Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 106, NO. D13, PAGES 14,539–14,550, JULY 16, 2001

Air mass factor formulation for spectroscopic measurements from satellites: Application to formaldehyde retrievals from the Global Ozone Monitoring Experiment

Paul I. Palmer,[1] Daniel J. Jacob,[1] Kelly Chance,[2] Randall V. Martin,[1] Robert J. D. Spurr,[2] Thomas P. Kurosu,[2] Isabelle Bey,[1] Robert Yantosca,[1] Arlene Fiore,[1] and Qinbin Li[1]

Abstract. We present a new formulation for the air mass factor (AMF) to convert slant column measurements of optically thin atmospheric species from space into total vertical columns. Because of atmospheric scattering, the AMF depends on the vertical distribution of the species. We formulate the AMF as the integral of the relative vertical distribution (shape factor) of the species over the depth of the atmosphere, weighted by altitude-dependent coefficients (scattering weights) computed independently from a radiative transfer model. The scattering weights are readily tabulated, and one can then obtain the AMF for any observation scene by using shape factors from a three dimensional (3-D) atmospheric chemistry model for the period of observation. This approach subsequently allows objective evaluation of the 3-D model with the observed vertical columns, since the shape factor and the vertical column in the model represent two independent pieces of information. We demonstrate the AMF method by using slant column measurements of formaldehyde at 346 nm from the Global Ozone Monitoring Experiment satellite instrument over North America during July 1996. Shape factors are computed with the Global Earth Observing System CHEMistry (GEOS-CHEM) global 3-D model and are checked for consistency with the few available aircraft measurements. Scattering weights increase by an order of magnitude from the surface to the upper troposphere. The AMFs are typically 20–40% less over continents than over the oceans and are approximately half the values calculated in the absence of scattering. Model-induced errors in the AMF are estimated to be ~10%. The GEOS-CHEM model captures 50% and 60% of the variances in the observed slant and vertical columns, respectively. Comparison of the simulated and observed vertical columns allows assessment of model bias.

1. Introduction

The retrieval of vertical column densities of atmospheric species from solar backscatter measurements made by satellite instruments involves a two-step procedure. First, slant column densities are determined by fitting modeled atmospheric spectra to the observed absorption spectra. Second, an air mass factor (AMF) is used to convert from slant columns to total vertical columns. The AMF is defined as the ratio of the slant column to the vertical column, and depends on the radiative transfer properties of the atmosphere.

In practice, the second step of the procedure often limits the accuracy of the retrieved vertical column quantity. Much previous work has relied on a geometric AMF, which is a function of the viewing angles of the satellite instrument but does not account for atmospheric scattering. However, atmospheric scattering is critical for the interpretation of solar spectra at UV-visible wavelengths. Some authors have endeavored to include this effect in the AMF calculation [e.g, *Rozanov et al.*, 1998; *Stammes and Koelemeijer*, 1999; *Buchwitz et al.*, 2000].

A major complication of accounting for scattering in the calculation of the AMF is that the AMF becomes sensitive to the vertical distribution of the atmospheric species being analyzed. The spectral fitting itself usually does not provide information on this distribution, which can be variable. One approach to account for this variability, and its possible correlation with the column quantity, is to use an atmospheric chemistry model; but one is then concerned about the practicality of implementation and also about model contamination of the observations.

We resolve this difficulty here with an AMF formulation that decouples the vertical dependence of the observational sensitivity to the atmospheric species of interest (calculated with a radiative transfer model) from the shape of the

[1]Division of Engineering and Applied Sciences, and Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts

[2]Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts.

Paper number 2000JD900772.
0148-0227/01/2000JD900772$09.00

vertical profile of the species concentration (calculated with an atmospheric chemistry model). This decoupling increases the flexibility of the vertical column retrieval, as both components of the formulation can be studied individually. The observational sensitivities obtained from the radiative transfer calculations can be tabulated, thus allowing an expedient calculation of the AMF for any vertical profile of the species. Our method also ensures consistency in the comparison of observed vertical columns with the columns predicted from the atmospheric chemistry model, when the same model is used to constrain the shape of the vertical profile of the species in the AMF calculation. The method presented here is applicable to any atmospheric slant column observation from space with the requirement that the species of interest be optically thin. This includes slant columns retrieved using direct fitting of measured radiances [e.g., *Chance et al.*, 2000] or differential optical absorption spectroscopy (DOAS) [*Platt*, 1994].

We apply our AMF formulation to fitted formaldehyde (HCHO) column measurements from the Global Ozone Monitoring Experiment (GOME) instrument [*Burrows et al.*, 1999] on board the European Remote Sensing 2 satellite, launched in 1995. The satellite is in a Sun-synchronous orbit, crossing the equator at 1030 local solar time (LT) in the descending node. *Chance et al.* [2000] fitted the GOME spectra in the 337.35 to 356.12-nm wavelength microwindow to obtain slant columns of HCHO with a sensitivity of 4.0×10^{15} molecules cm^{-2}. Slant column quantities of HCHO provide an excellent application of our AMF methodology, since HCHO is mostly in the boundary layer where atmospheric scattering is particularly important.

Formaldehyde is produced in the atmosphere by oxidation of hydrocarbons [*Lee et al.*, 1998]. In the remote atmosphere, CH_4 is the dominant HCHO precursor. In continental boundary layers, reactive nonmethane hydrocarbons (NMHCs) are often dominant. The sinks of HCHO are photolysis and reaction with the hydroxyl radical (OH), resulting in a lifetime of the order of hours. Measurements of HCHO columns in continental atmospheres offer a diagnostic of NMHC emissions, which are of great importance for understanding the formation of ozone (O_3) [*Sillman*, 1999] and of organic aerosol [*Griffin et al.*, 1999]. In the eastern United States in summer, where biogenic isoprene is the dominant HCHO precursor [*Lee et al.*, 1998], HCHO column measurements can provide a measure of isoprene emission.

The new AMF formulation is derived in a general form in section 2, and is applied in section 3 to HCHO retrievals from GOME. We conclude the paper with a discussion of results and future work.

2. Air Mass Factor Formulation

The Beer-Lambert law of extinction provides the basis for the spectroscopic measurement of the atmospheric column of an absorbing chemical species, hereinafter referred to as the absorber:

$$I_B = I_{B_0} e^{-\tau_s}, \tag{1}$$

where I_B is the backscattered intensity observed by the satellite instrument, I_{B_0} is the backscattered intensity that would be observed in the absence of the absorber, and τ_s is the optical thickness for the slant column absorber. In practice, τ_s is determined from fitting I_B to a synthetic spectrum beginning with I_{B_0} and the absorption cross sections.

The AMF is the ratio of the slant column abundance Ω_s of the absorber (i.e., that viewed by the satellite in the measured radiance spectrum) to the vertical column abundance Ω_v:

$$\mathrm{AMF} = \frac{\Omega_s}{\Omega_v} = \frac{\tau_s}{\tau_v}, \tag{2}$$

where τ_v is the optical thickness for the vertical column and Ω is in units of molecules cm^{-2}.

In the absence of atmospheric scattering, a geometric AMF ($\mathrm{AMF_G}$) can be defined, which is a simple function of the solar zenith angle θ_s and of the satellite viewing angle θ_v:

$$\mathrm{AMF_G} = \sec\theta_s + \sec\theta_v. \tag{3}$$

We define an effective solar zenith angle θ_E which yields the same slant column if the satellite instrument were viewing the nadir (θ_v=0):

$$\sec\theta_E = \sec\theta_s + \sec\theta_v - 1 \tag{4}$$

so that we can rewrite (3) as

$$\mathrm{AMF_G} = \frac{1 + \cos\theta_E}{\cos\theta_E}. \tag{5}$$

In the UV-visible region of the spectrum, Rayleigh and Mie scattering interfere with geometric viewing of the atmosphere. To include scattering in the AMF calculation we start from the Beer-Lambert law of extinction (equation (1)) and replace it into (2) to obtain an expression for the "scattering" AMF (hereinafter referred to simply as the AMF):

$$\mathrm{AMF} = \frac{\ln I_{B_0} - \ln I_B}{\tau_v}. \tag{6}$$

Let $\Delta I = I_B - I_{B_0}$ represent the change in backscattered intensity due to the absorber. For optically thin absorption, $\Delta I / I_B \ll 1$, so that the AMF is given by

$$\mathrm{AMF} = \frac{\ln\left(1 - \frac{\Delta I}{I_B}\right)}{\tau_v} \simeq -\frac{1}{\tau_v}\frac{\Delta I}{I_B}. \tag{7}$$

We now express ΔI as the integral of the absorption resulting from optical depth increments $\mathrm{d}\tau(z)$ of the absorber over the vertical extent of the atmosphere:

$$\Delta I = \int_0^{\tau_v} \frac{\partial I_B}{\partial \tau} \mathrm{d}\tau. \tag{8}$$

Replacing in (7):

$$\mathrm{AMF} = -\frac{1}{\tau_v}\int_0^{\tau_v} \frac{\partial \ln I_B}{\partial \tau} \mathrm{d}\tau, \tag{9}$$

which can be rewitten:

$$\mathrm{AMF} = -\frac{1}{\tau_v}\int_0^\infty \frac{\partial \ln I_B}{\partial \tau}\alpha(z)n(z)\mathrm{d}z, \qquad (10)$$

where $\alpha(z)$ and $n(z)$ are the absorption cross section (m^2 molecule^{-1}) and number density (molecules m^{-3}), respectively, of the absorber.

Let us define an effective absorption cross section α_e for the absorber in the atmospheric column:

$$\alpha_e = \frac{1}{\Omega_v}\int_0^\infty \alpha(z)n(z)\mathrm{d}z, \qquad (11)$$

which represents an average cross section weighted by the vertical distribution of the absorber in the column; this averaging is intended to account for the (generally weak) dependence of $\alpha(z)$ on temperature and pressure. By definition $\tau_v = \Omega_v\alpha_e$, and by replacing this into (10), we obtain

$$\mathrm{AMF} = -\int_0^\infty \frac{\partial \ln I_\mathrm{B}}{\partial \tau}\frac{\alpha(z)}{\alpha_e}\frac{n(z)}{\Omega_v}\mathrm{d}z. \qquad (12)$$

We now define scattering weights $w(z)$ to describe the sensitivity of the backscattered spectrum to the abundance of the absorber at altitude z:

$$w(z) = -\frac{1}{\mathrm{AMF_G}}\frac{\alpha(z)}{\alpha_e}\frac{\partial \ln I_\mathrm{B}}{\partial \tau}. \qquad (13)$$

and a vertical shape factor $S_z(z)$ for the absorber, representing a normalized vertical profile of number density

$$S_z(z) = \frac{n(z)}{\Omega_v}. \qquad (14)$$

Then the AMF can be rewritten as

$$\mathrm{AMF} = \mathrm{AMF_G}\int_0^\infty w(z)S_z(z)\mathrm{d}z. \qquad (15)$$

In this expression, $w(z)$ is a function of atmospheric scattering and can be calculated with a radiative transfer model, while $S_z(z)$ can be derived from an atmospheric chemistry model or some other prior knowledge of the vertical distribution of the absorber. A major advantage of the expression is to decouple the contributions from the radiative transfer model and from the atmospheric chemistry model, thus allowing flexibility in the AMF calculation. The $\mathrm{AMF_G}$ is introduced here to provide normalization of the scattering weights, such that $w(z) = 1$ in a non scattering atmosphere. For an optically thin absorber, $w(z)$ is independent of the column abundance or the vertical distribution of the species. In a Rayleigh scattering, clear-sky atmosphere $w(z)$ is dependent only on θ_E, wavelength λ, surface albedo A, and surface pressure P_S. Values of $w(z)$ can be readily tabulated and then used in combination with any model-specified vertical shape factor for local calculations of the AMF. Mie scattering by aerosols and clouds introduces complications in the calculation of $w(z)$ as discussed in the next sections, but the resulting values can still be tabulated as a function of the aerosol and cloud optical depths.

The decoupling of the scattering weights and the vertical shape factor is not strictly achieved by (15), because the calculation of α_e (equation (11)) involves an assumed shape factor. An alternate definition of the scattering weights and shape factor (equation (16)) would achieve a strict decoupling.

$$w'(z) = -\frac{1}{\mathrm{AMF_G}}\frac{\partial \ln I_\mathrm{B}}{\partial \tau}, \qquad (16a)$$

$$S'_z(z) = \frac{\alpha(z)n(z)}{\int_0^\infty \alpha(z)n(z)\mathrm{d}z} \qquad (16b)$$

$$\mathrm{AMF} = \mathrm{AMF_G}\int_0^\infty w'(z)S'_z(z)\mathrm{d}z. \qquad (16c)$$

In this formulation the vertical shape factor $S'_z(z)$ represents a normalized vertical profile of optical thickness. In practice, the vertical dependence of α is small, and for the implementation of (11) in the next section we use a typical HCHO shape factor (uniform mixing ratio in the lowest 2 km and exponential decrease at higher altitude with a scale height of 2.7 km).

Global atmospheric chemistry models generally use the terrain-following sigma (σ) vertical coordinate. Pressure is related to σ by $P = \sigma(P_S - P_T) + P_T$, where P_T represents the pressure at the upper boundary of the model. To convert (15) to the σ coordinate, we apply the hydrostatic relation:

$$\mathrm{d}\sigma = -\frac{\rho_a g}{P_S - P_T}\mathrm{d}z, \qquad (17)$$

where ρ_a is the air mass density and g is the acceleration due to gravity. Substitution into (12) yields an expression for the AMF in the σ coordinate:

$$\mathrm{AMF} = \mathrm{AMF_G}\int_0^1 w(\sigma)S_\sigma(\sigma)\mathrm{d}\sigma. \qquad (18)$$

In this expression, $S_\sigma(\sigma)$ is now a vertical shape factor representing a normalized vertical profile of mixing ratio:

$$S_\sigma(\sigma) = \frac{\Omega_a}{\Omega_v}C(\sigma), \qquad (19)$$

where Ω_a is the vertical column of air from P_S to P_T and $C(\sigma)$ is the mixing ratio of the absorber. This shape factor $S_\sigma(\sigma)$ is dimensionless, as opposed to $S_z(z)$, which has the dimension of m^{-1}.

3. Application to GOME Retrievals of Formaldehyde

We describe in this section the practical implementation of (18) to the retrieval of HCHO columns from the GOME instrument. We use HCHO observations over North America for July 1996 [*Chance et al.*, 2000].

3.1. Calculation of Scattering Weights

We calculate the scattering weights in (13) using the linearized discrete ordinate radiative transfer (LIDORT) model [*Spurr et al.*, 2001]. LIDORT solves the radiative transfer

equation in a multilayer atmosphere with multiple scattering using the discrete ordinate method [*Chandrasekhar*, 1960]. The model contains a full internal perturbation analysis of the intensity field, allowing $\partial \ln I_B/\partial\tau$ in (13) to be derived to the same level of accuracy specified for the intensity. Although LIDORT has been designed primarily as a general forward model for nonlinear atmospheric retrieval problems, calculation of the scattering weights is a straightforward application of the model.

For the present application we calculate the scattering weights on an altitude coordinate with vertical resolution of 0.5 km up to 18 km altitude and lower resolution at higher altitudes up to the assumed top of the atmosphere (65 km). We then map the results on to the σ coordinate. The surface albedo is treated as Lambertian. Rayleigh scattering cross sections are calculated using the formulae of *Chance and Spurr* [1997]. Vertical profiles of the UV absorbers O_3 and NO_2 are taken from a midlatitude summer U.S. Standard Atmosphere [1976]. Aerosol Mie scattering is included in a sensitivity calculation described below. Cloud effects are not included.

The scattering weights calculated in this manner for a Rayleigh scattering atmosphere are dependent on λ, θ_E, P_S, and A. We have tabulated these dependences for three different wavelengths in the HCHO fitting window (339.04, 346.04, and 352.76 nm), seven different effective solar zenith angles (15, 25, 35, 45, 55, 65, and 75°), five UV surface albedo values (0.01, 0.07, 0.15, 0.8, and 0.9), and three surface pressure values (1000, 900, and 800 hPa).

The sensitivities of w to λ and P_S were found to be negligible over the above ranges of interest. We elaborate here on the sensitivities to θ_E and A, which are more important. Figure 1 illustrates these sensitivities for a typical HCHO shape factor (uniform mixing ratio up to 2 km above the surface and exponential decrease at higher altitudes with a 2.7-km scale height).

Increasing values of the surface albedo A allow more solar radiation in the lower atmosphere (where most of the HCHO resides) to be reflected to space, thus increasing the observational sensitivity to the HCHO column. Except for surfaces covered by snow and ice, the UV surface albedo is <10% [*Herman and Celarier*, 1997]. At high values of surface albedo, atmospheric scattering of radiation reflected from the surface increases the photon path length beyond that defined by the viewing geometry, so that the AMF is larger than the AMF_G (Figure 1a).

Figure 1b shows that although the geometric optical path increases as θ_E increases, scattering along that path limits the actual gain in sensitivity, so that the sensitivity of the AMF to θ_E is, in fact, small (20% increase in AMF as θ_E increases from 0° to 65°). We observe deeper into the atmosphere when θ_E is small.

We examined the sensitivity of the AMF to aerosol scattering for the typical HCHO shape factor described above by prescribing an aerosol column optical thickness and distributing this optical depth vertically following the same shape factor as for HCHO (as may be roughly expected for an aerosol with a boundary layer source). We assumed a single-scattering albedo of 0.96 [*Dickerson et al.*, 1997]. Aerosol optical thicknesses at 346 nm over North America are generally much less than unity, but they can occasionally reach in excess of 2 during pollution episodes over the eastern United States in summer [*Dickerson et al.*, 1997]. An aerosol optical thickness of 2 increases the AMF by 40% relative to an aerosol-free atmosphere (Figure 1c); scattering aerosols collocated with HCHO increase the sensitivity to the HCHO column.

Overall, the results in Figure 1 show that the AMF values are approximately half of AMF_G for typical conditions, in the absence of snow or ice cover. Scattering thus decreases the sensitivity of GOME to atmospheric HCHO by about a factor of 2 in the 337.35–356.12 nm fitting window.

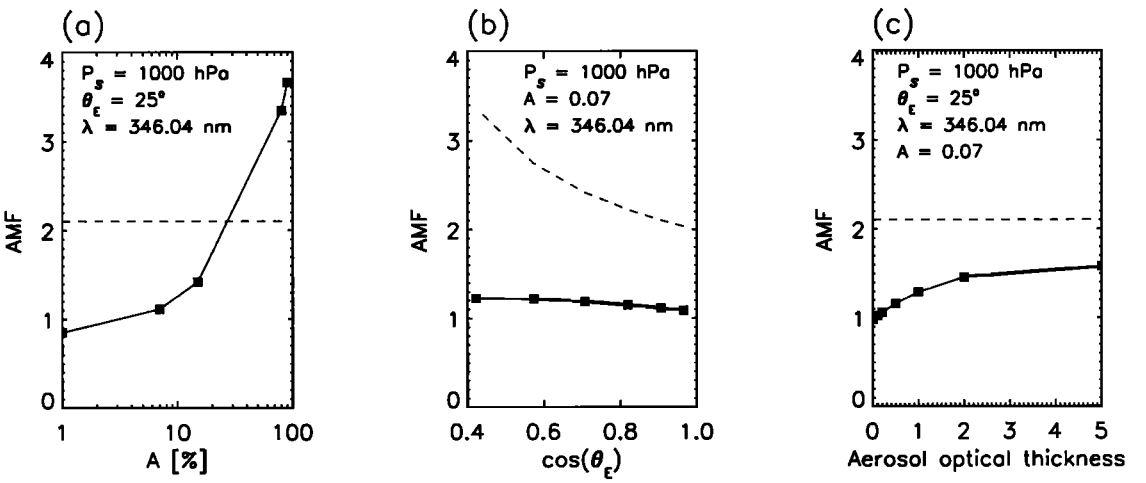

Figure 1. Air mass factor (AMF) sensitivity to (a) surface albedo A, (b) effective solar zenith angle θ_E, and (c) aerosol optical thickness. The AMF is calculated for a wavelength λ of 346.04 nm, a surface pressure P_S of 1000 hPa, and a typical vertical shape factor for HCHO (uniform mixing ratio up to 2 km above the surface, decreasing exponentially at higher altitudes with a scale height of 2.7 km). The dashed lines show the geometric AMF (AMF_G).

3.2. Calculation of Vertical Shape Factors

If a vertically uniform HCHO mixing ratio were assumed in the AMF calculation [$S_\sigma(\sigma)$=1], then typical GOME viewing conditions would yield AMF≃1.4. In fact the HCHO mixing ratio decreases with altitude [*Lee et al.*, 1998; *Singh et al.*, 2000] owing to near-surface sources of short-lived NMHC precursors, the temperature dependence of the source from CH_4 oxidation, and the UV dependence of the photolysis sink. As a result the AMF is <1.4 and varies depending on local NMHC sources, boundary layer depth, photochemical activity, and other factors. To resolve this variability we use the Global Earth Observing System CHEMistry (GEOS-CHEM) global 3-D model of tropospheric chemistry (*I. Bey et al.*, Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, submitted to *Journal of Geophysical Research*, 2000) (hereinafter referred to as submitted manuscript, 2000) to provide values of $S_\sigma(\sigma)$ for every GOME observation scene.

The GEOS-CHEM global 3-D model of tropospheric chemistry is driven by assimilated meteorological data updated every 3 hours from the NASA Data Assimilation Office GEOS model [*Schubert et al.*, 1993]. It builds on previous global 3-D models of tropospheric chemistry at Harvard University [e.g., *Wang et al.*, 1998a,b,c; *Horowitz and Jacob*, 1999]. The model version used here has 26 vertical levels on a sigma coordinate (surface to 0.1 hPa) and a horizontal resolution of 2° in latitude and 2.5° in longitude. The lowest model levels are centered at 50, 250, 600, and 1100 m above ground. The model includes a detailed description of tropospheric O_3-NO_x-hydrocarbon chemistry and solves the chemical evolution of ~120 species with a fast Gear solver [*Jacobson and Turco*, 1994]. The chemical mechanism is that of *Horowitz and Jacob* [1999] with recent updates and includes full chemistry for five NMHCs (ethane, propane, lumped >C_3 alkanes, lumped >C_2 alkenes, and isoprene). There are 24 chemical tracers (including HCHO) transported in the model. Photolysis rates are computed using the Fast-J radiative transfer algorithm [*Wild et al.*, 2000], which includes Rayleigh scattering and Mie scattering by clouds. We use monthly averaged UV surface albedo fields from *Herman and Celarier* [1997].

Emissions of NMHCs, CO, and NO_x in the model largely follow the schemes of *Wang et al.* [1998a] with some updates (*I. Bey et al*, submitted manuscript, 2000). Emission of isoprene, which is of particular interest here, depends on local ecosystem type, leaf area index, solar irradiance, and temperature, following the scheme of *Guenther et al.* [1995] with minor modifications [*Wang et al.*, 1998a; *I. Bey et al.*, submitted manuscript, 2000].

A general evaluation of the GEOS-CHEM model with observations for O_3 and related species is presented by *I. Bey et al.* (submitted manuscript, 2000). Of particular interest here is the ability of the model to simulate the vertical distribution of HCHO. We examine for this purpose the few in situ vertical profile measurements of HCHO made from aircraft over our North American domain, from the Southern Oxidants Study (SOS) 1995 [*Lee et al.*, 1998] and from the North Atlantic Regional Experiment (NARE) 1997 (*A. Fried et al.*, manuscript in preparation, 2000). The SOS and NARE measurements (samples of which are shown in Figure 2) provide vertical profiles in continental and marine atmospheres, respectively.

The SOS campaign was conducted in the summer of 1995 over the southeastern United States. Isoprene was the dominant source of HCHO in the boundary layer [*Lee et al.*, 1998]. Figures 2a and 2b show that much of the HCHO column is in the boundary layer below 2-km altitude, reflecting the strong source from oxidation of isoprene and the relatively short lifetime of HCHO against photolysis (~ 4 hours in daytime).

The NARE 1997 aircraft intensive took place over the Canadian Maritime Provinces and adjacent regions of the western North Atlantic Ocean. The observations shown in Figure 2c are representative of a marine atmosphere where CH_4 is the principal source of HCHO: the decrease of HCHO concentration with altitude is less steep than over the continent and reflects principally the temperature dependence of CH_4 oxidation. Figure 2d is also a marine atmosphere but features an elevated layer of HCHO at 1-km altitude from continental outflow.

The GEOS-CHEM model results compared to the observations in Figure 2 were obtained using assimilated meteorological data for the specific days of observation and with a sufficient initialization. The GEOS-CHEM model reproduces the continental boundary layer enhancement of HCHO measured during the SOS campaign (Figure 2a and 2b), although the decline in the free troposphere is steeper than observed. The same discrepancy is apparent in the marine profiles of Figure 2c and appears to reflect a general model underestimate of HCHO levels in the lower free troposphere (*A. Fried et al.*, manuscript in preparation, 2000) The model reproduces well the event of continental outflow over the North Atlantic at 1-km altitude in Figure 2d. *Singh et al.* [2000] previously reported a comparison of GEOS-CHEM model vertical profiles for HCHO with aircraft observations over the North Atlantic during the SONEX campaign in October–November 1997. The comparison showed agreement within a factor of 2 and the vertical shape factor in the model was similar to that in the observations.

The ability of the GEOS-CHEM model to provide accurate AMFs can be tested using the aircraft observations. Model profiles for the flight days were compared to all 14 profiles in the SOS data set and all seven profiles in the NARE 1997 data set. For every model and observed profile a shape factor $S(\sigma)$ was computed for the vertical levels described by the observations. Using a standard $w(\sigma)$ profile (A = 7%, θ_E = 25°, λ = 346.04 nm, P_S = 1000 hPa, clear sky and aerosol free), the AMF for each profile was calculated. If the aircraft data are assumed to represent the real atmosphere perfectly, an uncertainty can be assigned to the AMFs obtained from the GEOS-CHEM model. For the SOS aircraft data set the model AMF has a mean bias of -10%±9%, owing to the steep decline of the

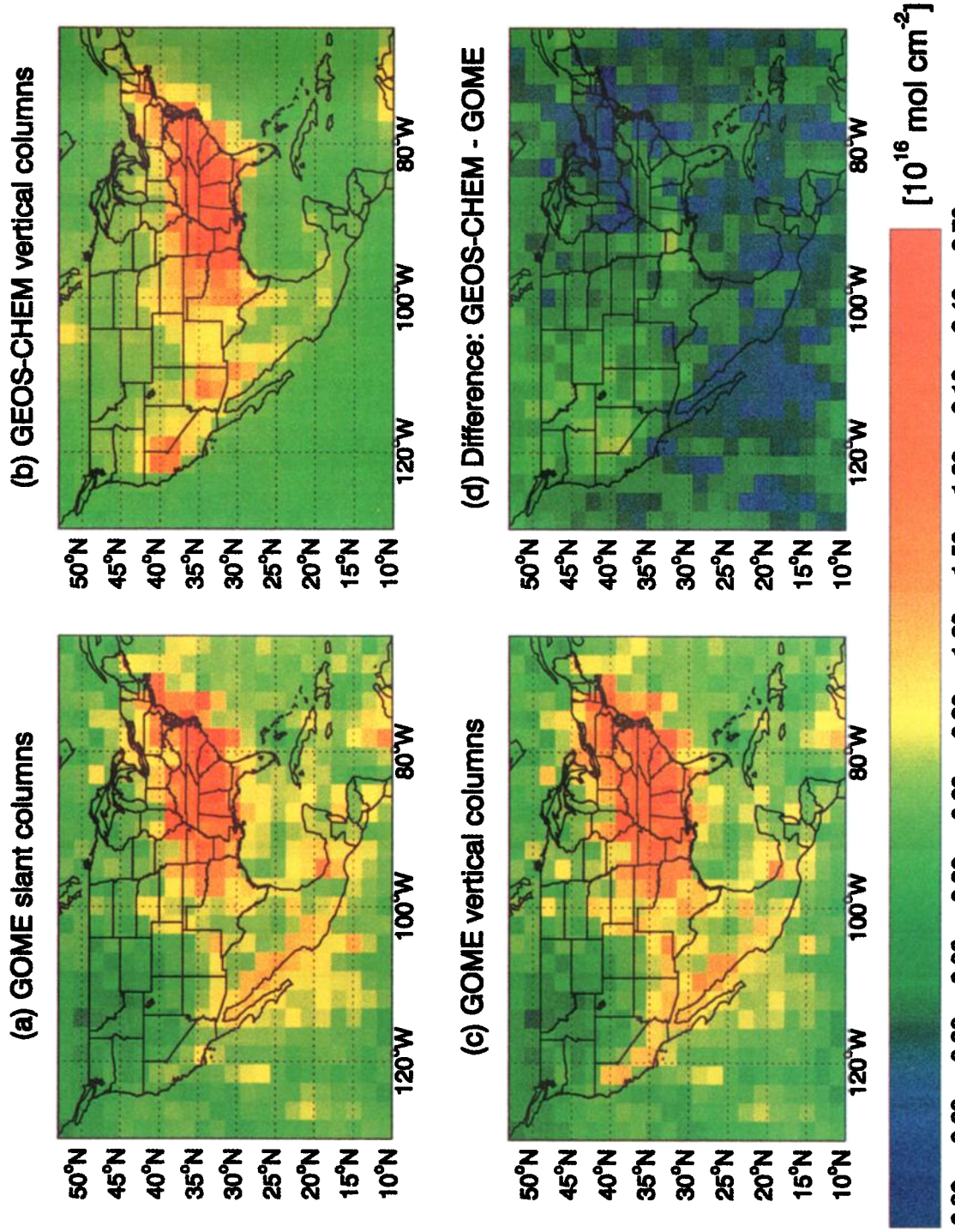

Plate 1. Mean formaldehyde columns over the United States for July 1996 on a 2° × 2.5° grid. (a) Observed slant columns from GOME (described by *Chance et al.* [2000]). (b) Vertical columns from the GEOS-CHEM global 3-D model. (c) Observed vertical columns from GOME after application of the AMF. (d) Differences between the modeled and observed vertical columns. Both model and observations are for 1000–1200 LT and for cloud cover <40%.

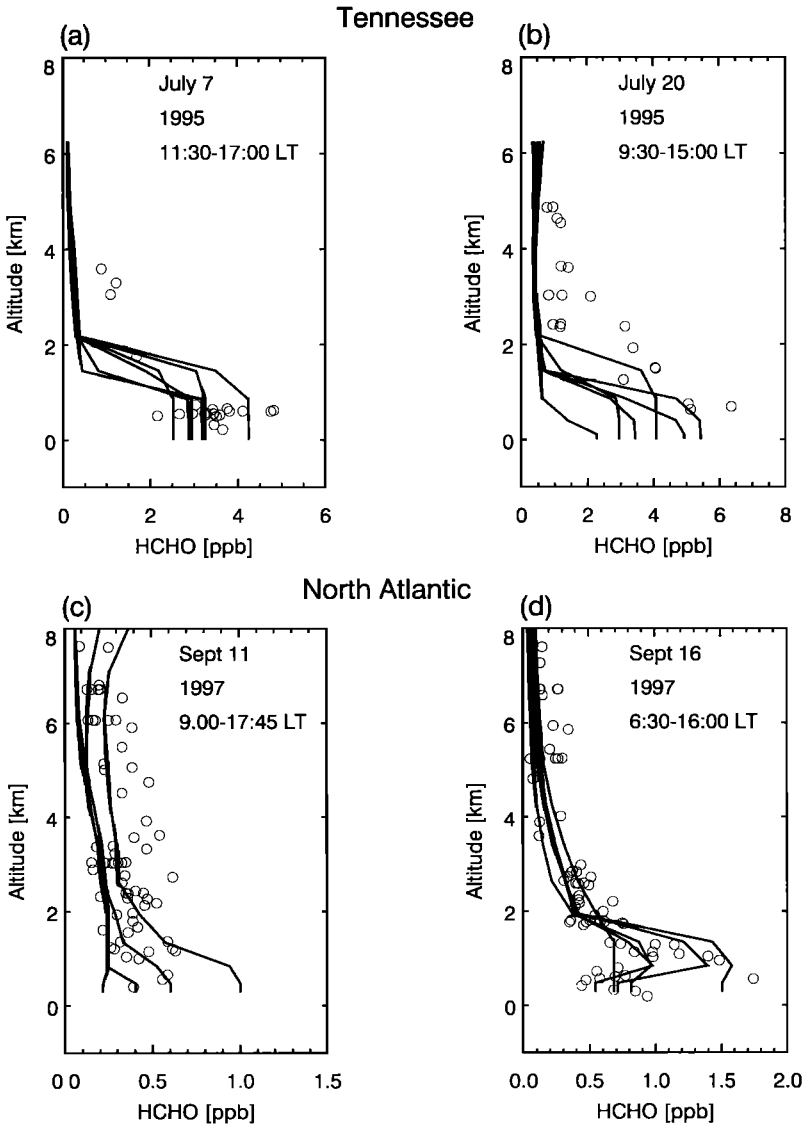

Figure 2. Comparisons of HCHO concentrations between the GEOS-CHEM model (lines) and aircraft observations (circles). Data are from two flights of the Southern Oxidants Study 1995 over Tennessee [*Lee et al.*, 1998] for (a) July 7 and (b) July 20, and from two flights of the North Atlantic Regional Experiment 1997 over the western North Atlantic (*A. Fried et al.*, manuscript in preparation, 2000) for (c) September 11 and (d) September 16. Different lines represent GEOS-CHEM model profiles for the ensemble of 2° × 2.5° grid boxes sampled by the aircraft flight tracks and for the specific flight days.

model shape factor in the lower free troposphere relative to the observations. For the NARE data set the model AMF has a mean bias of 0%±6%. We thus estimate that use of the GEOS-CHEM model to specify vertical shape factors of HCHO induces an error in the AMF calculation of the order of 10%.

Overall, the comparison of GEOS-CHEM model results for $S_\sigma(\sigma)$ with available in situ observations shows that the model provides useful information for constraining the retrievals of HCHO columns from GOME, capturing in particular the observed decline of concentrations between the boundary layer and the free troposphere and the difference in the slope of this decline between continental and marine atmospheres.

3.3. Implementation

Figure 3 shows vertical profiles of the scattering weights $w(\sigma)$ and the shape factors $S_\sigma(\sigma)$ for HCHO in sample GOME scenes over the the ocean (North Pacific) and over the eastern United States (Tennessee). Scattering weights are similar in both cases, because the solar zenith angles are almost the same and UV albedos are low. We see from the vertical profile of the scattering weights that observations are about an order of magnitude more sensitive to a given mixing ratio increment in the upper troposphere than near the surface, because penetration of UV radiation to the lower atmosphere is inhibited by scattering.

The shape factor over the North Pacific, where CH_4 is the dominant HCHO precursor, is fairly uniform up to

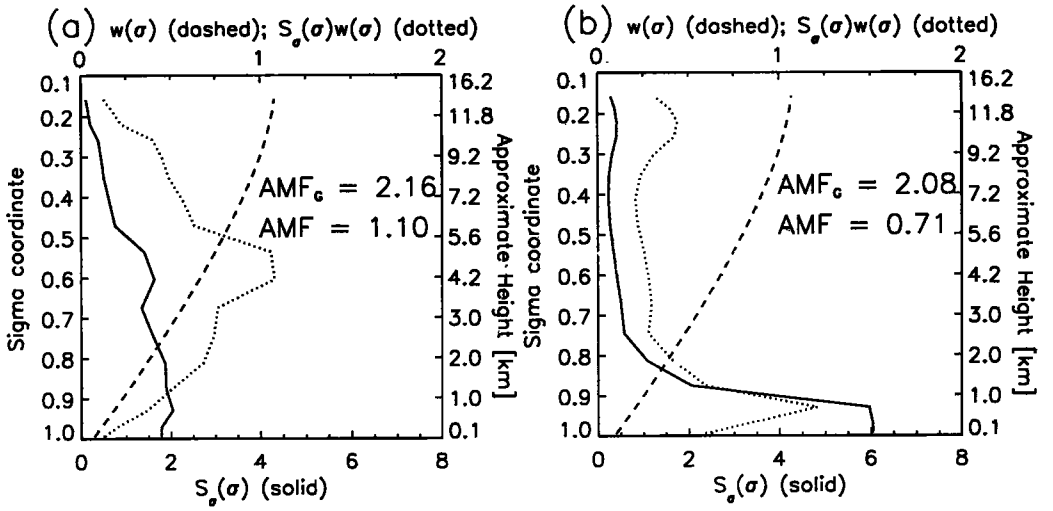

Figure 3. Scattering weights $w(\sigma)$ (dashed line) and vertical shape factors $S_\sigma(\sigma)$ (solid line) for HCHO as a function of altitude (σ coordinate) for two illustrative Global Ozone Monitoring Experiment (GOME) scenes over (a) the North Pacific and (b) Tennessee. The product $S_\sigma(\sigma)w(\sigma)$, which provides a measure of the HCHO signal seen by GOME, is shown as a dotted line. Values of the air mass factor (AMF) and the geometric air mass factor (AMF_G) are also shown.

4 km. The product $w(\sigma)S_\sigma(\sigma)$ shows that most of the HCHO seen in the GOME fitting spectral window is in the midtroposphere. The shape factor over Tennessee, where there is a large boundary layer source of HCHO from isoprene oxidation, is sharply peaked in the lower troposphere, in contrast to the shape factor over the ocean. Most of the HCHO measured in the GOME fitting spectral window $[w(\sigma)S_\sigma(\sigma)]$ is below $\sigma = 0.7$, i.e, below 3 km altitude. The AMF is 0.71, as compared to an AMF of 1.1 for the North Pacific scene; in this example, GOME is ~35% less sensitive to the HCHO column over Tennessee than over the North Pacific.

Figure 4a shows the mean geographical distribution of AMFs for July 1996. AMFs for much of the continental

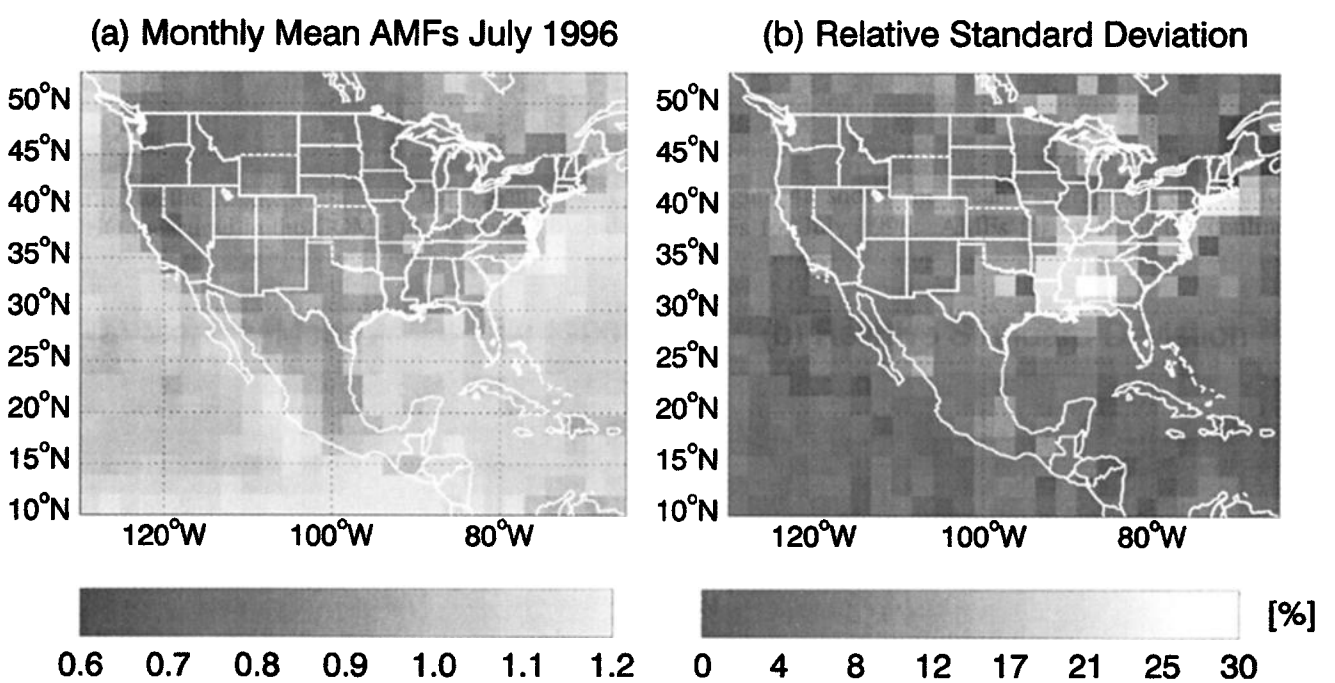

Figure 4. Distribution of AMFs for HCHO column retrievals over North America in July 1996. An AMF was calculated for each GOME observation scene using local values of the shape factor $S_\sigma(\sigma)$ from the GEOS-CHEM model and for cloud cover <40%. (a) Monthly mean AMF values and (b) corresponding relative standard deviations, expressed as percentages of the means.

United States are less than unity (0.70−1); in contrast, most of the ocean has AMF values of more than unity (1−1.2). The lower AMF values over the continent reflect the steeper shape factors (Figure 3). The particularly small AMF values over California are due to a shallow boundary layer, resulting from strong subtropical subsidence, combined with a strong surface source of HCHO from biogenic hydrocarbons. Further inland, in arid regions of the western United States, one would expect relatively high AMFs owing to deep boundary layers (up to 5 km) and low biogenic NMHC emissions. In fact, the AMF values are slightly less than over the eastern United States, an artifact introduced by the 3-hour temporal resolution of mixed-layer height (MLH) in the GEOS meteorological data used in the GEOS-CHEM model. At the 1030 LT sampling time of GOME, the MLHs from the GEOS data correspond to 1000−1300 LT at 75°W as compared to 0800−1100 LT at 105°W. During midmorning in summer the MLH typically grows by several hundred meters per hour, so that updating the MLH only every 3 hours can introduce bias in the model shape factor.

Figure 4b shows the relative standard deviation of the AMFs calculated from the temporal variability of the July 1996 values about the local means of Figure 4a. Regions with the largest AMF variability (15−25%) are over the southeastern United States and off the northeastern coast. The large AMF variability over the southeastern United States is due to variability of isoprene emissions driven in the model by changes in surface temperature and solar radiation [*Guenther et al.*, 1995]. Off the northeastern coast of the U.S. the AMF variability is mainly due to events of continental outflow, as seen in the NARE 1997 observations (Figure 2). Overall, however, the relative standard deviations of the AMFs are small (averaging 8% for the domain in Figure 4), implying that temporal variability in the model behavior has little influence on the retrieval.

3.4. Results

Plate 1 shows the HCHO columns obtained by applying our AMF formulation to GOME observations over North America (10−60°N, 65−130°W) for July 1996. The GEOS-CHEM simulation used to obtain vertical shape factors used assimilated meteorological observations for July 1996 and was sampled for the specific locations and days of the individual GOME observations. Also shown in Plate 1 are the corresponding mean GEOS-CHEM columns. The observed HCHO field is filtered to exclude data where cloud fraction (a GOME operational data product) is >0.4 [*Chance et al.*, 2000]; the same filter is applied to the model field. Figure 5 shows the frequency distribution of cloud fractions for the 124746 GOME observations made in July 1996 over the North America domain defined by the geographical range of Plate 1. By filtering out observations with cloud fractions >0.4, we retain 62% of the total number of observations. The cloud fraction threshold of 0.4 represents a trade-off between ensuring quality of the data used while retaining a reasonable number of observations. We find that using alternate cloud fraction thresholds of 0.3 or 0.5 does not alter significantly the mean HCHO slant column statistics over either continental or oceanic regions. We also find little correlation between the slant columns and cloud fraction (r = -0.0015 for cloud fractions <0.5). It thus appears that the choice of the cloud threshold does not induce bias in the monthly mean data reported in Plate 1, despite the evident effect of clouds on the scattering properties of the atmosphere.

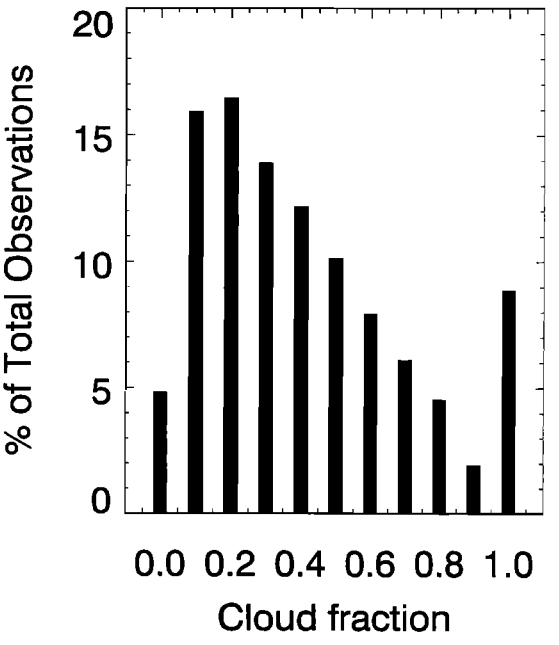

Figure 5. Frequency distribution of cloud fractions for the ensemble of GOME observations during July 1996 over the North America domain defined by the geographical range of Plate 1. There are 124,746 observations included in this frequency distribution.

The slant columns observed by GOME feature a strong enhancement over the southeastern United States which is captured by the model and reflects the strong isoprene emissions in that region. The Pearson correlation coefficient between the slant columns and the GEOS-CHEM model fields for the data shown in Plate 1a and 1b is r = 0.70. Considering observations only with values greater than the HCHO fitting uncertainty of 4×10^{15} molecules cm^{-2} does not alter significantly this correlation. The GEOS-CHEM model thus captures half of the spatial variance in the observed HCHO slant columns before any model information has been applied to the observations through the AMF correction. The sharp gradient between the land and the ocean shown by the model columns (Plate 1b) is not as clear in the observed slant columns, in part because of occasional high events in the observations over the oceans. Temporal variability of observed slant columns is typically >50% for North America, much larger than the 8% variability of the AMF correction (Figure 4b). Most of the temporal variability in the vertical HCHO column, after the AMF correction is applied, is thus due to the slant column fitting rather than to the AMF correction.

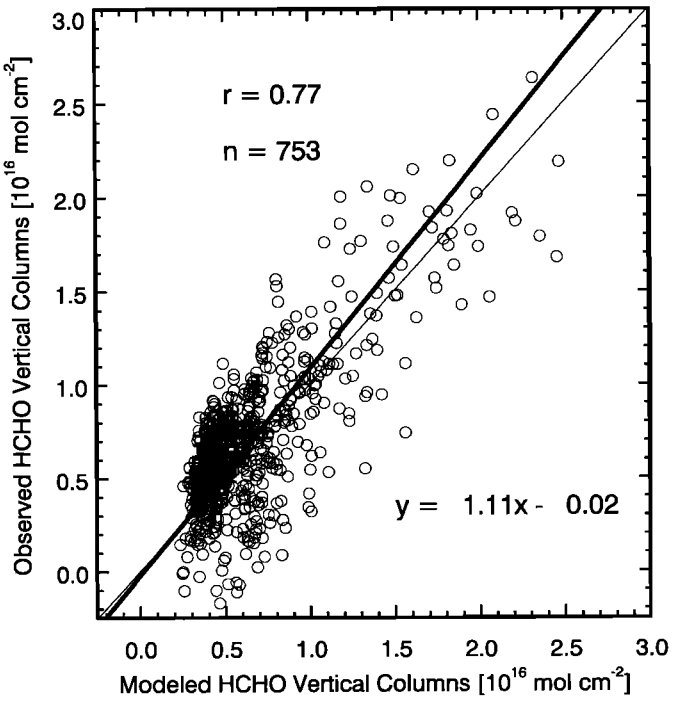

Figure 6. Scatter plot of observed (GOME) and modeled (GEOS-CHEM) monthly mean HCHO vertical columns in July 1996 over the 2° × 2.5° grid shown by Plate 1. The data plotted here are those of Plate 1c (GOME) and those of Plate 1b (GEOS-CHEM). The thick and thin solid lines represent the reduced major axis line [*Hirsch and Gilroy*, 1984] and the $y = x$ relation, respectively. The Pearson correlation coefficient r and the number of elements n used to compute r are also shown.

The vertical columns obtained after the application of the AMFs (Plate 1c) retain the spatial structure of the fitted slant columns, as expected, with typical differences of 20% relative to the slant columns (Figure 4a). Pattern differences between the slant and vertical columns reflect the introduction of information from the GEOS-CHEM model to define the AMF. Enhanced HCHO near 40°N over the west coast of the United States in the vertical columns (not shown by the observed slant columns) is due to shallow mixed layers and represents an extreme example of the fitted slant columns being amplified by the AMF.

Figure 6 shows a scatter plot of the July 1996 mean HCHO vertical columns as retrieved from GOME and as simulated by the GEOS-CHEM model. The correlation of the modeled and observed vertical columns is $r = 0.77$. This apparent improvement over the correlation with the slant columns ($r = 0.70$) reflects, in part, model information passed on to the observed vertical columns through the AMF. Higher HCHO columns in the model reflect boundary layer enhancements and hence lower AMFs, thus bringing the original spatial structure of the fitted slant columns closer to that simulated by the model.

The use of model fields in the AMF calculation thus plays a role in the enhancement of the correlation between the modeled and observed vertical columns. To test the degree to which the modeled and observed vertical columns can be thought of as independent, we considered a hypothetical field of uniform slant columns over the North America domain, applied the AMFs of Figure 4a to convert to vertical columns, and compared the resulting horizontal structure of the vertical columns with the GEOS-CHEM model fields of Plate 1b. We found that the introduction of the AMF describes 16% ($r = 0.4$) of the spatial variance shown by the modeled monthly mean vertical columns. This value agrees well with the increase in the fraction of the spatial variance captured by the model for the observed vertical columns (59%) as compared to the observed slant columns (49%).

A more objective test of the model skill for simulating HCHO is the absolute difference between the modeled and observed vertical columns, since the introduction of the AMF provides no information on the absolute values of the model columns. A comparison is shown in Plate 1d and also in the scatter plot of Figure 6. Observed vertical columns tend to be higher than the model for much of the range of the observations ($0.3-2.7 \times 10^{16}$ molecules cm^{-2}). A notable exception is the southeastern United States where the model overestimates HCHO by ~30% (8×10^{15} molecules cm^{-2}). Observations over the oceans are consistently higher than the model, but this bias is often within the spectral fitting uncertainty of 4×10^{15} molecules cm^{-2}.

4. Discussion

Spectral measurement of the column concentration of an atmospheric absorber from a nadir-viewing satellite instrument provides one piece of information in the form of a slant column abundance. To convert this quantity into a geophysically useful vertical column quantity, an air mass factor is required, which depends on atmospheric scattering and on the vertical distribution of the absorber. In this paper we have presented a novel method to obtain the AMF for an optically thin absorber. The method decouples the vertical distribution of the absorber (calculated using an atmospheric chemistry model) from the observational sensitivity to scattering (calculated using a radiative transfer model). Our AMF formulation represents an integral over the depth of the atmosphere of the relative vertical distribution of the absorber (shape factor) weighted by sensitivity coefficients from the radiative transfer model (scattering weights). This decoupling of the AMF calculation serves two main purposes. First, it increases flexibility by allowing both modeling components of the AMF calculation (radiative transfer and atmospheric chemistry) to be studied separately. The scattering weights can be computed independently of the chemistry model and then tabulated, allowing the AMF to be readily derived for any vertical profile. Second, using shape factors from an atmospheric chemistry model to formulate the AMF ensures consistency for subsequent evaluation of the model with the retrieved vertical columns.

As an illustration, we have applied the method to HCHO slant column observations from GOME over North America in July 1996. We have used the GEOS-CHEM 3-D global

chemistry model driven by assimilated meteorological data for July 1996 to provide vertical shape factors for every observation scene, after showing, through comparisons with the sparse aircraft observations in continental and marine atmospheres, that the model has some success in reproducing observed shape factors. The model resolves much of the spatial structure exhibited by the observed slant columns. The monthly mean HCHO field from the GEOS-CHEM model captures half of the observed variance shown by the fitted slant columns ($r = 0.70$) before introduction of any model information through the AMF. The model captures more of the variance ($r = 0.77$) of the observed vertical columns after the AMF application, although this improvement largely reflects a negative correlation between the AMF and the vertical column. The model overestimates HCHO vertical columns by ~30% over the southeastern United States, where the HCHO column is essentially a tracer of isoprene emission. It generally underestimates observations elsewhere. Further work will focus on more detailed model evaluation and on comparison of the GOME data with in situ measurements. We expect that this work will provide improved constraints on NMHC emission inventories, in particular for isoprene.

Further work will also focus on improving cloud and aerosol effects on the AMF determination. Clouds mask HCHO from GOME. We have avoided dealing with the effects of clouds in this work by filtering the GOME data to exclude observations with cloud fractions >40%. This threshold does not seem to introduce significant contamination in the retrievals, as inferred from comparison of results with cloud thresholds of 30% or 50%. Better definition of cloud fractions and cloud heights could perhaps yield some information on the HCHO vertical distribution from the slant column retrievals.

Aerosols may either enhance or reduce the GOME sentivity to HCHO, depending on their optical properties and their vertical distribution relative to that of HCHO. Results from a sensitivity analysis presented in this paper indicate that boundary layer aerosols collocated with HCHO and with a single-scattering albedo of 0.96 cause an increase in the AMF. The effect is <10% for typical values of aerosol optical thickness ($\tau_{346\ nm} < 1$) but may be as high as 40% under particularly polluted and hazy conditions ($\tau_{346\ nm} \simeq 2$). Inclusion of aerosols in the atmospheric chemistry model would allow for accouting of their effect on the AMF in a manner consistent with the model-derived HCHO shape factor.

The focus of this paper was the development of the AMF methodology, and we presented only a limited application to GOME HCHO observations over North America in summer. Extension of the HCHO retrievals to the global scale and several years should provide new insights on the spatial and seasonal distribution of isoprene emissions and on episodic emissions from biomass burning [*Thomas et al.*, 1998; *Holzinger et al.*, 1999].

Acknowledgments. This research is supported at Harvard University and the Smithsonian Astrophysical Observatory by the NASA Atmospheric Chemistry Modeling and Analysis Program. Additional support was provided by Smithsonian Institution internal funds. R.V. Martin was supported by a National Science Foundation Graduate Fellowship. We thank P. Cameron-Smith for useful discussions of the ideas presented. We thank two anonymous reviewers who provided thorough and thoughtful comments.

References

Buchwitz, M., V. V. Rozanov, and J. P. Burrows, A near-infrared optimized DOAS method for the fast global retrieval of atmospheric CH_4, CO, CO_2, H_2O, and N_2O total column amounts from SCIAMACHY Envisat-1 nadir radiances, *J. Geophys. Res.*, *105*, 15,231–15,246, 2000.

Burrows, J P., et al., The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results, *J Atmos. Sci.*, *56*, 151–175, 1999.

Chance, K , and R. J. D. Spurr, Ring effect studies: Rayleigh scattering, including molecular parameters for rotational Raman scattering, and the Fraunhofer spectrum, *Appl. Opt.*, *36*, 5224–5230, 1997

Chance, K., P. I. Palmer, R. J. D. Spurr, R. V. Martin, T. P. Kurosu, and D. J. Jacob, Satellite observations of formaldehyde over North America from GOME, *Geophys. Res. Lett.*, *27*, 3461–3464, 2000.

Chandrasekhar, S., *Radiative Transfer*, Dover Publications Inc., N.Y., 1960.

Dickerson, R. R., S. Kondragunta, G. Stenchikov, K. L. Civerolo, B. G. Doddridge, and B. N. Holben, The impact of aerosols on solar ultraviolet radiation and photochemical smog, *Science*, *278*, 827–830, 1997.

Griffin, R. J , D. R. Cocker III, R. C. Flagan, and J. H. Seinfeld, Organic aerosol formation from the oxidation of biogenic hydrocarbons, *J. Geophys. Res.*, *104*, 3555–3567, 1999.

Guenther, A , et al., A global model of natural volatile organic compound emissions, *J. Geophys. Res.*, *100*, 8873–8892, 1995.

Herman, J. R., and E. A. Celarier, Earth surface reflectivity climatology at 340-380 nm from TOMS data, *J. Geophys. Res.*, *102*, 28,003–28,011, 1997.

Hirsch, R. M., and E. J. Gilroy, Methods of fitting a straight line to data: examples in water resourses, *Water Res. Bull.*, *20*, 705–711, 1984.

Holzinger, R., C. Warneke, A. Hansel, A. Jordan, and W. Lindinger, Biomass burning as a source of formaldehyde, acetaldehyde, methanol, acetone, acetonitrile, and hydrogen cyanide, *Geophys. Res. Lett.*, *26*, 1161–1164, 1999.

Horowitz, L W., and D. J. Jacob, Global impact of fossil fuel combustion on atmospheric NO_x, *J. Geophys. Res.*, *104*, 23,823–23,840, 1999.

Jacobson, M. Z., and R. P. Turco, SMVGEAR: A sparse-matrix, vectorized Gear code for atmospheric models, *Atmos. Environ.*, *28*, 273–284, 1994.

Lee, Y.-N., et al., Atmospheric chemistry and distribution of formaldehyde and several multioxygenated carbonyl compounds during the 1995 Nashville/Middle Tennessee Ozone Study, *J. Geophys. Res.*, *103*, 22,449–22,462, 1998.

Platt, U., *Differential Optical Absorption Spectroscopy (DOAS)*, in Air Monitoring by Spectroscopic Techniques, pp. 27–84, John Wiley, New York, 1994.

Rozanov, V. V., T. Kurosu, and J. P. Burrows, Retrieval of atmospheric constituents in the UV-Visible: A new quasi-analytical approach for the calculation of weighting functions, *J. Quant. Spectrosc. Radiat. Transfer*, *60*, 277–299, 1998.

Schubert, S. D., R. B. Rood, and J. Pfaendtner, An assimilated data set for Earth Science applications, *Bull. Am. Meteorol. Soc.*, *74*, 2331–2342, 1993.

Sillman, S., The relation between ozone, NO_x and hydrocarbons in urban and polluted rural environments, *Atmos. Environ.*, *33*, 1821–1845, 1999.

Singh, H. B., et al., Distribution and fate of selected oxygenated organic species in the troposphere and lower stratosphere over the Atlantic, *J. Geophys. Res.*, *105*, 3795–3805, 2000.

Spurr, R. J. D., T. P. Kurosu, and K. V. Chance, A linearized discrete ordinate radiative transfer model for atmospheric remote sensing retrieval, *J. Quant. Spectrosc. Radiat. Transfer*, in press, 2001.

Stammes, P., and R. B. A. Koelemeijer, Error analysis of the GOME ozone column retrieval method, in *Proceedings of ESAMS'99 - European Symposium on Atmospheric Measurements from Space*, *Publ. WPP-161*, pp. 201–207, European Space Agency, 1999.

Thomas, W., E. Hegels, S. Slijkhuis, R. Spurr, and K. Chance, Detection of biomass burning combustion products in Southeast Asia from backscatter data taken by the GOME spectrometer, *Geophys. Res. Lett.*, *25*, 1317–1320, 1998.

Wang, Y., D. J. Jacob, and J. A. Logan, Global simulation of tropospheric O_3-NO_x-hydrocarbon chemistry, 1, Model formulation, *J. Geophys. Res.*, *103*, 10,713–10,726, 1998a.

Wang, Y., D. J. Jacob, and J. A. Logan, Global simulation of tropospheric O_3-NO_x-hydrocarbon chemistry, 3, Origin of tropospheric ozone and effects of nonmethane hydrocarbons, *J. Geophys. Res.*, *103*, 10,757–10,768, 1998b.

Wang, Y., J. A. Logan, and D. J. Jacob, Global simulation of tropospheric O_3-NO_x-hydrocarbon chemistry, 2, Model evaluation and global ozone budget, *J. Geophys. Res.*, *103*, 10,727–10,756, 1998c.

Wild, O., X. Zhu, and M. J. Prather, Fast-J: Accurate simulations of in- and below-cloud photolysis in tropospheric chemistry models, *J. Atmos. Chem.*, *37*, 245–282, 2000.

I. Bey, A. Fiore, D. J. Jacob, Q. Li, R. V. Martin, P. I. Palmer, and R. Yantosca, Division of Engineering and Applied Sciences, and Department of Earth and Planetary Sciences, Harvard University, Pierce Hall, 29 Oxford Street, Cambridge, MA 02138 (bey, amf; djj; qli; rvm; pip; bmy@io.harvard.edu)

K. Chance, T. P. Kurosu, and R. J. D. Spurr, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138 (kchance; tkurosu; rspurr@cfa.harvard.edu)

(Received August 2, 2000; revised October 27, 2000;
accepted November 30, 2000.)