Four–dimensional variational chemical assimilation of CRISTA stratospheric measurements

Quentin Errera and Dominique Fonteyn

Institut d'Aéronomie Spatiale de Belgique, Brussels, Belgium

Abstract. This paper presents a chemical assimilation system based on the variational method. It has been applied to an off-line stratospheric three–dimensional chemical transport model. This technique determines the best model initial conditions that minimize the model errors when compared to a set of observations for a predefined time window. The system is applied to the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) measurements for the November 5–11, 1994, mission. Observed species are O_3, CH_4, HNO_3, $ClONO_2$, N_2O_5, N_2O, and CFC–11. For an assimilation period of 12 hours, CRISTA observations are well estimated by the assimilation system. These results vary according to the species considered and depend on the CRISTA's observational error. In particular, assimilated ozone differ from CRISTA by less than 5%. Analyses have been compared with independent observations in order to validate the assimilation system. It was found that observations by the Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment and the Halogen Occultation Experiment (HALOE) agree well with the analyses. These comparisons suggested systematic differences between CRISTA and HALOE or ATMOS. Unconstrained constituents like NO_x and HCl are influenced through chemical coupling and show small discrepancies with ATMOS and HALOE data.

1. Introduction

A scientific description of our environment originates in the collection of observations and their synthesis in models. However, atmospheric observations provide information on our environment at specific times and locations. On the other hand, numerical models are based on numerous approximations due to the system's complexity and the requirement to comply with limited computer resources. Therefore exhaustive knowledge of the state of the atmosphere is impossible to achieve. In the eighties the need for improved weather prediction motivated the development of forecast systems that optimize the integration of observations into atmospheric models [*Ghil and Ide*, 1997]. This method, known as data assimilation, compensates the sparsity and inaccuracy of the observations and helps to provide a dynamically consistent picture of the atmosphere. Since then, data assimilation has been applied to other type of models, e.g., stratospheric chemical models. An overview of data assimilation in meteorology can be found in the work of *Ghil et al.* [1997].

Chemical data assimilation of the stratosphere, using the variational approach was introduced by *Fisher and Lary* [1995]. More recently, *Khattatov et al.* [1999] presented a discussion of the properties and results of both the variational technique and the extended Kalman filter. These studies, based on trajectory models, have shown that non-measured constituents can be determined through chemical coupling between measured and nonmeasured species. Using chemical trajectory models, the chemical processes are, in general, decoupled from the transport process. This reduces the problem size in the assimilation system, in contrast to assimilation in the Eulerian framework. *Elbern et al.* [1997] have applied the four–dimensional variational data assimilation technique to an Eulerian chemical transport model of the troposphere. They found that only a few measured key species convey sufficient information to improve the analysis of species coupled with the observed species.

In this context, we present a variational chemical data assimilation system, based on a Eulerian chemical transport model of the stratosphere. Using the four–dimensional variational data assimilation method, three sources of information can be optimally combined [*Elbern and Schmidt*, 1999]: (1) the "a priori" knowledge of the chemical state of the atmosphere, called first guess or background state, and the associated covariance statistics, (2) the model equations describing the evolution of the chemical state, providing a link between observed model components and unobserved components, and (3) the actual observations and the associated covariance statistics. An estimation of the chemical state (analysis) can be made on a mathematical basis given the reliability of these information sets relative to each other.

The assimilation system is applied to the observations of the first mission of the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA). Section 2 outlines the variational method, the stratospheric model on which it is applied, as well as the CRISTA data. The assimilation setup is discussed in section 3. In section 4 the analyses are compared with the CRISTA data and with independent observations. Finally, section 5 summarizes the results and discusses their implication.

2. Assimilation System

2.1. Variational Data Assimilation Method

The objective of the variational method is to find the initial condition of the model which allows it to approximate a set of observations. This can be done by minimizing the misfit between model values and observations, given by the objective or cost function J [*Talagrand and Courtier*, 1987] (using the notation of *Ide et al.* [1997])

$$J[\mathbf{x}(t_0)] = \frac{1}{2}[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]^T\mathbf{B}^{-1}(t_0)[\mathbf{x}(t_0) - \mathbf{x}^b(t_0)]$$

$$+\frac{1}{2}\sum_{i=0}^{N}(\mathbf{y}^o(t_i) - H[\mathbf{x}(t_i)])^T\mathbf{R}(t_i)^{-1}(\mathbf{y}^o(t_i) - H[\mathbf{x}(t_i)]),$$
$$(1)$$

where $\mathbf{x}(t_i)$ represents the model state vector at time t_i, $\mathbf{x}^b(t_0)$ is the first guess, and $\mathbf{B}$ is the error covariance matrix of $\mathbf{x}^b(t_0)$. Vectors $\mathbf{y}^o(t_i)$ and $\mathbf{R}(t_i)$ are respectively the observation state vector and the error covariance matrix associated with the observations at t_i. The operator H maps the model state $\mathbf{x}(t_i)$ into the observation space. The first term of the right–hand side, called the background term, ensures the uniqueness of the minimum. Furthermore, this term improves the accuracy of the solution by including prior information.

Efficient minimization algorithms require the knowledge of the gradient of J with respect to the initial condition, $\nabla J(\mathbf{x}_0)$, where $\nabla J(\mathbf{x}_0)$ is calculated with an algorithm that uses the adjoint model. In order to simplify the notation, let H be equal to the identity operator, such that $H[\mathbf{x}(t_i)] = \mathbf{x}(t_i)$. Furthermore, since the calculation of the gradient of the background term is straightforward, we will drop it in the following demonstration. If we apply a small perturbation on the initial state $\mathbf{x}(t_0)$, $\delta\mathbf{x}(t_0)$, the first–order perturbation on J, δJ will be given by

$$\delta J = \nabla J \, \delta\mathbf{x}(t_0). \qquad (2)$$

However, δJ can also be found from the definition of $\mathbf{x}$ in (0)

$$\delta J = \sum_{i=0}^{N}[\mathbf{y}^o(t_i) - \mathbf{x}(t_i)]^T\mathbf{R}(t_i)^{-1} \, \delta\mathbf{x}(t_i). \qquad (3)$$

Let M be the model operator. The evolution of the state $\mathbf{x}$ is given by

$$\frac{d\mathbf{x}(t)}{dt} = M[\mathbf{x}(t)]. \qquad (4)$$

Then

$$\frac{d\delta\mathbf{x}(t)}{dt} = M[\mathbf{x}(t) + \delta\mathbf{x}(t)] - M[\mathbf{x}(t)] = \mathbf{M}(t) \, \delta\mathbf{x}(t), \quad (5)$$

where $\mathbf{M}$ is the linear tangent operator (i.e., the Jacobian) of the model associated with M at time t. The evolution of a small perturbation of the initial state, $\delta\mathbf{x}(t_0)$ can be estimated by successive application of the operator $\mathbf{M}$ on that perturbation and one obtains

$$\delta\mathbf{x}(t_n) = \mathbf{M}(t_n, t_{n-1})\mathbf{M}(t_{n-1}, t_{n-2})\ldots$$
$$\mathbf{M}(t_1, t_0) \, \delta\mathbf{x}(t_0), \qquad (6)$$

where $\mathbf{M}(t_i, t_{i-1})$ is supposed to be sufficiently accurate for the stepwise integration of $\mathbf{x}(t_i)$, given $\mathbf{x}(t_{i-1})$. By equating (2) and (3) one finds

$$\nabla J \, \delta\mathbf{x}(t_0) = \sum_{i=0}^{N}(\mathbf{y}^o(t_i) - \mathbf{x}(t_i))^T\mathbf{R}(t_i)^{-1} \qquad (7)$$
$$\mathbf{M}(t_i, t_{i-1})\mathbf{M}(t_{i-1}, t_{i-2})\ldots\mathbf{M}(t_1, t_0) \, \delta\mathbf{x}(t_0).$$

This last equation gives the general method to calculate ∇J with the tangent linear model (TLM), but it requires a number of forward calculations of the TLM equal to the size of the initial state $\mathbf{x}(t_0)$, i.e., $\sim 4 \times 10^6$ in our case. Note that the model value $\mathbf{x}(t_i)$ must be known for each time step t_i.

A more efficient method uses the adjoint model. By definition, the adjoint operator $\mathbf{M}^*$ of the linear operator $\mathbf{M}$ satisfies [*Talagrand and Courtier*, 1987]

$$[\mathbf{y}^o(t) - \mathbf{x}(t)]^T\mathbf{R}(t)^{-1} \quad \mathbf{M}(t)\delta\mathbf{x}(t) = \qquad (8)$$
$$\mathbf{M}^*(t)[\mathbf{y}^o(t) - \mathbf{x}(t)]^T\mathbf{R}(t)^{-1} \quad \delta\mathbf{x}(t),$$

where $\mathbf{M}^*$ correspond to the transpose matrix of $\mathbf{M}$, i.e., $\mathbf{M}^T$. Applying repeatedly the properties in (8) on (7), one finds

$$\nabla J = \sum_{i=0}^{N}\mathbf{M}^T(t_0, t_1)\mathbf{M}^T(t_1, t_2)\ldots\mathbf{M}^T(t_{i-1}, t_i)$$
$$\mathbf{R}(t_i)^{-1}[\mathbf{y}^o(t_i) - \mathbf{x}(t_i)]. \qquad (9)$$

Hence, contrary to (7), (9) needs only one backward calculation to estimate ∇J. However, as for the TLM, the adjoint calculation requires the knowledge of the model state for each time step inside the assimilation period.

The minimization of the objective function by typical algorithms is an iterative process schematically shown by the following steps: (1) Calculate $J(\mathbf{x}_0)$, with a forward model run; (2) calculate $\nabla J(\mathbf{x}_0)$, with a backward adjoint model run; and (3) using $\mathbf{x}_0$, $\nabla J(\mathbf{x}_0)$ and an minimization algorithm, check the test criteria. If they are not satisfied, find the new initial conditions of the model and return to step 1. Else, stop.

2.2. Stratospheric Model and Its Adjoint

The core of the assimilation system, the three–dimensional (3–D) off-line chemical transport model, was developed for this study. The evolution of stratospheric species is calculated using operator splitting applied to the transport and the chemistry, with a time step of 1800 s. The dynamic is driven by the winds and temperature precalculated by

Table 1. Chemical Species Used in the Model

Category	Species
Transported and chemically active	O_3, O, $O(^1D)$, H, OH, H_2O_2, HO_2, HNO_3, HNO_4, N_2O_5, NO, NO_2, NO_3, Br, Br_2, BrCl, BrO, $BrONO_2$, HBr, HBrc, HOBr, CH_2O, CH_3, CH_3O, CH_3O_2, CH_3OOH, HCO, Cl, Cl_2, Cl_2O_2, $ClNO_2$, ClO, $ClONO_2$, ClOO, OClO, HOCl, HCl, HClc
Only transported	N_2O, CH_4, CFC–11
Fixed	O_2, N_2, H_2O, H_2, CO, CO_2

the U.K. Meteorological Office (UKMO) for the Upper Atmosphere Research Satellite (UARS) project [*Swinbank and O'Neill*, 1994]. They are updated every 24 hours at noon and are linearly interpolated at the time step of the model. They are defined at 22 vertical levels from 1000 up to 0.3 hPa (i.e., $10^{\frac{i}{6}-\frac{1}{2}}$, i=0,... 21) and have a horizontal resolution of 3.75° in longitude by 2.5° in latitude (i.e., 96 × 72 horizontal grid points). The model has the same resolution. However, we have reduced the vertical domain to 15 levels from 68.2 to 0.3 hPa in order to reduce the computer time. At these boundaries the vertical winds were set to zero.

Forty–one molecular species are calculated by the model and reported in Table 1. Long–lived species are fixed in the model, except for CH_4, N_2O, and CFC–11, which are transported. The advection of the volume mixing ratios is resolved using a semi–Lagrangian transport scheme [*Smolarkeiwicz and Rash*, 1991] with a time step of 1800 s. The chemical interaction is through 106 gas–phase reactions, 29 photoreactions, and 9 heterogeneous reactions given in Tables 2, 3, and 4. Since no polar stratospheric clouds were formed during the CRISTA 1 mission, only the heterogeneous reaction on sulfate aerosols was considered. The numerical algorithm used to integrate the chemical system of differential equations is a fourth–order Rosenbrock solver [*Hairer and Wanner*, 1991], with no family assumption. The method is efficiently implemented by reducing the linear algebra cost following *Sandu et al.* [1996, 1997]. This algorithm uses an internal time step, which is modulated ac-

Table 2. Gas–Phase Reactions Included in the Model[a]

No.	Reactions
1	$O + O_2 + M \rightarrow O_3 + M$
2	$O + O_3 \rightarrow 2O_2$
3	$O1D + N_2 \rightarrow O + N_2$
4	$O1D + O_2 \rightarrow O + O_2$
5	$O1D + O_3 \rightarrow 2O_2$
6	$O1D + O_3 \rightarrow O + O + O_2$
7	$O1D + H_2O \rightarrow 2OH$
8	$O1D + H_2 \rightarrow OH + H$
9	$O1D + CH_4 \rightarrow CH_2O + H_2$
10	$O1D + CH_4 \rightarrow CH_3 + OH$
11	$O1D + N_2O \rightarrow O_2 + N_2$
12	$O1D + N_2O \rightarrow NO + NO$
13	$H + O_2 + M \rightarrow HO_2 + M$
14	$H + O_3 \rightarrow OH + O_2$
15	$H_2 + OH \rightarrow H_2O + H$
16	$OH + O_3 \rightarrow HO_2 + O_2$
17	$OH + O \rightarrow O_2 + H$
18	$OH + OH \rightarrow H_2O + O$
19	$OH + OH + M \rightarrow H_2O_2 + M$
20	$HO_2 + O \rightarrow OH + O_2$
21	$HO_2 + O_3 \rightarrow OH + 2O_2$
22	$H + HO_2 \rightarrow 2OH$
23	$H + HO_2 \rightarrow H_2O + O$
24	$H + HO_2 \rightarrow H_2 + O_2$
25	$HO_2 + OH \rightarrow H_2O + O_2$
26	$HO_2 + HO_2 \rightarrow H_2O_2 + O_2$
27	$HO_2 + HO_2 + M \rightarrow H_2O_2 + O_2 + M$
28	$H_2O_2 + OH \rightarrow H_2O + HO_2$
29	$H_2O_2 + O \rightarrow OH + HO_2$
30	$NO + O_3 \rightarrow NO_2 + O_2$
31	$NO + HO_2 \rightarrow NO_2 + OH$
32	$NO_2 + O \rightarrow NO + O_2$
33	$NO_2 + O_3 \rightarrow NO_3 + O_2$
34	$NO_2 + OH + M \rightarrow HNO_3 + M$
35	$NO_2 + HO_2 + M \rightarrow HNO_4 + M$
36	$NO_3 + O \rightarrow O_2 + NO_2$
37	$NO_3 + NO \rightarrow 2NO_2$
38	$NO_3 + NO_2 + M \rightarrow N_2O_5 + M$
39	$N_2O_5 \rightarrow NO_2 + NO_3$
40	$HNO_3 + OH \rightarrow H_2O + NO_3$
41	$HNO_4 + OH \rightarrow H_2O + NO_2 + O_2$
42	$HNO_4 \rightarrow HO_2 + NO_2$
43	$Cl + O_2 + M \rightarrow ClOO + M$
44	$Cl + O_3 \rightarrow ClO + O_2$
45	$Cl + H_2 \rightarrow HCl + H$

Table 2. (continued)

No.	Reactions
46	$Cl + CH_4 \rightarrow HCl + CH_3$
47	$Cl + CH_2O \rightarrow HCl + HCO$
48	$Cl + HO_2 \rightarrow HCl + O_2$
49	$Cl + HO_2 \rightarrow OH + ClO$
50	$Cl + H_2O_2 \rightarrow HCl + HO_2$
51	$Cl + HOCl \rightarrow Cl_2 + OH$
52	$Cl + HOCl \rightarrow ClO + HCl$
53	$Cl + OClO \rightarrow ClO + ClO$
54	$Cl + ClOO \rightarrow Cl_2 + O_2$
55	$Cl + ClOO \rightarrow ClO + ClO$
56	$ClO + O \rightarrow Cl + O_2$
57	$ClO + OH \rightarrow HO_2 + Cl$
58	$ClO + OH \rightarrow HCl + O_2$
59	$ClO + HO_2 \rightarrow O_2 + HOCl$
60	$ClO + NO \rightarrow NO_2 + Cl$
61	$ClO + NO_2 + M \rightarrow ClONO_2$
62	$ClO + ClO \rightarrow Cl + OClO$
63	$ClO + ClO \rightarrow Cl + ClOO$
64	$ClO + ClO \rightarrow Cl_2 + O_2$
65	$ClO + ClO \rightarrow Cl_2O_2$
66	$ClOO \rightarrow Cl + O_2$
67	$ClO + NO_3 \rightarrow ClOO + NO_2$
68	$Cl_2O_2 \rightarrow 2ClO$
69	$HCl + OH \rightarrow H_2O + Cl$
70	$HCl + O \rightarrow OH + Cl$
71	$OClO + O \rightarrow ClO + O_2$
72	$OClO + OH \rightarrow HOCl + O_2$
73	$OClO + NO \rightarrow ClO + NO_2$
74	$HOCl + O \rightarrow ClO + OH$
75	$HOCl + OH \rightarrow H_2O + ClO$
76	$Cl_2 + OH \rightarrow HOCl + Cl$
77	$ClONO_2 + O \rightarrow ClO + NO_3$
78	$ClONO_2 + OH \rightarrow HOCl + NO_3$
79	$ClONO_2 + Cl \rightarrow Cl_2 + NO_3$
80	$Br + O_3 \rightarrow BrO + O_2$
81	$Br + HO_2 \rightarrow HBr + O_2$
82	$Br + CH_2O \rightarrow HBr + HCO$
83	$Br + OClO \rightarrow BrO + ClO$
84	$BrO + O \rightarrow Br + O_2$
85	$BrO + HO_2 \rightarrow HOBr + O_2$
86	$BrO + NO \rightarrow Br + NO_2$
87	$BrO + NO_2 \rightarrow BrONO_2$
88	$BrO + ClO \rightarrow Br + OClO$
89	$BrO + ClO \rightarrow Br + ClOO$
90	$BrO + ClO \rightarrow BrCl + O_2$
91	$BrO + BrO \rightarrow 2Br + O_2$
92	$BrO + BrO \rightarrow Br_2 + O_2$
93	$HBr + OH \rightarrow Br + H_2O$
94	$HBr + O \rightarrow Br + OH$
95	$HOBr + O \rightarrow BrO + OH$
96	$Br_2 + OH \rightarrow HOBr + Br$
97	$CO + OH \rightarrow H + CO_2$
98	$CH_4 + OH \rightarrow CH_3 + H_2O$
99	$CH_2O + OH \rightarrow HCO + H_2O$
100	$CH_2O + O \rightarrow HCO + OH$
101	$HCO + O_2 \rightarrow CO + HO_2$
102	$CH_3 + O_2 \rightarrow CH_3O_2$
103	$CH_3O + O_2 \rightarrow CH_2O + HO_2$
104	$CH_3O_2 + NO \rightarrow CH_3O + NO_2$
105	$CH_3O_2 + HO_2 \rightarrow CH_3OOH + O_2$
106	$CH_3OOH + OH \rightarrow CH_3O_2 + H_2O$

[a] Reaction rates are taken from *DeMore et al.* [1997].

cording to an error control mechanism. An external time step of 1800 s is used for the reaction rates calculation. The photodissociation coefficients are interpolated for each internal time step from a lookup table as a function of altitude, zenith angle, and O_3 column. The lookup table is precalculated using the two-stream delta Eddington method [*Toon et al.*, 1989] . The reaction rates and cross sections are taken from *DeMore et al.* [1997]. The liquid aerosol surface area density was derived from the aerosol extinction coefficient distribution at 1.02 μm with a climatological relation introduced by *Thomasson et al.* [1997]. This distribution was calculated with one month of the Stratospheric Aerosol and Gas Experiment (SAGE) II measurements following the method presented by *Franssens et al.* [2000] and then zonally averaged.

The backward in time calculation of the adjoint model requires the knowledge of the model state at each time step. For this reason, each species is stored every external time step in the forward run, after the transport and chemistry calculations. The adjoint of the advection algorithm is implemented using the automatic differentiation software TAMC [*Giering*, 1997]. For the chemistry the adjoint of the chemical equations was set up by TAMC. This system of equation was then solved with the same numerical algorithm as the forward case, following *Carmichael et al.* [1997]. Since this solver has a self–adaptive time step, we should, ideally, calculate the model state at each internal time step. This may lead to an important increase in the computing requirements. A simple solution was tested and adopted. Each intermediate state is determined from the stored states with a linear interpolation in time. This approximation was compared with the full backward model calculation in which the forward model recalculated the intermediate model states. With our numerical solver and our chemistry this approximation leads to less than 1% error in the derivative calculation, and no significant differences were present between the two assimilation results. This approximation leads to a speed up of a factor 2 for the adjoint chemistry calculation.

Finally, the minimization of the objective function uses the quasi–Newton algorithm M1QN3 from INRIA [*Gilbet and Lemarechal*, 1989]. In this procedure, we have inputed $log_e[\mathbf{x}(t_0)]$ and $\nabla_{ln(\mathbf{x}_0)} J(\mathbf{x}_0)$ instead of $\mathbf{x}(t_0)$ and $\nabla_{\mathbf{x}_0} J(\mathbf{x}_0)$ in order to assure the positiveness of the new initial condition [*Fisher and Lary*, 1995].

2.3. CRISTA Observations

The data used in this study are taken from the Cryogenic Infrared Spectrometers and Telescopes for the Atmosphere (CRISTA) experiment during the ATLAS 3 mission, between November 4 and 11, 1994. CRISTA measures the Earth limb infrared emission from which vertical profiles of O_3, HNO_3, $ClONO_2$, N_2O_5, CH_4, N_2O, and CFC–11 are inverted. This observation technique allows to obtain one scan of the limb every 25 s during daytime as well as during nighttime. Moreover, CRISTA has three telescopes working simultaneously, with their viewing direction separated by 18°. This leads to an horizontal resolution of 200 km × 650 km at the Equator and increasing with the latitude. For more details on the CRISTA instruments and their data, see *Offermann et al.* [1999] and *Riese et al.* [1999].

During that mission, CRISTA took over 50,000 profiles of trace gases with a latitude coverage from −57° to +67°, nearly all at the same local time. A global coverage of the equatorial region is obtained after 12 hours and a global coverage near the pole is completed in 24 hours. Plate 1b shows

Table 3. Photolytic Reactions Included in the Model[a]

No.	Reactions
1	$O_2 + h\nu \rightarrow 2O$
2	$O_3 + h\nu \rightarrow O + O_2$
3	$O_3 + h\nu \rightarrow O1D + O_2$
4	$HO_2 + h\nu \rightarrow OH + O$
5	$H_2O_2 + h\nu \rightarrow 2OH$
6	$NO_2 + h\nu \rightarrow NO + O$
7	$NO_3 + h\nu \rightarrow NO_2 + O$
8	$NO_3 + h\nu \rightarrow NO + O_2$
9	$N_2O_5 + h\nu \rightarrow NO_2 + NO_3$
10	$HNO_3 + h\nu \rightarrow OH + NO_2$
11	$HNO_4 + h\nu \rightarrow OH + NO_3$
12	$HNO_4 + h\nu \rightarrow HO_2 + NO_2$
13	$Cl_2 + h\nu \rightarrow 2Cl$
14	$OClO + h\nu \rightarrow O + ClO$
15	$Cl_2O_2 + h\nu \rightarrow Cl + ClOO$
16	$HOCl + h\nu \rightarrow OH + Cl$
17	$ClONO_2 + h\nu \rightarrow Cl + NO_3$
18	$ClONO_2 + h\nu \rightarrow Cl + NO_2 + O$
19	$ClNO_2 + h\nu \rightarrow Cl + NO_2$
20	$BrCl + h\nu \rightarrow Br + Cl$
21	$BrO + h\nu \rightarrow Br + O$
22	$HOBr + h\nu \rightarrow Br + OH$
23	$BrONO_2 + h\nu \rightarrow Br + NO_3$
24	$CH_2O + h\nu \rightarrow HCO + H$
25	$CH_2O + h\nu \rightarrow CO + H_2$
26	$CH_3OOH + h\nu \rightarrow CH_3O + OH$
27	$ClOO + h\nu \rightarrow O + ClO$
28	$ClO + h\nu \rightarrow O + Cl$
29	$Br_2 + h\nu \rightarrow 2Br$

[a] Cross–sections are taken from *DeMore et al.* [1997].

Table 4. Heterogeneous Reactions Included in the Model[a]

No.	Reactions
1	$ClONO_2 + H_2O(c) \rightarrow HOCl + HNO_3(c)$
2	$CLONO_2 + HCl(c) \rightarrow Cl_2 + HNO_3$
3	$N_2O_5 + H_2O(c) \rightarrow 2HNO_3(c)$
4	$N_2O_5 + HCl(c) \rightarrow ClNO_2 + HNO_3(c)$
5	$HOCl + HCl(c) \rightarrow Cl_2 + H_2O(c)$
6	$BrONO_2 + HBr(c) \rightarrow HOBr + HNO3$
7	$HOCl + HCl(c) \rightarrow Cl_2 + H_2O(c)$
8	$HOBr + HBr(c) \rightarrow Br_2 + H_2O$
9	$BrONO_2 + HClC \rightarrow BrCl + HNO_3$

[a] Reaction rates are taken from *DeMore et al.* [1997].

Table 5. CRISTA 1 Version 3 Systematic and Random Errors[a] for O_3, N_2O_5, N_2O, $ClONO_2$, and CH_4

Trace Gas	Altitude, km								
	20[b]	20[c]	25	30	35	40	45	55	58
O_3 systematic	20	12	11	–	10	–	11	15	17
O_3 random	8	2.5	2.0	–	2.0	–	2.5	4.0	6.0
$ClONO_2$ systematic	55	55	32	–	25	–	–	–	–
$ClONO_2$ random	25	6	3	–	10	–	–	–	–
CH_4 systematic	–	–	–	30	26	24	24	18	–
CH_4 random	–	–	–	12	8.5	8.5	7.5	7.5	–
N_2O systematic	–	–	26	23	26	38	–	–	–
N_2O random	–	–	3	3.5	6.5	7.5	–	–	–
N_2O_5 systematic	–	–	19	21	21	17	–	–	–
N_2O_5 random	–	–	5	5	5	5	–	–	–

[a] Errors are in percent.

[b] High latitudes.

[c] Equatorial latitudes.

Table 6. Same as Table 5 but for HNO_3 and CFC–11

Trace Gas	Altitude, km								
	20^a	20^b	20^c	25^a	25^b	25^c	30^a	30^b	30^c
HNO_3 systematic	12	0.3 ppbv	12	10	14	11	11	8	11
HNO_3 random	1.9	16	2.0	2.0	3.1	2.2	2.9	2.4	3.2
CFC–11 systematic	13	6	16	–	–	–	–	–	–
CFC–11 random	3.4	1.6	4.1	–	–	–	–	–	–

[a] At 40°N.

[b] Tropical latitudes.

[c] At 40°S.

the volume mixing ratio of HNO_3 at 31.2 hPa on November 11, gridded at the model resolution (see below for the description of the gridding operation). Fine horizontal dynamical processes can be seen, such as the tropical extrusion over South America [*Offermann et al.*, 1999].

Version 3 of the CRISTA 1 level 2 data was used in this study. The CRISTA systematic and random errors are given [*Riese et al.*, 1999] in Tables 5 and 6, as a function of the altitude and, for several species, with a distinction for different latitude regions. Unfortunately, the error given above does not take into account the regional variability of the concentration for some species, i.e., error of representativeness. For example, on Plate 1b, uniform error percentage is assumed for HNO_3 in the Southern Hemisphere, regardless of the location of the data point, e.g., inside or outside the South American streamer.

In this study, the CRISTA observations are vertically gridded to the model levels with a linear interpolation. We mapped the horizontal position of each data point on the nearest model grid point to simplify the calculation of the model state in the observation space. The covariance error matrix was defined using an linear interpolation of the Tables 4 and 5 to the vertical levels of the model. Only random errors were taken into account, and the data were not corrected for their systematic error. The observation error covariance matrix **R** is also diagonal, implying that correlations between species are not taken into account.

3. Assimilation setup

3.1. Analysis Strategy

The assimilation begins on November 5, 1994 (first day of uninterrupted data availability in the CRISTA 1 mission). The assimilation period is 12 hours. For a beter exploitation of the chemical coupling between observed and nonobserved species, an assimilation period longer than 24 hours is preferrable. However, this requires too much computer time [*Fisher and Lary*, 1995]. For an assimilation period of 12 hours the number of available observations is in the range $1–2×10^5$. A maximum of 60 iterations for the minimization algorithm was imposed for the first day of assimilation. This number was reduced to 30 for the following assimilation periods. This corresponds to a decrease of the objective function of less then or equal to 0.015 between the two last iterations. Using 8 processors on a CRAY SV1, the CPU time is 3950 s for one iteration, i.e., one calculation of the forward model, the adjoint model and the finding of the new initial condition, with a parallelization efficiency of 85%.

Every assimilation provides a new set of initial conditions for the considered period. We call "analysis" the model results during the assimilation period. After that window, the model results are named "forecast." Except for the November 5 morning assimilation period, all assimilations were initialized using the last model state from the previous analysis. The assimilation of the morning of November 5 has been initialized from a 4 day model simulation, started on November 1 at 0000 UT. This simulation was initialized using zonal

means of the chemical species taken from the interactive 2D model of the middle–atmosphere SOCRATES [*Huang et al.*, 1998] for the same period.

3.2. Background Covariance Matrix

The ratio between the corresponding elements of the observations error covariance matrix **R** and the background error covariance matrix **B** controls the influence of the observations on the new initial conditions *Elbern and Schmidt*, 1999 [@]. Since the initial guess is derived from a 2–D model climatology and since CRISTA requires 24 hours to obtain a global coverage of the stratosphere, a different matrix **B** was used for the first assimilated day than for the rest of the mission. For the two periods of November 5 the error of the first guess was set to 500%. For the other periods the first guess error has been reduced to 50%. In this case, the information from the observations is maximally introduced in the analysis. The matrix **B** was set up as simply as possible, all species having the same relative background error, all uncorrelated. This implies that spatial correlations and the correlation between chemically coupled species are not taken into account. The background covariance matrix is therefore diagonal.

4. Results

A typical result of the assimilation of CRISTA is presented on Plate 1, where we show both analyzed and observed HNO_3 at 31.2 hPa. The analysis is shown at noon on November 11, 1994, on Plate 1a, while Plate 1b shows a composite map of the observations over 24 hours. One can see a very good agreement between analysis and CRISTA. For example, the three tropical extrusions are finely reproduced by the analysis [*offermann et al.*, 1999].

In subsection 4.1 we will discuss how the analyses estimate the CRISTA data. Then the analyses will be compared with independent observations, taken by HALOE and ATMOS, in order to validate our assimilation system.

Since the analysis on November 5 uses a different definition of the matrix **B** than subsequent analyses, the following discussion includes only results from the November 6–11 period.
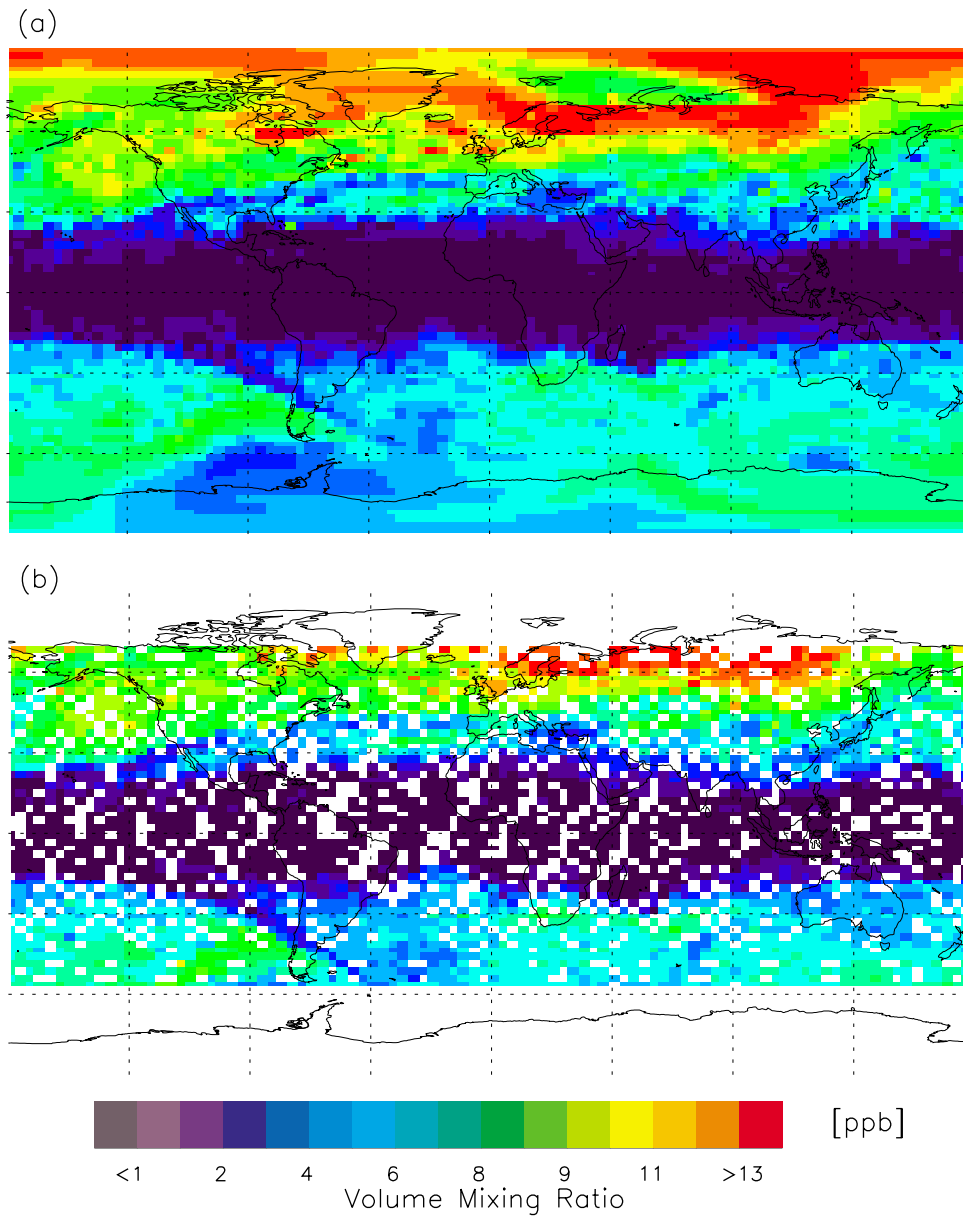
(a)

(b)

Plate 1. (a) Analyzed HNO₃ on November 11 at noon and (b) CRISTA HNO₃ on November 11 for the whole day, at 31.2 hPa.
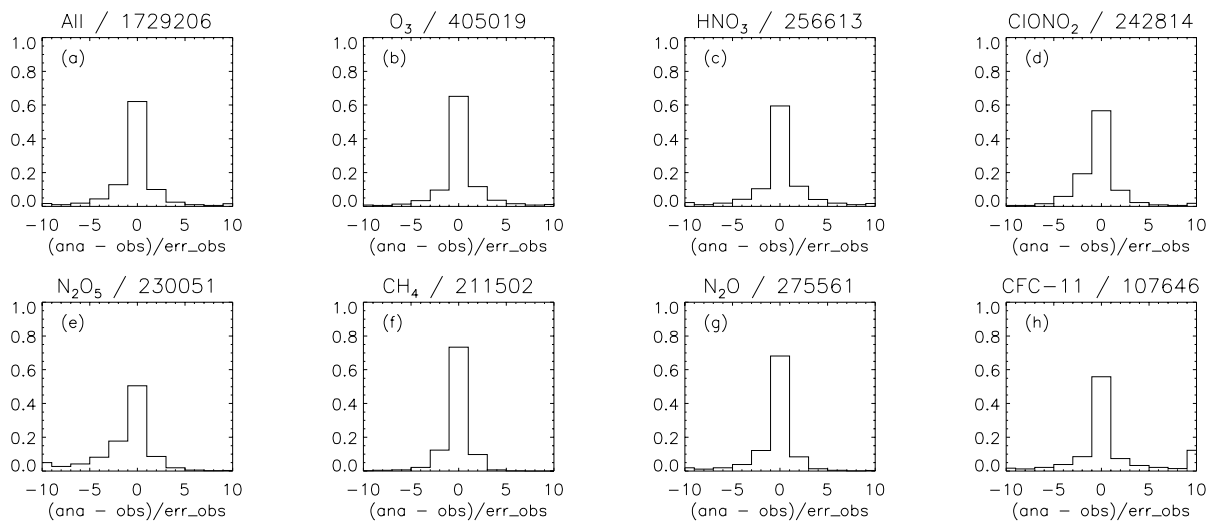
Figure 1. Distribution of the deviation between analysis and CRISTA, relative to the CRISTA random error, for the whole November 6–11 period: (a) all the species, (b–h) individual species.
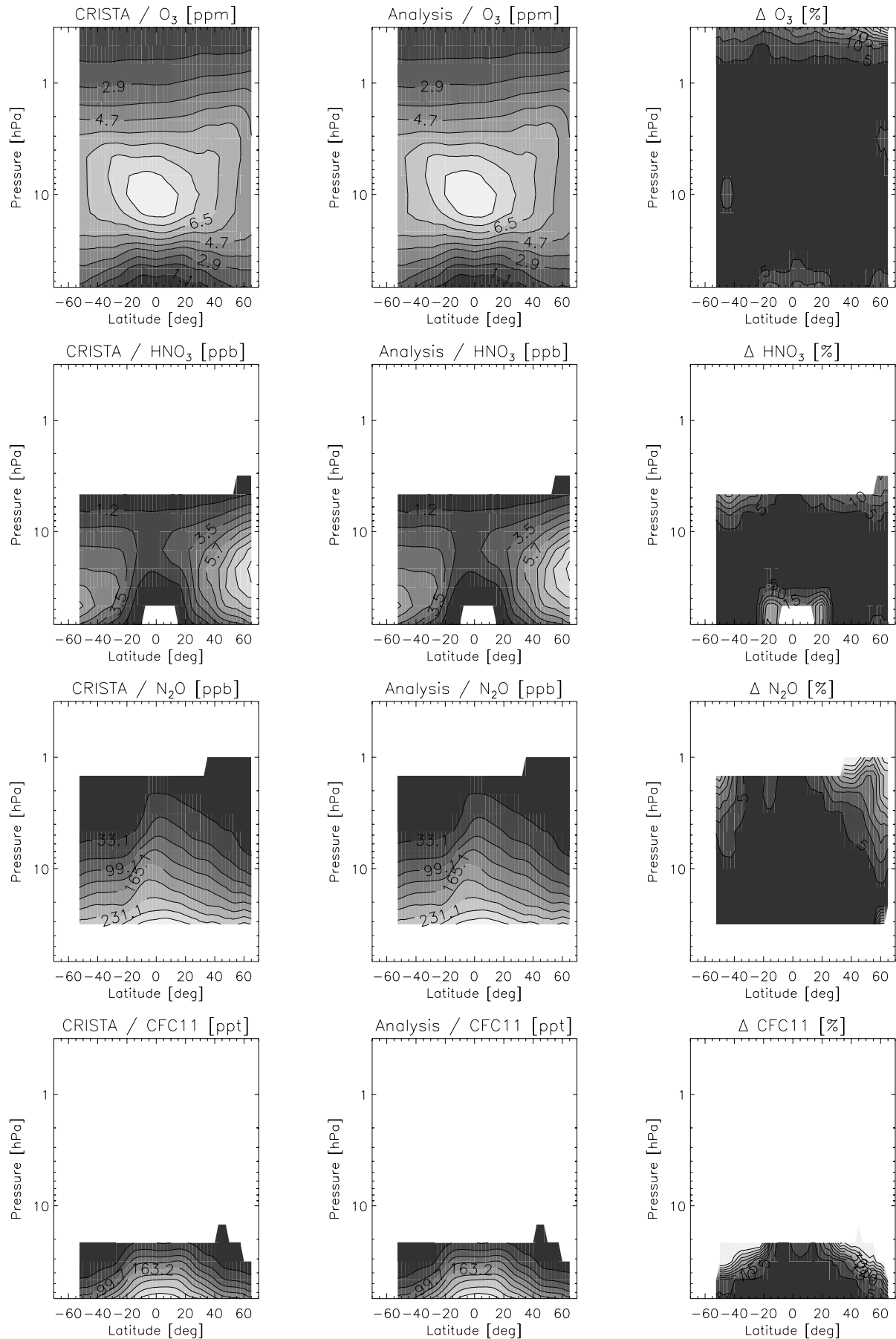
Figure 2. O_3, HNO_3, N_2O, CFC–11, CH_4, $ClONO_2$ and N_2O_5 (left) CRISTA zonal mean, the (middle) corresponding analysis, and the (right) percentage relative differences, for the assimilation of November 9, 1994, between 1200 and 2400 UT.
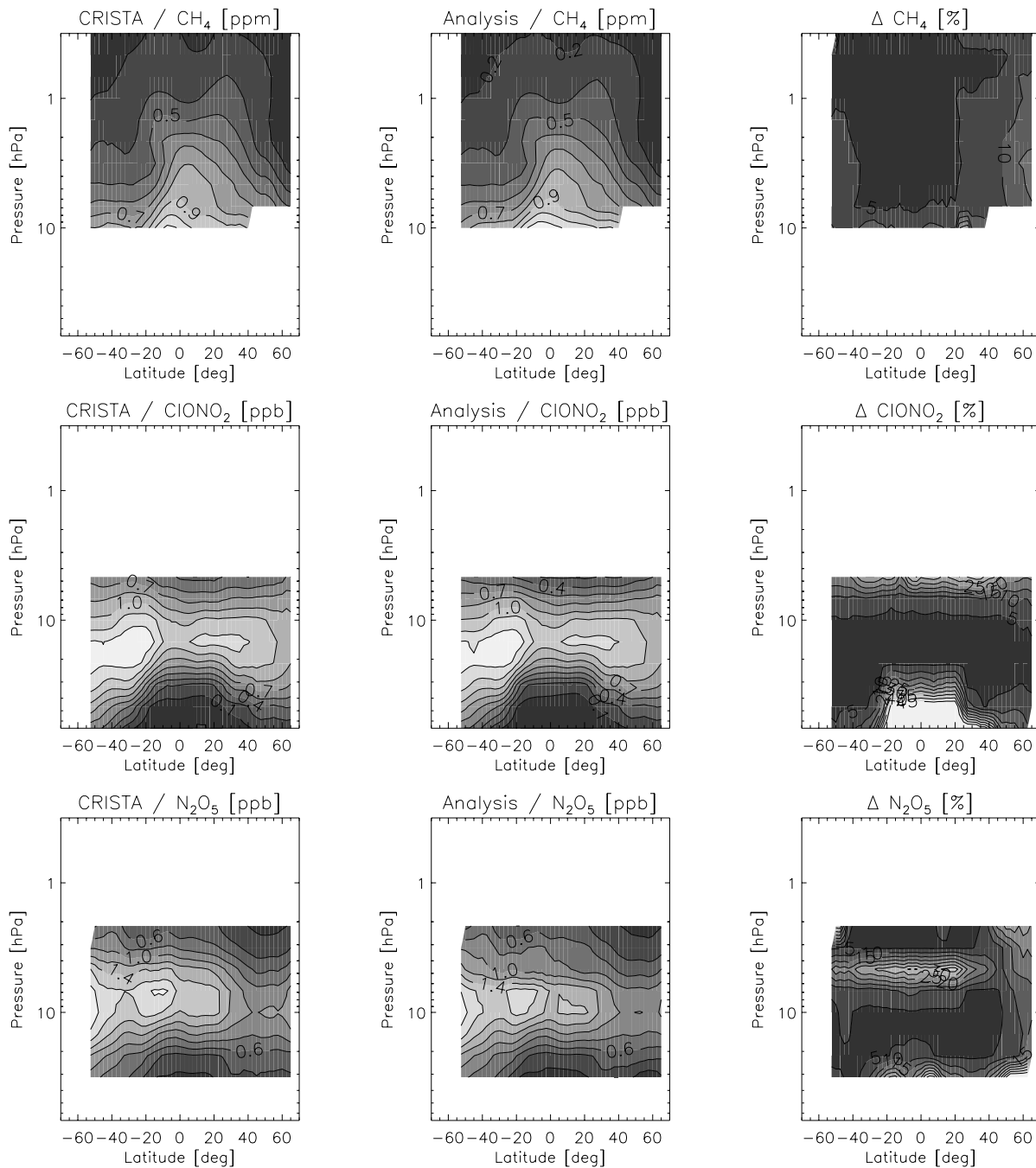
FIGURE 2 (continued)

4.1. Analysis Versus CRISTA

In Figure 1 we show the distributions of the deviations between the CRISTA data and the analyses, relative to the CRISTA random error ($\mathbf{R}$), for the whole mission. Figure 1a aggregates the deviations of all species, while Figures 1b–1h are relative to the individual species. Histograms are normalized by the total number of events, written at the top of each graph. In Figure 1a we see that deviations are narrowly centered on zero. More than 60% of the data are estimated inside the CRISTA error range. However, this is not a convincing result given the very simple definition of matrix $\mathbf{R}$. For example, CH_4 presents deviations smaller than CFC–11 mostly because the CH_4 random error is more than twice the CFC–11 random error. More interesting is

the asymmetry of the distribution of the $ClONO_2$ and N_2O_5 deviations (on Figure 1d and 1e). Since $ClONO_2$ and N_2O_5 as well as O_3 and HNO_3 are chemically coupled, they cannot be taken as independent variables. For this reason, the use of the CRISTA total error (random and systematic) in the definition of $\mathbf{R}$ could reduce the asymmetries presented in Figures 1d and 1e.

In order to provide a more detailed view of these deviations, Figure 2 displays the CRISTA zonal mean and the corresponding analysis of these species, for the morning of November 9 assimilation period. We also give the zonal mean of the relative differences (in percent), between CRISTA and the analysis. We see that large deviations occur when the amount of the considered species is very low, e.g., O_3 above 1 hPa, HNO_3 and $ClONO_2$ in the tropical regions, N_2O in
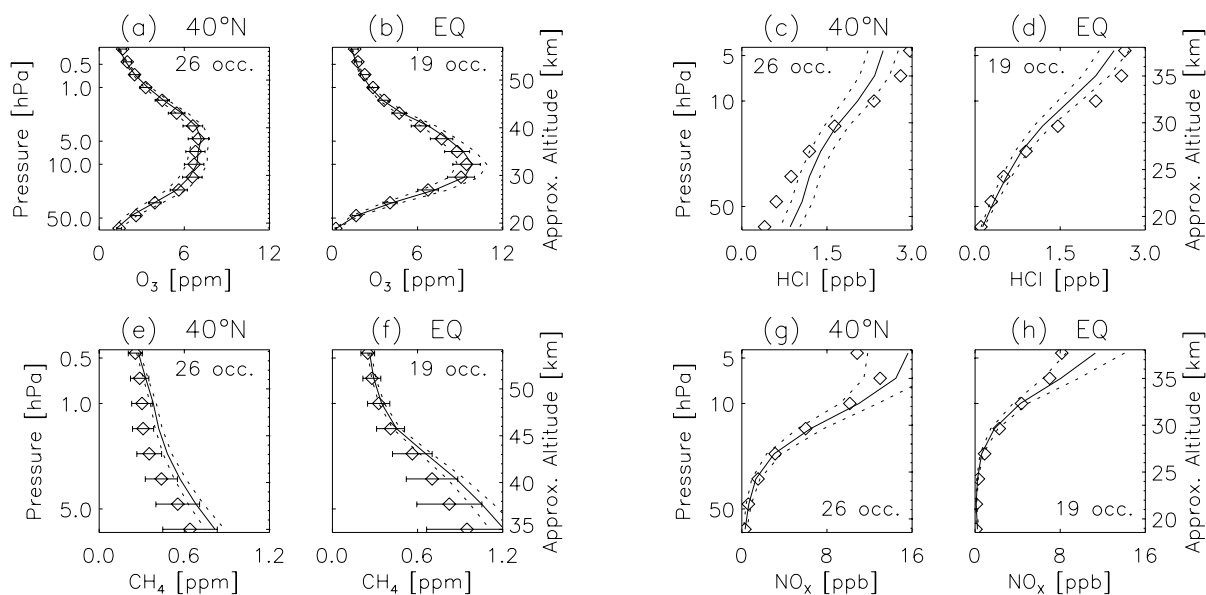
Figure 3. Comparison between HALOE (solid line) and analysis (diamond) for O_3 (a and b), HCL (c and d) , CH_4 (e and f), and NO_x (g and h) at different latitudes, 40°N and the Equator (EQ) (see text for details). The dotted lines represent the mean HALOE uncertainties, while the error bars on the analysis represent the CRISTA systematic uncertainties (for O_3 and CH_4).

the polar region near 2–5 hPa, CFC–11 above 30 hPa at midlatitudes, and $ClONO_2$ and N_2O_5 at their lower and upper levels. In Figure 2 a large deviation ($\sim$30%) can be seen between analysis and observations for N_2O_5, at 4.6 hPa, which could come from the retrieval process of N_2O_5 (Martin Riese, personal communication, 1999). Nevertheless, we generally see a very good agreement between analysis and data.

4.2. Comparison With HALOE Measurements

Considering the very good agreement between the CRISTA observations and the analyses, we have compared them with the Halogen Occultation Experiment (HALOE) [*Russell et al.*, [1993] observations, version 19. It uses the occultation measurement technique and retrieves the vertical profiles of numerous trace gases of the stratosphere, in particular, O_3, NO, NO_2, HCl, and CH_4. HALOE has performed around 180 occultations from 10°S to 40°N during the November 6–11 period.

Average vertical profiles of HALOE concentrations are compared with the assimilated fields in Figure 3, for two different latitudes (40°N and 0°), on November 11. The number of used profiles is given in Figure 3. This averaging takes the total error on the measurements into account by a weighting procedure, providing also the uncertainties on the average [*Taylor*, [1982] (the dotted lines on Figure 3). The averaged analysis profiles are determined from the assimilated fields at time steps and locations nearest to the times and positions of the corresponding HALOE measurements. The error bars associated to the analysis in Figure 3 represent the systematic CRISTA errors (only for the species measured by CRISTA). This allows us to estimate the possible systematic differences between the analysis and HALOE data. Note the altitude range, which corresponds to the range where these species are measured by CRISTA. For NO_x and HCl we took the same levels as for HNO_3 and $ClONO_2$, the species influencing NO_x and HCl, respectively.

Differences between the analyzed O_3 and HALOE are less than 7%. These discrepancies are in the error range of CRISTA and HALOE. This comparison confirms the

good quality of the CRISTA O_3, since this result would not be possible without high–quality ozone measurements. The same comparison for methane shows that the analysis underestimates HALOE by around $\sim$25%. It follows that the CRISTA CH_4 concentrations seems to be underestimated since (1) this discrepancy cannot be due to the transport representation in the model and (2) the HALOE CH_4 data have been intercompared with correlative data with an agreement of 15% [*Park et al.*, [1996]. Note that this is consistent with the CRISTA systematic error for methane.

HALOE measurements of NO, NO_2, and HCl provide a good opportunity to test how these constituents are influ-
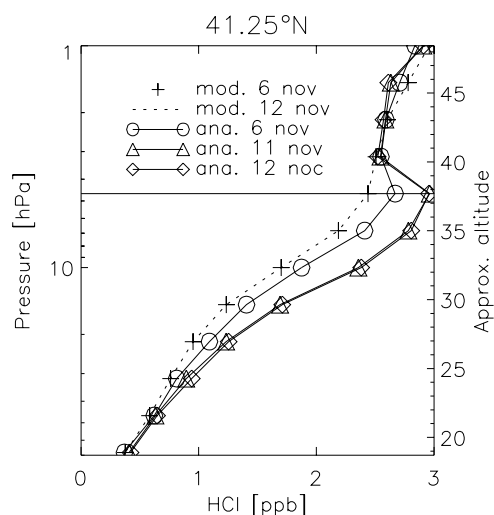

Figure 4. Comparison between zonal HCl taken from two model simulations and three analyses at 41.25°N. Model results are for November 6 (plus) and 12 (dotted line) and analysis are for November 6 (solid line, circle), 11 (solid line, triangle) and 12 (solid line, diamond); all at 0000 UT. The horizontal line at 4.6 hPa corresponds to the upper limit of the $ClONO_2$ measurements by CRISTA.
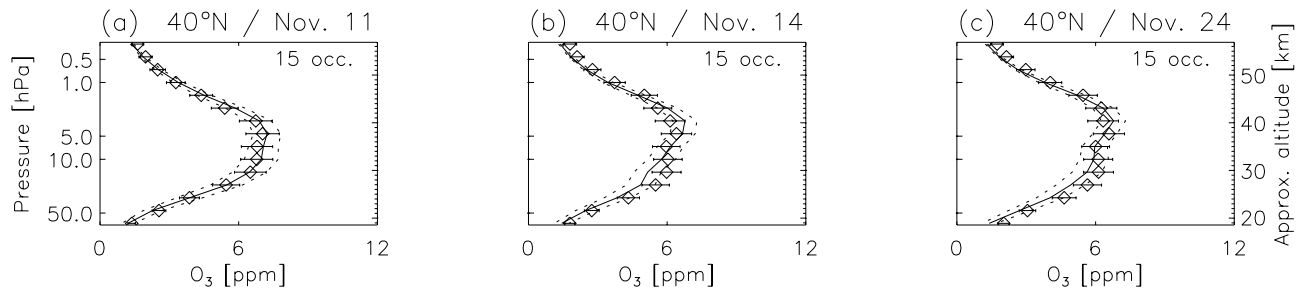
Figure 5. Same as Figure 3 for ozone on November 11, 14, and 24, 1994. On November 11, HALOE is compared with the analyzed O_3 while on November 14 and 24, HALOE is compared with the forecasted O_3.

enced by the constrained species (O_3, HNO_3, $ClONO_2$, and N_2O_5) through chemical coupling. We have compared NO_x ($NO+NO_2$) instead of NO or NO_2 in order to avoid a possible large error at sunset and sunrise, due to a possible time shift of maximum 15 min between the data and the analysis. In Figures 3g and 3h the analyzed NO_x is seen to be in good agreement with HALOE NO_x. We find the discrepancy to be less than 15% at 40°N and less than 25% at the Equator except at the 4.6 hPa pressure level. This is probably due to the better agreement between the analysis and CRISTA at midlatitudes than at the Equator, for HNO_3 and $ClONO_2$ (see Figure 2). Also, the large deviation at 4.6 hPa could be related to the N_2O_5 analysis problem at this level.

Before comparing the analyzed HCl with HALOE, we would like to estimate the influence of $ClONO_2$ on HCl. In Figure 4 we have compared the zonal HCl at 40° N from different analyses and model simulations. The three analyses correspond to the CRISTA data assimilation on November 6, 11, and 12, 1994 at 0000 UT. The model results from a free simulation started on November 1 with the SOCRATES initialization (see section 3.2) are also shown, corresponding to November 6 and 12 at 0000 UT. The influence of $ClONO_2$ is clear: (1) we see that HCl is not influenced at altitudes where $ClONO_2$ is not measured by CRISTA and (2) while the zonal HCl does not change between November 6 and November 12 in the model simulation, we see an increase of HCl in the analysis, which is stabilized after 5 days of assimilation of CRISTA. For this last reason, the comparison of HCl with HALOE has been done at the end of the mission. The comparison shows a good agreement, around 25% (Figures 3c and 3d). The analyzed HCl suggests a larger vertical gradient than in the HALOE observations.

Starting from the analysis of November 11, a forecast until November 25 has been made. The forecasted ozone is compared with the HALOE ozone to investigate a possible drift in the model. Figure 5 shows the 24–hour mean HALOE ozone profiles for November 11, 14, and 24 with colocated analysis/forecast ozone profiles. From figure 5 it is clear that the forecasted ozone does not evolve noticeably in time and remains close to the HALOE observations. In general, the forecast of November 24 compares better to HALOE than does the forecast on November 14. This indicates that although the model is highly consistent with HALOE, the modelled ozone variability is not entirely consistent with HALOE. However, a single forecast is not sufficient to quantify possible model deficiencies. Nevertheless, this forecast did show a good ozone prediction. Further-

more, this good agreement between the model and HALOE ozone confirms the reliability of the transport scheme implementation and the O_3 chemistry used in the model.

4.3. Comparison With ATMOS Measurements

During the ATLAS 3 mission the Atmospheric Trace Molecule Spectroscopy (ATMOS) experiment was pre-sent on the space shuttle [*Gunson et al.*, [1996]]. This occultation experiment has retrieved the vertical profiles of O_3, CH_4, HCl, NO, NO_2, HNO_3, $ClONO_2$, and N_2O_5, between 10° and 40° N (for the sunrise measurements). Since the com-
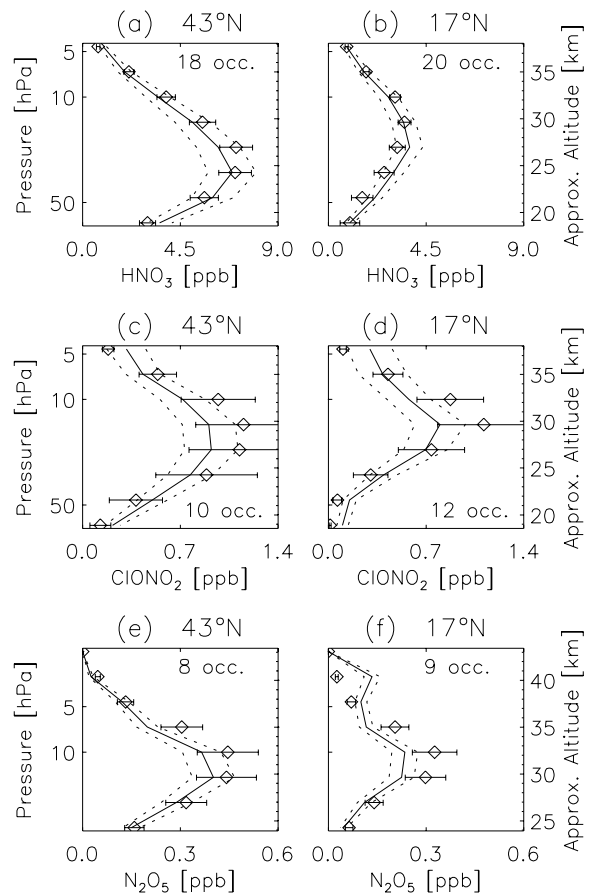

Figure 6. Comparison between ATMOS (solid line) and analysis (diamond) for HNO_3, $ClONO_2$, and N_2O_5 at different latitudes, 40°N and 17°N (see text for details). The dotted lines represent the mean ATMOS uncertainties, while the error bars on the analysis represent the CRISTA systematic uncertainties.

parison between the analysis and ATMOS for O_3, CH_4, HCl and NO_x confirms the conclusions found in the previous section, they will not be presented here.

Average ATMOS (version 2) profiles and their corresponding analysis are plotted in Figure 6 for HNO_3, $ClONO_2$, and N_2O_5. The number of profiles used is given in Figure 6.

The analyzed HNO_3 agrees well with ATMOS, generally better than 15%, and the analysis is in the error range of ATMOS. We found larger deviations in the comparisons for $ClONO_2$ and N_2O_5, but these measurements are known to be less accurate [*Abrams et al.*, [1996]], as suggested by the error bars. At their maximum concentrations the analysis overestimates ATMOS by 25% for $ClONO_2$ and 17% for N_2O_5. Nevertheless, note the overlap of the error bars. These differences between analysis $ClONO_2$ and N_2O_5 and the ATMOS data could be due to systematic differences between ATMOS and CRISTA for these observations. Indeed, we have seen in section 4.1 that the analysis follows very well the CRISTA data for these species at these altitudes.

5. Summary and Discussion

We have presented a four–dimensional chemical assimilation system of the stratosphere applied to the CRISTA 1 observations, using the variational method. The analyses generally estimate the CRISTA data within the CRISTA random error. Larger relative differences were found in regions where a given species is less abundant. Taking into account the simple error description and the simple implementation of the CRISTA observations, the agreement between the analysis and the observational data is very good. Further improvements in the data assimilation system should include a more accurate error representation and possibly a more detailed error covariance matrix allowing for spatial and species correlations.

The good agreement between the analysis and independent measurements by HALOE and ATMOS has validated the assimilation system. The assimilated concentrations are always within the error bars. These comparisons have demonstrated that the analysis is very representative of the CRISTA data. This is essentially due to the small weight assigned to the first guess in the minimization procedure. It is therefore possible to use the analysis in order to derive systematic differences between CRISTA and other instruments.

The comparison of analyzed HCl and NO_x with HALOE has shown that unmeasured species can be predicted through chemical coupling with the measured species (O_3, HNO_3, $ClONO_2$, and N_2O_5 here). We have seen that $\sim$5 days of assimilation are necessary to obtain a stable HCl using an assimilation period of 12 hours. Using a longer assimilation than 12 hours, e.g., 24 hours or more, would have resulted in fewer days of assimilation to obtain an HCl consistent with $ClONO_2$. Unfortunately, such a period is unpractical, owing to computer limitations. Computation time is probably the major limitation of the 4D–VAR method. In conclusion, this assimilation system can provide an important support for data validation and interpretation in the future, e.g., for the Environment Satellite (ENVISAT) mission.

References

Abrams, M. C., et al., On the assessment and uncertainty of atmospheric trace gas burden measurements with high resolution infrared solar occultation spectra from space by the ATMOS experiment, *Geophys. Res. Lett.*, *23*, 2337–2340, 1996.

Carmichael, G. R., A. Sandu, and F. A. Potra, Sensitivity analysis for atmospheric chemistry models via automatic differentiation, *Atmos. Environ.*, *31*, 475–489, 1997.

DeMore, W. B., S. P. Sander, D. M. Golden, R. F. Hampson, M. J. Kurylo, C. J. Howard, A. R. Ravishankara, C. E. Kolb, and M. J. Molina, Chemical kinetics and photochemical data for use in stratospheric modeling, evaluation number 12, *Publ. 97–4*, Jet Propul. Lab., Pasadena, Calif., 1997.

Elbern, H., and H. Schmidt, A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling, *J. Geophys. Res.*, *104*, 18,583–18,599, 1999.

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

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

Franssens, G., D. Fonteyn, M. DeMazière, and D. Fussen, A comparison between interpolation and assimilation as cartography methods for the SAGE–II aerosol product, in *Inverse Methods in Global Biogeochemical Cycles*, Geophys. Monogr. Ser., vol. 114, edited by P. Kasibhatla et al., pp. 155–169, AGU, Washington, D.C., 2000.

Ghil, M., and K. Ide, Introduction, in *Data Assimilation in Meteorology and Oceanography: Theory and Practice*, edited by M. Ghil et al., pp. i–iii. The Meteorol. Soc. of Jpn., Tokyo, 1997.

Ghil, M., K. Ide, A. Bennett, P. Courtier, M. Kimoto, M. Nagata, M. Saiki, and N. Sato. (ed.), *Data Assimilation in Meteorology and Oceanography: Theory and Practice*, The Meteorol. Soc. of Jpn., Tokyo, 1997.

Giering, R. *Tangent Linear and Adjoint Model Compiler: User Manual*, Cent. for Global Change Sci., MIT Press, Cambridge, Mass., 1997.

Gilbert, J.-C., and C. Lemarechal, Some numerical experiments with variable storage quasi–newton algorithms, *Math. Prog.*, *B25*, 407–435, 1989.

Gunson, M. R., et al., The atmospheric trace molecule spectroscopy (ATMOS) experiment: Deployment on the atlas space shuttle missions, *Geophys. Res. Lett.*, *23*, 2333–2336, 1996.

Hairer, E., and G. Wanner, *Solving Ordinary Differential Equations II. Stiff and Differential - Algebraic Problems*, Springer, New York, 1991.

Huang, T., et al., Description of SOCRATES - A chemical dynamical radiative two-dimensional model, *Tech. Rep. TN-440+EDD*, 94 pp., Natl. Cent. for Atmos. Res., Boulder, Colo. 1998.

Ide, K., P. Courtier, M. Ghil, and A. Lorenc, Unified notation for data assimilation: Operational sequential and variational, *J. Meteorol. Soc. Jpn.*, *75*, 181–189, 1997.

Khattatov, B. V., J. C. Gille, L. V. Lyjak, G. P. Brasseur, V. L. Dvortsov, A. E. Roche, and J. W. Water, Assimilation of photochemically active species and a case analysis of UARS data, *J. Geophys. Res.*, *104*, 18,715–18,737, 1999.

Offermann, D., K. U. Grossmann, P. Barthol, P. Knieling, M. Riese, and R. Trant, The cryogenic infrared spectrometer and telescopes for the atmosphere (CRISTA) experiment and middle atmosphere variability, *J. Geophys. Res.*, *104*, 16,311–16,327, 1999.

Park, J. H., et al., Validation of halogen occultation experiment CH4 measurements from the UARS, *J. Geophys. Res.*, *101*, 10,183–10,203, 1996.

Riese, M., R. Spang, P. Preusse, M. Ern, M. Jarisch, D. Offermann, and K. U. Grossmann, Cryogenic infrared spectrometers and telescopes for the atmosphere (CRISTA) data processing and atmospheric temperature and trace gas retrieval, *J. Geophys. Res.*, *104*, 16,349–16,369, 1999.

Russell, J. M., et al., The halogen occultation experiment, *J. Geophys. Res.*, *98*, 10,777–10,797, 1993.

Sandu, A., F. A. Potra, G. R. Carmichael, and V. Damian, Efficient implementation of fully implicit methods for atmospheric chemical kinetics, *J. Comput. Phys.*, *129*, 101–110, 1996.

Sandu, A., J. G. Verwer, J. G. Blom, E. J. Spee, G. R. Carmichael, and F. A. Potra, Benchmarking stiff ode solvers for atmospheric chemistry problems, – II, Rosembrock solvers, *Atmos. Environ.*, *31*, 3459–3472, 1997.

Smolarkiewicz, P. T., and P. J. Rash, Monotone advection on the sphere: An Eulerian versus semi–Lagrangian approach, *J. Atmos. Sci.*, *48*, 793–810, 1991.

Swinbank, R., and A. O'Neill, A stratosphere-troposphere data assimilation system, *Mon. Weather Rev.*, *122*, 686–702, 1994.

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

Taylor, J. R. *An Introduction to Error Analysis*, 1st ed., 270 pp., Oxford Univ. Press, New York, 1982.

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

Toon, O. B., C. P. Mckay, and T. P. Ackerman, Rapid calculation of radiative heating rates and photodissociation rates in homogeneous multiple scattering atmospheres, *J. Geophys. Res.*, *94*, 16,287–16,301, 1989.

Q. Errera and D. Fonteyn, Institut d'Aéronomie Spatiale de Belgique, 3 avenue circulaire, 1180 Brussels, Belgium. (quentin.errera@oma.be; dominiq.fonteyn@oma.be)