2 + NO_2$	C21
(R86)	$NO_2 + O_3 \rightarrow NO_3 + O_2$	C22
(R87)	$NO_2 + NO_3 \rightarrow NO + NO_2 + O_2$	C23
(R88)	$NO_3 + NO_3 \rightarrow NO_2 + NO_2 + O_2$	C24
(R89)	$CH_2O + O(^3P) \rightarrow OH + HCO$	D4
(R90)	$CH_4 + OH \rightarrow H_2O + CH_3$	D11
(R91)	$CH_2O + OH \rightarrow H_2O + HCO$	D14
(R92)	$CH_3OOH + OH \rightarrow CH_3O_2 + H_2O$	D16
(R93)	$CH_3OOH + OH \rightarrow CH_2O + OH + H_2O$	D16
(R94)	$CH_3O_2 + HO_2 \rightarrow CH_3OOH + O_2$	D35
(R95)	$HCO + O_2 \rightarrow CO + HO_2$	D44
(R96)	$CH_3O + O_2 \rightarrow CH_2O + HO_2$	D46
(R97)	$CH_3O_2 + NO \rightarrow CH_3O + NO_2$	D51
(R98)	$ClO + O(^3P) \rightarrow Cl + O_2$	F1
(R99)	$OCIO + O(^3P) \rightarrow ClO + O_2$	F2
(R100)	$HCl + O(^3P) \rightarrow OH + Cl$	F4
(R101)	$HOCl + O(^3P) \rightarrow ClO + OH$	F5
(R102)	$ClONO_2 + O(^3P) \rightarrow ClO + NO_3$	F6
(R103)	$OCIO + O_3 \rightarrow O_2 + O_2 + ClO$	F7
(R104)	$Cl_2 + OH \rightarrow HOCl + Cl$	F9
(R105)	$ClO + OH \rightarrow Cl + HO_2$	F10
(R106)	$ClO + OH \rightarrow HCl + O_2$	F10
(R107)	$OCIO + OH \rightarrow HOCl + O_2$	F11
(R108)	$HCl + OH \rightarrow Cl + H_2O$	F12

**Table 2.** (continued)

Reaction Number	Reaction <sup>a</sup>	Reaction Signature From <i>Sander et al.</i> [2006]
(R109)	$\text{HOCl} + \text{OH} \rightarrow \text{ClO} + \text{H}_2\text{O}$	F13
(R110)	$\text{ClONO}_2 + \text{OH} \rightarrow \text{HOCl} + \text{NO}_3$	F15
(R111)	$\text{CH}_3\text{Cl} + \text{OH} \rightarrow \text{H}_2\text{O}$	F16
(R112)	$\text{CHF}_2\text{Cl} + \text{OH} \rightarrow \text{H}_2\text{O}$	F22
(R113)	$\text{CH}_3\text{CCl}_3 + \text{OH} \rightarrow \text{H}_2\text{O}$	F26
(R114)	$\text{Cl} + \text{HO}_2 \rightarrow \text{HCl} + \text{O}_2$	F45
(R115)	$\text{Cl} + \text{HO}_2 \rightarrow \text{OH} + \text{ClO}$	F45
(R116)	$\text{ClO} + \text{HO}_2 \rightarrow \text{HOCl} + \text{O}_2$	F46
(R117)	$\text{OCIO} + \text{NO} \rightarrow \text{ClO} + \text{NO}_2$	F48
(R118)	$\text{Cl} + \text{O}_3 \rightarrow \text{ClO} + \text{O}_2$	F52
(R119)	$\text{Cl} + \text{H}_2 \rightarrow \text{HCl} + \text{H}$	F53
(R120)	$\text{Cl} + \text{H}_2\text{O}_2 \rightarrow \text{HCl} + \text{HO}_2$	F54
(R121)	$\text{Cl} + \text{NO}_3 \rightarrow \text{ClO} + \text{NO}_2$	F55
(R122)	$\text{Cl} + \text{CH}_4 \rightarrow \text{HCl} + \text{CH}_3$	F59
(R123)	$\text{CH}_2\text{O} + \text{Cl} \rightarrow \text{HCl} + \text{HCO}$	F61
(R124)	$\text{Cl} + \text{OCIO} \rightarrow \text{ClO} + \text{ClO}$	F79
(R125)	$\text{Cl} + \text{Cl}_2\text{O}_2 \rightarrow \text{Cl} + \text{Cl}_2 + \text{O}_2$	F82
(R126)	$\text{Cl} + \text{HOCl} \rightarrow \text{Cl}_2 + \text{OH}$	F83
(R127)	$\text{Cl} + \text{HOCl} \rightarrow \text{HCl} + \text{ClO}$	F83
(R128)	$\text{Cl} + \text{ClONO}_2 \rightarrow \text{Cl}_2 + \text{NO}_3$	F85
(R129)	$\text{ClO} + \text{NO} \rightarrow \text{Cl} + \text{NO}_2$	F111
(R130)	$\text{ClO} + \text{NO}_3 \rightarrow \text{Cl} + \text{NO}_2 + \text{O}_2$	F112
(R131)	$\text{CH}_3\text{O}_2 + \text{ClO} \rightarrow \text{CH}_3\text{O} + \text{Cl} + \text{O}_2$	F117
(R132)	$\text{ClO} + \text{ClO} \rightarrow \text{Cl} + \text{OCIO}$	F118
(R133)	$\text{ClO} + \text{ClO} \rightarrow \text{Cl}_2 + \text{O}_2$	F118
(R134)	$\text{ClO} + \text{ClO} \rightarrow \text{Cl} + \text{Cl} + \text{O}_2$	F118
(R135)	$\text{BrO} + \text{O}(^3\text{P}) \rightarrow \text{Br} + \text{O}_2$	G1
(R136)	$\text{HBr} + \text{O}(^3\text{P}) \rightarrow \text{Br} + \text{OH}$	G2
(R137)	$\text{HOBr} + \text{O}(^3\text{P}) \rightarrow \text{OH} + \text{BrO}$	G3
(R138)	$\text{CH}_3\text{Br} + \text{OH} \rightarrow \text{H}_2\text{O}$	G3
(R139)	$\text{BrO} + \text{OH} \rightarrow \text{Br} + \text{HO}_2$	G6
(R140)	$\text{HBr} + \text{OH} \rightarrow \text{Br} + \text{H}_2\text{O}$	G7
(R141)	$\text{Br} + \text{HO}_2 \rightarrow \text{HBr} + \text{O}_2$	G24
(R142)	$\text{BrO} + \text{HO}_2 \rightarrow \text{HOBr} + \text{O}_2$	G25
(R143)	$\text{Br} + \text{O}_3 \rightarrow \text{BrO} + \text{O}_2$	G31
(R144)	$\text{CH}_2\text{O} + \text{Br} \rightarrow \text{HBr} + \text{HCO}$	G34
(R145)	$\text{Br} + \text{OCIO} \rightarrow \text{BrO} + \text{ClO}$	G35
(R146)	$\text{BrO} + \text{NO} \rightarrow \text{Br} + \text{NO}_2$	G39
(R147)	$\text{BrO} + \text{ClO} \rightarrow \text{Br} + \text{OCIO}$	G41
(R148)	$\text{BrO} + \text{ClO} \rightarrow \text{Br} + \text{Cl} + \text{O}_2$	G41
(R149)	$\text{BrO} + \text{ClO} \rightarrow \text{BrCl} + \text{O}_2$	G41
(R150)	$\text{BrO} + \text{BrO} \rightarrow \text{Br} + \text{Br} + \text{O}_2$	G42
(R151)	$\text{O}(^3\text{P}) + \text{O}_2 \rightarrow \text{O}_3$	A1
(R152)	$\text{N}_2 + \text{O}(^1\text{D}) \rightarrow \text{N}_2\text{O}$	A2
(R153)	$\text{H} + \text{O}_2 \rightarrow \text{HO}_2$	B1
(R154)	$\text{OH} + \text{OH} \rightarrow \text{H}_2\text{O}_2$	B2
(R155)	$\text{NO} + \text{O}(^3\text{P}) \rightarrow \text{NO}_2$	C1
(R156)	$\text{NO}_2 + \text{O}(^3\text{P}) \rightarrow \text{NO}_3$	C2
(R157)	$\text{NO}_2 + \text{OH} \rightarrow \text{HNO}_3$	C4
(R158)	$\text{NO}_2 + \text{HO}_2 \rightarrow \text{HNO}_4$	C5
(R159)	$\text{NO}_2 + \text{NO}_3 \rightarrow \text{N}_2\text{O}_5$	C6
(R160)	$\text{N}_2\text{O}_5 \rightarrow \text{NO}_3 + \text{NO}_2$	Table 3-5
(R161)	$\text{HNO}_4 \rightarrow \text{HO}_2 + \text{NO}_2$	Table 3-2
(R162)	$\text{CO} + \text{OH} \rightarrow \text{H} + \text{products}$	D1
(R163)	$\text{CH}_3 + \text{O}_2 \rightarrow \text{CH}_3\text{O}_2$	D2
(R164)	$\text{ClO} + \text{NO}_2 \rightarrow \text{ClONO}_2$	F8
(R165)	$\text{ClO} + \text{ClO} \rightarrow \text{Cl}_2\text{O}_2$	F10
(R166)	$\text{Cl}_2\text{O}_2 \rightarrow \text{ClO} + \text{ClO}$	Table 3-14
(R167)	$\text{BrO} + \text{NO}_2 \rightarrow \text{BrONO}_2$	G2

<sup>a</sup>The term “products” represents constituents that are not considered in the reaction scheme.

number of iterations range between 12 and 16, with the former being sufficient for case studies spanning a sequence of consecutive days, where benefits of good background fields from earlier assimilation can be drawn.

[39] As another practical advantage of the icosahedral grid structure, it should be mentioned that the treatment of the

adjoint of the semi-Lagrangian transport is particularly simple, since there are no pole singularities. Problems reported by *Tanguay and Polavarapu* [1999] at the poles and actions for their reduction are inapplicable.

[40] During the course of adjoint model integration, the gradient due to observations  $\mathbf{y}_{n+1}$ , which are available within



**Table 3.** Heterogeneous Reactions Included in the SACADA Reaction Scheme<sup>a</sup>

Reaction Number	Reaction	Uptake Coefficient		
		Liquid/STS	NAT	Ice
(R168)	$\text{BrONO}_2 + \text{H}_2\text{O}(\text{c}) \rightarrow \text{HOBr} + \text{HNO}_3$	$f(t, p_{\text{H}_2\text{O}})^{\text{b}}$		0.26
(R169)	$\text{N}_2\text{O}_5 + \text{H}_2\text{O}(\text{c}) \rightarrow \text{HNO}_3 + \text{HNO}_3$	$f(t, p_{\text{H}_2\text{O}})^{\text{b}}$	0.0004	0.02
(R170)	$\text{ClONO}_2 + \text{H}_2\text{O}(\text{c}) \rightarrow \text{HNO}_3 + \text{HOCl}$	$f(t, p_{\text{H}_2\text{O}}, p_{\text{HCl}})^{\text{c}}$	0.004	0.3
(R171)	$\text{ClONO}_2 + \text{HCl}(\text{c}) \rightarrow \text{Cl}_2 + \text{HNO}_3$	$f(t, p_{\text{H}_2\text{O}}, p_{\text{HCl}})^{\text{c}}$	0.2	0.3
(R172)	$\text{HOCl} + \text{HCl}(\text{c}) \rightarrow \text{Cl}_2 + \text{H}_2\text{O}$	$f(t, p_{\text{H}_2\text{O}}, p_{\text{HCl}})^{\text{c}}$	0.1	0.2
(R173)	$\text{N}_2\text{O}_5 + \text{HCl}(\text{c}) \rightarrow \text{HNO}_3 + \text{products}$		0.003	0.03
(R174)	$\text{HOBr} + \text{HCl}(\text{c}) \rightarrow \text{BrCl} + \text{H}_2\text{O}$	0.01		0.3
(R175)	$\text{ClONO}_2 + \text{HBr}(\text{c}) \rightarrow \text{BrCl} + \text{HNO}_3$		0.3	0.3
(R176)	$\text{HOCl} + \text{HBr}(\text{c}) \rightarrow \text{BrCl} + \text{H}_2\text{O}$			0.05
(R177)	$\text{BrONO}_2 + \text{HCl}(\text{c}) \rightarrow \text{BrCl} + \text{HNO}_3$	0.3		0.3

<sup>a</sup>The notation “(c)” indicates a species in the condensed (liquid or solid) phase. The term “products” represents constituents that are not considered in the reaction scheme.

<sup>b</sup>As recommended by Sander *et al.* [2006].

<sup>c</sup>From Shi *et al.* [2001], as recommended by Sander *et al.* [2006].

the time interval  $[t_n, t_{n+2}]$ , is added to the adjoint variable  $\mathbf{c}^*$  ( $t_{n+1}$ ), which then is propagated backward in time by means of the adjoint model,

$$\mathbf{c}^*(t_{n-1}) = [\mathbf{M}^{h*} \circ \mathbf{M}^{v*} \circ \mathbf{M}^{e*}](t_n) \cdot (\mathbf{c}^*(t_{n+1}) + \mathbf{H}^T \mathbf{R}^{-1} [\mathbf{y}_{n+1} - \mathbf{H}\mathbf{c}(t_{n+1})]). \quad (10)$$

Finally, at  $n - 1 = 0$ , we get the sought after gradient  $\nabla_{\mathbf{c}_0} J^0 = \mathbf{c}^*(t_0)$ .

[41] As a simplification, the adjoint formulation does not include the dependence of reaction rates on concentration levels. This concerns the rates of photolysis reactions, which depend on the overhead ozone column, as well as heterogeneous reaction rates, which are calculated from particle size and composition. Extensive tests proved that the use of this slightly inexact adjoint does not hamper the minimization procedure.

[42] For the minimization of the cost function, the quasi-Newton limited memory L-BFGS (Limited-memory Broyden Fletcher Goldfarb Shanno) algorithm, which was devised by Nocedal [1980] and Liu and Nocedal [1989], was used. The entire model setup is parallelized. On a PC cluster with 16 processors (AMD Opteron with Infiniband connectivity) roughly representing the state of general purpose technology at the time of writing, the 4D-var assimilation of observational data collected during 24 h can be accomplished in less than 4 h, demonstrating the required computational efficiency of the setup.

## 4.2. Formulation of the Cost Function

[43] The (de)correlation length or radius of influence  $L$  is the control parameter for the spatial information spread induced by the BECM. Extending the correlation lengths, the condition number of  $\mathbf{B}$  increases rapidly [see, e.g., Elbern and Schmidt, 1999], and preconditioning of the minimization problem is mandatory. With the aid of the diffusion approach [Weaver and Courtier, 2001], the BECM is easily used for preconditioning with the transformation to variable  $\mathbf{v}_0$ ,

$$\mathbf{v}_0 := \mathbf{B}^{-1/2} \delta \mathbf{x}_0, \quad (11)$$

where  $\delta \mathbf{x}_0 := \mathbf{x}_0 - \mathbf{x}^b$ , and  $\mathbf{B} = \mathbf{B}^{1/2} \mathbf{B}^{T/2}$ . Here,  $\mathbf{B}^{T/2}$  means the transposed square root of  $\mathbf{B}$ . The cost function remains invariant under this transformation, i.e.,  $J(\mathbf{x}_0) = J(\mathbf{v}_0)$ , while the gradient now reads

$$\nabla_{\mathbf{v}_0} (J^b + J^0) = \mathbf{v}_0 + \mathbf{B}^{T/2} \nabla_{\mathbf{x}_0} J^0. \quad (12)$$

After the application of the minimization routine, the new output  $\mathbf{v}_0$  is saved for the next iteration step, while the initial values read  $\mathbf{x}_0 = \mathbf{B}^{1/2} \mathbf{v}_0 + \mathbf{x}^b$ . Selecting  $\mathbf{x}_0 = \mathbf{x}^b$  as a first guess atmospheric state, results in  $\mathbf{v}_0 = 0$  for the first iteration. Consequently, the transformation (11), and hence  $\mathbf{B}^{-1/2}$ , is never explicitly needed.

## 4.3. Correlation Modeling Using a Diffusion Approach

[44] In practice, BECM modeling by the diffusion paradigm [Weaver and Courtier, 2001] satisfies three fundamental needs: (1) it substitutes an intractably large matrix by an operator of linear complexity only, while ensuring positive definiteness and symmetry, (2) it allows for flexible design of anisotropic and inhomogeneous correlation lengths, and (3) it provides a square root decomposition for preconditioning. Given the prominence of this technique in this study, the method is outlined below.

[45] Weaver and Courtier [2001] make use of the fundamental solution  $\psi(z, t)$  at location  $z$  and for time  $t$  of the diffusion equation, for clarity reading in one-dimensional form,

$$\frac{\partial \psi}{\partial t} - \kappa \frac{\partial^2 \psi}{\partial z^2} = 0, \quad (13)$$

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

$$\psi(z, t) = \frac{1}{\sqrt{4\pi\kappa t}} \int \exp\left(-\frac{(z-z')^2}{4\kappa t}\right) \psi(z', 0) dz'. \quad (14)$$

After normalization a Gaussian function defines a valid correlation function with

$$2\kappa t =: L^2 \quad (15)$$

acting as the square of a correlation length scale.

[46] The same considerations apply for the solution of the two-dimensional diffusion equation on the sphere. *Weaver and Courtier* [2001] further demonstrated that a quasi-Gaussian covariance operator can be approximated by a diffusion operator  $\mathbf{L}$ , which solves (13) numerically. The BECM can be factorized by the standard deviation matrix  $\Sigma$  and the correlation matrix  $\mathbf{C}$ , reading

$$\mathbf{B} = \Sigma \mathbf{C} \Sigma. \quad (16)$$

#### 4.3.1. Discretized Formulation

[47] In a formalized notation, the solution  $\psi$  of the discretized diffusion equation after  $M$  time steps can be written as

$$\psi(t_M) = (\mathbf{I} + \kappa \Delta t \mathbf{D})^M \psi(t_0) =: \mathbf{L} \psi(t_0), \quad (17)$$

with  $\mathbf{D}$  denoting a discrete representation of the Laplace operator.

[48] In discretized form the Laplace operator is self-adjoint with respect to a  $\mathbf{W}$  metric and therefore satisfies  $\langle \mathbf{W} \mathbf{D} \psi_1, \psi_2 \rangle = \langle \psi_1, \mathbf{W} \mathbf{D} \psi_2 \rangle = \langle \mathbf{D}^T \mathbf{W} \psi_1, \psi_2 \rangle$ , where  $\mathbf{W} := \text{diag}(\Delta s_1, \dots, \Delta s_n)$  is a diagonal matrix containing the volume of the corresponding grid cells, and  $\langle \cdot, \cdot \rangle$  denotes the canonical scalar product. The square-root formulation of the discrete diffusion operator  $\mathbf{L}$  needed for preconditioning is obtained from the fact that

$$\mathbf{L} \mathbf{W}^{-1} = \mathbf{L}^{1/2} \mathbf{W}^{-1} \mathbf{L}^{T/2} = \mathbf{W}^{-1} \mathbf{L}^T. \quad (18)$$

However, diffusion integration by  $\mathbf{L}$  does not satisfy the autocorrelation condition  $\rho_{i,i} = 1$ . To ensure this, a diagonal normalization matrix  $\Lambda$  is introduced, defining

$$\begin{aligned} \mathbf{C} &:= \Lambda \mathbf{L}^{1/2} \mathbf{W}^{-1} \mathbf{L}^{T/2} \Lambda \\ &= (\Lambda \mathbf{L}^{1/2} \mathbf{W}^{-1/2}) (\mathbf{W}^{-1/2} \mathbf{L}^{T/2} \Lambda) = \mathbf{C}^{1/2} \mathbf{C}^{T/2}. \end{aligned} \quad (19)$$

In the idealized case a homogeneous and isotropic radius of influence,  $\text{diag} \Lambda = \sqrt{2\pi} L^2$  preserves the normalization of a Gaussian distribution. Relaxing homogeneity and isotropy simplifications, as is of interest here, other techniques must be invoked (see section 4.3.2).

[49] To apply the diffusion equations as correlation models a three-dimensional diffusion operator  $\mathbf{L}$  and its square root decomposition have to be constructed, by applying the vertical and horizontal diffusion operator  $\mathbf{L}_h$  and  $\mathbf{L}_v$  alternately.

#### 4.3.2. Flow-Dependent Inhomogeneous and Anisotropic Correlations

[50] Radii of influence are introduced to background error covariance modeling not only to achieve smooth fields, but also to spread information of observations within a reasonable and statistically sustained distance. This is based on the assumption that fairly homogeneous air masses with similar concentration levels of chemical compounds show high correlation values. As air masses are subject to atmospheric dynamics, for better exploitation of measurements it is desirable to devise a flow-dependent background error covariance parameterization. Due to the linkage between model dynamics and background 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

$$P = (\nabla \times \mathbf{v} + 2\boldsymbol{\Omega}) \cdot \nabla \theta = \zeta \cdot \nabla \theta \approx \mathbf{k} \cdot \zeta \frac{\partial \theta}{\partial p}, \quad (20)$$

where  $\theta$  is the potential temperature,  $\mathbf{k}$  the vertical unit vector, and  $\zeta$  the vector of absolute vorticity. 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, the approximation in (20) to be accurate within 15%. As the potential vorticity is conserved for adiabatic and frictionless motion, it can be concluded that the chemical composition of individual air parcels is correlated along lines of constant potential vorticity, as long as diabatic processes are weak enough. *Sankey and Shepherd* [2003] investigated the correlation between different long-lived species, and found it dependent on height and latitude. Since in our context the correlation between PV and species is only invoked locally and the assimilation interval is not extended beyond 24 h, our assumptions appear to be justifiable. It is clearly desirable to attribute each species individual correlation lengths. However, due to lack of related knowledge, our present approach rests on the assumption that the chemical scenario and conditions are sufficiently uniform in a contiguous air mass of equal potential vorticity level.

[51] A disadvantage concerning the use of PV for the purpose of discriminating different air masses is the fact, that the term  $\partial \theta / \partial p$  varies exponentially with height. Therefore it was decided to use a modified potential vorticity, which is defined according to *Lait* [1994] as

$$\Pi = P \left( \frac{\theta}{\theta_0} \right)^{-\frac{1+\zeta}{\zeta}}, \quad (21)$$

with  $\zeta = R/c_p$ , where  $R$  is the gas constant for air and  $c_p$  is the specific heat at constant pressure. The dimensionless scaling factor  $(\theta/\theta_0)^{\frac{1+\zeta}{\zeta}}$  (with  $\theta_0$  arbitrarily chosen as 420 K) removes most of the altitude dependence of  $P$  while preserving its conservation properties (see *Lait* [1994] for details). An example for zonal mean PV modified following (21) for 18 November, 2002, is presented in Figure 3, where vertical dependencies are considerably attenuated.

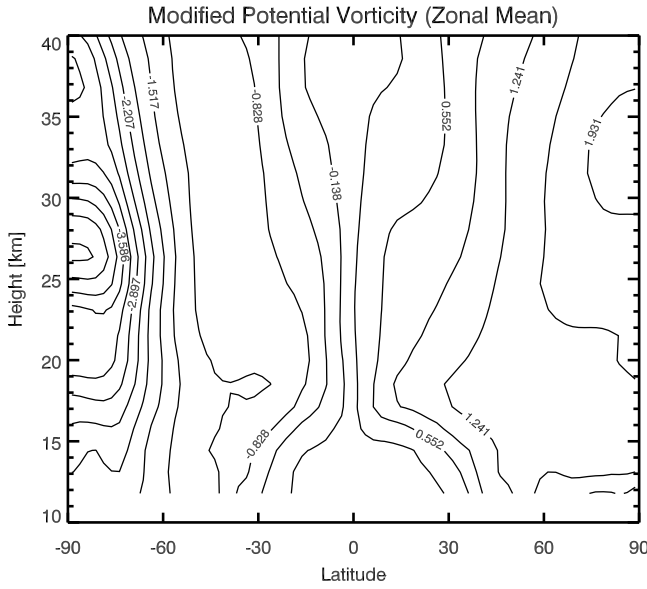
[52] To account for anisotropic and inhomogeneous background error correlations a symmetric coordinate stretching tensor  $\mathbf{S}$  is introduced in the horizontal two-dimensional diffusion equation.

$$\frac{\partial \psi}{\partial t} - \kappa_h \text{div}(\mathbf{S} \text{grad}_h \psi) = 0. \quad (22)$$

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:

$$\begin{aligned} \mathbf{S} &= \mathbf{T}(\alpha) \tilde{\mathbf{S}} \mathbf{T}^T(\alpha) \\ &= \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \tilde{s}_1 & 0 \\ 0 & \tilde{s}_2 \end{pmatrix} \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}, \end{aligned} \quad (23)$$

where  $\alpha$  is the rotation angle.



**Figure 3.** Zonal mean of modified potential vorticity field ( $10^{-6} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$ ) according to equation (21) for 18 September 2002.

[53] The discrete three-dimensional diffusion operator takes the same form as the isotropic scheme, with the exception that the horizontal Laplacian operator  $\mathbf{D}_h$  is replaced by a discrete representation of  $\text{div}(\mathbf{S} \text{ grad}_h)$ .

#### 4.3.3. Normalization Matrix

[54] An algorithm, which calculates an approximation of the normalization factors, has been proposed by *Weaver and Courtier* [2001]: An ensemble of  $Q$  random vectors  $\mathbf{v}_q$  having zero mean and unit variance is generated, and the transformed ensemble  $\tilde{\mathbf{v}}_q := \mathbf{L}^{1/2} \mathbf{W}^{-1/2} \mathbf{v}_q$  is calculated. Since  $E[\mathbf{v}] = 0$  and  $E[\mathbf{v}\mathbf{v}^T] = \mathbf{I}$ , where  $E[\cdot]$  stands for expectation value, the diagonal elements  $\tilde{l}_i$  may be estimated from

$$\begin{aligned} \mathbf{L}\mathbf{W}^{-1} &\approx \frac{1}{Q} \sum_{q=1}^Q \tilde{\mathbf{v}}_q \tilde{\mathbf{v}}_q^T \\ &= \frac{1}{Q} \sum_{q=1}^Q \left( \mathbf{L}^{1/2} \mathbf{W}^{1/2} \mathbf{v}_q \right) \left( \mathbf{L}^{1/2} \mathbf{W}^{1/2} \mathbf{v}_q \right)^T. \end{aligned} \quad (24)$$

The standard deviation of the  $\tilde{l}_i$  derived on the basis of (23) can be shown to be  $1/\sqrt{2Q}$  if  $\mathbf{v}$  is a Gaussian random vector [see *Weaver and Courtier*, 2001, and references therein]. Hence, an efficient way to approximate the normalization matrix is given by setting  $\mathbf{\Lambda} = \text{diag}(1/\sqrt{\tilde{l}_1}, \dots, 1/\sqrt{\tilde{l}_n})$ .

#### 4.3.4. BECM Implementation

[55] Two background error covariance parameterizations, an isotropic and a generalized scheme based on (22) have been developed for the SACADA system. Normalization factors for the anisotropic scheme are computed according to the random method (23). The distribution of relative differences between the random normalization factors and the exact solution is shown in Figure 4 for different numbers of ensemble members  $Q$ . Standard deviations of the three tested realizations of random normalization factors are in good accordance with the theoretical value  $1/\sqrt{2Q}$ . For  $Q = 5000$ , which was the ensemble size selected for subsequent tests,

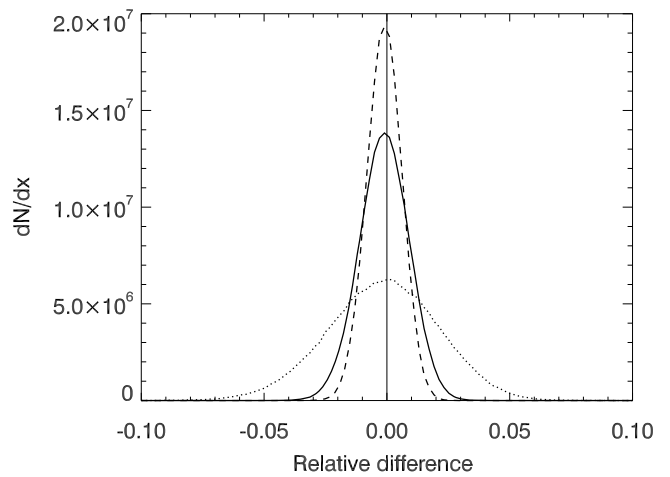
the resulting standard deviation is 1% and hence, the probability is less than 1% that a single normalization factor differs more than 3% from the exact value. The computational effort that has to be spent on the generation of an ensemble of this size depends on the diffusion length scale  $L_h$  and the stretching factors  $\tilde{s}_1, \tilde{s}_2$ , but is generally comparable to the time consumed by one iteration of the assimilation algorithm.

[56] The rotation angle  $\alpha$  is calculated at each grid point as the angle between the gradient of the potential vorticity and the direction to the north of the local coordinate system. Hence, according to (22),  $\mathbf{T}^T \nabla_h \psi$  is the gradient of an arbitrary scalar field  $\psi$  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, there is a linear decrease of  $\tilde{s}_1$  and  $\tilde{s}_2$  between  $|\nabla P|_{\text{max}}$  and  $|\nabla P|_{\text{max}}/5$ , where  $|\nabla P|_{\text{max}}$  is the absolute value of the maximum gradient of modified PV over the model domain. At locations where the PV gradient is smaller than 20% of the maximum PV value, no coordinate stretching is applied.

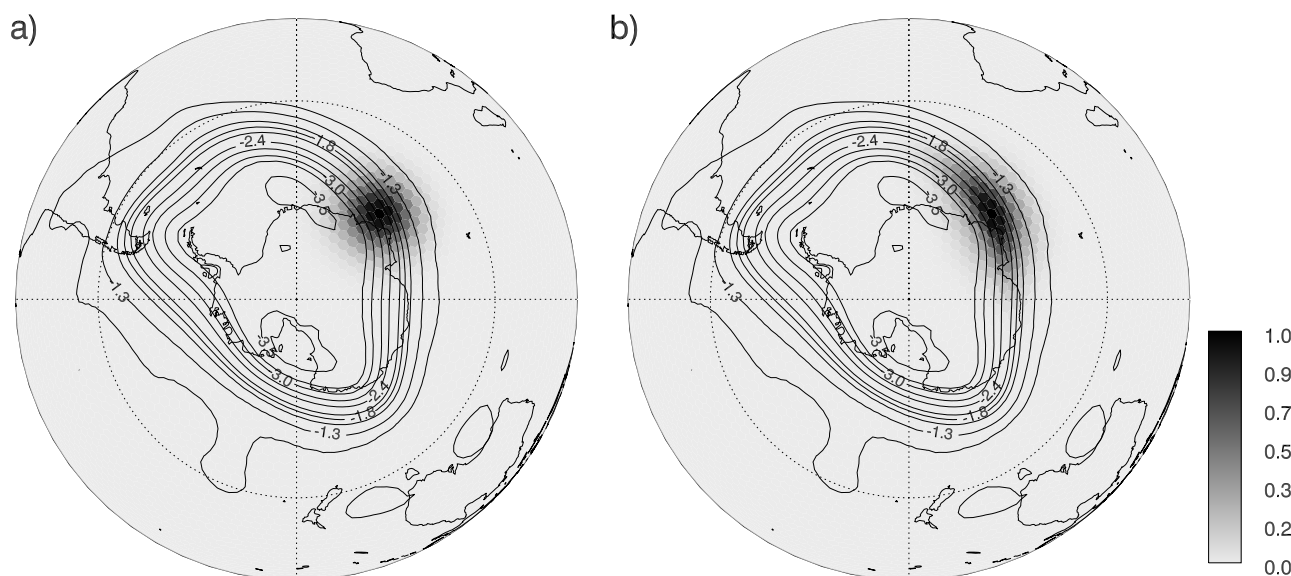
[57] As an example, the horizontal correlations generated by the two schemes are shown in Figure 5. Isotropic correlations with a horizontal length scale of 600 km are displayed at the top, and the outcome of the anisotropic scheme using the settings  $L_h = 600 \text{ km}$ ,  $\tilde{s}_1 = 4$  and  $\tilde{s}_2 = 0.25$  is shown at the bottom. The meteorological situation from 18 September 2002 was taken to calculate the anisotropic correlations; the corresponding PV field is shown by isopleths in Figure 5. The larger correlation between grid cells along the edge of the polar vortex is clearly visible.

## 5. Data Description

[58] The SACADA assimilation system has been tested and evaluated by means of two case studies. Data from three



**Figure 4.** Distribution of relative differences between random and exact normalization factors for 1000 (dotted line), 5000 (solid line) and 10,000 (dashed line) ensemble members  $Q$ ; the standard deviations of these three distributions are 2.25%, 1.0%, and 0.71%, respectively.



**Figure 5.** (a) Isotropic and (b) PV-dependent anisotropic extension of influence radii for a single observation at the southern vortex edge. Grey shading indicates correlation, and isopleths indicate modified PV units ( $10^{-6} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$ ) on the 43.7 hPa level for 18 September 2002.

different limb viewing instruments, namely MIPAS, HALOE, and SAGE II, have been used for this purpose. MIPAS has been the data source for assimilation case studies presented in this work. Data from SAGE II and HALOE have been withheld for use as independent (not assimilated) control data sets.

### 5.1. MIPAS

[59] MIPAS is a Fourier-transform spectrometer measuring high-resolution emission spectra in the midinfrared from 4.1 to  $14.7 \mu\text{m}$  wavelength in a limb viewing mode. A detailed description of the instrument design is given by *Fischer and Oelhaf* [1996]. In 2002, MIPAS was placed into a Sun-synchronous polar orbit with an inclination of  $98.55^\circ$ . 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 1000 local time. 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. The trace gas profiles which are operationally delivered by ESA as described by *Ridolfi et al.* [2000] and *Raspollini et al.* [2006] as version 4.61/4.62 are assimilated in this system validation study. Profiles of  $\text{O}_3$ ,  $\text{NO}_2$ ,  $\text{CH}_4$ ,  $\text{HNO}_3$ ,  $\text{H}_2\text{O}$  and  $\text{N}_2\text{O}$  are contained in these data products and assimilated.

### 5.2. SAGE II and HALOE

[60] The SAGE II instrument was launched in October 1984 on board the Earth Radiation Budget Satellite (ERBS) spacecraft. The instrument operated in the UV, visible and near-infrared (385–1020 nm) wavelength region employing the solar occultation technique leading to a maximum of 15 profiles per day after technical problems in July 2000. Due to the orbit geometry, tangent point locations vary slowly from

$70^\circ\text{N}$  to  $70^\circ\text{S}$  within approximately one month. Retrieved profiles of aerosol extinction,  $\text{O}_3$ ,  $\text{NO}_2$  and  $\text{H}_2\text{O}$  are provided.

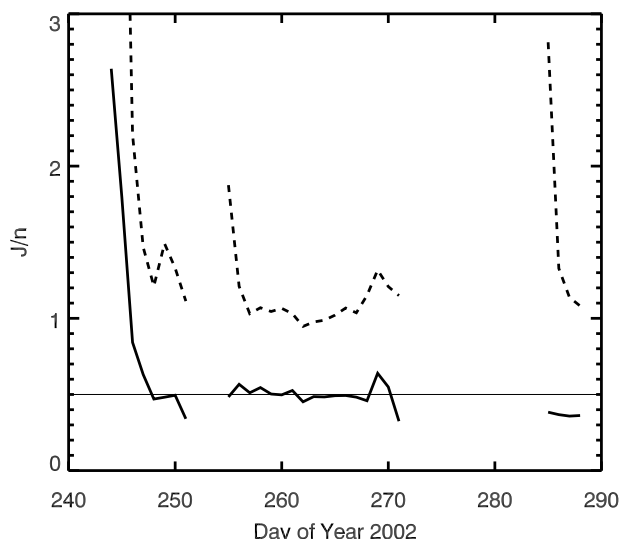
[61] HALOE was in operation from October 1991 to November 2005 on board the Upper Atmosphere Research Satellite (UARS). The sensor performed solar occultation measurements at sunrise and sunset in the infrared wavelength region and profiles of  $\text{O}_3$ ,  $\text{HCl}$ ,  $\text{CH}_4$ ,  $\text{H}_2\text{O}$ ,  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{HF}$ , 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  $80^\circ\text{N}$  to  $80^\circ\text{S}$  within approximately one month. The latest data product release (version 19) from the NASA Langley Research Center has been used for validation.

## 6. Case Studies

[62] The performance achieved with the 4D-var system both in terms of analysis skill and numerical efficiency is demonstrated by operation during two selected episodes. The two time intervals in 2002 and 2003 cover 1 September to 15 October 2002, hereafter referred to as CS1, with two data gaps (9 to 11 September and 29 September to 11 October), and 21 October to 30 November 2003, referred to as CS2. During CS1, at the end of September 2002, a major stratospheric warming occurred with the southern polar vortex, which caused the antarctic polar vortex to split into two fragments around 25 September [e.g., *Allen et al.*, 2003; *Bencherif et al.*, 2007].

### 6.1. Experimental Setup and Cost Function Reduction

[63] The 4D-Var analysis window is 24 h long, and the meteorological simulation is initialized at the start of each window with fields from European Centre for Medium Range Weather Forecasting (ECMWF) analyses. The analyzed chemical fields from the end of one assimilation window are



**Figure 6.** Evolution of the normalized cost function values (cost divided by the number of observations)  $J_p^b$  (dashed line) and  $J_p^a$  (solid line) for the time span 1 September to 15 October 2002.

used as the background for the next assimilation. At the beginning of each case study chemical initial values provided by the two-dimensional SOCRATES model [Brasseur *et al.*, 1995] are adopted. After interpolation of these data onto the icosahedral grid, a 48 h model spin-up run was performed for chemical relaxation, followed by a second spin-up phase of 6 days with data assimilation. During the spin-up phase the PSC scheme has been turned off until day five. Both episodes have been analyzed twice: In the first experiment traditional isotropic covariances with  $L_h = 600$  km radius of influence were applied, and in the second experiment PV-structure controlled anisotropic covariances were used. In the latter case local coordinate stretching factors have been set to  $\tilde{s}_1 = 4$  and  $\tilde{s}_2 = 0.25$ , corresponding to halving the correlation length in the direction aligned with the local PV gradient (300 km), and doubling along PV isopleths (1200 km). Values for background error standard deviation are taken to be  $\epsilon_b = 40\%$  and  $L_v = 1.5$  km has been chosen as the settings for the case study periods. Control model runs without data assimilation starting from the final analysis of the spin-up assimilation have been performed for both case studies.

[64] The observation error covariance matrix  $\mathbf{R}$  has been simplified to diagonal form with the variances taken from the respective data products. As observations with unrealistically small errors receive too much weight within the data assimilation procedure, it has been decided to assume some minimum relative error for the MIPAS data products. Different values for the minimum relative error margin have been tested and the results presented here, have been obtained with MIPAS retrieval error margins increased to a minimum of 10% relative error.

[65] In 4D-var, the background errors of the chemical species, that is the standard deviations  $\sqrt{\text{diag}(\mathbf{B})}$ , must be selected a priori, but are amenable to posteriori by  $\chi^2$  validation following Talagrand [1998]. The statistical evaluation is considered also in part II of this study by a novel and

more focused approach. In order to test this condition, the cost function values are normalized by the number of available observations  $p$  to render the results comparable among different days. Following Talagrand [1998], at the minimum of the cost function (1), that is for the sought after most probable chemical state or analysis  $\mathbf{x}^a$ , the value of  $J_p^a := J(\mathbf{x}^a)/p = 1/2$  should be expected, if properly defined covariance matrices  $\mathbf{R}$  and  $\mathbf{B}$  had been used.

[66] Figure 6 presents the evolution of the costs of both the background and assimilation based runs during the sequence of days of episode CS1, also displaying the effect of temporal data gaps. While the costs of the first days during the assimilation spin-up period exhibit large deviations from the target value of  $J_p^a = 1/2$ , a fast approximation can be observed toward the desired value.

[67] During the spin-up assimilation of 6 days, the normalized background costs, where the background term in the cost function is  $J^b = 0$ ,

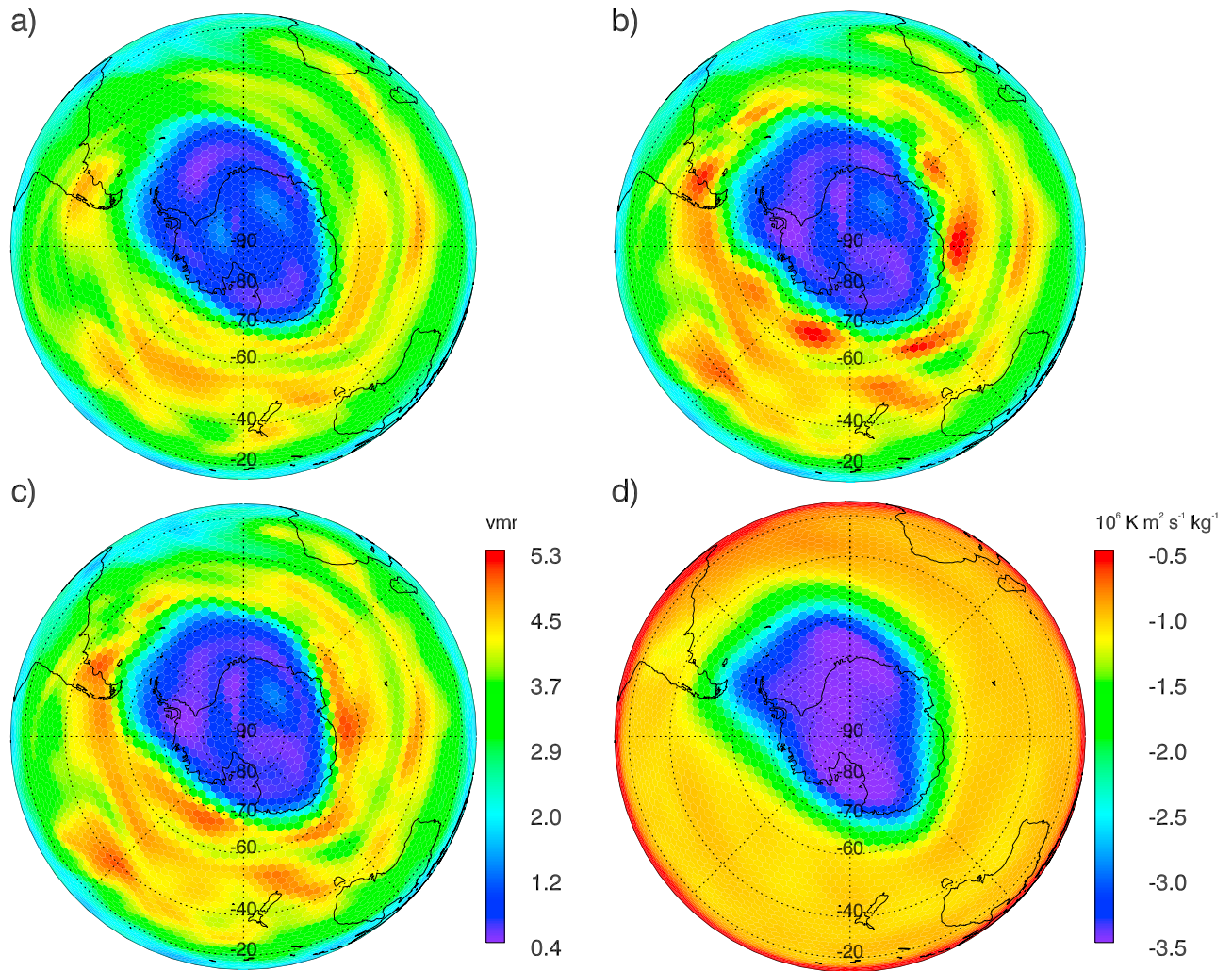
$$J_p^b := J(\mathbf{x}^b)/p \\ = \frac{1}{2p} \sum_{i=0}^N [H(M_i(\mathbf{x}^b)) - \mathbf{y}_i]^T \mathbf{R}^{-1} [H(M_i(\mathbf{x}^b)) - \mathbf{y}_i]$$

rapidly decrease within the first few days. Note that  $J_p^b$  gives the mean square difference between the background based forward run and observations, weighted by the observation error covariances. This quantity can be regarded as a quadratic measure of 24 h forecast skill, depending largely on the quality of the previous analysis. During the rest of the period the cost function is reduced by a factor of approximately 2–3, with little day-to-day variation.

[68] The impact of the data assimilation procedure compared to free control forecasts can be estimated by data gaps of 3 and 14 days. To bridge these gaps, the model has been operated in forecast mode starting from the last available analysis. Consequently,  $J_p^b$  at 12 September 2002 and 12 October 2002 reflect the forecast skill of a 4 day and a 2 week prediction, respectively. The factor by which  $J_p^b$  is larger for the 4 day forecast in comparison to the  $J_p^b$  value at other days is less than 2, and even after bridging the data gap of 14 days this factor is not larger than 3. This demonstrates the beneficial impact of data assimilation.

## 6.2. Flow-Dependent BECM Parameterization

[69] The benefits of the anisotropic background error covariance formulation is estimated by the elongated dynamical structures along polar vortex edges, where “collars” of elevated ozone mixing ratios occur. The impact of the flow-dependent, anisotropic and inhomogeneous background error covariance formulation versus the isotropic formulation can be observed from Figure 7. The ozone analyses for 18 September 2002 at 44 hPa obtained with the isotropic BECM parameterization can be compared with the analysis inferred by anisotropic formulation, together with the potential vorticity and the 24 h forecast. Prevailing zonally elongated structures are clearly exhibited by the model forecast. While these structures are emulated by the anisotropic BECM formulation, the isotropic version distorts filamentary structures at the vortex edge. Given the polar orbit of ENVISAT and many



**Figure 7.** (a) Ozone 24 h forecast for 18 September 2002 at 44 hPa, (b) ozone analysis with isotropic BECM, and (c) anisotropic BECM. (d) Corresponding PV distribution.

other typical environmental satellites, the zonal extension of influence radii has a desirable complementing effect by propagating information in the zonal direction, between the satellite tracks. A second example taken from case study CS2 at 55 hPa on 23 October 2003, is given in Figure 8.

[70] For independent validation, two SAGE II profile retrievals close to the southern polar vortex edge were available for this day, the position of which are indicated by a square and triangle symbol in Figure 8. Figure 9 exhibits comparisons between assimilation results from homogeneous and isotropic versus inhomogeneous and anisotropic correlation modeling. In both cases, it can be seen that there are substantially smaller differences to SAGE profiles with the inhomogeneous and anisotropic correlation model in comparison with the homogeneous and isotropic correlation approach.

[71] It should be noted that the parameters for the diffusion scheme are still tentatively estimated: The basic correlation length scales  $L_h$  and  $L_v$  as well as the relative background error  $\epsilon_b$  are, while easily variable, held constant in the model domain and the coordinate stretching factors  $\tilde{s}_1$  and  $\tilde{s}_2$  in (22) are calculated using a simple dependence on the gradient of potential vorticity.

[72] Nevertheless, it can be clearly observed that the elongated gradient structures of the ozone field along the edge of the polar vortex are maintained if the anisotropic BECM formulation is applied, and the collar of high ozone volume mixing ratios around the vortex does show a typical streamer structure. Hence, the benefit of the anisotropic BECM formulation is visible, at least in the region around the polar vortex.

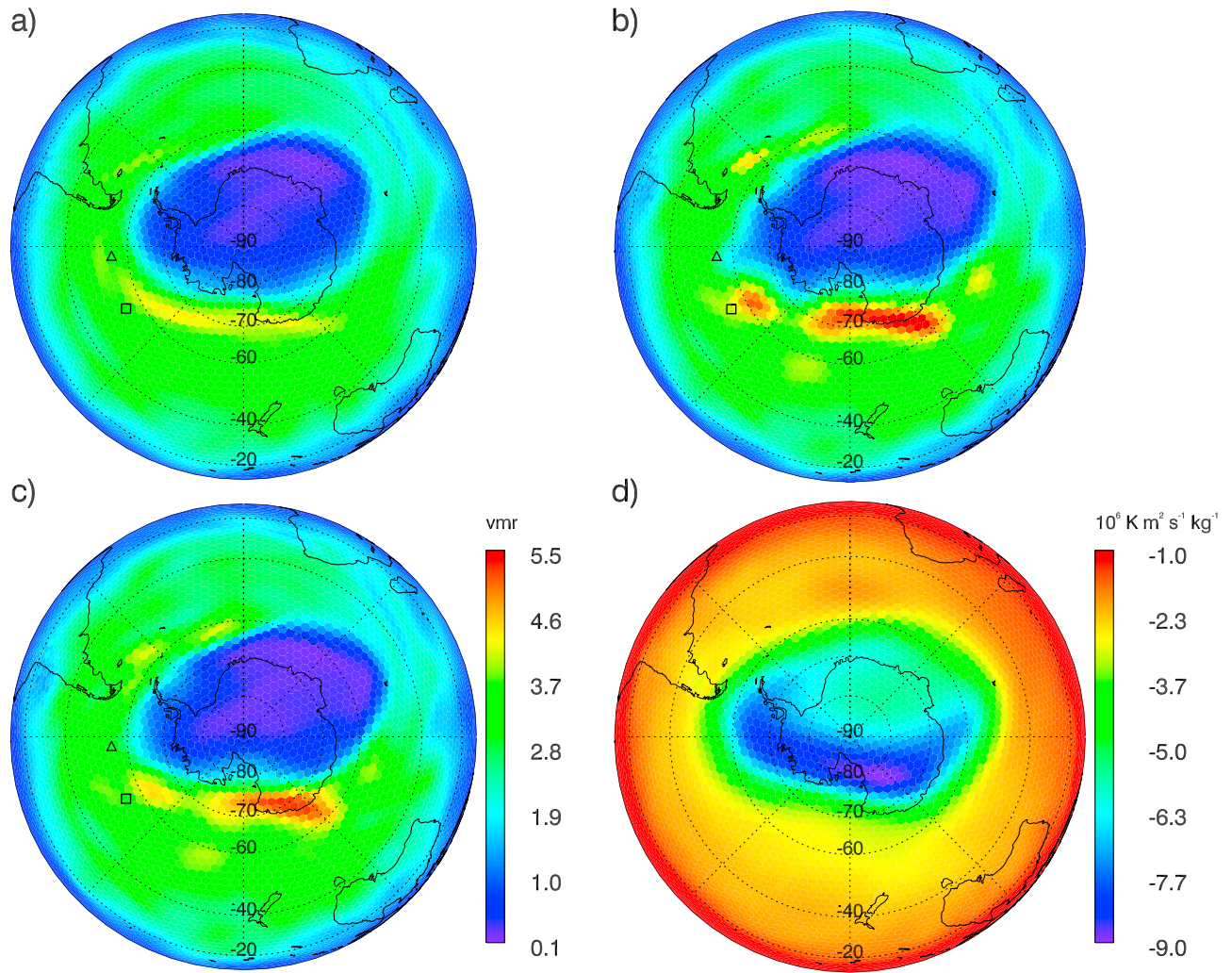
### 6.3. Statistical Evaluation

[73] The validation comprises assimilation results after model and data assimilation spin-up runs. Emphasis will be placed on ozone, complemented by  $\text{CH}_4$ ,  $\text{N}_2\text{O}$ ,  $\text{NO}_2$ ,  $\text{H}_2\text{O}$ , and  $\text{HNO}_3$ . The following validation is separated into consistency tests with respect to assimilated MIPAS retrievals and, in section 6.3.2, validation with independent data from SAGE II and HALOE.

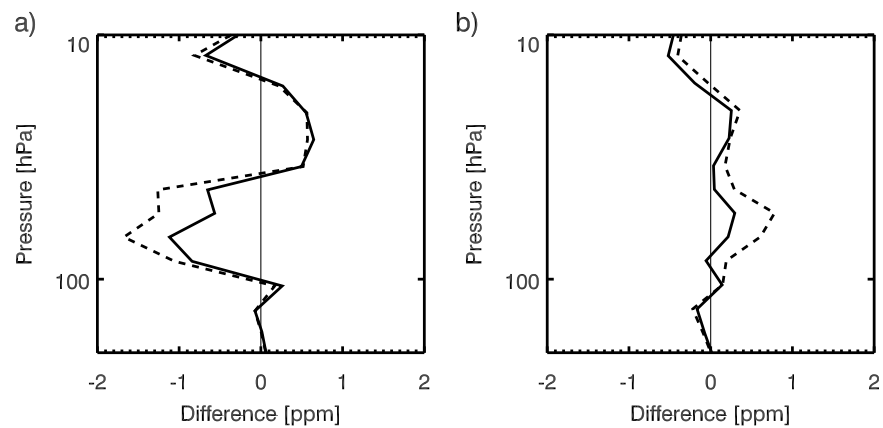
#### 6.3.1. Self-Consistency Validation of Analyses by Assimilated MIPAS Retrievals

[74] Figure 10 presents four examples of averaged volume mixing ratio profiles for CS1, including  $\text{O}_3$ ,  $\text{CH}_4$ ,  $\text{H}_2\text{O}$ , and  $\text{NO}_2$ , the latter for nighttime conditions. Averaged profiles

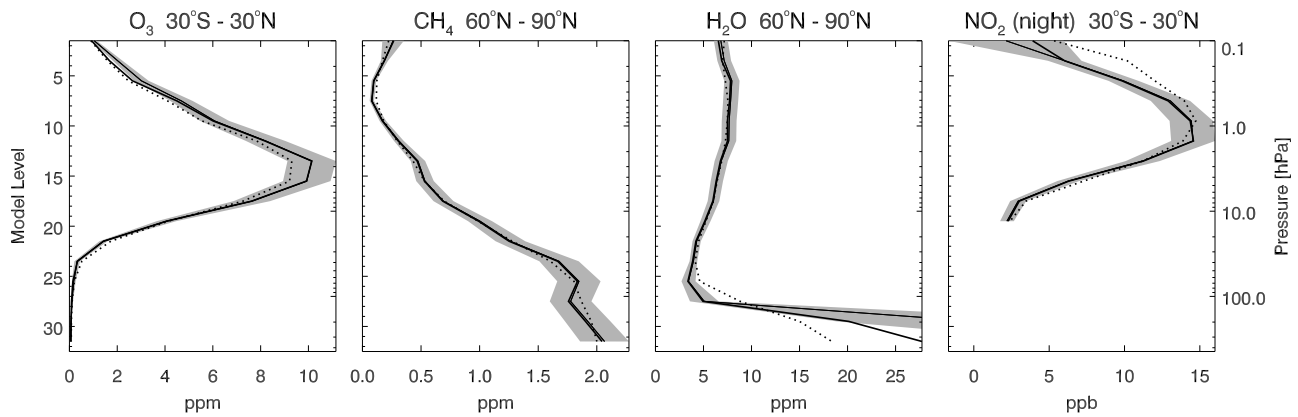




**Figure 8.** As in Figure 7 but for 23 October 2003 at 55 hPa. The square and triangle in Figures 8a, 8b, and 8c mark the location of two SAGE II profiles used for evaluation of the assimilation results as shown in Figure 9.



**Figure 9.** Analyzed ozone obtained with the flow-dependent BECM scheme minus SAGE II retrieved ozone (thick solid line) and analyzed ozone obtained with the isotropic BECM minus SAGE II (dashed line). The profiles are located at (a) 55°S 94°W (triangle in Figure 8) and at (b) 55°S 118°W (square in Figure 8).



**Figure 10.** Mean profiles of  $\text{O}_3$ ,  $\text{CH}_4$ ,  $\text{H}_2\text{O}$ , and  $\text{NO}_2$  averaged over time span 8 September to 15 October 2002 and latitude belts as indicated above each plot. Bold solid lines indicate averaged analysis results, dotted lines indicate control run average, and thin solid lines represent average retrieval profiles, embedded in shaded area indicating mean observation errors including a minimum of 10% as lower threshold as defined in **R**.

of analysis results, the control run and MIPAS retrieval results are depicted by averages of the 8 November to 15 October 2002 time span over latitude belts as indicated. Shaded error margins are obtained from the retrieval files, but are augmented to a minimum of 10% as lower threshold defined for **R**. The selected averaged altitude intervals of 3–5 km are compatible with the vertical resolution of the MIPAS instrument. Overall, the analysis results are closely aligned to the centerline of the retrievals, with some exceptions. In the case of ozone, the control run indicates a model proclivity to underestimate the retrieved ozone concentrations for middle stratospheric and lower mesospheric height levels, which is corrected by the assimilation procedure. This is also seen in BASCOE analyses [Errera *et al.*, 2008]. In the case of methane, analysis and control run differences are small, but indicating a small secondary maximum at 70 hPa. Similarly, for water vapor, significant discrepancies between control simulations and retrievals and analyses in the lower stratosphere, indicate false water vapor updrafts through the tropopause, presumably due to numerical diffusion. This effect is markedly reduced by data assimilation, albeit not eliminated. As an example of a species with a pronounced diurnal cycle,  $\text{NO}_2$  profiles are presented, exhibiting significant discrepancies between control run and retrieval in the lower mesosphere. Again, the assimilation results indicate a close agreement between retrievals and analyses.

[75] Figure 11 displays the relative difference profiles of assimilation results minus assimilated MIPAS retrievals, normalized by retrieved concentration levels:  $(A - O)/O$ . The associated standard deviation profiles are depicted by symmetric graphs as well. As additional information, differences between observations and control run as well as first guess/background run profiles are included.

[76] Ozone analyses and background values are generally in good agreement with retrievals, as can be seen in Figure 11a. Major relative deviations from retrievals can be observed at lower stratospheric and lower mesospheric height levels. Nevertheless, in view of the prevalent low concentration levels at those height levels, differences are marginal in absolute terms. However, the tendency for modeled lower

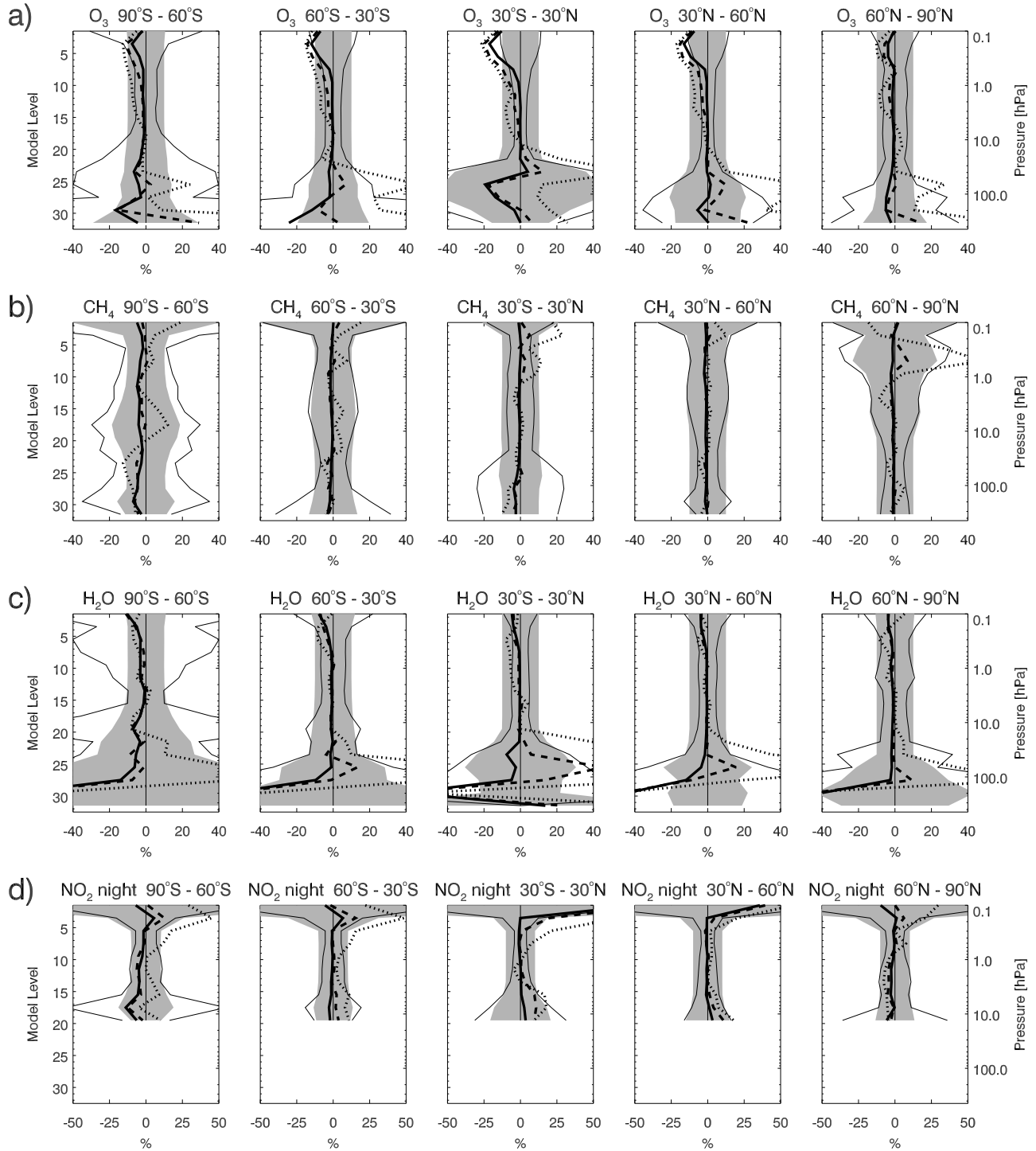
mesospheric ozone concentrations being lower than retrieved by MIPAS is visible throughout all latitude belts, especially at tropical latitudes, indicating a systematic discrepancy between MIPAS and the model. An example of the beneficial effect of data assimilation can be seen from the striking improvement of southern polar vortex dependent ozone analyses at about 50 hPa. Here, and at other latitude belts, the control run sequence exhibits a marked peak, designating a difference between the control run on the one hand, and MIPAS retrievals, analyses and short-term forecasts on the other hand.

[77] The assimilation of MIPAS methane retrievals renders substantial improvements after data assimilation in both polar regions. Figure 11b exhibits significant deviations between analyses or assimilation based forecasts and the control run throughout all height levels in the southern polar region, and in the northern polar region above the strato-pause. All other latitudinal belts indicate far better agreement, with insignificant deviations from MIPAS for analyses and background runs.

[78] Water vapor data assimilation results are given in Figure 11c. While the MIPAS retrieval assimilation displays a good to excellent agreement with both assimilation results and also the control run at most height levels, visible deviations in the lower mesosphere, and pronounced deviations in the tropopause and lowest stratosphere region occur.

[79] Mean profiles relative to MIPAS retrievals of  $\text{NO}_2$  have been separated into nighttime and daytime profiles, as  $\text{NO}_2$  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  $100^\circ$ . To exclude twilight observations, a nighttime profile was defined with zenith angles greater than  $110^\circ$ , while daytime observations were classified with zenith angles smaller than  $90^\circ$ . The assimilation of photochemically reactive constituents is generally more challenging than the assimilation of more inert gases, as the tangent-linear approximation, which is implied by the use of the adjoint model, may become poor, especially if an observation took place during daylight and the analysis at the beginning of the assimilation window is



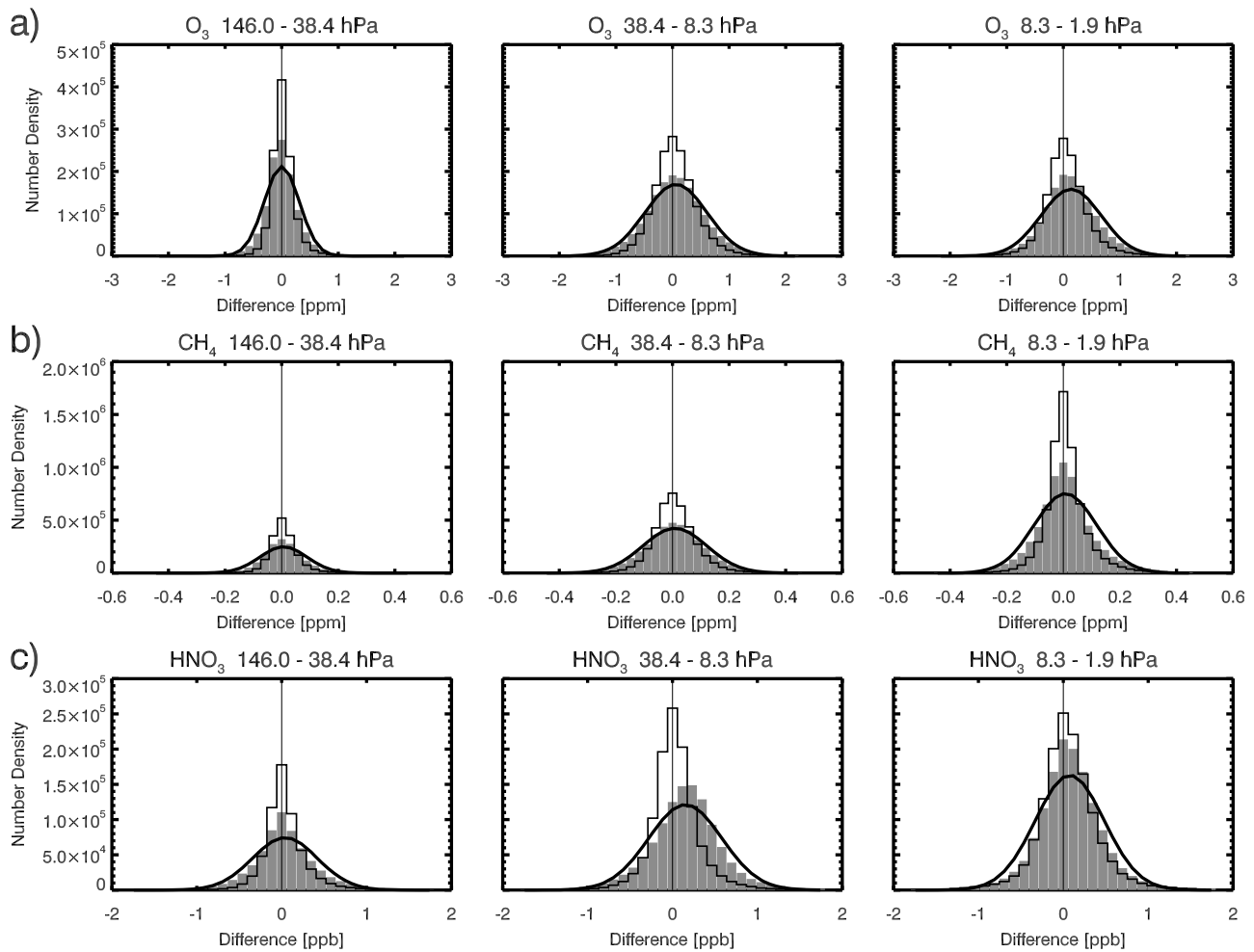


**Figure 11.** Mean normalized differences between profiles of analysis (A), control run (C), and background run (B)  $((A - O)/O$ , bold solid line;  $(C - O)/O$ , dotted line; and  $(B - O)/O$ , dashed line) minus retrievals O from MIPAS for (a)  $O_3$ , (b)  $CH_4$ , (c)  $H_2O$ , and (d)  $NO_2$  averaged over time span 8 September to 15 October and latitude belts as indicated above each plot. Symmetric thin solid lines represent  $(A - O)/O$  standard deviations, while shaded areas indicate mean retrieval errors, with error margins increased to a minimum of 10% relative error, as applied for R.

valid for local nighttime conditions, or vice versa. Nevertheless, the analyses of both, nighttime and daytime  $NO_2$  show a good agreement with observations, as is displayed for nocturnal  $NO_2$  in Figure 11. The latitude belt between  $30^\circ S$ –

$60^\circ N$  is an exception at highest model levels. Overall, these results agree with those found by Errera *et al.* [2008].

[80] Adopting the assumption of Gaussian probability density functions (PDF) of errors and the resulting quadratic



**Figure 12.** Density distributions of (O - B) differences (full grey histograms), (O - A) (outline histogram), and exact Gaussian distribution with the same mean and standard deviation as the (O - B) differences (solid line) for 28 October to 30 November 2003 for (a)  $O_3$ , (b)  $CH_4$ , and (c)  $HNO_3$ . Species and height levels are as indicated above each plot.

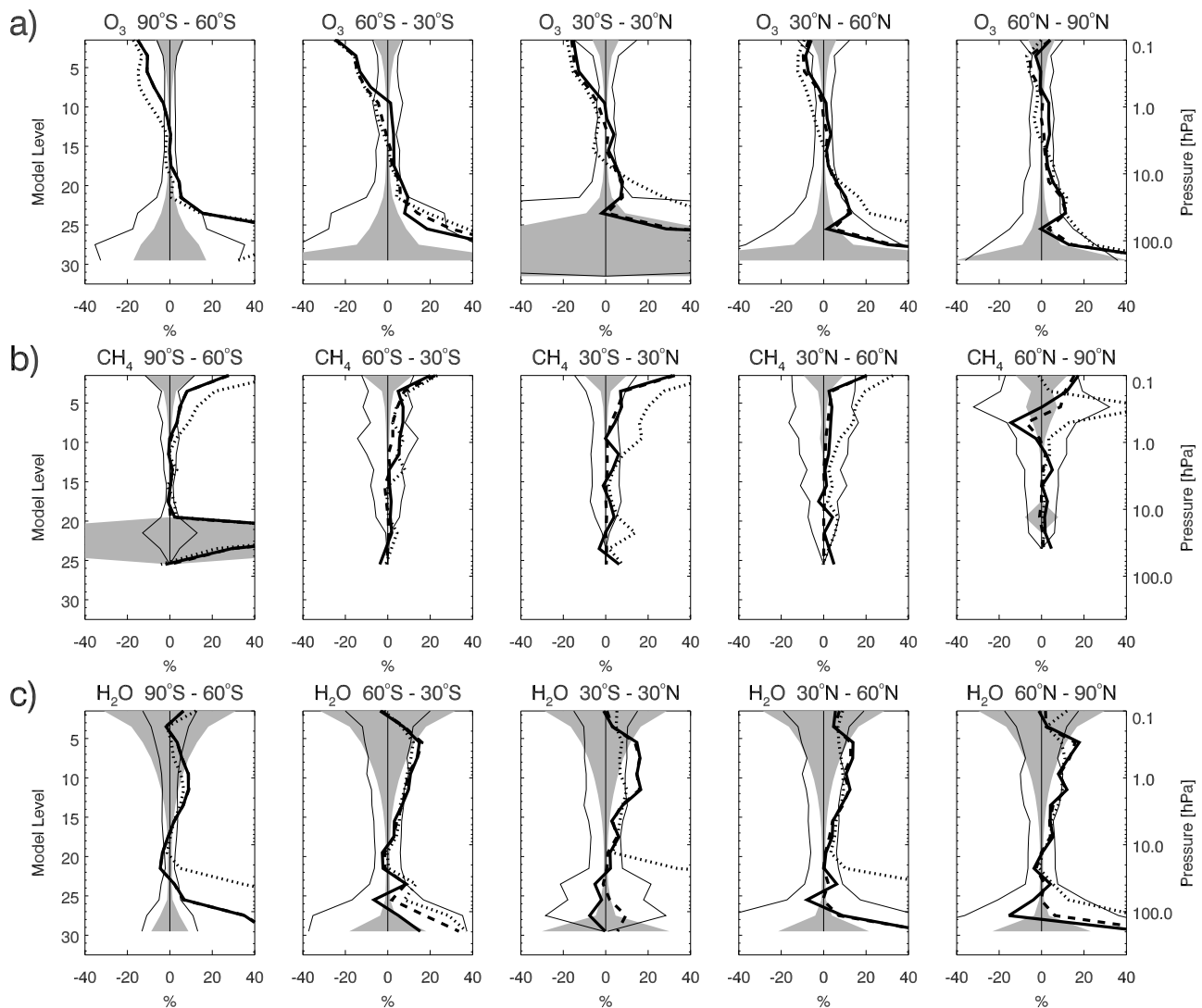
form of the cost function (1) implies that also the PDF of analysis errors is Gaussian. These underlying statistical assumptions constitute necessary conditions for the correctness of the inferred analyses. The practical justification can be considered by inspecting the differences between observations and background (O - B) as well as observations and analysis (O - A) values.

[81] Figure 12 exhibits an estimate of the PDF of (O - B) and (O - A) differences for ozone, methane, and nitric acid in CS2, by three height intervals, spanning 146–2 hPa. Upon inserting the graph of an exact Gaussian with maximum located at the mean (O - B) value and the corresponding standard deviation, a direct comparison reveals the degree of agreement between the assumption and practical results. Note that absolute biases can only be identified by independent measurements of significantly higher quality than assimilated data. Relative biases can be directly inferred by displacements of histogram peaks from the zero line. Histograms narrower than the exact Gaussian while higher at the maximum indicate a noticeable number of outliers. If there is a marked difference between the positions of the histogram

maximum and the exact Gaussian maximum, a single sided prevalence of outliers is indicated.

[82] In general we find that most distributions are peaked around zero and are approximately symmetric, whereas a notable bias occurs for  $HNO_3$  in the 38–8 hPa range. The fact that the distributions are generally more peaked than a Gaussian with the same variance, is a well known phenomenon. It is caused primarily by observational data with gross errors [see, e.g., Kalnay, 2003]. While these outliers are included for this statistical exposition here with full quadratic weight, the data assimilation result may be only marginally influenced. The reason for this is that outlier retrievals are typically attributed with much larger errors. Therefore, the impact of the outliers on the assimilation result is markedly reduced.

[83] It is concluded, that the underlying Gaussian assumptions are not violated to an extent which causes problems with the analyses, such that the application of 4D-var is not any more justified. Improvements can be expected from a more advanced identification scheme of retrieval faults by quality control.



**Figure 13.** Mean  $(A - O)/O$  differences between profiles of analysis, control run, and background run relative to retrievals from HALOE for (a)  $O_3$ , (b)  $CH_4$ , and (c)  $H_2O$  averaged over time span 8 September to 15 October and latitude belts as indicated above each plot. Plotting conventions are as for Figure 11 but without minimum threshold value of relative errors.

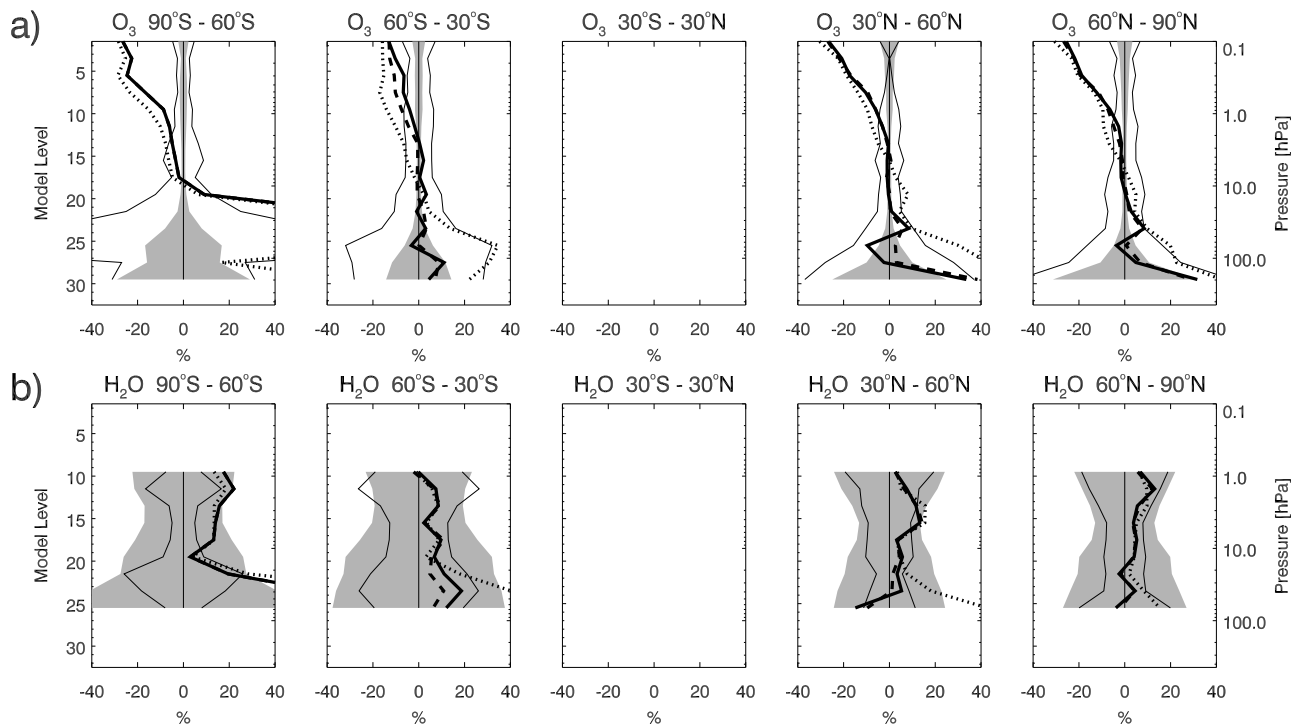
### 6.3.2. Independent Evaluation by HALOE and SAGE II

[84] In order to validate the MIPAS based assimilation results by independent data, retrievals of HALOE and SAGE II were used, but withheld from assimilation. It must be noted that both HALOE and SAGE II are occultation instruments, which provide data profiles crossing the terminator line. Therefore HALOE and SAGE II retrievals are geolocated far away from MIPAS footprints, except in polar regions.

[85] Figure 13 shows comparisons of analyses, background based runs and control runs with HALOE retrievals for  $O_3$ ,  $CH_4$ ,  $H_2O$ , and  $NO_2$ , by the same format as for Figure 11. For ozone above 1 hPa, the model proclivity for underestimation is more pronounced in Figure 13a than shown for MIPAS retrievals in Figure 11a. In contrast, for the lower stratosphere, control run and assimilation results are mostly larger

than HALOE retrievals, maintaining in better agreement with MIPAS retrievals. Improvements by data assimilation can be claimed, as there is a closer fit of analysis profiles and background profiles with HALOE, than the control run. Again, the forecasts based on the background initial values benefit from assimilation results of previous days. A study by Cortesi *et al.* [2007] reveals that displayed systematic discrepancies in the lower stratosphere appear to result mainly from corresponding systematic differences between HALOE and MIPAS retrievals, at least within the 100–1 hPa height range, where MIPAS retrievals are mainly higher than HALOE retrievals.

[86] Comparing assimilation results for  $CH_4$  with HALOE retrievals (Figure 13b), a general agreement can be found throughout the stratosphere. Notable differences occur in the lower mesosphere, with analysis values being smaller. This can also be stated for the lower stratospheric southern



**Figure 14.** Mean  $(A - O)/O$  differences between profiles of analysis, control run, and background run relative to retrievals from SAGE2 for (a)  $O_3$  and (b)  $H_2O$  averaged over time span 8 September to 15 October and latitude belts as indicated above each plot. Plotting conventions are as for Figure 11 but without minimum threshold value of relative errors.

polar region. Given the excellent consistency between MIPAS and analyses, the discrepancies with HALOE result from the assimilation of MIPAS data.

[87] HALOE water vapor observations are systematically smaller than assimilation results in the middle and even more in the upper stratosphere (Figure 13c). This is in contrast to the general agreement found between MIPAS retrievals and both analyses and control runs, and reflects the differences between MIPAS and HALOE retrievals in those height levels. Similar or even enhanced discrepancies between HALOE and MIPAS retrievals occur below 100 hPa. For this low elevation however, estimated retrieval errors of both sensors are high, such that there is little impact on the analyses by those retrievals.

[88] For a comparison of MIPAS based assimilation results with SAGE II data, only retrieved ozone and water vapor profiles are available. Figure 14a displays the results for ozone, where the same scheme of deviations from MIPAS based analyses is exhibited as for HALOE: Aloft of 10 hPa the  $(A - O)/O$  differences are increasingly negative with height, while becoming positive in the lower stratosphere. For the southern polar region  $[90^\circ S - 60^\circ S]$  ozone below 10 hPa shows a large discrepancy between MIPAS and SAGE II. It should be noted again that the estimated retrieval errors in the lower stratosphere are large. Nevertheless, the differences between analyses and SAGE II retrievals are mostly smaller than for the control run and the first guess run. This indicates the sustained information gain by the assimilation procedure, despite systematic differences between MIPAS on the one

hand and the occultation sensors SAGE II and HALOE on the other hand.

[89] SAGE II water vapor  $(A - O)/O$  differences are similar to HALOE, where values aloft of 20 hPa are higher than MIPAS retrievals and assimilation results. The use of SAGE II retrievals are explicitly not recommended at lowest stratospheric height levels and therefore cut off in Figure 14a).

## 7. Conclusions

[90] The novel SACADA 4D-var system for operational assimilation of stratospheric observations has been developed ab initio, aiming to provide a middle atmospheric chemistry analysis tool, which is efficient in terms of processing available retrieval information and use of computational resources. The common grid structure and the common time steps of the meteorological driver module GME with the chemistry data assimilation section avoids spatial and temporal interpolation of the meteorological fields with associated information loss and error generation. In particular, a consistent representation of vertical wind fields is available for the solution of the advection-reaction equation.

[91] Moreover, the numerical efficiency benefits considerably from the icosahedral grid, where the computational costs are reduced by about 30%, compared to CTMs employing a traditional latitude-longitude grid. In addition, the semi-Lagrange horizontal transport scheme applicable for short simulation intervals, leads to an excellent efficiency of the

new system, where the number of transported constituents is large. While for adjoints of semi-Lagrangian schemes on traditional grids considerable efforts are required to maintain accuracy at the poles, the icosahedral grid is not affected. The efficient system design enables the application of the computationally costly 4D-var technique some eight times faster than real time, and opens options for further grid refinements and chemical mechanism extensions.

[92] Particular efforts were devoted to implement spatial correlations with the BECM. This problem was solved by introducing the diffusion approach following *Weaver and Courtier* [2001]. A modified potential vorticity was adopted to identify air mass structures, allowing to estimate anisotropic and inhomogeneous correlation lengths accordingly. The implementation presented in this work assumes larger background error correlations along isopleths of potential vorticity in regions where large gradients of potential vorticity prevail. At the polar vortex edge, analyses of chemical constituents appear to be more consistent with the dynamics than those made with a homogeneous and isotropic formulation, which ignores dynamic patterns.

[93] The diffusion approach proved to have several advantages, as it (1) saves the storage of a full BECM, which is replaced by an operator with the same statistical properties, (2) does not need the assumption of separability of horizontal and vertical correlations, as related diffusion operators are alternately applied during the three-dimensional integration, (3) allows for an easy preconditioning of the minimization problem by straightforward calculation of the square root of the BECM, (4) is numerically efficient, as the complexity of the calculation is linear with the dimension of the model grid, (5) is amenable for flow-dependent design of spatial covariances, allowing for pronounced variations of the correlation lengths, and (6) can be easily adapted to nonstandard grid design as the icosahedral grid, without any difficulties.

[94] A suite of two case studies, comprising 1 September to 15 October 2002, and 21 October to 30 November 2003 served for validation of the assimilation system. EnviSat-MIPAS data products have been assimilated and the assimilation results have been validated with independent (not assimilated) data from SAGE II and HALOE. A sequence of tests proved the quality of the assimilation results. The a posteriori validation of normalized costs demonstrated a fast convergence toward the optimal value  $J_p^a \approx 1/2$  after 4 days of spin-up. Display of  $(O - A)$  and  $(O - B)$  density distributions reveal nearly bias free analyses with significantly reduced variance of PDF. The PV controlled dynamical BECM formulation exhibits advantages in areas with pronounced PV gradients. Comparisons of MIPAS based assimilation results with not assimilated SAGE II and HALOE retrievals revealed significantly improved analyses, albeit in the limits of agreements between the infrared sounder and the occultation sensors. It could be shown that the application of the assimilation system also improved short-range forecasts, as demonstrated by the background based model runs.

[95] In a companion paper (J. Schwinger and H. Elbern, Chemical state estimation for the middle atmosphere by 4-dimensional variational data assimilation: 2. A posteriori validation of error statistics in observation space, submitted

to *Journal of Geophysical Research*, 2010), an in-depth a posteriori validation is provided, where the mutual consistency of the background and observation error covariance matrices will be established and an analysis error assessment in observation space will be presented.

[96] The SACADA system is running an operational service providing daily analyses in near real time.

[97] **Acknowledgments.** The authors are highly indebted to the German Weather Service and D. Majewski for giving access to GME code and providing advice. W. Joppich, S. Pott, and H.-G. Reschke, SCAI, Fraunhofer Society, gave a lot of support in adapting the vertical grid structure of GME to the needs of stratospheric modeling. J. Hendricks, DLR, provided advice on the use of the chemical mechanism including heterogeneous chemistry to the SACADA system. We are very grateful to D. Poppe, ICG-2, Research Centre Jülich, and E.-P. Röth, University Essen and ICG-2, FZ Jülich, for a critical final review of the extended version of the chemistry mechanism, and to Anne Smith, NCAR, for provision of photolysis rates. MIPAS data have been processed and provided by ESA. We are grateful to G. Brasseur, NCAR, and A. Sandu, Virginia Tech, for giving access to SOCRATES and KPP software, respectively. SAGE II data were obtained from NASA Langley Research Centre, and HALOE data were obtained from Hampton University, Virginia, and NASA Langley Research Center. We are further indebted to the SACADA team, most notably M. Riese and L. Hoffmann, ICG 1, FZ Jülich, T. von Clarmann, IMK, KIT, and H. Bovensmann, IFE, University of Bremen, for manifold discussions on satellite retrieval error characteristics. The meteorological data for driving GME were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF). Computational resources were provided by University of Cologne's computer centre RRZK and the Jülich Supercomputing Centre. This work was funded by the German Federal Ministry of Education and Research in the frame of the funding program AFO 2000 with the grant FZK 07ATF48. The authors want to thank three anonymous reviewers, who helped to improve the manuscript.

## References

- Allen, D. R., R. M. Bevilacqua, G. E. Nedoluha, C. E. Randall, and G. L. Manney (2003), Unusual stratospheric transport and mixing during the 2002 Antarctic winter, *Geophys. Res. Lett.*, **30**(12), 1599, doi:10.1029/2003GL017117.
- Bannister, R. N. (2008a), A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances, *Q. J. R. Meteorol. Soc.*, **134**(637), 1951–1970.
- Bannister, R. N. (2008b), A review of forecast error covariance statistics in atmospheric variational data assimilation. II: Modelling the forecast error covariance statistics, *Q. J. R. Meteorol. Soc.*, **134**(637), 1971–1996.
- Bencherif, H., L. E. Amraoui, N. Semane, S. Massart, D. V. Charyulu, A. Hauchecorne, and V.-H. Peuch (2007), Examination of the 2002 major warming in the southern hemisphere using ground-based and Odin/SMR assimilated data: Stratospheric ozone distributions and tropic/mid-latitude exchange, *Can. J. Phys.*, **85**(11), 1287–1300.
- Brasseur, G., M. H. Hitchman, S. Walters, M. Dymek, E. Falise, and M. Pirre (1995), An interactive chemical dynamical radiative two-dimensional model of the middle atmosphere, *J. Geophys. Res.*, **95**(D5), 5639–5655.
- Carslaw, K. S., B. Luo, and T. Peter (1995), An analytic expression for the composition of aqueous  $\text{HNO}_3$ - $\text{H}_2\text{SO}_4$  stratospheric aerosols including gas phase removal of  $\text{HNO}_3$ , *Geophys. Res. Lett.*, **22**(14), 1877–1880.
- Carslaw, K. S., T. Peter, and R. Müller (1997), Uncertainties in reactive uptake coefficients for solid stratospheric particles—2. Effect on ozone depletion, *Geophys. Res. Lett.*, **24**(14), 1747–1750.
- Cohn, S. E. (1997), An introduction to estimation theory, *J. Meteorol. Soc. Jpn.*, **75**, 257–288.
- Cortesi, U., et al. (2007), Geophysical validation of MIPAS-ENVISAT operational ozone data, *Atmos. Chem. Phys.*, **7**, 4807–4867.
- Deshler, T., M. E. Hervig, D. J. Hofmann, J. M. Rosen, and J. B. Liley (2003), Thirty years of in situ stratospheric aerosol size distribution measurements from Laramie, Wyoming (41°N), using balloon-borne instruments, *J. Geophys. Res.*, **108**(D5), 4167, doi:10.1029/2002JD002514.
- Desroziers, G., L. Berre, B. Chapnik, and P. Poli (2005), Diagnosis of observation, background and analysis-error statistics in observation space, *Q. J. R. Meteorol. Soc.*, **131**(613), 3385–3396.

- Elbern, H., and H. Schmidt (1999), A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, *J. Geophys. Res.*, **104**(D15), 18,583–18,598.
- Elbern, H., and H. Schmidt (2001), Ozone episode analysis by four-dimensional variational chemistry data assimilation, *J. Geophys. Res.*, **106**(D4), 3569–3590.
- Elbern, H., H. Schmidt, and A. Ebel (1997), Variational data assimilation for tropospheric chemistry modeling, *J. Geophys. Res.*, **102**(D13), 15,967–15,985.
- Elbern, H., J. Hendricks, and A. Ebel (1998), A climatology of tropopause folds by global analyses, *Theor. Appl. Climatol.*, **59**, 181–200.
- Elbern, H., A. Strunk, H. Schmidt, and O. Talagrand (2007), Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, **7**, 1–59.
- Errera, Q., and D. Fonteyn (2001), Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, *J. Geophys. Res.*, **106**(D11), 12,253–12,265.
- Errera, Q., F. Daerden, S. Chabrilat, J. C. Lambert, W. A. Lahoz, S. Viscardy, S. Bonjean, and D. Fonteyn (2008), 4D-Var assimilation of MIPAS chemical observations: Ozone and nitrogen dioxide analyses, *Atmos. Chem. Phys.*, **8**, 6169–6187.
- Fierli, F., A. Hauchecorne, S. Bekki, B. Théodore, and O. F. d'Andon (2002), Data assimilation of stratospheric ozone using a high-resolution transport model, *Geophys. Res. Lett.*, **29**(10), 1381, doi:10.1029/2001GL014272.
- Fischer, H., and H. Oelhaf (1996), Remote sensing of vertical profiles of atmospheric trace constituents with MIPAS limb-emission spectrometers, *Appl. Opt.*, **35**(16), 2787–2796.
- Fisher, M., and P. Courtier (1995), Estimating the covariance matrices of analysis and forecast error in variational data assimilation, Tech. Memo. 220, 26 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, U. K.
- Fisher, M., and D. Lary (1995), Lagrangian four-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, **121**(527), 1681–1704.
- Geer, A. J., et al. (2006), The ASSET intercomparison of ozone analyses: Method and first results, *Atmos. Chem. Phys.*, **6**, 5445–5474.
- Giering, R. (1999), Tangent linear and adjoint model compiler: Users manual, TAMC version 5.2, report, Mass. Inst. of Technol., Cambridge.
- Hanea, R. G., G. J. M. Velders, and A. Heemink (2004), Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport, *J. Geophys. Res.*, **109**, D10302, doi:10.1029/2003JD004283.
- Hanson, D., and K. Mauersberger (1988), Laboratory studies of the nitric acid trihydrate: Implications for the south polar stratosphere, *Geophys. Res. Lett.*, **15**(8), 855–858.
- Hanson, D. R., A. R. Ravishankara, and S. Solomon (1994), Heterogeneous reactions in sulfuric acid aerosols: A framework for model calculations, *J. Geophys. Res.*, **99**(D2), 3615–3629.
- Hendricks, J., E. Lippert, H. Petry, and A. Ebel (1999), Heterogeneous reactions on and in sulfate aerosols: Implications for the chemistry of the midlatitude tropopause region, *J. Geophys. Res.*, **104**(D5), 5531–5550.
- Hendricks, J., F. Baier, G. Günter, B. C. Krüger, and A. Ebel (2001), Stratospheric ozone depletion during the 1995–1996 Arctic winter: 3-D simulations on the potential role of different PSC types, *Ann. Geophys.*, **19**, 163–1181.
- Hölzemann, J., H. Elbern, and A. Ebel (2001), PSAS and 4D-var data assimilation for chemical state analysis by urban and rural observation sites, *Phys. Chem. Earth*, **26**, 807–812.
- Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc (1997), Unified notation for data assimilation: operational, sequential and variational, *J. Meteorol. Soc. Jpn.*, **75**, 181–189.
- Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, *J. Basic Eng.*, **82D**, 35–45.
- Kalnay, E. (2003), *Atmospheric Modeling, Data Assimilation and Predictability*, Cambridge Univ. Press, Cambridge, U. K.
- Lahoz, W. A., Q. Errera, R. Swinbank, and D. Fonteyn (2007), Data assimilation of stratospheric constituents: A review, *Atmos. Chem. Phys.*, **7**, 5745–5773.
- Lait, L. R. (1994), An alternative form for potential vorticity, *J. Atmos. Sci.*, **51**(12), 1754–1759.
- Liu, D. C., and J. Nocedal (1989), On the limited memory BFGS method for large scale optimization, *Math. Program.*, **45**, 503–528.
- Luo, B. P., K. S. Carslaw, T. Peter, and S. L. Clegg (1995), Vapor pressures of H<sub>2</sub>SO<sub>4</sub>/HNO<sub>3</sub>/HCl/HBr/H<sub>2</sub>O solutions to low stratospheric temperatures, *Geophys. Res. Lett.*, **22**(3), 247–250.
- Majewski, D., D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner (2001), The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests, *Mon. Weather Rev.*, **130**, 319–338.
- Marti, J., and K. Mauersberger (1993), A survey and new measurements of ice vapor pressure at temperatures between 170 and 250 K, *Geophys. Res. Lett.*, **20**(5), 363–366.
- McRae, G. J., W. R. Goodin, and J. H. Seinfeld (1982), Numerical solution of the atmospheric diffusion equation for chemically reacting flows, *J. Comput. Phys.*, **45**, 1–42.
- Nocedal, J. (1980), Updating quasi-Newton matrices with limited storage, *Math. Comput.*, **35**(151), 773–782.
- Otte, T. L., N. L. Seaman, and D. R. Stauffer (2001), A heuristic study on the importance of anisotropic error distributions in data assimilation, *Mon. Weather Rev.*, **129**, 766–783.
- Purser, R. J., W.-S. Wu, D. F. Parrish, and N. M. Roberts (2003), Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances, *Mon. Weather Rev.*, **131**, 1536–1548.
- Raspollini, P., et al. (2006), MIPAS level 2 operational analysis, *Atmos. Chem. Phys.*, **6**, 5605–5630.
- Ridolfi, M., et al. (2000), Optimized forward model and retrieval scheme for MIPAS near-real-time processing, *Appl. Opt.*, **39**(8), 1323–1340.
- Riishøjgaard, L. P. (1998), A direct way of specifying flow-dependent background error correlations for meteorological analysis systems, *Tellus, Ser. A*, **50**, 42–57.
- Sander, S. P., et al. (2006), Chemical kinetics and photochemical data for use in atmospheric studies, *JPL Publ.*, **06-2**, 523 pp.
- Sandu, A., J. G. Verwer, M. V. Loon, G. R. Carmichael, F. A. Potra, D. Dabdub, and J. H. Seinfeld (1997), Benchmarking stiff ODE solvers for atmospheric chemistry problems—I. Implicit vs explicit, *Atmos. Environ.*, **31**(19), 3151–3166.
- Sankey, D., and T. G. Shepherd (2003), Correlations of long-lived chemical species in a middle atmosphere general circulation model, *J. Geophys. Res.*, **108**(D16), 4494, doi:10.1029/2002JD002799.
- Segers, A. J., H. J. Eskes, R. J. van der A, R. F. van Oss, and P. F. J. van Velthoven (2005), Assimilation of GOME ozone profiles and a global chemistry-transport model using a Kalman filter with anisotropic covariance, *Q. J. R. Meteorol. Soc.*, **131**(606), 477–502.
- Seinfeld, J. H., and S. N. Pandis (1998), *Atmospheric Chemistry and Physics*, Wiley-Intersci., Hoboken, N. J.
- Shi, Q., J. T. Jayne, C. E. Kolb, D. R. Worsnop, and P. Davidovits (2001), Kinetic model for reaction of ClONO<sub>2</sub> with H<sub>2</sub>O and HCl and HOCl with HCl in sulfuric acid solutions, *J. Geophys. Res.*, **106**(D20), 24,259–24,274.
- Simmons, A. J., and D. M. Burridge (1981), An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates, *Mon. Weather Rev.*, **109**, 758–766.
- Staniforth, A., and J. Côté (1991), Semi-Lagrangian integration schemes for atmospheric models—A review, *Mon. Weather Rev.*, **119**, 2206–2223.
- Talagrand, O. (1997), Assimilation of observations, an introduction, *J. Meteorol. Soc. Jpn.*, **75**, 191–209.
- Talagrand, O. (1998), A posteriori evaluation and verification of analysis and assimilation algorithms, paper presented at Workshop on Diagnosis of Data Assimilation Systems, Eur. Cent. for Medium-Range Weather Forecasts, Reading, U. K.
- Talagrand, O., and P. Courtier (1987), Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. R. Meteorol. Soc.*, **113**(478), 1311–1328.
- Tanguay, M., and S. Polavarapu (1999), The adjoint of the semi-Lagrangian treatment of the passive tracer equation, *Mon. Weather Rev.*, **127**, 551–564.
- Thépaut, J.-N., P. Courtier, G. Belaud, and G. Lemaître (1996), Dynamical structure functions in a four-dimensional variational assimilation: A case study, *Q. J. R. Meteorol. Soc.*, **122**(530), 535–561.
- Thiebaut, H. J. (1976), Anisotropic correlation functions for objective analysis, *Mon. Weather Rev.*, **104**, 994–1002.
- Thomason, L. W., L. R. Poole, and T. Deshler (1997), A global climatology of stratospheric aerosol surface area density deduced from Stratospheric Aerosol and Gas Experiment II measurements: 1984–1994, *J. Geophys. Res.*, **102**(D7), 8967–8976.
- van Loon, M., P. J. H. Builtjes, and A. J. Segers (2000), Data assimilation of ozone in the atmospheric transport chemistry model LOTOS, *Environ. Modell. Software*, **15**, 603–609.
- Verlaan, M., and A. W. Heemink (1995), Reduced rank square root filters for large scale data assimilation problems, paper presented at Second International Symposium on Assimilation of Observations in Meteorology and Oceanography, World Meteorol. Org., Tokyo.
- Verwer, J. G., E. J. Spee, J. G. Blom, and W. Hundsdorfer (1997), A second order Rosenbrock method applied to photochemical dispersion problems, *Tech. Rep. MAS R9717*, Centrum voor Wiskunde en Informatica, Amsterdam.

- Wang, K.-Y., D. J. Lary, D. E. Shallcross, S. M. Hall, and J. A. Pyle (2001), A review on the use of the adjoint method in four-dimensional atmospheric-chemistry data assimilation, *Q. J. R. Meteorol. Soc.*, **127** (576), 2181–2204.
- Weaver, A., and P. Courtier (2001), Correlation modelling on the sphere using a generalized diffusion equation, *Q. J. R. Meteorol. Soc.*, **127** (575), 1815–1846.
- World Meteorological Organization (1999), Scientific assessment of ozone depletion: 1998, *Tech. Rep. 44*, Geneva.
- Wu, W.-S., R. J. Purser, and D. F. Parrish (2002), Three-dimensional variational analysis with spatially inhomogeneous covariances, *Mon. Weather Rev.*, **130**, 2905–2916.
- Yanenko, N. N. (1971), *The Method of Fractional Steps: Solution of Problems of Mathematical Physics in Several Variables*, Springer, New York.
- 
- R. Botchorishvili, Department of Exact and Natural Sciences, Tbilisi State University, Chavchavadze av. 3, 0128 Tbilisi, Georgia. (ramaz@tsu.ge)
- H. Elbern and J. Schwinger, Rhenish Institute for Environmental Research, University of Cologne, Aachener Str. 209, D-50931 Köln, Germany. (he@eurad.uni-koeln.de; js@eurad.uni-koeln.de)