Sequential assimilation of stratospheric chemical observations in a three-dimensional model

M. P. Chipperfield
School of the Environment, University of Leeds, Leeds, UK

B. V. Khattatov
National Center for Atmospheric Research, Boulder, Colorado, USA

D. J. Lary
Department of Chemistry, University of Cambridge, Cambridge, UK

Received 18 January 2002; revised 14 June 2002; accepted 18 June 2002; published 12 November 2002.

[1] We describe a technique to assimilate chemical observations in a three-dimensional (3-D) chemical transport model (CTM). The method uses the established sequential technique of *Khattatov et al.* [2000], but here, it is applied simultaneously to many observed species. Following the assimilation, care is taken to preserve compact correlations between all modeled long-lived tracers and the total abundance of reactive families (e.g., inorganic chlorine). This way, the observations of long-lived tracers and family members constrain many other species in the model. In this paper, we apply the technique to the assimilation of O_3, CH_4, H_2O, and HCl from the Halogen Occultation Experiment (HALOE) in 1992. Despite the poor coverage of HALOE, the assimilation of species with long photochemical lifetimes is a useful global constraint on the model. Results of the assimilation model have been tested by comparison with Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS) profiles of O_3, CH_4, H_2O, HCl, and N_2O. Direct comparison of the assimilated species shows that the assimilation model performs better in reproducing the independent observations. Comparison of the nonassimilated species (N_2O) shows that assimilation has generally improved the comparison, especially in the midlatitude lower stratosphere. *INDEX TERMS:* 0340 Atmospheric Composition and Structure: Middle atmosphere—composition and chemistry; 0341 Atmospheric Composition and Structure: Middle atmosphere—constituent transport and chemistry (3334); 3334 Meteorology and Atmospheric Dynamics: Middle atmosphere dynamics (0341, 0342); 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; *KEYWORDS:* data assimilation, atmospheric chemistry, modeling, 3-D CTM

Citation: Chipperfield, M. P., B. V. Khattatov, and D. J. Lary, Sequential assimilation of stratospheric chemical observations in a three-dimensional model, *J. Geophys. Res.*, *107*(D21), 4585, doi:10.1029/2002JD002110, 2002.

1. Introduction

[2] The technique of data assimilation is used routinely in numerical weather prediction to create meteorological analyses. Over the past 5 years or so, there has been increasing interest in applying similar techniques to observations of chemical species in the atmosphere. The assimilation of such observations, and the creation of ''chemical analyses'' is expected to lead to better use of observations and to improvements in chemical models. The methods used for the assimilation of chemical observations can be divided into variational and sequential [e.g., *Khattatov et al.*, 1999].

[3] Variational chemical data assimilation was pioneered by *Fisher and Lary* [1995] who used a box model with a simplified photochemical scheme and assimilated observations along trajectories. Variational assimilation in a full chemistry model is computationally expensive, although it has now been applied in three dimensions for short simulations. *Errera and Fonteyn* [2001] used a variational method with a 3-D chemical transport model (CTM) to assimilate CRISTA data over 12 hour time windows during the course of a 6-day mission during 5–11 November 1994. They compared the assimilated fields with independent observations and so were able to comment on the systematic agreement between different instruments.

[4] *Khattatov et al.* [2000] gave a comprehensive discussion of the use of optimal interpolation and the Kalman filter in global chemical models. *Lyster et al.* [1997] described the first application of the (full) Kalman filter in global atmospheric chemistry. They assimilated Upper Atmosphere Research Satellite (UARS) observations of CH_4 (treated as an inert tracer) at a single potential temperature (θ) level. The simplicity of the problem considered

0148-0227/02/2002JD002110$09.00

permitted the use of the full Kalman filter though it appeared too computationally expensive for 3-D multispecies models. *Menard et al.* [2000] and *Menard and Chang* [2000] studied various approaches for the application of the Kalman filter for assimilating UARS data in a two-dimensional (2-D) model. In particular, they described a diagnostic process based on the χ-square method for adjusting free parameters. *Khattatov et al.* [2000] extended the methodology of *Menard et al.* [2000] and *Menard and Chang* [2000] to develop a sequential system based on the suboptimal Kalman filter.

[5] The scheme of *Khattatov et al.* [2000] has already been applied in several studies of atmospheric chemistry. An early variant of the scheme was used by *Levelt et al.* [1998] to assimilate UARS Microwave Limb Sounder (MLS) O_3 observations into a 3-D model with full stratospheric chemistry. After a 60-day simulation the model with ozone assimilation gave a better comparison with independent observations than the model without assimilation. *Lamarque et al.* [1999] assimilated satellite observations of CO into a global 3-D model with detailed tropospheric chemistry. Over the 10-day run the memory of the assimilation was limited to a few days (due to transport-induced drift) but other model species were significantly influenced. *Clerbaux et al.* [2001] also assimilated satellite CO observations (from the Interferometric Monitor for Greenhouse Gases (IMG)). They argued that data assimilation helped to highlight the difference between the model and observations and again showed that assimilated fields from the model gave good agreement with independent CO observations.

[6] In this study we combine a data assimilation scheme into our 3-D chemical transport model (CTM), SLIMCAT. This model has already been used extensively for studies of stratospheric chemistry over multiannual timescales [e.g., *Chipperfield*, 1999]. Therefore, we require an assimilation scheme which is efficient enough to allow multiannual integrations but will improve on the basic model by keeping the model constrained by observed chemical species.

[7] In this paper we have also used the sequential assimilation scheme of *Khattatov et al.* [2000]. We have extended on previous studies by using it to assimilate many (in this case 4) different chemical species simultaneously. Although the assimilation itself of each species is done independently, the nature of the species treated (e.g., long-lived tracers and members of chemical families) means that it is necessary to impose constraints on the assimilation to ensure self-consistency (and to ensure consistency with the nonassimilated fields). Also, by assimilating the observations into an established 3-D model, our aim is not to produce daily, global fields of assimilated species with as little model content as possible, but to use the observations to continually "nudge" the 3-D model toward reality.

[8] These consistency constraints raise the question of the philosophy of data assimilation. On one side, one could argue that observations should be assimilated into a model without any other constraint. In terms of chemical species one would then rely on the coupling of species though chemical reactions to ensure the nonassimilated species have realistic values. On the other hand, one could say that the assimilated fields from the model should not be allowed to violate some basic constraints which are well established in atmospheric chemistry. Some such constraints are easy to envisage, but the point at which one decides to limit these constraints is probably arbitrary. In general one can imagine that the 4D-Var assimilation of many simultaneous observations could be done in a model with no external constraints—there may be enough information in the observations to keep the important species limited to reasonable values. This is the procedure adopted in the 4D-Var work of Fonteyn and coworkers (D. Fonteyn, personal communication, 2001). In contrast, the long-term sequential assimilation of few species (as done here) imposes a greater need to apply certain constraints.

[9] It is worth noting here that the assimilation of ozone, which is often used as a test case because of abundant observations, is the most straightforward. As ozone contains only O atoms and is formed from O_2, which is present in a huge abundance, the atmospheric concentration of ozone can vary without any physical constraint. During assimilation there is no concern about conservation of O atoms or of limits imposed by the abundance of other species.

[10] Section 2 describes our 3-D model, the assimilation scheme and the experiments performed. Section 3 summarizes the HALOE data used in the assimilation. In section 4 we discuss our method for assimilating the observations and, in particular, how we impose constraints on nonassimilated species. In section 5 we discuss how the assimilation has improved the performance of the 3-D model for selected test cases. Our discussion and conclusions are given in section 6.

2. Model and Experiments

2.1. SLIMCAT 3-D CTM

[11] We have used the SLIMCAT off-line 3-D CTM which is described in detail by *Chipperfield* [1999]. Horizontal winds and temperatures are specified using meteorological analyses. Vertical advection is calculated from heating rates using the MIDRAD radiation scheme [*Shine*, 1987] and chemical tracers are advected by conservation of second-order moments [*Prather*, 1986]. The model has the most important species in the O_x, NO_y, Cl_y, Br_y, HO_x families along with a CH_4 oxidation scheme and long-lived tracers. The model has a detailed gas-phase stratospheric chemistry scheme as well as a treatment of heterogeneous chemistry of liquid and solid aerosols (for more information see *Chipperfield* [1999]). The version used here uses photochemical data from *DeMore et al.* [1997]. The model has been widely used in previous studies of stratospheric chemistry [e.g., *Chipperfield et al.*, 1996; *Chipperfield and Jones*, 1999].

2.2. Assimilation Scheme

[12] The sequential assimilation scheme used in this study is that of *Khattatov et al.* [2000]. Details of the scheme are described in Appendix A. For each of the assimilated species (in this case the 4 HALOE species O_3, CH_4, H_2O and HCl) an extra tracer was added to the model to advect the forecast error. The values of the scheme's tunable parameters used in this study are given in section 4.2.

2.3. Model Runs

[13] In the experiments discussed here the CTM was run with a horizontal resolution of $7.5^\circ \times 7.5^\circ$ (T15 Gaussian grid) and 18 isentropic levels from 330 K to 3000 K

Table 1. 3-D Model Assimilation Runs

Model run	Dates	Species Assimilated	Notes
CON	1 January 1992–30 June 1992	None	
HAL	1 January 1992–31 January 1992	O_3, CH_4, H_2O, HCl	
HALC	1 January 1992–30 June 1992	O_3, CH_4, H_2O, HCl	Constraints

(approximately 10 to 55 km). We have used 24-hourly United Kingdom Meteorological Office (UKMO) analyses [*Swinbank and O'Neill*, 1994] to force the model. A basic model run was initialized in October 1991 from a 2-D model and integrated until 31 December 1991. This output was then used to initialize a series of 3 assimilation runs which are summarized in Table 1. Run **CON** is a control run without assimilation. Run **HAL** is like **CON** but includes assimilation of HALOE O_3, CH_4, H_2O, and HCl. Run **HALC** also assimilates these species but includes constraints on non-assimilated species as discussed in section 4.1.

3. HALOE Data

[14] The Halogen Occultation Experiment (HALOE) [*Russell et al.*, 1993] provides solar occultation observations of a range of trace gases including O_3, CH_4, H_2O and HCl which are used in this study. Although this technique gives relatively poor coverage on any day (only measurements at a sunrise and at a sunset latitude—see Figure 1) these data are still useful for our assimilation purpose. This is because the species assimilated have long photochemical lifetimes, at least in some regions of the stratosphere. HALOE also measures NO and NO_2. However, the short photochemical lifetime of these species makes their assimilation in a sequential system meaningless.

[15] In this paper we have used version 19 data for O_3 [*Bruehl et al.*, 1996], HCl, [*Russell et al.*, 1996], H_2O, [*Harries et al.*, 1996], and CH_4 [*Park et al.*, 1996]. The estimated accuracies for these species for different altitudes were taken from these validation papers and are listed in Table 2.

4. Assimilation Methodology

[16] In this section we describe how we have implemented the sequential assimilation scheme in our 3-D CTM.

4.1. Assimilation Constraints

[17] Our first attempts to assimilate HALOE HCl directly into the 3-D CTM resulted in the modeled total inorganic chlorine (Cl_y) rapidly becoming unrealistic. The modeled values of Cl_y exceeded 4 ppbv, when we know from atmospheric chlorofluorocarbon (CFC) abundances that there is a limit of Cl_y of about 3.6 ppbv for 1990s conditions. Indeed, it is relatively straightforward for an atmospheric model to predict Cl_y based on known CFC abundances. In our model (which is used for long trend simulations, for example) the incorrect Cl_y would be very undesirable. Therefore, we need to impose a limit on the model Cl_y (and other inorganic chlorine species) when we assimilate HCl (see below for details).

[18] Another relationship between certain atmospheric species concerns the correlations that are observed between pairs of long-lived tracers [*Fahey et al.*, 1989; *Plumb and*

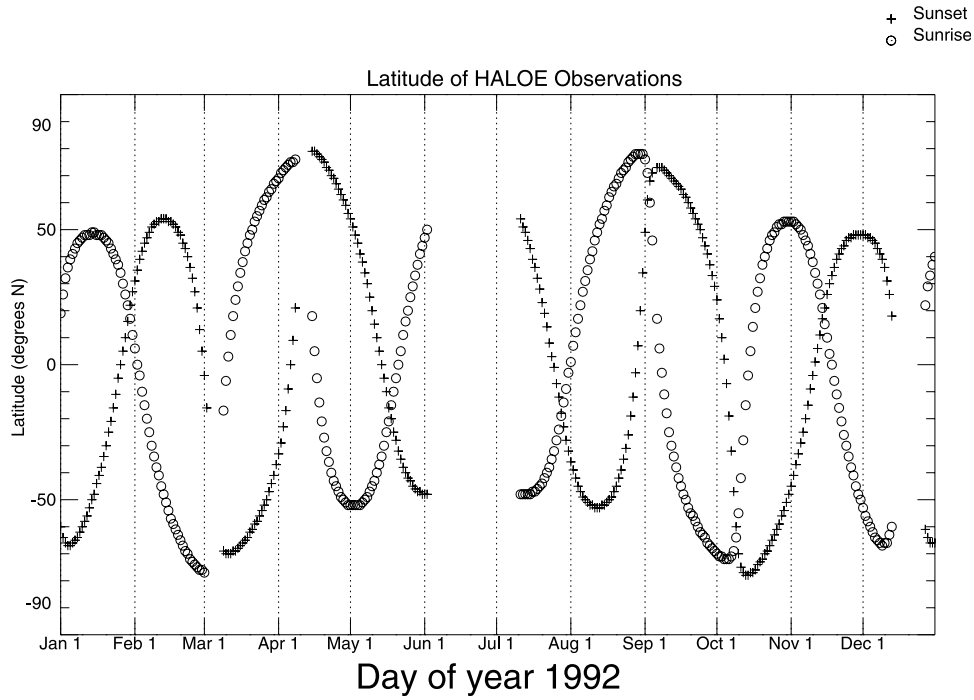

Figure 1. Coverage of HALOE sunrise and sunset observations for 1992. (Figure adapted from HALOE home webpage.)

Table 2. Accuracies (Fractional) for HALOE Data

Pressure, hPa	O_3	CH_4	H_2O	HCl
0.464	0.06	0.06	0.18	0.31
0.681	0.07	0.06	0.16	0.23
1.00	0.08	0.06	0.14	0.15
1.47	0.08	0.07	0.14	0.15
2.15	0.09	0.08	0.14	0.14
3.16	0.09	0.08	0.14	0.14
4.64	0.09	0.09	0.14	0.12
6.81	0.11	0.10	0.16	0.13
10.0	0.12	0.11	0.17	0.14
14.7	0.14	0.12	0.19	0.15
21.5	0.15	0.13	0.21	0.16
31.6	0.17	0.14	0.23	0.17
46.4	0.20	0.16	0.24	0.20
68.1	0.25	0.17	0.26	0.22
100	0.30	0.19	0.27	0.24

Ko, 1992]. Measurements of one long-lived tracer can therefore be used to derive other long-lived tracers. Figure 2 shows the model correlations of certain tracers from the model run without assimilation (run **CON**). In general the model displays the expected compact correlations for the long-lived tracers. In the case of CH_4 versus N_2O the model agrees well with the canonical straight-line fit estimated from ER-2 aircraft data [see *Kawa et al.*, 1993; *Waugh et al.*, 1997]. For NO_y versus N_2O, although the model displays a relatively compact correlation this version overestimates the observed NO_y abundance (dashed line in Figure 2c). For $CFCl_3$ versus N_2O the model curve is compact for each θ level, although the model produces a correlation which varies with altitude. This separation with θ in the model is probably related to a too slow vertical motion in the model. Throughout most of the stratosphere O_3 is not long-lived and overall the correlation plot with N_2O is not compact—it is included here for comparison.

[19] If HALOE CH_4 observations, for example, are assimilated into the model then we would expect compact correlations involving this tracer to break down. The dramatic extent to which this occurs is shown in Figure 3. The expected compact correlation has been replaced by one where the value of CH_4 for, say, 150 ppbv N_2O varies from 0.5 to 1.4 ppmv. Similar results are obtained for other correlations between CH_4 and nonassimilated long-lived tracers (e.g., CFCs, Cl_y, etc., not shown). Although, the

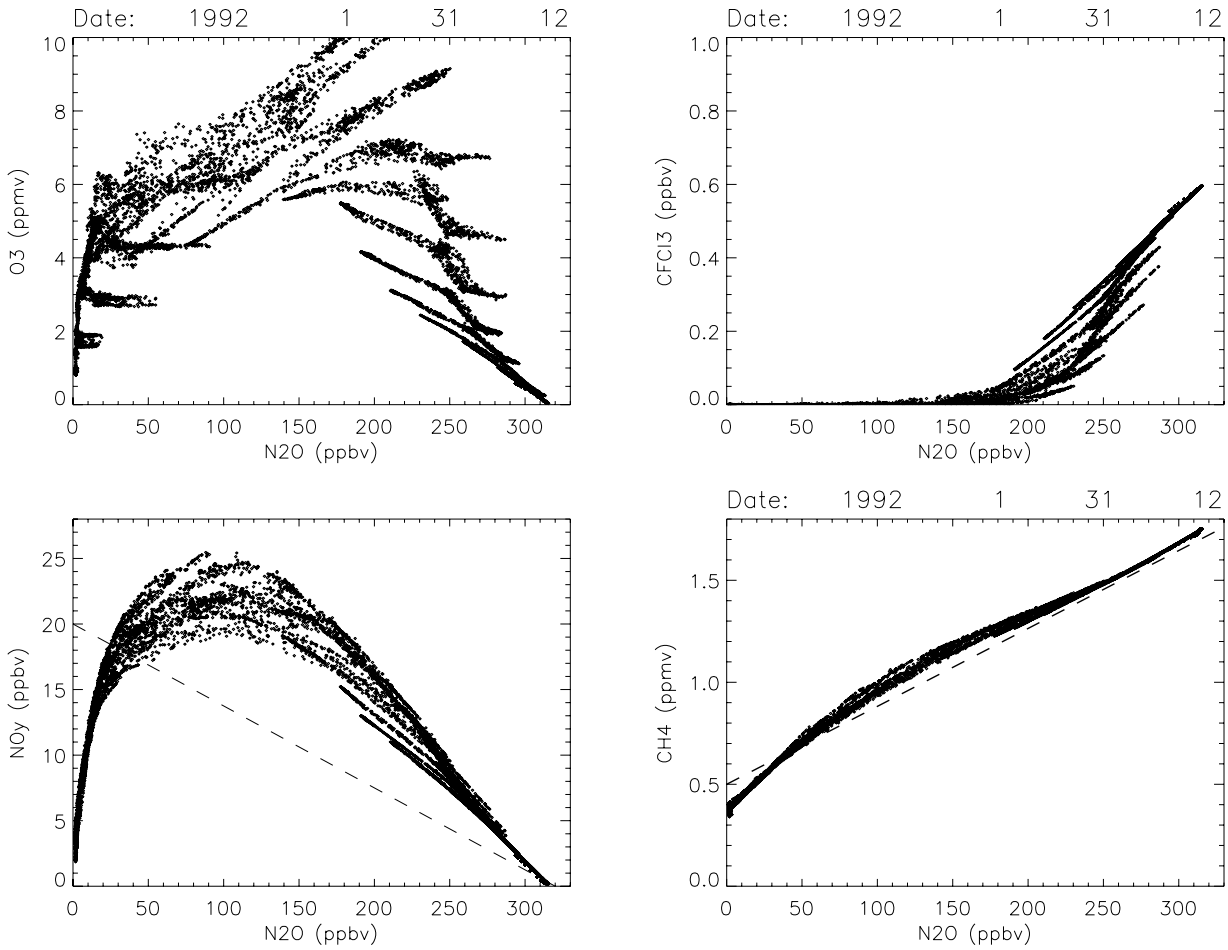

Figure 2. Correlation plots of tracers on 31 January 1992 from model run **CON** (without assimilation) for (a) O_3 versus N_2O, (b) $CFCl_3$ versus N_2O, (c) NO_y versus N_2O, and (d) CH_4 versus N_2O. The dashed line in panel (c) shows the fit NO_y(ppbv) = 20.0 − 0.0625*N_2O(ppbv) based on midlatitude balloon profiles and ER-2 data [see *Kondo et al.*, 1996]. The dashed line in panel (d) shows the fit N_2O(ppbv) = 262*CH_4(ppmv) − 131 from ER-2 data [see *Kawa et al.*, 1993].

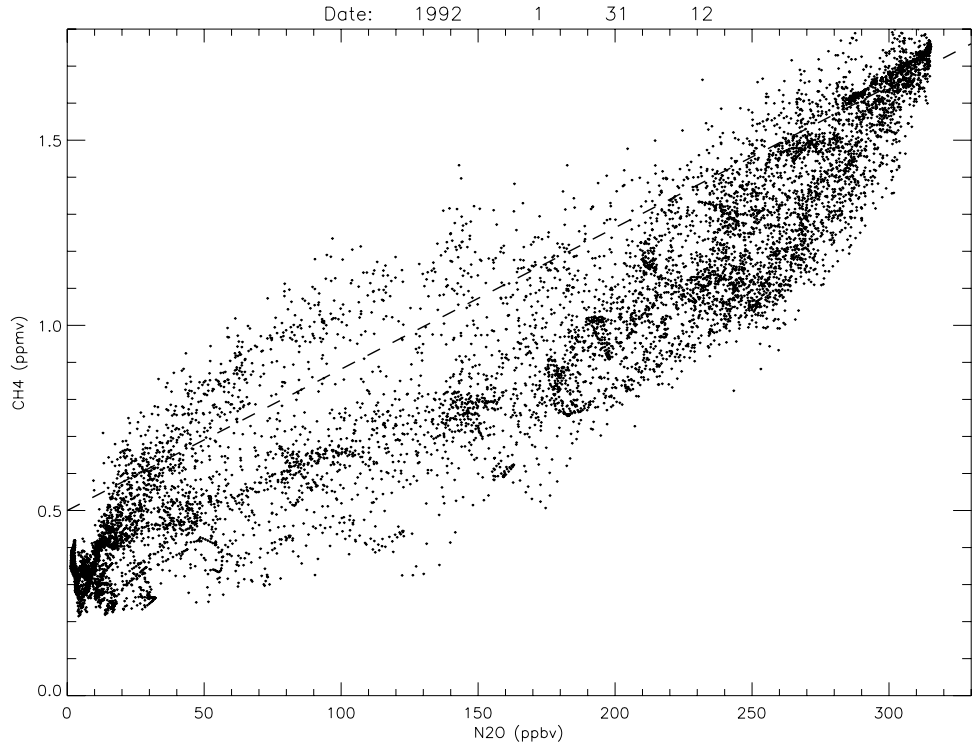

Figure 3. As Figure 2d but for run **HAL**.

assimilation of CH_4 will have resulted in a better modeled distribution of CH_4 (see section 5) the overall model performance will have degraded.

[20] Within the philosophy of data assimilation in this study (i.e., to constrain multiannual simulations of an established 3-D model), it is desirable both to assimilate observations and to maintain the correlations with nonassimilated long-lived tracers. One method would be to use a fit to the CH_4 versus N_2O correlation based on atmospheric observations (e.g., dashed line in Figure 2d) to infer a pseudo-observed N_2O. This would, in effect, be mixing the HALOE data with other (e.g., ER-2) observations. An alternative approach would be to use the model-predicted correlations to derive a new N_2O value for each assimilated CH_4 point. We have chosen this latter approach so that tracer correlations remain determined by the model chemistry (i.e., so that other observations can still be used to test the model's N_2O distribution).

[21] In run **HALC** we have added the following procedure to adjust the model long-lived species following the assimilation of HALOE CH_4 and HCl. (The assimilation of O_3 and H_2O are not used in the following correction, although if H_2O was not available from HALOE, the modeled field would also need to be corrected).

1. Determine the model gridboxes in which assimilation has changed the concentration of CH_4 by more than 0.1%. For these gridboxes steps 2–6 are then applied.

2. Average the global model fields of long-lived tracers ($CFCl_3$, CF_2Cl_2, N_2O and NO_y) before assimilation into bins of different CH_4 values. This averaging is done for each θ level i by using the levels $i - 1$, i and $i + 1$.

3. Use postassimilation value of CH_4, and the average model correlations deduced in (2), to find a new value for the nonassimilated long-lived tracers.

4. Based on the postassimilation values of $CFCl_3$ and CF_2Cl_2 (the two Cl-containing source gases in the model) derive a new distribution of Cl_y.

5. Repartition Cl_y species based on postassimilation values of Cl_y and HCl and preassimilation ratios of other species.

6. Repartition NO_y species based on postassimilation values of NO_y and preassimilation ratios of other species.

[22] An important consideration in these constraints (step (1)) is that in the absence of observations they do not modify the model fields. This is a desirable property of any assimilation scheme. Note that we apply the correlations with CH_4 as a strong constraint. In our case, as CH_4 is the only observed tracer used, this is reasonable. However, if two (or more) long-lived tracers were available we would not necessarily expect the inferred values of a nonobserved species from each of these observations to be consistent. In this case it would be more appropriate to apply a weak constraint.

[23] The result of including these constraints is shown in Figure 4. The CH_4 versus N_2O correlations is now more compact than Figure 3 and similar to the basic model in Figure 2; evidently for any gridbox in which CH_4 is changed by assimilation, the scheme adjusts N_2O to maintain the model-predicted correlation. The NO_y versus N_2O and $CFCl_3$ versus N_2O correlations are also similar to the basic model. However, these correlations show more curvature which has been slightly smoothed by the assimilation/correction procedure. Overall, the corrections applied to the long-lived tracers has resulted in a much better behaved model.

4.2. Assimilation Parameters

[24] The assimilation scheme contains a number of tunable parameters. The values for these were chosen using χ^2

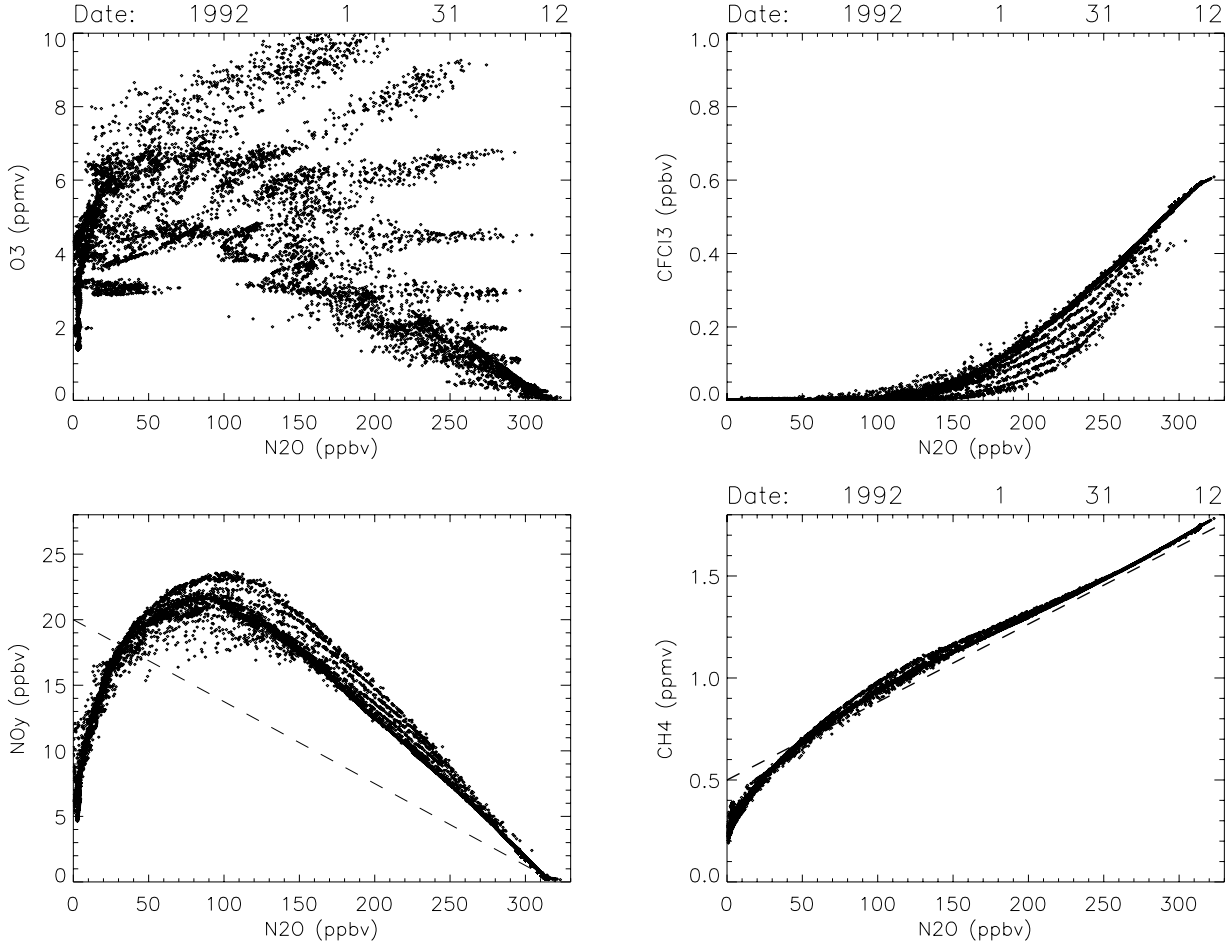

Figure 4. As Figure 2 but for run **HALC**.

diagnostics and OmF (observations − forecast) differences (as described by *Khattatov et al.* [2000]) in a series of 1-month test runs. The values chosen are given in Table 3, along with the values used by *Khattatov et al.* [2000] for assimilation of MLS O_3 data for comparison. The values of the normalized χ^2 diagnostics are shown in Figure 5 for the first 3 months of run **HALC**. As discussed by *Khattatov et al.* [2000], this diagnostic should ideally produce values near 1 and we have adjusted the parameters in Table 3 to achieve this. In run **HALC** we have used the published HALOE accuracies (given in Table 2) except for H_2O. When we used the published errors the χ^2 diagnostic (Figure 5b) gave values much smaller than 0.5. Therefore, we have reduced the H_2O errors used in the model to 30% of the published values to increase χ^2/N following the discussion given by *Khattatov et al.* [2000]. Even with this reduction in the error, the χ^2 diagnostic for H_2O is only around 0.6.

[25] For chemically inert tracers (no production or loss) the error growth rates (ϵ) should be identical for all species. In this case ϵ is simply an indicator of how good are the dynamical field and model numerics. These rates should be numerically different at different latitudes and altitudes. In our approach, we are trying to find a single number for ϵ that fits all geographical regions, which is an approximation. Therefore, for species which are not chemically inert and which depend on different chemistry in different regions the values of ϵ can differ, which is the case for the species given in Table 2. The value of ϵ for O_3 derived in this study is different from that derived by *Khattatov et al.* [2000], which could be due to different model dynamics and numerics.

4.3. Error Tracers

[26] Figure 6 shows the zonal mean cross-section of the normalized model error tracers for the 4 assimilated species on 31 January 1992. At this time the HALOE observations occur at latitudes around 20°N (see Figure 1). Consequently, this latitude corresponds to a minimum in the species errors. Ideally, for chemically inert tracers with the same observational errors these plots should show a similar pattern for all species. The differences between the

Table 3. Adjustable Parameters for Assimilation Scheme

	For MLS[a]	For HALOE			
Parameter	O_3	O_3	H_2O	CH_4	HCl
Error growth ϵ/hour	0.0135	0.005	0.001	0.01	0.005
Rel. Rep. error r	0.1	0.1	0.1	0.1	0.1

L_{xy} = 1000 km, L_z = 0.4 scale height. See Appendix A.
[a] *Khattatov et al.* [2000].

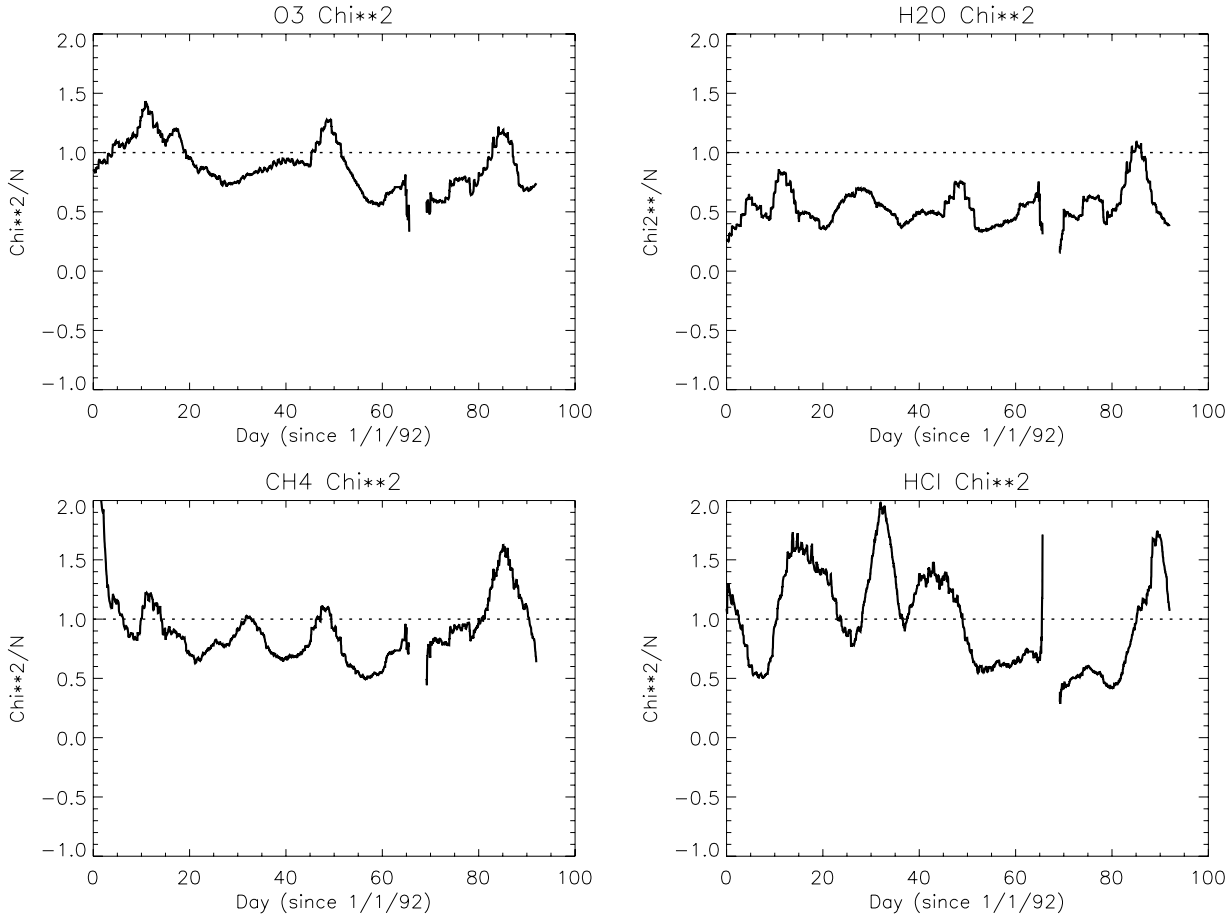

Figure 5. Plot of χ^2/N as a function of day for 1 January 1992 to 31 March 1992 from run **HALC** for (a) O_3, (b) H_2O, (c) CH_4 and (d) HCl. The plots show hourly output smoothed with a 100-hour running mean. The gap in early March indicates a 6-day period when no HALOE data was available.

species plots are due to the different photochemical lifetimes of the species in different regions and to the different instrumental errors for different species at different altitudes (see Table 2).

5. Results

[27] We now compare results from the model runs with and without assimilation to examine the benefit of assimilation on some aspects of the 3-D model performance.

5.1. Long-Lived Tracers

[28] Figure 7 shows the zonal mean cross-section of CH_4, H_2O and $2CH_4 + H_2O$ on 31 January 1992 from runs **CON** and **HALC**. The most evident effect of the assimilation on CH_4 is an increase in the tracer gradient in the subtropics, driven mainly by an increased descent in the midlatitudes and also more ascent in the tropics above about 5 hPa. A similar change is seen in the H_2O distribution. In the basic model CH_4 and H_2O are related by the bottom boundary condition $2CH_4 + H_2O = 7$ ppmv and the model CH_4 oxidation chemistry which is assumed to yield $2H_2O$ per CH_4 oxidized. This conservation can be seen in Figure 7e except in the lower stratosphere due to remnants of Antarctic dehydration. The sum from the HALOE data shows more fine structure and it varies between 6 and 7 ppmv. (Note that the contour interval in Figure 7f is only 0.1 ppmv, while a 10% error in a 7 ppmv quantity in can lead to a 1 ppmv variation).

5.2. Chlorine Species

[29] Figure 8 shows the zonal mean cross-section of HCl and total inorganic chlorine (Cl_y) on 31 January 1992 from runs **CON** and **HALC**. The constraints imposed on the modeled long-lived chlorine source gases $CFCl_3$ and CF_2Cl_2 (via their correlation with methane), and the subsequent balance of total chlorine between organic and inorganic forms, has produced a Cl_y distribution in run **HALC** which resembles the assimilated CH_4 (i.e., stronger descent at midlatitudes and stronger subtropical gradients). This additional Cl_y in the mid–high latitude lower stratosphere is mostly partitioned into HCl.

5.3. Ozone

[30] Figure 9 shows the zonal mean cross-section of O_3 on 31 January 1992 from runs **CON** and **HALC**. The assimilated model shows less elongated tracer isopleths in the mid stratosphere.

5.4. Comparison With ATMOS Data

[31] Figure 10 compares modeled profiles from runs **CON** and **HALC** with observations from the Atmospheric

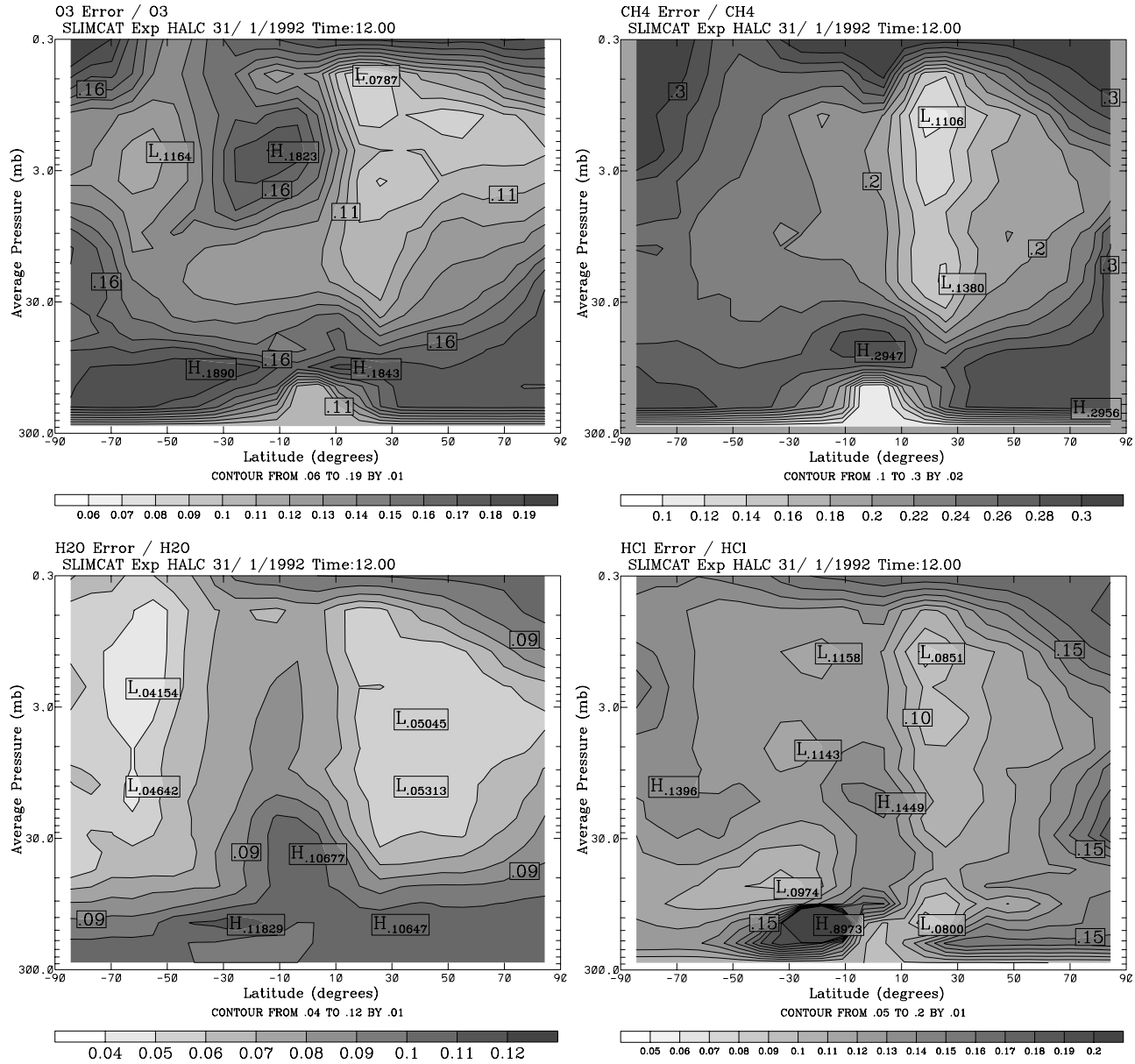

Figure 6. Zonal mean distributions of error tracers (normalized by the respective chemical tracers) for (a) O_3, (b) CH_4, (c) H_2O and (d) HCl. See color version of this figure at back of this issue.

Trace Molecule Spectroscopy Experiment (ATMOS) [*Gunson et al.*, 1996] during late March 1992. These observations cover the latitude range 17.5°N to 51.6°S and the model profiles from the nearest output time have been interpolated to the same location. For CH_4 in the southern midlatitudes (38.8°S to 51.6°S) the assimilated run (**HALC**) is more realistic, especially in the lower stratosphere. In **HALC** the isopleths of CH_4 show more descent (see Figure 7), in better agreement with the ATMOS observations. In the tropics and subtropics (7.8°S to 17.5°N) there is less difference between the two model runs and both simulations give a similar comparison with the observations.

[32] The assimilation procedure has modified the modeled N_2O distribution which is compared with ATMOS measurements in Figure 10. Again, in the southern hemisphere lower stratosphere the increased descent in the N_2O profiles gives better agreement with the ATMOS data. In the upper stratosphere, however, the assimilation run **HALC** sometimes does not simulate ATMOS N_2O well, although the CH_4 agreement is very good. Evidently in model run **HALC** the CH_4:N_2O correlation deviates from the ATMOS observations. This is illustrated in Figure 11a which shows ATMOS observations with relatively large values of N_2O (for 1 ppmv CH_4) compared to the model runs and the canonical fit from ER-2 data.

[33] The assimilation model shows less H_2O throughout the profile than model run. Based on the profile comparisons it is not possible to state that either model gives better agreement with the ATMOS data. It is clear that the ATMOS profiles show a lot of vertical structure which is not captured by the model. (The vertical resolution of the

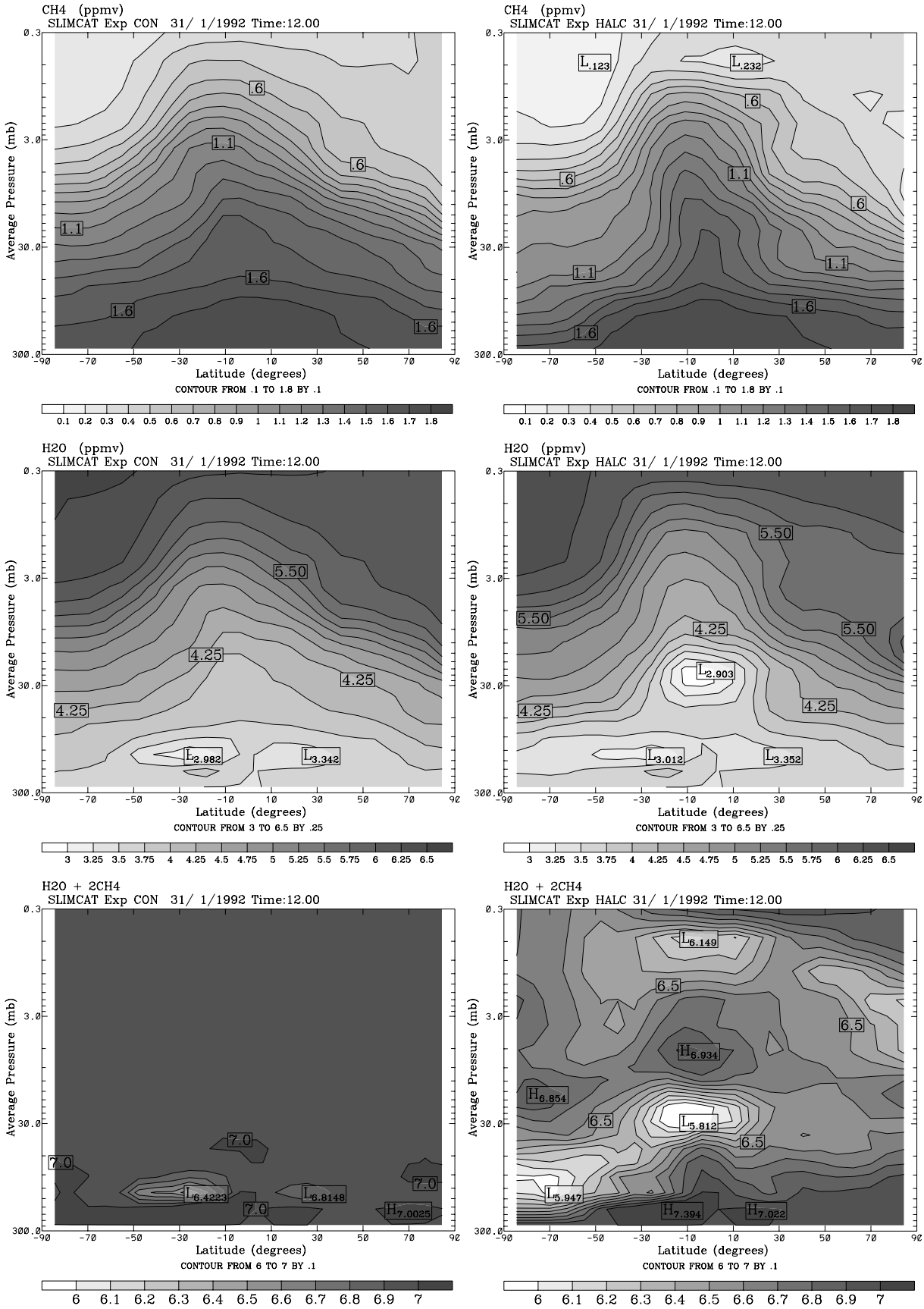

Figure 7. Zonal mean distributions of CH_4, H_2O, and $2CH_4 + H_2O$ (ppmv) on 31 January 1992 for run **CON** (left, no assimilation) and run **HALC** (right, with HALOE assimilation). See color version of this figure at back of this issue.

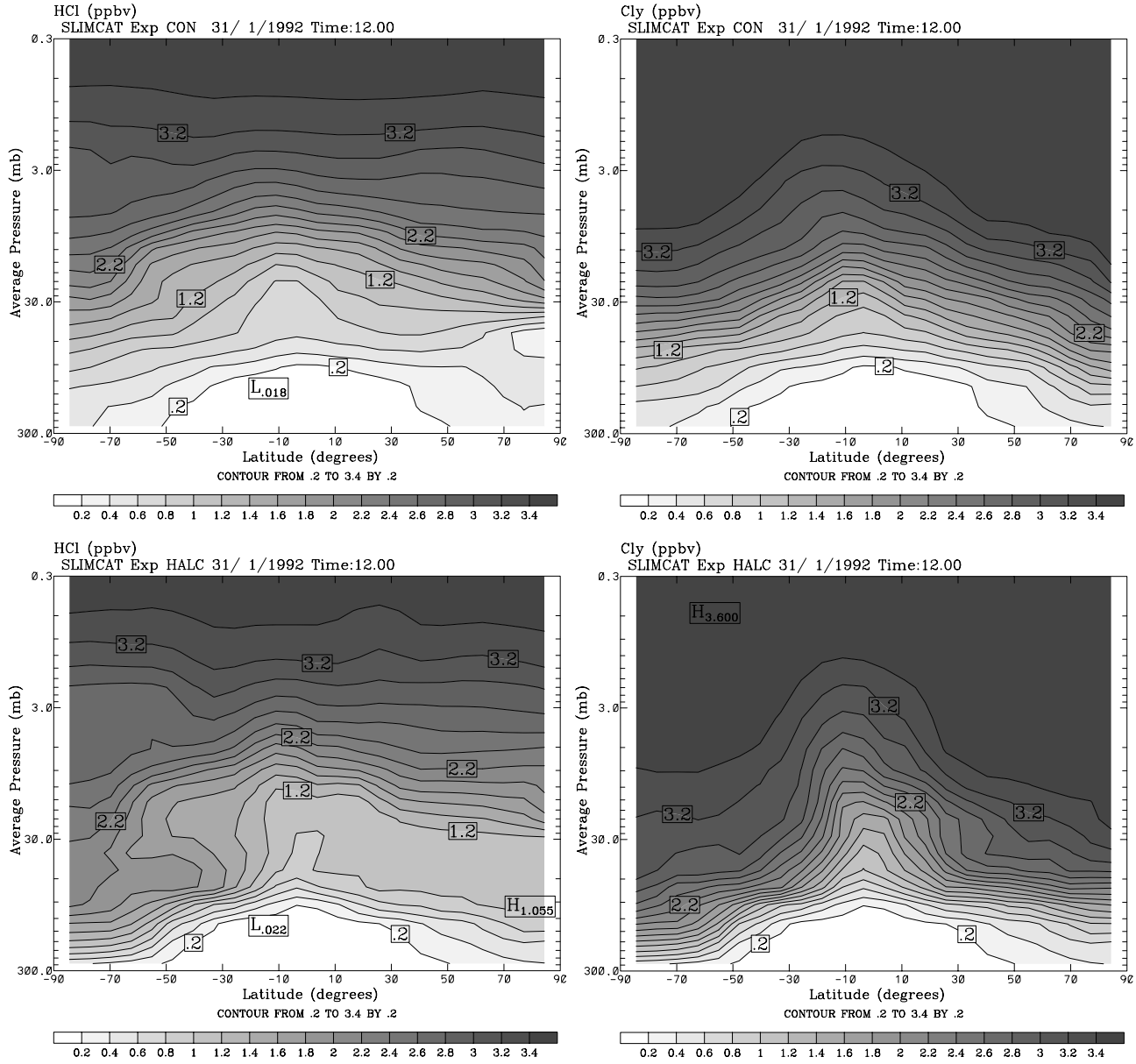

Figure 8. Zonal mean distributions of HCl and Cl_y (ppbv) on 31 January 1992 for runs (a) **CON** (above), and (b) **HALC** (below). See color version of this figure at back of this issue.

assimilation could be increased by decreasing the parameter L_z). Figure 11b shows the CH_4:H_2O correlation plot for the profiles shown in Figure 10. There is considerable scatter in the ATMOS data which precludes a critical test of the two model runs.

[34] For HCl the assimilation model shows increases in the lower stratosphere (due to increased Cl_y correlating with decreased CH_4) and some decreases at higher altitudes compared to run **CON**. This has improved the comparison slightly in the upper stratosphere. For the two ATMOS HCl profiles which extend to the lower stratosphere (38°S, 47°S) the assimilation model shows better agreement.

[35] For O_3 the basic model (run **CON**) already shows reasonable agreement with the observations. There is not much change in the assimilation run **HALC**. However, at 47°S there is an improvement in the detail of the profile: there is a small decrease in the maxima around 10 hPa and an increase in O_3 above this altitude.

6. Summary

[36] We have described a technique to assimilate chemical observations in the SLIMCAT three-dimensional (3-D) chemical transport model (CTM). The model has a detailed description of stratospheric chemistry and has been widely used and tested in past studies. We have used an established sequential assimilation scheme [*Khattatov et al.*, 2000] and extended it to apply it simultaneously to many observed species.

[37] A major improvement of our method is that following the assimilation step the scheme ensures the consistency between the assimilated species and between the assimilated and nonassimilated species. The consistency is imposed by

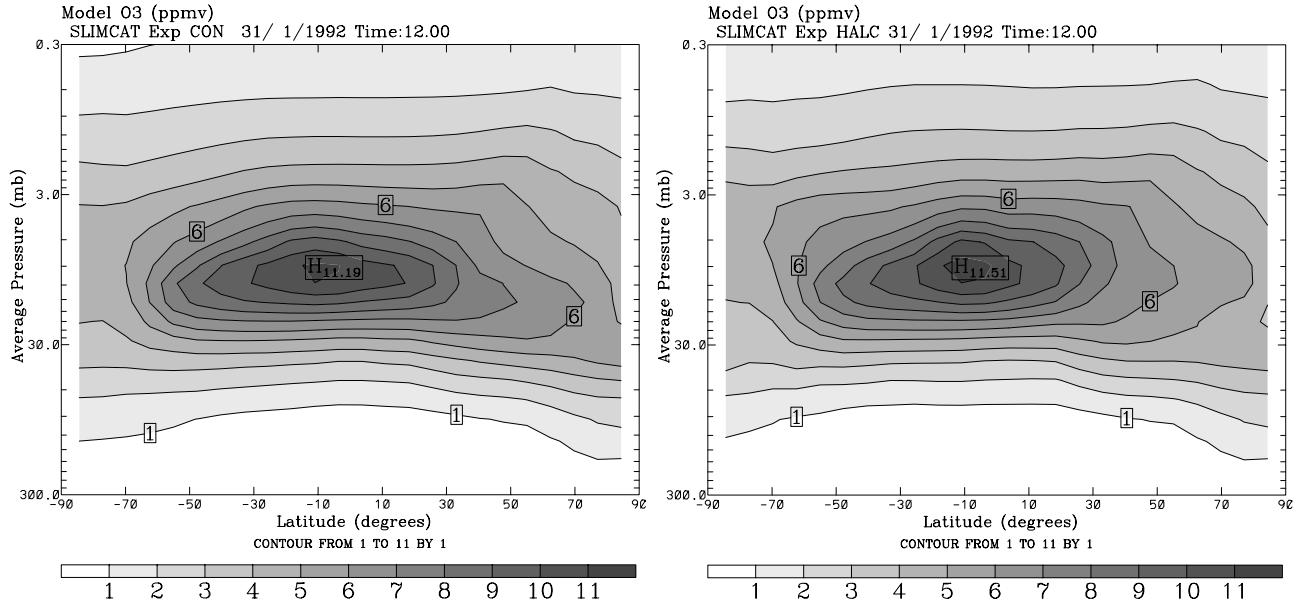

Figure 9. Zonal mean distributions of O_3 (ppmv) on 31 January 1992 for runs (a) **CON** (left), and (b) **HALC** (right). See color version of this figure at back of this issue.

using the compact correlations between long-lived tracers and the total abundance of chemical families. In this study we have applied the correlations using a single observed long-lived tracer (CH_4) as a strong constraint.

[38] In this study we have used the CTM to assimilate HALOE occultation observations of O_3, CH_4, H_2O and HCl. Even though the coverage of the occultation observations is limited on any day, where the photochemical lifetime of any observed species is long, the assimilation still provides a useful constraint on the model. This is because we are using the observations to perform slight adjustments to a realistic model, rather than requiring the assimilation to change the model fields drastically.

[39] A number of model simulations have been performed for early 1992. The assimilation of HALOE data has improved the model overall when compared with independent ATMOS observations for both assimilated and nonassimilated species. As well as generally better comparison of absolute magnitudes, the assimilated model shows more realistic tracer gradients in the subtropical lower stratosphere.

[40] The assimilation method described here is computationally cheap. The assimilation of the HALOE data (around 15 profiles per day) adds only a minor overhead to the full chemistry model. Therefore, it can be used in multiannual simulations to further improve our ability to model long-term changes.

Appendix A: Assimilation Scheme

[41] Our assimilation scheme is an efficient sequential assimilation scheme with estimate of analysis errors, based on *Khattatov et al.* [2000], and is described here.

A.1. Method

[42] The integration of model **M** gives concentrations of species **x** at a new time:

$$\mathbf{x_{t+\Delta t}} = \mathbf{M}(\mathbf{t}, \mathbf{x_t})$$

The observations **y** are generally available on a different grid which is related to **x** by linear operators **I** (horizontal interpolation) and **A** (“averaging kernel”).

$$\mathbf{y} = \mathbf{A}(\mathbf{I}(\mathbf{x}))$$

We can define an “observational operator” **H** such that:

$$\mathbf{y} = \mathbf{H}((\mathbf{x})) \tag{1}$$

The solution to (1) is:

$$\mathbf{x_t^a} = \mathbf{x_t} + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x_t})$$

The Kalman gain matrix **K** is given by:

$$\mathbf{K} = \mathbf{B}\mathbf{H}^\mathbf{T}\left(\mathbf{H}\mathbf{B_t}\mathbf{H}^\mathbf{T} + \mathbf{O} + \mathbf{R}\right)^{-1}$$

where, $\mathbf{B}_t$ is the forecast error covariance, **O** is the observation error covariance, and **R** is representativeness error covariance (errors of interpolation and discretization).

[43] The analysis error covariance is:

$$\mathbf{B_t^a} = \mathbf{B_t} - \mathbf{B_t}\mathbf{H}^\mathbf{T}\left(\mathbf{H}\mathbf{B_t}\mathbf{H}^\mathbf{T} + \mathbf{O} + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{B_t} \tag{2}$$

A.2. Treatment of Errors

[44] In the extended Kalman filter method, the evolution of error covariance is obtained using a linearization **L** of the original model **M**.

$$\mathbf{B_{t+\Delta t}} = \mathbf{L}\mathbf{B_t^a}\mathbf{L}^\mathbf{T} + \mathbf{Q} \tag{3}$$

where

$$\mathbf{L} = \frac{\mathbf{dx_{t+\Delta t}}}{\mathbf{dx_t}} \tag{4}$$

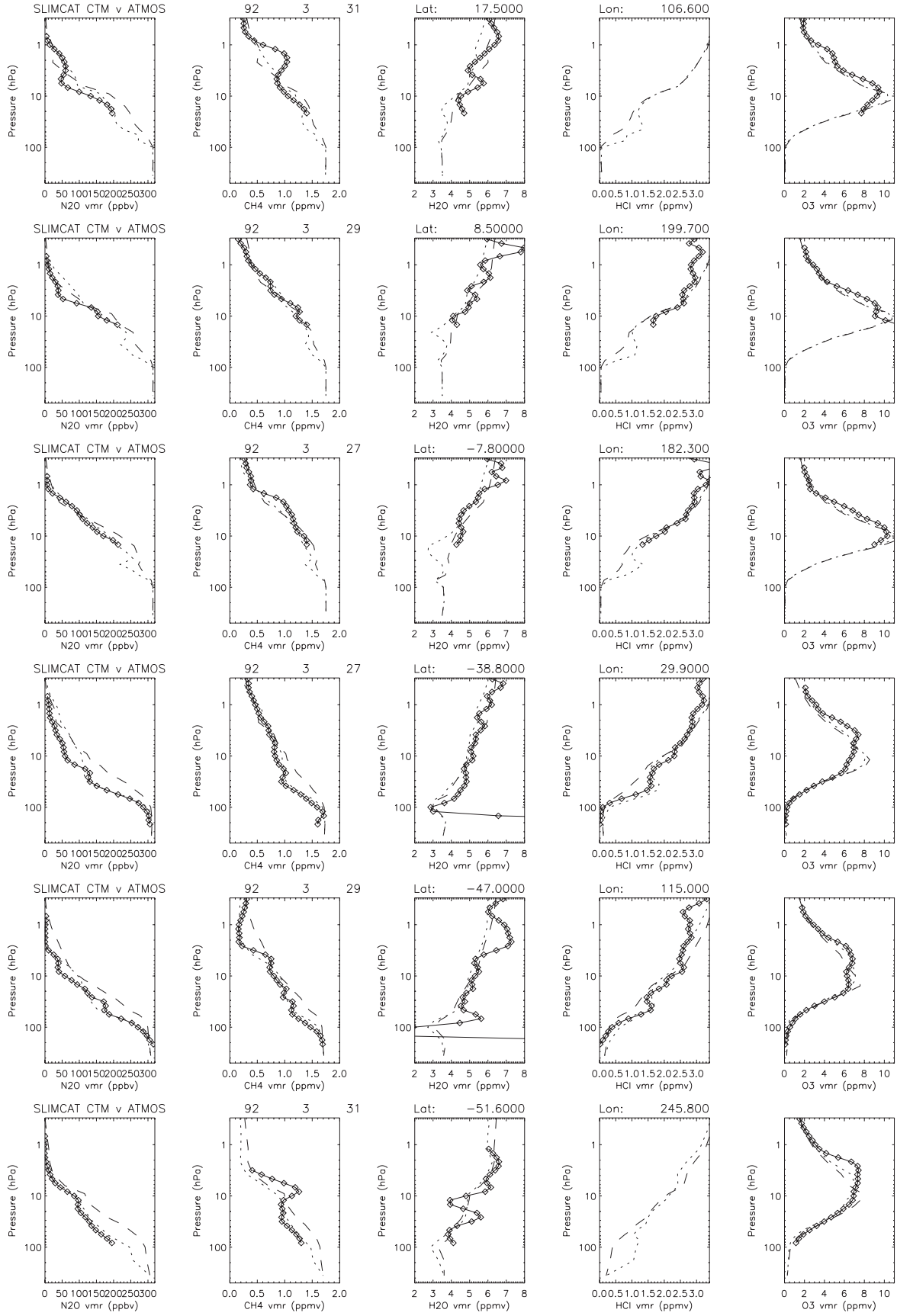

Figure 10. Comparison of version 3 ATMOS profiles of (a) N_2O, (b) CH_4, (c) H_2O (d) HCl and (e) O_3 with run **CON** (dashed line) and **HALC** (dotted line) for 6 profiles in late March 1992 between 17.5°N and 51.6°S.

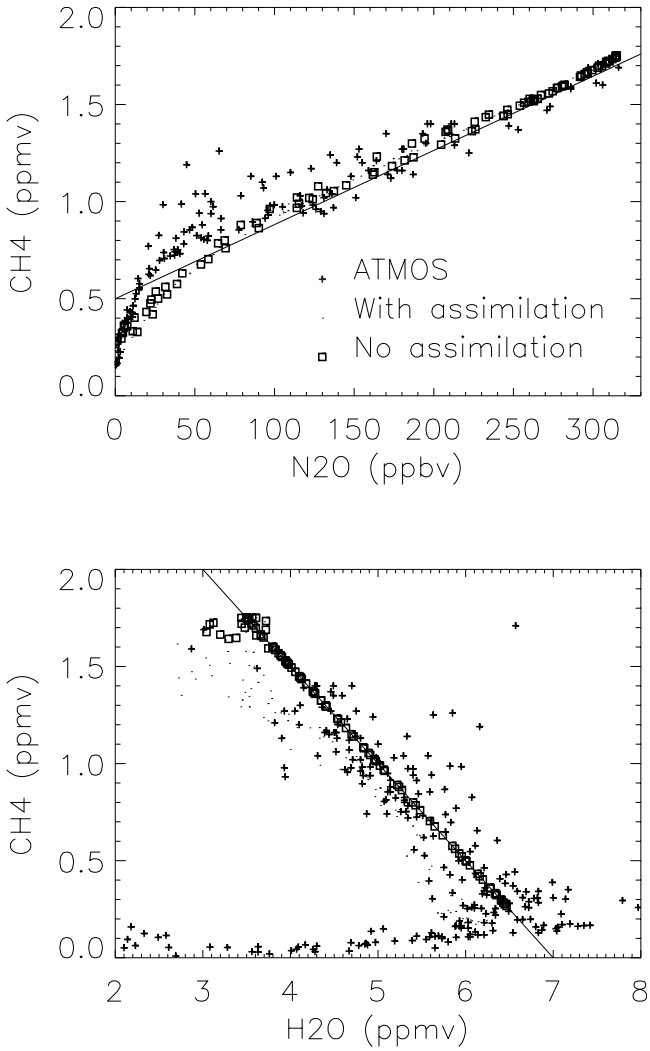

Figure 11. Correlation of (a) CH_4 versus N_2O and (b) CH_4 versus H_2O from the ATMOS profiles shown in Figure 10. Also shown are the model results from Figure 10 (**CON**—dots; **HALC**—squares). The solid line in panel (a) shows the fit $N_2O(ppbv) = 261.8*CH_4(ppmv) - 130.9$. The solid line in panel (b) shows the line $H_2O = 7.0 - 2CH_4$, which is the boundary condition used in the basic 3-D model (e.g., run **CON**).

[45] The sizes of **x**, **B**, **L** can be large which makes a 3-D analysis impossible. Therefore, in the 3-D scheme some simplifications are employed. The off-diagonal elements of **B** are obtained from:

$$b_{ij} = \sqrt{b_{ii}b_{jj}}\exp\left(-\frac{\Delta r_{xy}^2}{2L_{xy}^2}\right)\exp\left(-\frac{\Delta r_z^2}{2L_z^2}\right)$$

where Δr_{xy} and Δr_z represent horizontal and vertical distances between locations i and j.

[46] The time evolution of error covariance is parameterized as:

$$b_{ii}(t+\Delta t) = \tilde{b}_{ii}(t+\Delta t) + q_{ii}(t) \quad (5)$$

$$\tilde{b}_{ii}(t+\Delta t) = \mathbf{M}(b_{ii}(t)) \quad (6)$$

$$q_{ii}(t) = (\epsilon x_i(t+\Delta t)\Delta t)^2 \quad (7)$$

where ϵ is a tunable parameter.

[47] The observational error covariance **O** is assumed to be diagonal and for our HALOE assimilation the elements are set equal to the estimated absolute error at each pressure altitude (see section 4.2).

[48] The representativeness error covariance matrix **R** is assumed to be diagonal, with elements computed as:

$$r_{ii} = (ry_i)^2$$

where r is the relative representativeness error.

[49] The analysis error variance is computed directly from equation (2). Since clearly it impossible to implement matrix operations implied by (2) directly, only the diagonal elements of the covariance matrix (i.e., variances) are computed. The $\mathbf{HB_tH^T}$ term represents the background error covariance interpolated to the locations of observations as discussed by *Menard et al.* [2000] and thus can be computed easily. The matrix inversion is performed directly since the size of the matrices in observation space is fairly small. Once this is done, each line of $\mathbf{B_tH^T}$ is computed and stored and multiplied by the appropriate column of the $(\mathbf{HB_tH^T + O})^{-1}$ matrix.

[50] The adjustable parameters (r, ϵ) are chosen using χ^2 diagnostics and OmF (observations − forecast) differences (see Table 3).

[51] **Acknowledgments.** This work was initiated by a European Space Agency contract. We are grateful to Tobias Wehr for this support. This work was also supported by the U.K. Natural Environment Research Council. We thank the HALOE team for the use of the HALOE data which was obtained via the Cambridge Atmospheric Data Centre (CADC). We thank the ATMOS team for use of the ATMOS data.

References

Bruehl, C., et al., Halogen Occultation Experiment ozone channel validation, *J. Geophys. Res.*, *101*, 10,217–10,240, 1996.

Chipperfield, M. P., Multiannual simulations with a three-dimensional chemical transport model, *J. Geophys. Res.*, *104*, 1781–1805, 1999.

Chipperfield, M. P., and R. L. Jones, Relative influences of atmospheric chemistry and transport on Arctic O_3 trends, *Nature*, *400*, 551–554, 1999.

Chipperfield, M. P., M. L. Santee, L. Froidevaux, G. L. Manney, W. G. Read, J. W. Waters, A. E. Roche, and J. M. Russell, Analysis of UARS data in the southern polar vortex in September 1992 using a chemical transport model, *J. Geophys. Res.*, *101*, 18,861–18,881, 1996.

Clerbaux, C., J. Hadji-Lazaro, D. Hauglustaine, G. Megie, B. Khattatov, and J. F. Lamarque, Assimilation of carbon monoxide measured from satellite in a three-dimensional chemistry-transport model, *J. Geophys. Res.*, *106*, 15,385–15,394, 2001.

DeMore, W. B., et al., Chemical kinetics and photochemical data for use in stratospheric modeling, Evaluation no. 12, *JPL Publ. 97-4*, NASA Jet Propulsion Lab., Pasadena, Calif., 1997.

Errera, Q., and D. Fonteyn, Four-dimensional variational chemical assimilation of CRISTA stratospheric measurements, *J. Geophys. Res.*, *106*, 12,253–12,265, 2001.

Fahey, D. W., et al., Measurements of nitric oxide and total reactive nitrogen in the Antarctic stratosphere: Observations and chemical implications, *J. Geophys. Res.*, *94*, 16,665–16,681, 1989.

Fisher, M., and D. J. Lary, Lagrangian 4-dimensional variational data assimilation of chemical species, *Q. J. R. Meteorol. Soc.*, *121*, 1681–1704, 1995.

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.

Harries, J. E., J. M. Russell, A. F. Tuck, L. L. Gordley, P. Purcell, K. Stone, R. M. Bevilacqua, M. Gunson, G. Nedoluha, and W. A. Traub, Validation of measurements of water vapor from the Halogen Occultation Experiment (HALOE), *J. Geophys. Res.*, *101*, 10,205–10,216, 1996.

Kawa, S. R., R. A. Plumb, and U. Schmidt, Simultaneous observations of long-lived species, chap. H, The atmospheric effects of stratospheric aircraft: Report of the 1992 models and measurements workshop, *NASA Ref. Publ. 1292*, 352 pp., NASA Goddard Space Flight Center, Greenbelt, Md., 1993.

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

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

Kondo, Y., U. Schmidt, T. Sugita, A. Engel, M. Koike, P. Aimedieu, M. R. Gunson, and J. Rodriguez, NO_y correlation with N_2O and CH_4 in the midlatitude stratosphere, *Geophys. Res. Lett.*, *23*, 2369–2372, 1996.

Lamarque, J. F., B. V. Khattatov, J. C. Gille, and G. P. Brasseur, Assimilation of Measurement of Air Pollution from Space (MAPS) CO in a global three-dimensional model, *Geophys. Res. Lett.*, *104*, 26,209–26,218, 1999.

Levelt, P. F., B. V. Khattatov, J. C. Gille, G. P. Brasseur, X. X. Tie, and J. W. Waters, Assimilation of MLS ozone measurements in the global three-dimensional chemistry transport model ROSE, *Geophys. Res. Lett.*, *25*, 4493–4496, 1998.

Lyster, P. M., S. E. Cohn, R. Menard, L.-P. Chang, S.-J. Lin, and R. Olsen, An implementation of a two dimensional filter for atmospheric chemical constituent assimilation on massively parallel computers, *Mon. Weather Rev.*, *125*, 1674–1686, 1997.

Menard, R., and L.-P. Chang, Stratospheric assimilation of chemical tracer observations using a Kalman filter, 2, Chi-square validated results and analysis of variance and correlation dynamics, *Mon. Weather Rev.*, *128*, 2672–2686, 2000.

Menard, R., S. E. Cohn, L.-P. Chang, and P. M. Lyster, Stratospheric assimilation of chemical tracer observations using a Kalman filter, 1, Formulation, *Mon. Weather Rev.*, *128*, 2654–2671, 2000.

Park, J. H., et al., Validation of Halogen Occultation Experiment CH4 measurements from the UARS, *J. Geophys. Res.*, *101*, 10,217–10,240, 1996.

Plumb, R. A., and M. K. W. Ko, Interrelationships between mixing ratios of long-lived stratospheric constituents, *J. Geophys. Res.*, *97*, 10,145–10,156, 1992.

Prather, M. J., Numerical advection by conservation of second-order moments, *J. Geophys. Res.*, *91*, 6671–6681, 1986.

Russell, J. M., et al., The Halogen Occultation Experiment, *J. Geophys. Res.*, *98*, 10,777–10,797, 1993.

Russell, J. M., et al., Validation of hydrogen chloride measurements made by the Halogen Occultation Experiment from the UARS platform, *J. Geophys. Res.*, *101*, 10,151–10,162, 1996.

Shine, K. P., The middle atmosphere in the absence of dynamical heat fluxes, *Q. J. R. Meteorol. Soc.*, *113*, 603–633, 1987.

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

Waugh, D. W., et al., Mixing of polar vortex air into middle latitudes as revealed by tracer-tracer scatterplots, *J. Geophys. Res.*, *102*, 13,119–13,134, 1997.

M. P. Chipperfield, School of the Environment, University of Leeds, Woodhouse Lane, Leeds, LS2 9JT, UK. (martyn@env.leeds.ac.uk)

B. V. Khattatov, National Center for Atmospheric Research, Boulder, CO, USA.

D. J. Lary, Department of Chemistry, University of Cambridge, Cambridge, UK.

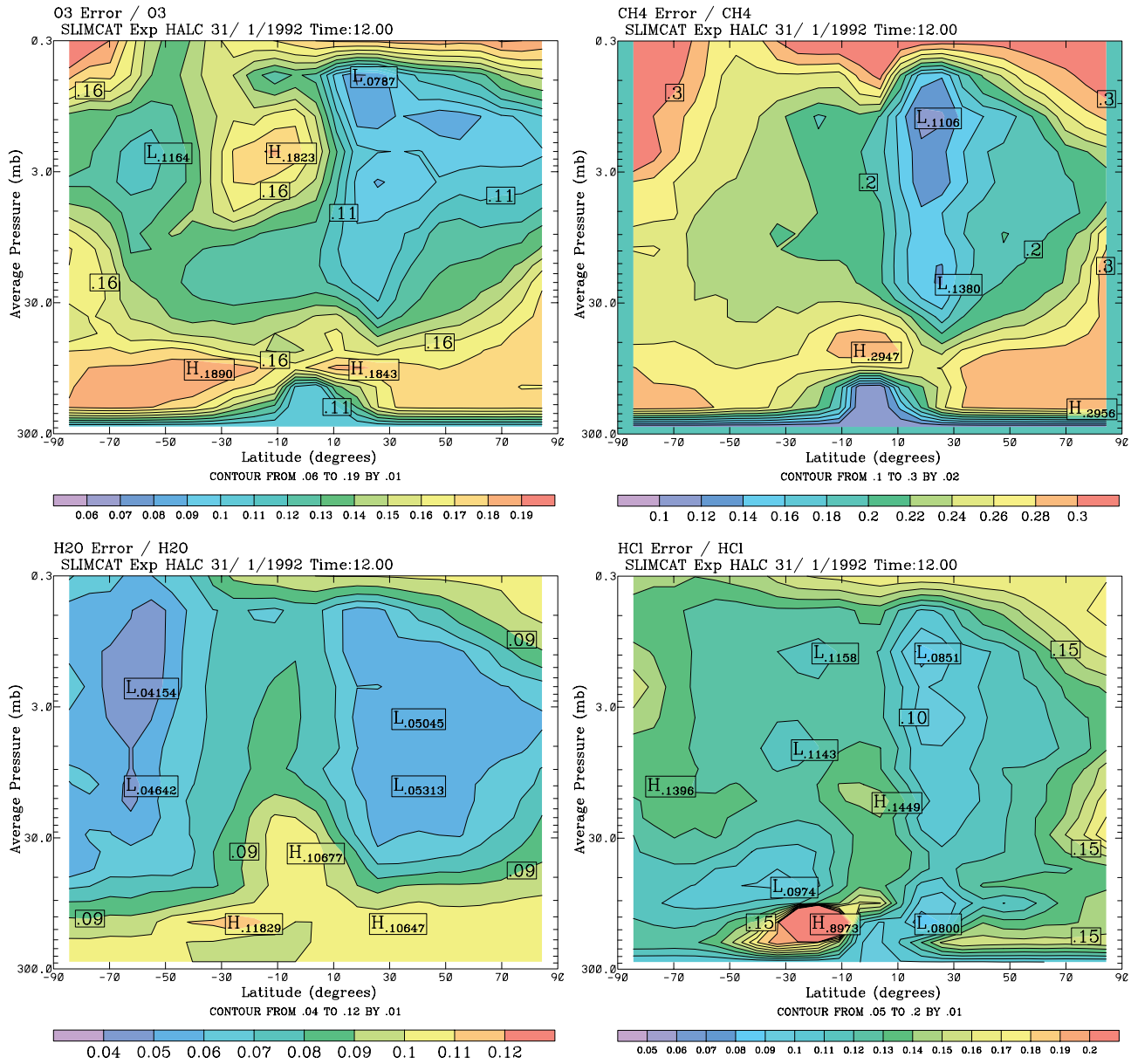

Figure 6. Zonal mean distributions of error tracers (normalized by the respective chemical tracers) for (a) O_3, (b) CH_4, (c) H_2O and (d) HCl.

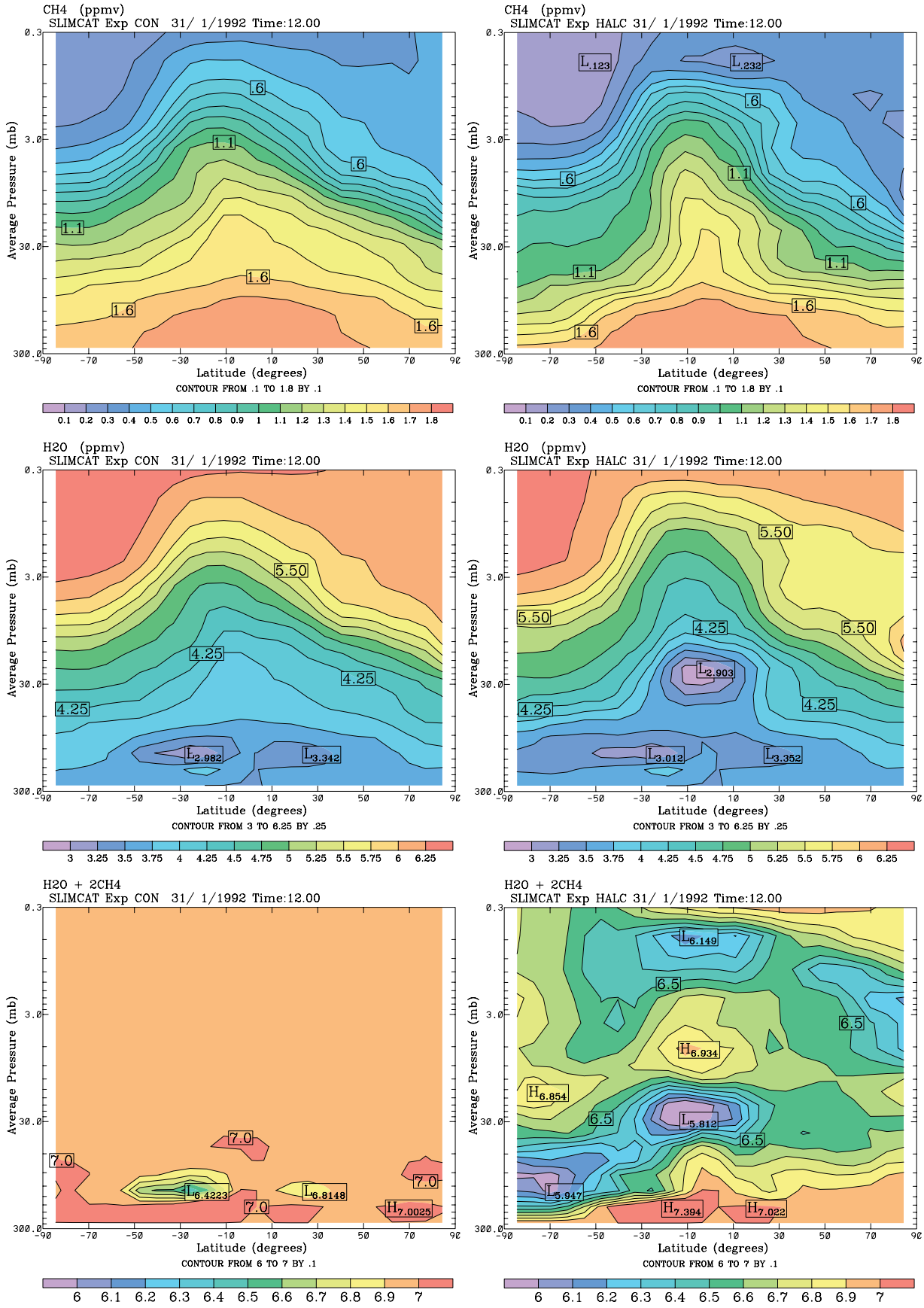

Figure 7. Zonal mean distributions of CH_4, H_2O, and $2CH_4 + H_2O$ (ppmv) on 31 January 1992 for run **CON** (left, no assimilation) and run **HALC** (right, with HALOE assimilation).

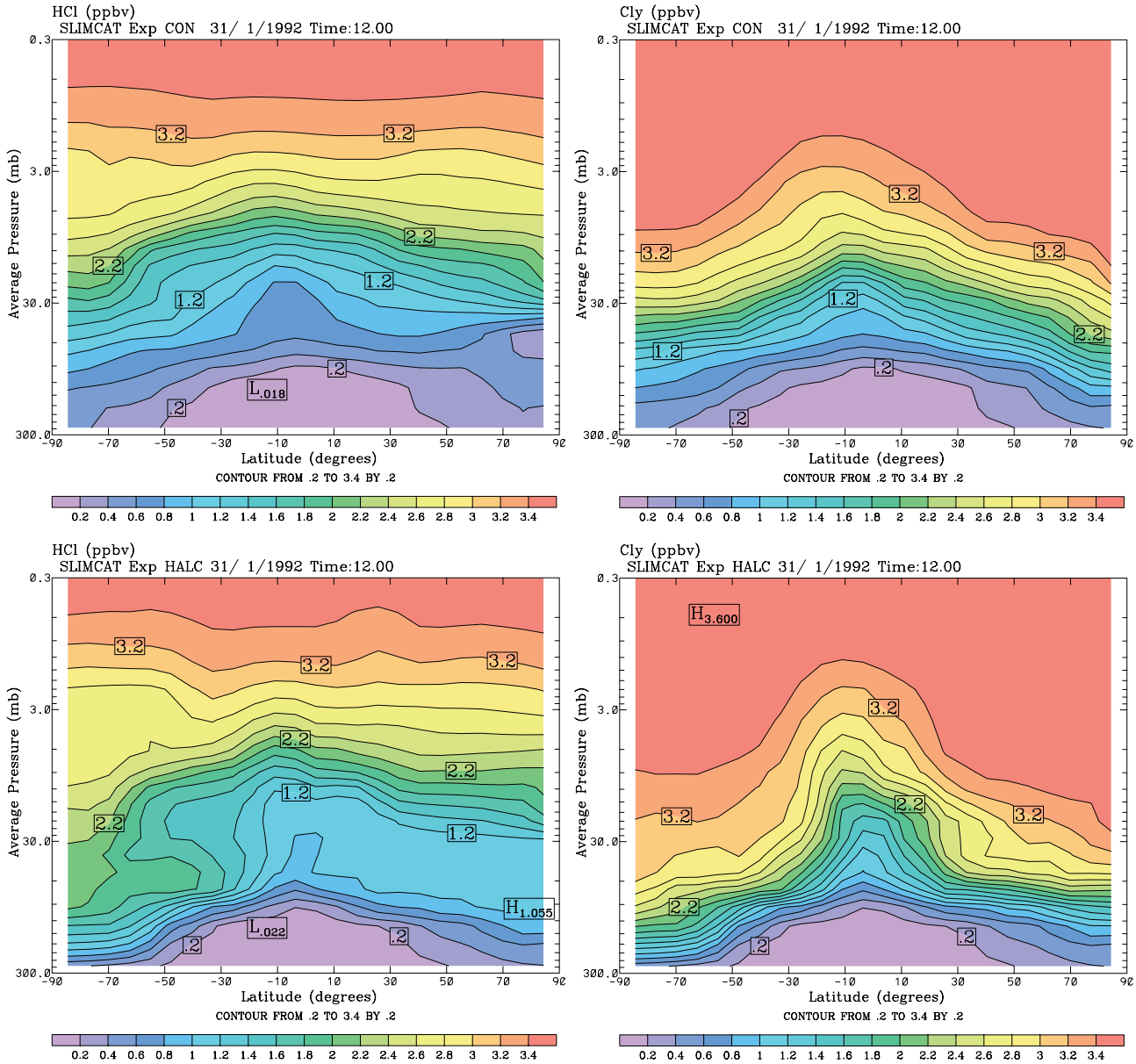

Figure 8. Zonal mean distributions of HCl and Cl_y (ppbv) on 31 January 1992 for runs (a) **CON** (above), and (b) **HALC** (below).

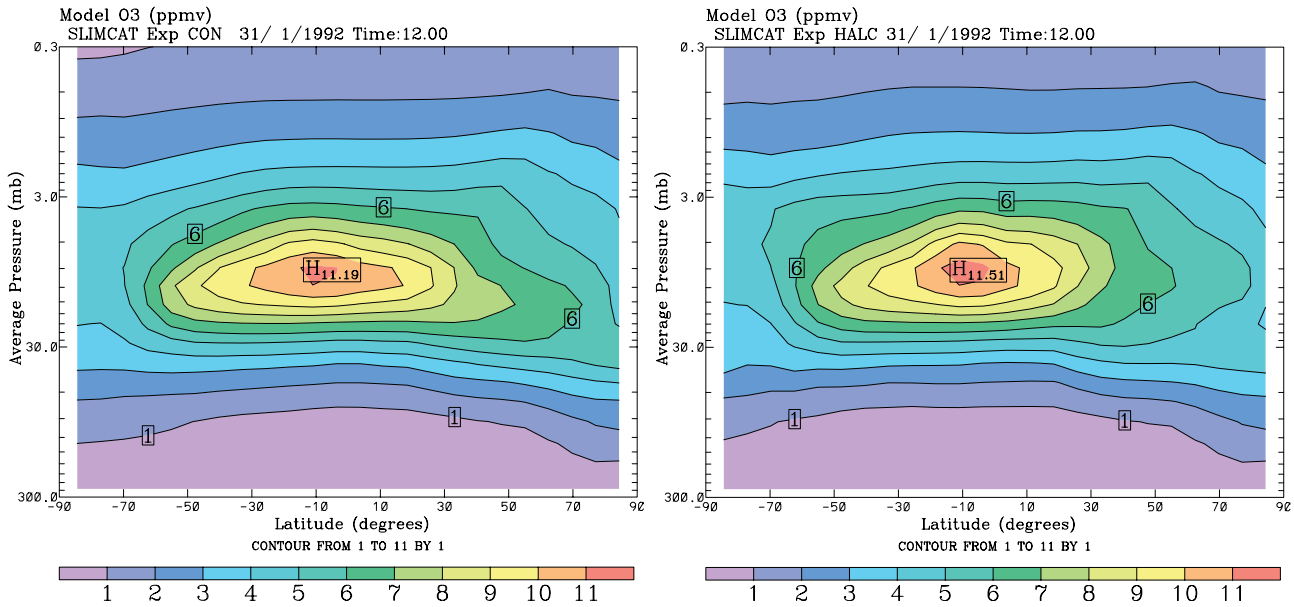

Figure 9. Zonal mean distributions of O_3 (ppmv) on 31 January 1992 for runs (a) **CON** (left), and (b) **HALC** (right).