Formaldehyde Distribution over North America: Implications for Satellite Retrievals of Formaldehyde Columns and Isoprene Emission

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation

Published Version
doi:10.1029/2005JD006853

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:3743792

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Formaldehyde distribution over North America: Implications for satellite retrievals of formaldehyde columns and isoprene emission

Dylan B. Millet,1 Daniel J. Jacob,1 Solène Turquety,1,2 Rynda C. Hudman,1 Shiliang Wu,1 Alan Fried,3 James Walega,3 Brian G. Heikes,4 Donald R. Blake,5 Hanwant B. Singh,6 Bruce E. Anderson,7 and Antony D. Clarke8

Received 5 November 2005; revised 20 January 2006; accepted 27 March 2006; published 8 September 2006.

[1] Formaldehyde (HCHO) columns measured from space provide constraints on emissions of volatile organic compounds (VOCs). Quantitative interpretation requires characterization of errors in HCHO column retrievals and relating these columns to VOC emissions. Retrieval error is mainly in the air mass factor (AMF) which relates fitted backscattered radiances to vertical columns and requires external information on HCHO, aerosols, and clouds. Here we use aircraft data collected over North America and the Atlantic to determine the local relationships between HCHO columns and VOC emissions, calculate AMFs for HCHO retrievals, assess the errors in deriving AMFs with a chemical transport model (GEOS-Chem), and draw conclusions regarding space-based mapping of VOC emissions. We show that isoprene drives observed HCHO column variability over North America; HCHO column data from space can thus be used effectively as a proxy for isoprene emission. From observed HCHO and isoprene profiles we find an HCHO molar yield from isoprene oxidation of 1.6 ± 0.5, consistent with current chemical mechanisms. Clouds are the primary error source in the AMF calculation; errors in the HCHO vertical profile and aerosols have comparatively little effect. The mean bias and 1σ uncertainty in the GEOS-Chem AMF calculation increase from <1% and 15% for clear skies to 17% and 24% for half-cloudy scenes. With fitting errors, this gives an overall 1σ error in HCHO satellite measurements of 25–31%. Retrieval errors, combined with uncertainties in the HCHO yield from isoprene oxidation, result in a 40% (1σ) error in inferring isoprene emissions from HCHO satellite measurements.


1. Introduction

[2] Atmospheric volatile organic compounds (VOCs) are precursors of tropospheric ozone and secondary organic aerosol, and play an important role in HOx and NOx radical cycling. Bottom-up VOC emission estimates, relying on extrapolation of point measurements to regional and larger scales, are inherently uncertain. Formaldehyde (HCHO), which is produced in high yield during the atmospheric oxidation of VOCs, absorbs in the near-UV and can be measured as a column integral from satellite-borne solar backscatter instruments [Burrows et al., 1999; Chance et al., 2000]. Such measurements offer the means to derive global, top-down constraints on surface emissions of VOCs. In order to do so reliably we need to quantify the errors associated with the column measurements and their relationship to precursor emissions. Here we use 69 aircraft vertical profiles over North America collected during the Intercontinental Chemical Transport Experiment (INTEX-A) aircraft campaign in summer 2004 (H. B. Singh et al., Overview of the summer 2004 Intercontinental Chemical Transport Experiment-North America (INTEX-A), submitted to Journal of Geophysical Research, 2006) to quantify the dominant errors in satellite retrievals of HCHO columns, and determine how the HCHO column and its variability can be interpreted in terms of the underlying reactive VOC emissions.

[3] While oxidation of methane is the main HCHO source in the remote atmosphere [Lowe and Schmidt, 1983; Singh et al., 2000; Heikes et al., 2001; Frost et al., 2002; Wagner et al., 2002], more reactive VOCs frequently dominate in

Copyright 2006 by the American Geophysical Union.
0148-0227/06/2005JD006853$09.00
the continental boundary layer [Lee et al., 1998; Fried et al., 2003]. The main HCHO sinks are reaction with the hydroxyl radical and photolysis [Levy, 1972], which combine to give an atmospheric lifetime of a few hours in daytime [Atkinson, 2000].

[4] Isoprene, a highly reactive compound which is the principal VOC emitted from vegetation [Zimmerman et al., 1978; Fehsenfeld et al., 1992; Fuentes et al., 2000; Guenther et al., 2000], is in particular a major source of HCHO during the growing season [Zimmerman et al., 1978; Hanst et al., 1980; Shepson et al., 1991; Lee et al., 1998].

Globally, emissions of isoprene are estimated to be around 500 Tg yr⁻¹ [Guenther et al., 1995; Wang and Shallcross, 2000; Potter et al., 2001; Levis et al., 2003; Sanderson et al., 2003; Naik et al., 2004; Guenther et al., 2006], substantially greater than the total VOC emissions from all anthropogenic sources [Singh and Zimmerman, 1992].

Even over densely populated and industrialized eastern North America, on regional scales isoprene far exceeds anthropogenic VOC emissions during the growing season [Jacob et al., 1993] and is a major source of surface ozone [Fiore et al., 2005] and also possibly organic aerosol [Jang et al., 2002; Czoschke et al., 2003; Limbeck et al., 2003; Claeys et al., 2004a, 2004b; Zhang et al., 2004; Lim et al., 2005].

[5] Despite considerable efforts in constructing bottom-up emission inventories for isoprene and other biogenic VOCs, important uncertainties remain because of the need to extrapolate over vegetation types, the complex effects of environmental stressors, and evolving land cover [Guenther et al., 2000]. Palmer et al. [2001, 2003] developed a top-down approach for inferring isoprene emission fluxes using space-based column measurements of HCHO, and applied it to derive emissions from North America using data from the Global Ozone Monitoring Experiment (GOME) satellite instrument. This work has since been extended to examine seasonal and interannual variability in North American isoprene emissions [Abbot et al., 2003; Palmer et al., 2006].

Recently, Shim et al. [2005] carried out a Bayesian inversion of GOME HCHO column measurements for different continental source regions, and derived global isoprene emissions 50% larger than the a priori estimate.

[6] Work to date in this area has used data from GOME, which has 40 km × 320 km resolution and global coverage every three days. The Ozone Monitoring Instrument (OMI), which was launched in 2004 aboard the NASA Aura satellite, provides daily global coverage and a footprint of 13 × 24 km. HCHO columns measured from OMI should enable us to quantify surface fluxes of VOCs at a far greater level of temporal and spatial detail than is possible with GOME. The validity of such analyses depends on the uncertainties associated with the retrieval of HCHO vertical columns from the satellite spectra and with their interpretation in terms of VOC emission. The dominant source of error in the retrieval is the air mass factor (AMF) [Palmer et al., 2006], which defines the relationship between the HCHO abundance along the viewing path of the satellite instrument ("slant column") and the vertical column amount [Noxon et al., 1979; Perliski and Solomon, 1993; Marquard et al., 2000; Palmer et al., 2001; Hild et al., 2002; Richter and Burrows, 2002; Boersma et al., 2004].

The AMF calculation requires external information on atmospheric scattering by air molecules, clouds and aerosols, on the shape of the HCHO vertical distribution, and on the UV albedo of the surface. Clouds, which increase instrument sensitivity to the absorber above the cloud while decreasing it below, represent a significant source of uncertainty in the computation of the AMF [Koelmeijeer and Stamnes, 1999; Velders et al., 2001; Richter and Burrows, 2002; Martin et al., 2003a; Boersma et al., 2004]. The interpretation of observed HCHO columns in terms of precursor VOC emissions also requires prior information on the relationship between the VOC surface flux and the resulting HCHO column amount.

[7] INTEX-A included numerous aircraft profiles over North America. The resulting data set, which includes measurements of HCHO together with VOCs, aerosol extinction, and cloud extinction, provides us with an excellent opportunity to go beyond previous work and quantify the errors involved in relating satellite HCHO measurements to VOC emissions. In the present study, we use the INTEX-A aircraft data and output from the GEOS-Chem chemical transport model (CTM) to (1) determine the important precursors contributing to HCHO columns and variability over North America, and quantify the relationships between precursor emissions and HCHO columns; (2) carry out a statistically meaningful and geographically extensive quantification of the errors in the AMF calculation; and (3) draw conclusions regarding the mapping of VOC emissions from space.

2. Background

2.1. Satellite Retrievals of HCHO Columns

[8] The retrieval of atmospheric HCHO column abundance using space-borne solar backscatter instruments can be performed by fitting the backscattered spectrum in the HCHO absorption window (337–356 nm) to modeled atmospheric spectra [e.g., Chance et al., 2000], or by differential optical absorption spectroscopy [e.g., Leue et al., 2001]. The resulting HCHO abundance integrated along the viewing path is called the slant column, and the ratio of the slant column to the actual vertical column is termed the AMF. In the case of a nonscattering atmosphere, the geometric air mass factor AMFG would be determined solely by the satellite viewing angle (θv) and the solar zenith angle (θs):

\[
AMF_G = \sec \theta_s + \sec \theta_v. \tag{1}
\]

This simple expression must be corrected for scattering by air molecules, clouds, and aerosols, which results in sensitivity to the vertical distribution of the absorbing gas and to the surface albedo. The correction factor can be expressed in the optically thin case as an integral of sensitivity over the depth of the atmosphere [Palmer et al., 2001]:

\[
AMF = \frac{AMF_G}{P_s} \int_{P_s}^0 w(P)S(P)dP, \tag{2}
\]

where \(P\) is pressure and \(P_s\) is the pressure at Earth’s surface. The scattering weights \(w(P)\) represent the sensitivity of the
backscattered radiation measured from space to the abundance of the absorber (here HCHO) at pressure \( P \), and are determined using a radiative transfer model. The shape factor \( S(P) \) is the normalized vertical profile of mixing ratio of the absorber, and is determined from typical atmospheric observations or an atmospheric CTM. Aerosol vertical profiles for the radiative transfer calculation are similarly specified from climatologies [Velders et al., 2001; Richter and Burrows, 2002; Beirle et al., 2004; Richter et al., 2004; Savage et al., 2004], or from the same CTM used to specify \( S(P) \) [Martin et al., 2003a; Jaegle et al., 2004]. Like clouds, aerosols can act to either increase or decrease the instrument sensitivity to HCHO, depending on the single scattering albedo of the aerosol and its vertical distribution relative to the absorber.

[9] Satellite viewing scenes are typically partly cloudy, particularly for an instrument with a large footprint such as GOME. The ability to resolve partly cloudy scenes in the retrieval is thus critical to the data coverage. In the case of GOME, broadband polarization monitoring devices provide information on cloud fraction, and spectral fitting in and around the oxygen A-band provides information on cloud top altitude and cloud optical depth [Kurosu et al., 1999; Koelmeijer et al., 2001]. Previous GOME retrievals for HCHO or NO\(_2\) have either assumed cloud effects to be negligible for pixels having cloud fractions below a given threshold [Palmer et al., 2001; Lauer et al., 2002; Richter and Burrows, 2002; Ladstatter-Weissenmayer et al., 2003; Palmer et al., 2003; Stohl et al., 2003; Beirle et al., 2004; Richter et al., 2004; Savage et al., 2004; Irie et al., 2005; Konovalov et al., 2005], applied a constant correction factor to cloudy scenes [Leue et al., 2001; Velders et al., 2001; Beirle et al., 2003], or calculated the radiative transfer through clouds explicitly using local cloud information [Martin et al., 2002a; Abbot et al., 2003; Martin et al., 2003a; Jaegle et al., 2004; Palmer et al., 2006].

[10] In the latter approach, developed by Martin et al. [2002a], the AMF for a partly cloudy scene is derived as the weighted sum of the values for the clear and cloudy subscenes:

\[
\text{AMF} = \frac{\text{AMF}_{\text{clear}} R_{\text{clear}} (1 - f) + \text{AMF}_{\text{cloud}} R_{\text{cloud}} f}{R_{\text{clear}} (1 - f) + R_{\text{cloud}} f} \quad (3)
\]

where \( f \) is the cloud fraction, and \( R_{\text{clear}} \) and \( R_{\text{cloud}} \) are the reflectivities of the clear and cloudy portions of the retrieval scene. The two subscenes are assumed to have the same HCHO shape factor \( S(P) \). Martin et al. [2002a] and the subsequent studies calculated the AMF for the cloudy subscene by using the GOME information on cloud top and optical depth, and distributing the cloud extinction vertically assuming an optical thickness increment of 8 for each 100 hPa below cloud top.

[11] The primary uncertainties in space-borne measurements of HCHO columns arise from the slant column fitting [Chance et al., 2000], which defines the instrumental detection limit, and errors in the AMF, which define a relative error for columns well above the detection limit [Palmer et al., 2001; Martin et al., 2002a; Boersma et al., 2004; Palmer et al., 2006]. The slant column fitting uncertainty is \( 4 \times 10^3 \) molecules cm\(^{-2} \) for GOME. The main sources of error in the AMF are the surface albedo, specification of the HCHO vertical profile, and aerosol and cloud effects, and quantification of this error is a focus of our paper.

### 2.2. Relating HCHO Columns to Precursor Emissions

[12] Column HCHO measurements from space can be used to derive precursor emission fluxes [Palmer et al., 2003], provided the associated HCHO production yields are known. At steady state and in the absence of horizontal transport, the HCHO column \( (\Omega_{\text{HCHO}} \text{ molecules cm}^{-2}) \) would be related to the emissions of precursors \( i \) by

\[
\Omega_{\text{HCHO}} = \frac{1}{k_{\text{HCHO}}} \sum_i k_i Y_i \Omega_i = \frac{1}{k_{\text{HCHO}}} \sum_i Y_i E_i, \quad (4)
\]

where \( k_{\text{HCHO}} \) and \( k_i \) are the column-average effective rate constants (s\(^{-1} \)) for chemical loss of HCHO and precursor \( i \), \( Y_i \) is the molar HCHO yield from the oxidation of species \( i \), and \( E_i \) is the emission flux. Horizontal transport smears this relationship, resulting in a spatial offset between the HCHO column and the location of precursor emission, and diluting the HCHO signal associated with the emission [Palmer et al., 2003]. For isoprene, which has a lifetime in summer of \( \sim 0.5 \) h and yields HCHO in its first generation of products, the smearing length scale is only \( 10-100 \) km, smaller than the GOME pixel size [Palmer et al., 2003]. For longer-lived VOCs or VOCs with delayed HCHO production, the smearing length scale may be sufficiently large to dilute the HCHO signal to below the fitting uncertainty [Palmer et al., 2006]. This smearing effect will limit gains in spatial resolution otherwise achievable with new satellite instruments unless inverse methods are developed that account for fine-scale VOC transport and chemistry.

[13] Inferring VOC emission fluxes from observed HCHO columns requires assumptions about the HCHO yield \( Y_i \). Palmer et al. [2006] computed the time-dependent HCHO yield from isoprene oxidation using two independent mechanisms, GEOS-Chem (http://www.as.harvard.edu/chemistry/trop/geos/index.html) [Horowitz et al., 1998; Evans and Jacob, 2005] and the Master Chemical Mechanism (MCM) v. 3.1 [Bloss et al., 2005]. They found the yield calculated by MCM to be 20–30% higher than that derived by GEOS-Chem. The resulting HCHO molar yields after one day under low-NO\(_x\) (0.1 ppb) conditions are 0.9 (GEOS-Chem) and 1.6 (MCM). The high-NO\(_x\) (1 ppb) molar yields are 1.9 (GEOS-Chem) and 2.4 (MCM).

### 3. Approach

[14] The previous section has highlighted a number of uncertainties in the derivation of VOC emissions from space-based HCHO column measurements. In this section we describe our use of the INTEX-A aircraft measurements to better quantify these uncertainties and thus improve the constraints on the top-down VOC emission estimates. We use the GEOS-Chem CTM as the source of external information here but our findings can be applied to any CTM-assisted retrieval.

#### 3.1. Measurements

[15] The primary objective of INTEX-A (1 July to 15 August 2004) was to observe the chemical outflow from North America and infer constraints on chemical sources
and export. Here we use measurements of HCHO and related tracers made aboard the NASA DC-8 aircraft (ceiling 12 km) over North America and the adjacent oceans during INTEX-A. The DC-8 aircraft flew 18 science flights between 1 July and 15 August with extensive vertical profiling from the boundary layer to the upper troposphere. All flights took place during daytime, typically from 1000 to 1800 local time. The flight tracks are shown in Figure 1.

HCHO measurements were carried out by two groups, from the National Center for Atmospheric Research (NCAR) and the University of Rhode Island (URI) (hereinafter referred to as the NCAR and URI measurements, respectively). HCHO was measured by the NCAR group using tunable diode laser absorption spectroscopy [Wert et al., 2003; Rollert et al., 2006; A. Fried et al., The role of convection in redistributing formaldehyde to the upper troposphere over North America and the North Atlantic during the summer 2004 INTEX campaign, manuscript in preparation, 2006, hereinafter referred to as Fried et al., manuscript in preparation, 2006]. The total calibration uncertainty is estimated to be ±12% (2σ), and the limit of detection (LOD) is 77 ppt for the first part of the mission and 66 ppt for the last 7 flights (31 July 2004 to 14 August 2004), for 1-min averaged data (both 2σ). The measurement precision is the same as the LOD over the HCHO mixing ratios measured during INTEX-A. The URI HCHO measurement was performed using aqueous collection and enzyme-fluorescence detection [Heikes et al., 2001]. The LOD is 50 ppt for the first part of the mission, and 25 ppt for the last 5 flights (6–14 August 2004). The total uncertainty in the measurement is estimated at ± (33 ppt + 0.15[HCHO mixing ratio]). Measurements from the two groups were highly correlated over the ensemble of the INTEX-A mission (R² = 0.89). However, a reduced major axis regression of the two data sets yields a slope of 0.69, with the URI data being lower than the NCAR data. A detailed measurement intercomparison is presented elsewhere [Fried et al., manuscript in preparation, 2006].

Oxygenated VOCs were measured with a sampling frequency of 2.5–5 min using cryogenic preconcentration, gas chromatographic (GC) separation, and detection by photoionization detector and reduction gas detector [Singh et al., 2004]. Detection limits range from 5 to 20 ppt, and the sensitivity and precision of measurement are approximately 20% and 10%. Hydrocarbons were measured by whole air sampling followed by cryogenic preconcentration, GC separation, and detection by flame ionization detector and mass selective detector [Colman et al., 2001; Blake et al., 2003]. Calibration was based on whole air standards (for <C₆ gases) and per-carbon response factors (for C₆–C₁₀ gases). The limit of detection is approximately 3 ppt for the species reported here. The measurement precision and overall accuracy are 1–3% and 2–10%, respectively, depending on the compound.

Aerosol scattering coefficients were measured at 450, 550, and 700 nm using two TSI 3563 three-wavelength integrating nephelometers. The value at 346 nm was estimated from the Angstrom exponent derived from the 450 nm and 550 nm measurements. Aerosol absorption coefficients at 470, 530 and 660 nm were measured using a pair of Radiance Research Particle Soot Absorption Photometers [Bond et al., 1999; Virkkula et al., 2005]; the absorption coefficient at 346 nm was assumed equal to that at 470 nm.

3.2. Model Description

Atmospheric distributions of HCHO and related tracers were simulated for the INTEX-A period using the GEOS-Chem global 3D CTM [Bey et al., 2001; Park et al., 2004]. The GEOS-Chem CTM (version 7.02, http://www.as.harvard.edu/chemistry/trop/geos/index.html) uses GEOS-4 assimilated meteorological data from the NASA Goddard Earth Observing System including winds, convective mass fluxes, mixing depths, temperature, precipitation, and surface properties. The data have 6-hour temporal resolution (3-hour for surface variables and mixing depths), 1° × 1.25° horizontal resolution, and 55 vertical layers. We degrade the horizontal resolution to 2° × 2.5° for input to GEOS-Chem.

The model includes detailed ozone-NOₓ-VOC chemistry coupled to aerosols. Global emissions are as described by Bey et al. [2001], with recent updates [Martin et al., 2002b; Park et al., 2004; Xiao et al., 2004]. Isoprene emissions are calculated using the GEIA inventory [Guenther et al., 1995]. Anthropogenic emissions from North America are based on the EPA NEI 1999 v.1 inventory (http://www.epa.gov/ttn/chief/net/1999inventory.html). The aerosol simulation is as described by Park et al. [2003, 2004], and includes sulfate-nitrate-ammonium, organic and black carbon, size-resolved soil dust, and size-resolved sea salt. Aerosol optical thicknesses are calculated from the mass concentrations and optical properties for each aerosol type, as a function of local relative humidity, following Martin et al. [2003b]. Previous studies have compared GEOS-Chem results with observations over North America and the North Atlantic for HCHO [Singh et al., 2004].
et al., 2000; Palmer et al., 2001; Fiore et al., 2002; Martin et al., 2004] and aerosols [Park et al., 2003, 2004; A. van Donkelaar et al., Model evidence for a significant source of secondary organic aerosol from isoprene, submitted to Atmospheric Environment, 2006]. The current simulation was conducted at $2^\circ \times 2.5^\circ$ horizontal resolution and with 30 vertical layers. Results are presented here for July and August 2004, and follow a 1 year spin-up.


4. Atmospheric Distribution of HCHO Over North America

4.1. Vertical Distributions

We calculated HCHO vertical columns from observed and simulated HCHO mixing ratios during the DC-8 vertical profiles. Extensive vertical profiling from the boundary layer (~300 m above surface) to the upper troposphere (~10 km) was conducted during the mission.

4.2. Vertical Columns

Figure 2. Mean simulated (lines) and observed (symbols) HCHO vertical distributions during INTEX-A, calculated for 100 hPa vertical bins. Error bars represent observed standard deviations (the numbers of points are indicated on the right). The vertical coordinates for the URI observations are offset slightly for visibility. Here and elsewhere, the model is sampled along the flight tracks and for the time of the measurements.
Here, we define as a vertical profile any flight segment meeting the following criteria: (1) observations extending from below 600 m (1000 m for marine profiles) to above 8 km radar altitude, (2) horizontal drift of less than 3° latitude × 4° longitude, and (3) at least 15 valid measurements. Mixing ratios above and below the profile were estimated by extending the values obtained at the highest and lowest altitudes uniformly to the tropopause and to Earth’s surface. Mixing ratios in the stratosphere are negligible. Modeled columns calculated with these assumptions agree well with the corresponding model calculations for the full columns (slope = 0.97, R² = 0.96). We obtain in this manner 69 total profiles with a mean horizontal drift of ~190 km. Missing observations reduce the number of profiles to 36 for the NCAR HCHO data set and 13 for the URI HCHO data set.

[24] Figure 3 shows the resulting HCHO columns. Measured values range from 0.4 to 3.1 × 10¹⁶ molecules cm⁻² over continental North America, and from 0.4 to 0.8 × 10¹⁶ molecules cm⁻² over the ocean. Again, the bias between the two sets of measurements is manifest. Both modeled and measured columns are highest over the southeast United States, reflecting elevated isoprene emission [Lee et al., 1998] as discussed previously by Palmer et al. [2003]. Prior in situ observations have been too sparse to clearly define this maximum, but it is clearly revealed by the INTEX-A data. Scatterplots of simulated versus observed HCHO vertical columns are displayed in Figure 4. The model captures 70% of the variability in the observed NCAR columns, and 42% of that in the URI columns. The modeled HCHO columns have a bias (given by the slope of the regression line) of +4% compared to NCAR and +34% compared to URI.

5. Relating HCHO Columns to Reactive VOC Emissions

[25] In the following sections we use the data from the INTEX-A aircraft profiles to determine how column HCHO data from space can be interpreted in terms of the underlying reactive VOC emissions. Our first step is to determine which parent VOCs drive the variability in the HCHO column. For this purpose, we compute column integrated HCHO production rates from the precursor VOCs measured
Table 1. Formaldehyde Production Yields

<table>
<thead>
<tr>
<th>Species</th>
<th>Molar HCHO Yield$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biogenics</td>
<td></td>
</tr>
<tr>
<td>Isoprene</td>
<td>2.3</td>
</tr>
<tr>
<td>Alpha-pinene</td>
<td>1.9$^b$</td>
</tr>
<tr>
<td>Beta-pinene</td>
<td>1.7$^c$</td>
</tr>
<tr>
<td>Alkanes</td>
<td></td>
</tr>
<tr>
<td>Methane</td>
<td>1.0</td>
</tr>
<tr>
<td>Ethane</td>
<td>1.1</td>
</tr>
<tr>
<td>Propane</td>
<td>0.6</td>
</tr>
<tr>
<td>≥C4 alkanes</td>
<td>2.0</td>
</tr>
<tr>
<td>Alkenes</td>
<td></td>
</tr>
<tr>
<td>Ethene</td>
<td>1.8</td>
</tr>
<tr>
<td>≥C3 alkenes</td>
<td>2.0</td>
</tr>
<tr>
<td>Aromatics</td>
<td></td>
</tr>
<tr>
<td>Toluene</td>
<td>1.2$^b$</td>
</tr>
<tr>
<td>m-xylene</td>
<td>2.1$^b$</td>
</tr>
<tr>
<td>p-xylene</td>
<td>2.0$^b$</td>
</tr>
<tr>
<td>o-xylene</td>
<td>2.2$^c$</td>
</tr>
<tr>
<td>Ethylbenzene</td>
<td>2.1$^c$</td>
</tr>
<tr>
<td>C9 aromatics</td>
<td>2.3$^c$</td>
</tr>
<tr>
<td>Oxygenated VOCs (OVOCs)</td>
<td></td>
</tr>
<tr>
<td>Acetone</td>
<td>2.0</td>
</tr>
<tr>
<td>Methanol</td>
<td>1.0</td>
</tr>
<tr>
<td>Ethanol</td>
<td>1.0</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>1.0</td>
</tr>
</tbody>
</table>

$^a$Yields are from the GEOS-Chem chemical mechanism (http://www.as.harvard.edu/chemistry/trop/geo/index.html) except where noted. See text for details.

$^b$From MCMv3.1 [Bloss et al., 2005; Palmer et al., 2006].

$^c$Estimated assuming the same per-carbon yield as m-xylene.

aboard the aircraft and relate those to the measured HCHO columns. Here and for the remainder of the paper we restrict our analysis to the NCAR HCHO data set owing to its factor of ~2 higher data coverage.

[26] HCHO yields (high-NO$_x$ conditions) for all measured precursor VOCs with significant emissions are shown in Table 1. Yields and rates are obtained from the GEOS-Chem chemical mechanism, except where noted, and represent cumulative yields from the successive stages of oxidation of the parent compound until a product with a lifetime longer than a few hours is reached. The dependence of HCHO yields on NO$_x$ is discussed by Palmer et al. [2003, 2006]; low-NO$_x$ conditions leading to organic peroxide formation have little effect on ultimate yields if the organic peroxides decompose to regenerate radicals, as is commonly assumed, but the HCHO production is delayed.

[27] Column integrated HCHO production rates ($P_{HCHO}$, molecules cm$^{-2}$ s$^{-1}$) from VOC oxidation by OH were computed from the DC-8 vertical profiles for different classes of measured precursors using rate constant data from Sander et al. [2002] and Atkinson et al. [2004]. HCHO yields from Table 1, and local OH concentrations from the GEOS-Chem model (Hudman et al., submitted manuscript, 2006). VOC vertical profiles were extrapolated below and above the aircraft in the same way as HCHO. Results for isoprene, oxygenated volatile organic compounds (OVOCs), anthropogenic nonmethane organic compounds (ANMHC), the monoterpenes $\alpha$- and $\beta$-pinene, and methane are displayed as probability density functions in Figure 5. The slant column fitting error for the satellite instrument ($4 \times 10^{13}$ molecules cm$^{-2}$ for GOME [Chance et al., 2000]), divided by the HCHO column lifetime (~2 hours), gives a lower limit from the magnitude of the HCHO source that can be detected from space ($\sim 0.6 \times 10^{12}$ molecules cm$^{-2}$ s$^{-1}$). Methane and the OVOCs are ubiquitous in the atmosphere, and account for the majority of the total HCHO production. However, the variability in the resulting HCHO production is low (standard deviations of 0.3 and $0.2 \times 10^{12}$ molecules cm$^{-2}$ s$^{-1}$), and the column integrated HCHO production rate is always less than twice the nominal satellite detection limit of $0.6 \times 10^{12}$ molecules cm$^{-2}$ s$^{-1}$. These compounds therefore provide a HCHO column background over the study domain with no detectable variability. ANMHCs and monoterpenes are negligible under all conditions encountered during the vertical profiles. Column HCHO production from isoprene, on the other hand, with a variability (standard deviation $2.3 \times 10^{12}$ molecules cm$^{-2}$ s$^{-1}$) more than a factor of five greater than that due to any other VOC group, reaches levels well above the minimum level detectable from space. We conclude therefore that detectable variability in the HCHO column over North America in summer is driven primarily by isoprene emission, though with the caveat that the DC-8 sampling strategy did not include profiles directly over cities. While there were extensive boreal fires in Alaska and northern Canada during the study period [Pfister et al., 2005], because of the short HCHO lifetime they did not significantly impact the column integral HCHO during the vertical profiles.

[28] The HCHO column is strongly correlated ($R^2 = 0.60$) with $P_{HCHO}$ from isoprene, and not with $P_{HCHO}$ from other precursors (with the exception of monoterpenes). Even if isoprene were the sole source of HCHO, the correlation between isoprene emissions and HCHO columns would still be degraded by the smearing effect of horizontal transport. The observed correlation is similar to that simulated by the GEOS-Chem model [Palmer et al., 2003]. While the HCHO column is also correlated with HCHO production from the monoterpenes $\alpha$- and $\beta$-pinene ($R^2 = 0.65$), the highest observed $P_{HCHO}$ from these compounds is a factor of 10 less than the satellite detection limit of $0.6 \times 10^{12}$ molecules cm$^{-2}$ s$^{-1}$. The observed correlation is likely due to collocation of monoterpenes and isoprene emissions. It has been suggested [Kurpis and Goldstein, 2003; Di Carlo et al., 2004; Goldstein et al., 2004; Holzinger et al., 2005] that reactive biogenic emissions from forests may include large amounts of unmeasured, possibly terpenoid, species. The reactivity-weighted abundance of these unmeasured compounds would have to be approximately 100 times greater than the sum of $\alpha$-plus $\beta$-pinene to generate values of $P_{HCHO}$ comparable to those observed for isoprene, assuming comparable HCHO production yields. The fact that the simulated HCHO is in good agreement with the NCAR data (and is higher than the URI data), together with the fact that the distribution of HCHO columns over North America correlates with isoprene, not terpene, emission patterns [Palmer et al., 2006], indicates that any inherent bias in the approach due to unmeasured reactive terpenes is small.

[29] Figure 6 shows the relationship between observed HCHO columns and $P_{HCHO}$ for the different precursors. Again we see that methane and the OVOCs give rise to a significant background $P_{HCHO}$ of ~1–1.5 $\times 10^{12}$ molecules cm$^{-2}$ s$^{-1}$, but not to variability that that would be detectable from space. When the HCHO column exceeds the fitting uncertainty by a sufficient margin to provide a useful signal to the satellite instrument, changes are driven by isoprene. We conclude that space-borne measurements of HCHO
columns can reliably be used as a direct proxy for isoprene emissions over North America. This calculation provides an observational basis for previous studies, which have relied on modeled HCHO-isoprene relationships.

6. HCHO Yield From Isoprene

[30] Using HCHO column data from space as a proxy for isoprene emission requires quantification of the relationship between the two. This has been done previously using GEOS-Chem model output [Palmer et al., 2003, 2006] and we use here the INTEX-A vertical profiles as a test of this approach. From equation (4), the slope of a linear regression of column HCHO ($\Omega_{HCHO}$) versus column isoprene ($\Omega_{ISOP}$), normalized by the ratio of the effective loss rate constants $k_{HCHO}$ and $k_{ISOP}$, represents the molar yield of HCHO production from isoprene oxidation. Figure 7a (black symbols) shows a scatterplot of modeled $k_{HCHO}\Omega_{HCHO}$ versus $k_{ISOP}\Omega_{ISOP}$ over the spatial domain encompassed by the continental and nearshore aircraft profiles ($27.82–49.80^\circ$N; $59.81–98.96^\circ$W) and averaged over the INTEX-A timeframe. The reduced major axis slope, 1.84, is consistent.

Figure 5. Probability density function of HCHO column production rates from different measured VOCs during INTEX-A vertical profiles: isoprene, oxygenated volatile organic compounds (OVOCs), anthropogenic nonmethane hydrocarbons (ANMHCs), methane, and monoterpenes. The dominant OVOCs for HCHO production are methanol and acetaldehyde, while the dominant ANMHCs are ethane, butane, and isopentane. Note the different $x$ axis scale in the fifth plot.
with the nominal yield in Table 1 (2.3), given that the latter value assumes the high-NO limit and that the integrated yield is a time-dependent quantity. Figure 7b (black symbols) shows modeled $W_{\text{HCHO}}$ versus $W_{\text{ISOP}}$ for the same region and time frame, with a slope of 3.99. The ratio of the slopes from the two plots (2.2) corresponds to the mean ratio $k_{\text{ISOP}}/k_{\text{HCHO}}$.

[31] Plots of $\Omega_{\text{HCHO}}$ versus $\Omega_{\text{ISOP}}$ calculated from concentrations measured aboard the DC-8 aircraft or simulated along the flight track during the continental and nearshore vertical profiles are shown in Figures 7b (model, red symbols) and 7c (measurements). Using the model value $k_{\text{ISOP}}/k_{\text{HCHO}} = 2.2$, which should be reliable (errors in model OH partly cancel in the ratio), the observed $\Omega_{\text{HCHO}} - \Omega_{\text{ISOP}}$ slope implies an average molar HCHO yield from isoprene oxidation of $1.63 \pm 0.26$, compared to the modeled value of $1.66 \pm 0.27$. Uncertainties reflect the standard error of the regression. Error estimates computed using jackknife resampling are slightly higher (0.33 and 0.52 for the modeled and measured values, respectively). We conclude from the INTEX-A data that the GEOS-Chem HCHO yield from isoprene oxidation is correct to within 30%.

7. Uncertainty in the Air Mass Factor

7.1. AMF Simulation

[32] In this section we employ the extensive mapping of HCHO over North America from the INTEX-A mission to quantify the uncertainties and bias in the AMF calculation. To do so, we calculate air mass factors separately on the...
basis of measurements and on the basis of model results, for each of the DC-8 vertical profiles during INTEX-A. Assuming that the measurements perfectly represent the atmosphere, the comparison statistics between the measured and modeled AMFs give a measure of the corresponding error in retrieved satellite HCHO vertical columns.

33 Measured and modeled air mass factors were calculated from equation (2) for a nadir viewing geometry. Shape factors, $S(P)$, were determined using either measured or modeled HCHO mixing ratios. Extrapolation of mixing ratios above and below the profile was done in the same way as for the column estimates (section 4.2). Measured and modeled shape factors averaged over all the continental and oceanic vertical profiles are displayed in Figure 8. The GEOS-Chem model accurately captures the mean shape of the vertical profile, including the steep drop-off above the continental boundary layer.

34 Scattering weights, $w(P)$, for each profile were computed using the Linearized Discrete Ordinate Radiative Transfer (LIDORT) model [Spurr et al., 2001], and include scattering by air molecules, aerosols and clouds. Surface UV albedos are from a climatological database based on GOME observations [Koelemeijer et al., 2003]. Aerosol effects on the measured AMF values were accounted for using local aerosol scattering and absorption measured at 10–60 s resolution aboard the aircraft (A. Clarke et al., Biomass burning and pollution aerosol over North America: Organic components and their influence on spectral optical properties and humidification response, submitted to Journal of Geophysical Research, 2006; Y. Shinozuka et al., Aircraft profiles of aerosol microphysics and optical properties over North America: Aerosol optical depth and its association with PM2.5 and water uptake, submitted to Journal of Geophysical Research, 2006). The median single scattering albedo at 346 nm during the aircraft profiles is 0.88 (0.1–0.9 quantiles: 0.76–0.97). However, the aerosols encountered near the surface ($P > 800$ hPa), where the majority of the aerosol (and HCHO) column resides, were predominantly scattering (median single scattering albedo 0.97). The measured aerosol absorption coefficient has a large percentage of missing values, which we fill in by applying the mean single-scattering albedo to the measured scattering coefficient. Integrating the measured aerosol extinction over the individual vertical profiles yields aerosol optical thicknesses (AOTs) at 346 nm ranging from 0.05 to 0.83, with a mean of 0.26 (Figure 9). The corresponding modeled AOTs computed from GEOS-Chem and used in the modeled AMF calculation range from a minimum of

Figure 8. Measured (black) and modeled (red) vertical profiles of satellite instrument sensitivity (scattering weights, $w$), normalized HCHO concentration (shape factor, $S$), and aerosol extinction ($b_{\text{EXT}}$, $\text{Mm}^{-1}$), averaged over the DC-8 continental and oceanic vertical profiles during INTEX-A.
0.03 to a maximum of 0.66 with a mean of 0.22 (Figure 9). A reduced major axis regression of the modeled versus measured AOTs yields a slope of 0.73, and a coefficient of determination ($R^2$) of 0.41. The model reproduces the general vertical shape of the measured aerosol extinction, in particular the sharp decrease above the continental boundary layer which is similar to that observed for HCHO (Figure 8). However, the modeled extinction is biased high in the marine boundary layer and biased low in the continental boundary layer.

[35] The impact of clouds on the measured AMFs was included using in situ cloud extinction measurements made during the vertical profiles, as shown in Figure 9. Only 16 of the profiles have cloud optical thicknesses greater than unity because the DC-8 flight strategy favored clear-sky profiling. For the modeled AMF, we assume that there is accurate information available regarding the cloud top height and optical thickness, since those parameters can be retrieved from satellite instruments such as GOME [Martin et al., 2002a]. Koelemeijer et al. [2001], comparing two different cloud retrieval schemes for GOME, report average differences of 0.04 and 65 hPa for cloud fraction and cloud top pressure. Comparing four cloud fraction retrievals along four GOME tracks, Tuinder et al. [2004] found the mean difference between products to range from 2 to 25%. Here we take the cloud top and total optical thickness information from the aircraft cloud extinction data, and distribute the cloud optical thickness vertically below cloud top following Martin et al. [2002a] by assuming an optical thickness of 8 per 100 hPa of cloud. The cloud top height is taken as the maximum altitude above which the cloud optical thickness is greater than unity (detection limit from GOME; T. P. Kurosu, personal communication, 2005). Profiles having an integrated cloud optical thickness less than one were treated as being cloud-free for the modeled AMF calculation, but not for the measured AMF calculation.

[36] The resulting mean vertical profiles of scattering weights $w(P)$ are shown in Figure 8 for continental and oceanic scenes. The vertical distribution reflects the increasing sensitivity of the satellite instrument with altitude, and deviates from a smooth curve because of cloud and aerosol scattering. As we see, model assumptions regarding aerosols and clouds do not incur significant error in the mean $w(P)$ profile, although this could reflect the prevalence of clear-sky scenes. A more specific assessment of the error for cloudy scenes is presented below.

[37] Measured and modeled AMFs for the ensemble of INTEX-A vertical profiles are mapped in Figure 10 and compared in Figure 11. A reduced major axis regression of the modeled versus measured AMF gives a slope of 0.78 and an $R^2$ of 0.45 (Figure 11). This includes all clear and cloudy profiles. As the overall AMF for a partly cloudy scene is given by a weighted average of the clear and cloudy values (equation (3)), we can assess the errors in AMF$_{clear}$ and AMF$_{cloud}$ separately using the clear and cloudy profiles, and that in the overall AMF as a function of the cloud fraction $f$. In what follows, AMF comparisons are given in terms of harmonic means, since the dependence of the HCHO vertical column on the slant column is defined by the inverse of the AMF. Biases in the modeled AMF are calculated as $(\text{AMF}_\text{mod} - \text{AMF}_\text{meas})/\max(\text{AMF}_\text{mod}, \text{AMF}_\text{meas})$; the AMF
bias is opposite in sign to the resulting effect on the retrieved HCHO vertical column.

7.2. AMF<sub>clear</sub>

[38] For all clear-sky profiles (N = 31), the measured and modeled AMFs range from 0.72 to 1.69 and from 1.00 to 1.71, respectively (Figure 11 and Table 2). The harmonic means (1.21 and 1.23) are identical to within one standard error. The mean bias in the modeled AMF is less than 1%; the standard deviation of the bias, which gives a measure of the precision of the retrieval, is 17%. To the extent that the model error is random, the uncertainty in time-averaged vertical profiles of HCHO will decrease with increasing observations.

[39] The AMF uncertainty for continental profiles under clear-sky conditions is comparable to that for the entire data set (SD of the bias: 18%), while that for oceanic profiles is lower (SD of the bias: 10%). The mean bias in the modeled AMF is −2% (SE: 4%) over the continent, and +7% (SE: 4%) over the ocean.

7.3. AMF<sub>cloud</sub>

[40] Of the 69 aircraft profiles, 16 have cloudy skies (cloud optical thickness greater than one). Of these, only 3 have a sufficient number of HCHO and aerosol extinction measurements to compute air mass factors. In order to make the best use of the available data, and in view of the importance of determining cloud effects, we calculated cloudy AMFs by applying the 16 cloud profiles to each of the 34 vertical profiles with adequate HCHO and aerosol data. This gives 544 values of AMF<sub>cloud</sub> (Figure 11 and Table 2). The measured and modeled distributions of AMF<sub>cloud</sub> show two distinct modes, the first at AMF values less than one and the second at values greater than one.

Figure 10. Measured and modeled HCHO air mass factors for the DC-8 vertical profiles during INTEX-A.

Figure 11. (top) Scatterplot of modeled versus measured air mass factors computed for the DC-8 vertical profiles. Clear-sky and cloudy profiles are represented by points and asterisks, respectively. Probability density functions of (middle) AMF<sub>clear</sub> and (bottom) AMF<sub>cloud</sub>. The pdf of AMF<sub>cloud</sub> is based on 544 data points as described in the text.
AMF Comparison Statistics

<table>
<thead>
<tr>
<th></th>
<th>Harmonic Mean</th>
<th>Harmonic Mean</th>
<th>Model Bias, %</th>
<th>Model Bias, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AMF_clear</td>
<td>AMF_cloud</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All profiles</td>
<td>1.21</td>
<td>0.22</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Measured</td>
<td>1.23</td>
<td>0.19</td>
<td>0%</td>
<td>17%</td>
</tr>
<tr>
<td>Modeled</td>
<td>1.18</td>
<td>0.24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Continental and nearshore profiles</td>
<td>1.24</td>
<td>0.16</td>
<td>-2%</td>
<td>18%</td>
</tr>
<tr>
<td>Measured</td>
<td>1.18</td>
<td>0.24</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Modeled</td>
<td>1.34</td>
<td>0.05</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Oceanic profiles</td>
<td>1.43</td>
<td>0.16</td>
<td>+7%</td>
<td>10%</td>
</tr>
<tr>
<td>Measured</td>
<td>0.13</td>
<td>0.75</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Modeled</td>
<td>0.44</td>
<td>1.08</td>
<td>+46%</td>
<td>39%</td>
</tr>
</tbody>
</table>

*Measured and modeled AMFs calculated from in situ measurements and chemical transport model results for each of the DC-8 vertical profiles during INTEX-A.

[4] For each of the 34 vertical profiles with adequate data coverage, values of AMF were calculated using all of the 16 cloud profiles encountered during the entire study.

[5] The harmonic mean for AMF is skewed by values close to zero. The median measured and modeled values are 0.35 and 0.76, respectively.

(Figure 11). This reflects the tendency of low clouds to increase the measurement sensitivity, and of high clouds to decrease the sensitivity. The mean bias in the modeled AMF is +46% (compared to <1% for AMF), and the standard deviation of the bias is 39% (compared to 17% for AMF). Clouds are therefore the largest source of error in the AMF calculation.

7.4. AMF for Partly Cloudy Scenes

[41] A useful parameter for the satellite retrieval is the cloud fraction above which the AMF error becomes unacceptably large. Here, we employ the 544 measured values of AMF to derive a measure of the AMF error as a function of the cloud fraction of the retrieval scene. The weighted average AMF was calculated from equation (3) for cloud fractions ranging from zero to one. Figure 12 shows the bias and the standard deviation of the bias in the modeled AMF as a function of the cloud fraction of the scene. The mean bias from the 544 partly cloudy AMF calculations (solid black line in Figure 12) is 10%, 14%, 17%, and 21% at cloud fractions of 30%, 40%, 50%, and 60%. The standard deviation in the bias reflects the precision in the AMF calculation, with values of 19%, 21%, 23%, and 25% at cloud fractions of 30%, 40%, 50%, and 60% (red line in Figure 12). On the basis of this result, we recommend discarding scenes with 50% cloud coverage or more. This does not include uncertainty arising from the surface albedo, which is considered separately below.

[42] In addition to clouds, the other potentially important sources of model uncertainty in the AMF calculation are the shape factor, aerosols, and the surface albedo. In order to identify potential areas for improvement in the AMF calculation, the effects of these are assessed individually using sensitivity calculations described below.

7.5. Surface Albedo

[43] Surface albedos for AMF calculations may be obtained from climatological databases derived from satellite measurements [Martin et al., 2003a; Beirle et al., 2004; Boersma et al., 2004; Martin et al., 2004; Konovalov et al., 2005; Palmer et al., 2006]. The precision of the surface albedo database derived from GOME spectra is estimated at 0.02 [Koelemeijer et al., 2003]. We assessed the resulting error in the HCHO air mass factor by recalculating the modeled clear-sky AMFs with the UV albedos uniformly increased and decreased by 0.02. The resulting AMF values have a mean bias of +5% in the first case, and -5% in the second. We therefore estimate the 1σ uncertainty introduced by the surface albedo at 5%.

[44] Adding this quantity in quadrature to the standard deviation of the AMF bias calculated above as a function of cloud fraction, results in an overall 1σ uncertainty due to the AMF which increases from 15% for clear skies, to 18%, 20%, 22%, and 24% at f = 0.2, 0.3, 0.4, and 0.5.

7.6. Aerosols

[45] In order to examine the importance of aerosols for the AMF calculation, and the extent to which the GEOS-Chem model captures this effect, the measured and modeled AMFs were calculated for each vertical profile assuming aerosol-free conditions. For the purposes of this aerosol sensitivity study, we use the 19 out of 34 profiles where the aerosol absorption was measured rather than estimated. On average, the presence of aerosols increases the measured AMF by 14% relative to the aerosol-free scenario. The effect is substantially larger over the North American continent (16% increase) than over the ocean (2% increase) because of higher aerosol loadings (mean column optical thickness of 0.3 versus 0.1). The modeled AMF using the GEOS-Chem aerosol profile information also shows a positive sensitivity to aerosols, somewhat stronger than observations (22% over continents, 10% over the ocean).

7.7. Shape Factor

[46] Errors introduced in the AMF because of the use of the modeled HCHO shape factor were assessed by using the
measured HCHO vertical profile in the calculation of the modeled AMF. In the mean, errors in the model shape factor change the bias in the modeled AMF from +5% to −2% over the continent and from +3% to +7% over the ocean. Over the continent, the positive bias induced in the AMF by the modeled aerosol (+5%) is therefore masked by the negative bias induced by the modeled shape factor (−7%).

7.8. AMF Variability

[47] Using CTM simulations of HCHO vertical profiles for individual scenes to calculate AMFs ensures consistency when comparing the resulting observed vertical columns to those simulated by the same CTM. For the more general purpose of displaying observed vertical columns, however, there is advantage to using a single representative HCHO profile, since this ensures that variability in the observations is real and not introduced by the model.

[48] Overall, the modeled AMF captures 45% of the variability in the measured AMF (Figure 11). This is mostly driven by cloud scenes. The variability in the clear-sky AMF is low (relative standard deviation, RSD = 0.15), which indicates that factors such as mixing height do not introduce significant variability into the AMF. The coefficient of determination between the inverse of the measured and modeled AMFs increases with cloud fraction to $R^2(AMF_{\text{mod}}^{-1}, AMF_{\text{meas}}^{-1}) = 0.61$ at $f = 0.5$. The noise that is introduced in the retrieved columns from the modeled AMF, given by $(1 - R^2(AMF_{\text{meas}}^{-1}, AMF_{\text{mod}}^{-1})) \times \text{RSD}(AMF_{\text{mod}}^{-1})$, increases from 0.12 at $f = 0$ to 0.17 at $f = 0.5$.

8. Conclusions

[49] We used extensive aircraft vertical profiling over North America during the INTEX-A mission in summer 2004 to quantify the errors in retrieving and interpreting HCHO column data from space. By correlating the aircraft observations of HCHO columns with the column HCHO production rates inferred from concurrent VOC measurements, we showed that variability in the HCHO column over North America in summer is mainly determined by isoprene emission. For the ensemble of the INTEX-A profiles, none of the other VOCs contributed to HCHO at a level that would be detected from space, with the caveat that the DC-8 sampling strategy did not include profiles directly over cities. Satellite retrievals of HCHO columns can therefore be used reliably as a proxy for isoprene emissions over North America. In addition to providing independent constraints on emission inventories, these data offer the opportunity to examine the sensitivity of isoprene emissions to environmental drivers, assess the magnitude and implications of interannual variability in biogenic emissions [Abbott et al., 2003; Palmer et al., 2006], and study the effects of human influences such as logging and land use change on emissions of isoprene.

[50] Relating HCHO columns to isoprene emissions requires accurate knowledge of the yield of HCHO from isoprene oxidation. From correlation of measured HCHO and isoprene columns measured from the aircraft, we estimate a molar HCHO yield of $1.6 \pm 0.5$. This value is consistent with current chemical mechanisms used in CTMs, and in particular in the GEOS-Chem CTM used in past interpretation of GOME HCHO column data. The observed correlation between HCHO and isoprene columns has an $R^2$ of 0.60, again consistent with GEOS-Chem and indicating some horizontal smearing due to the time lag between isoprene emission and HCHO production.

[51] The primary source of error in HCHO satellite retrievals is the air mass factor (AMF), which defines the relationship between measured radiances and HCHO vertical columns. The standard approach for computing the AMF is to use local vertical profile information from a CTM, and we have used GEOS-Chem for this purpose in the past. Here we compared the AMFs calculated from the observed vertical profiles of HCHO, aerosol extinction, and cloud extinction to those calculated using GEOS-Chem model HCHO and aerosol profiles combined with the cloud information. One would expect to get from space (cloud top and total optical thickness). Aerosols increase the AMF over North America by 16% on average and are thus important to include in the AMF calculation.

[52] Our analysis shows that clouds are the main source of error in the model AMF calculation. The mean bias in the model AMF increases from <1% under clear-sky conditions to 7% under 50% cloudy conditions. The residual 1σ error (after subtraction of the mean bias) is 15% under clear-sky conditions and 24% under 50% cloudy conditions. The UV surface albedo used in our AMF calculation results in an AMF uncertainty of ±5%. Combining these quantities in quadrature with a fitting uncertainty of $4 \times 10^{15}$ molecules cm$^{-2}$, we arrive at an overall 1σ uncertainty in retrieved HCHO vertical columns which increases from 25% at $f = 0$ to 31% at $f = 0.5$, for a slant column of $2 \times 10^{16}$ molecules cm$^{-2}$. We recommend discarding retrieval scenes with greater than 50% cloud cover. The fraction of the total data coverage this represents will depend on the size of the satellite footprint; GOME scenes with >40% cloud cover represent approximately 40% of the data coverage over North America in summer [Abbott et al., 2003].

[53] In the absence of clouds, AMF variability is low (RSD = 0.15). We find that the artificial variability that is introduced in HCHO column retrievals from the use of AMFs modeled using the GEOS-Chem CTM is 12–17% when the cloud fraction is less than 0.5.

[54] How accurately we can infer isoprene emissions from HCHO column measurements made from space depends on the retrieval errors, as well as uncertainties in the HCHO yield, errors in the HCHO loss rate, and uncertainties associated with converting the HCHO column at the satellite overpass time to a diurnal average. Modeled HCHO yields from isoprene oxidation can differ by 30% between models at a given NO$_x$ level [Palmer et al., 2006], with differences highest at low NO$_x$. The HCHO yield calculated in the present work also has an estimated uncertainty of 30%. Errors associated with the HCHO loss rate and diurnal cycle are likely to be minor in comparison. The uncertainty in the HCHO production yield, combined in quadrature with the retrieval errors calculated above, results in a 1σ uncertainty in isoprene emissions derived from satellite measurements of HCHO columns of 39–43% (again for a slant column abundance of $2 \times 10^{16}$ molecules cm$^{-2}$). This level of uncertainty compares favorably to that associated with extrapolating leaf and plant-level emission data [Guenther et al., 2000]. The overall approach therefore
offers a useful and independent means of inferring surface emissions of isoprene. [55] In other parts of the world, processes such as biomass burning [Thomas et al., 1998; Burrows et al., 1999; Spichtinger et al., 2004; Meyer-Arnak et al., 2005] and anthropogenic emissions (T.-M. Fu et al., manuscript in preparation, 2006) can also make significant and detectable contributions to column HCHO. New satellite instruments such as OMI, aboard Aura, and GOME-2, to be launched aboard the MetOp satellites, should enable mapping of biogenic, urban and biomass burning VOC emissions with much improved spatial and temporal coverage compared to GOME. The results presented here offer a foundation for future such analyses.

[56] Acknowledgments. Financial support for this research was provided by the NASA Atmospheric Chemistry Modeling and Analysis Program and by the NOAA Climate and Global Change Postdoctoral Fellowship Program (DBM). The authors thank Bob Yantosca, Paul Palmer, and Thomas Kurosas for their help and the entire INTEX-A science team for their efforts.

References
Koelmeijer, R. B. A., J. F. de Haan, and P. Stammes (2003), A database of spectral surface reflectivity in the range 335–772 nm derived from


B. E. Anderson, Atmospheric Sciences Division, NASA Langley Research Center, Hampton, VA 23681, USA.

D. R. Blake, Department of Chemistry, University of California, Irvine, CA 92697, USA.

A. D. Clarke, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Honolulu, HI 96822, USA.

A. Fried and J. Walega, Earth Observing Laboratory, National Center for Atmospheric Research, Boulder, CO 80307, USA.

B. G. Heikes, Graduate School of Oceanography, University of Rhode Island, Narragansett, RI 02882, USA.

R. C. Hudman, D. J. Jacob, D. B. Millet, and S. Wu, Department of Earth and Planetary Sciences and Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. (dbm@io.harvard.edu)

H. B. Singh, NASA Ames Research Center, Moffett Field, CA 94035, USA.

S. Turquety, Service d’Aéronomie, Institut Pierre-Simon Laplace, Université Pierre et Marie Curie, Boîte 102, 4, place Jussieu, F-75252 Paris Cedex 05, France.