Algorithm for NO_2 Vertical Column Retrieval From the Ozone Monitoring Instrument

Eric J. Bucsela, Edward A. Celarier, Mark O. Wenig, James F. Gleason, J. Pepijn Veefkind, K. Folkert Boersma, and Ellen J. Brinksma

***Abstract*—We describe the operational algorithm for the retrieval of stratospheric, tropospheric, and total column densities of nitrogen dioxide (NO_2) from earthshine radiances measured by the Ozone Monitoring Instrument (OMI), aboard the EOS-Aura satellite. The algorithm uses the DOAS method for the retrieval of slant column NO_2 densities. Air mass factors (AMFs) calculated from a stratospheric NO_2 profile are used to make initial estimates of the vertical column density. Using data collected over a 24-h period, a smooth estimate of the global stratospheric field is constructed. Where the initial vertical column densities exceed the estimated stratospheric field, we infer the presence of tropospheric NO_2, and recalculate the vertical column density (VCD) using an AMF calculated from an assumed tropospheric NO_2 profile. The parameters that control the operational algorithm were selected with the aid of a set of data assembled from stratospheric and tropospheric chemical transport models. We apply the optimized algorithm to OMI data and present global maps of NO_2 VCDs for the first time.**

***Index Terms*—Algorithm, nitrogen dioxide (NO_2), Ozone Monitoring Instrument (OMI), troposphere.**

I. Introduction

MEASUREMENTS of nitrogen dioxide (NO_2) are important to the understanding of tropospheric and stratospheric chemistry, particularly in relation to ozone production and loss. NO_2 takes part in catalytic destruction of ozone in the stratosphere [1], and anthropogenic NO_2 emissions are precursors for tropospheric ozone production, largely through reactions with hydrocarbons, e.g., [2]. Brewer *et al.* [3] made the first ground-based measurements of stratospheric NO_2, and extensive analysis of stratospheric NO_2 behavior and distribution was undertaken by Noxon [4]–[6], [45] and Solomon and Garcia [7]. Data from the Global Ozone Monitoring Experiment (GOME), deployed in 1995, have been used to retrieve global NO_2 column amounts, which have been used to study the behavior of stratospheric NO_2 [8]. Early results from GOME, showing enhanced NO_2 over the populated areas of the Eastern United States and Europe, were presented by Burrows *et al.* [9], who attributed the enhancements to urban tropospheric pollution. Leue *et al.* [10] and Richter and Burrows [11] attempted to quantify the tropospheric amounts. Comparisons between GOME tropospheric NO_2 and models have been carried out by Velders *et al.* [12], Martin *et al.* [13], Lauer *et al.* [14], and Heland *et al.* [15] made the first comparisons with *in situ* aircraft measurements. A new generation of satellite instruments now provides measurements of trace gases, including NO_2, at spatial resolutions that exceed GOME resolutions by factors of seven or more. One of these is the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) [16]. Martin *et al.* [17] have recently analyzed SCIAMACHY NO_2 data along with aircraft measurements to constrain NO_x emission inventories. The Ozone Monitoring Instrument (OMI), on the Earth Observing System (EOS) Aura satellite, has better spatial and temporal resolution than SCIAMACHY and is the subject of the current study.

Satellite-based Earth radiance measurements yield trace gas slant column densities (SCDs), which depend on not only the density of the gas, but on numerous other measurement parameters. Since the quantity of interest is the vertical column density (VCD), one must convert the SCD into the VCD by dividing the SCD by the *air mass factor* (AMF). The AMFs are calculated using radiative transfer models that account for optical geometry, surface reflectivity, cloud and aerosol properties, and the vertical distribution of the absorbing trace gas. For optically thin trace gases in the stratosphere and upper troposphere, the AMF depends almost entirely on the geometry alone. In the case of NO_2, which is a weak absorber and not widely distributed in the troposphere, a stratospheric AMF can be used to obtain a first-order approximation of the VCD. However, although this method is valid over much of the Earth, it underestimates total column densities in areas with significant boundary layer NO_2. Thus, more accurate analyzes of satellite NO_2 data require subtraction of the estimated stratospheric NO_2 before evaluation of the tropospheric component. Variations on this general approach have been used effectively with GOME data [10]–[13]. The correction procedures consist of two steps: 1) recognition of geographic regions that contain significant tropospheric pollution and 2) accurate evaluation of the AMF in these polluted regions. We present the considerations involved in both of these steps.

Algorithms to identify polluted regions have relied on the fact that most tropospheric NO_2 enhancements occur over land and industrially developed regions and that geographic variation in the tropospheric NO_2 occurs on smaller distance scales than that of stratospheric NO_2. Many investigators [11], [13], [14] use the *reference sector method*, in which the stratospheric component of the NO_2 column in any latitude band is approximated by the total NO_2 column value at the corresponding latitude in

Manuscript received April 29, 2005; revised October 27, 2005.

E. J. Bucsela, E. A. Celarier, M. O. Wenig, and J. F. Gleason are with the NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: eric.bucsela@gsfc.nasa.gov).

J. P. Veefkind, K. F. Boersma, and E. J. Brinksma are with the Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands (e-mail: veefkind@knmi.nl).

Digital Object Identifier 10.1109/TGRS.2005.863715

the central Pacific. Wenig *et al.* [18], Velders *et al.* [12], and Leue *et al.* [10] used an image-processing technique that assumes only smooth and low-amplitude latitudinal and longitudinal variations of stratospheric NO_2 column densities. In either approach, the tropospheric NO_2 column at any location is evaluated using the difference between the total and estimated background components. The difference is then corrected using an AMF derived from *a priori* vertical profiles. Studies in which various profiles have been used to derive tropospheric AMFs reveal that the AMFs are significantly sensitive to NO_2 profile shape, though not to the actual NO_2 amount [13], [19].

In this paper, we present the operational algorithm that is used to process measurements made by OMI. We examine a variety of candidate techniques to effect the separation of stratospheric and tropospheric NO_2 and to calculate the corrected AMF. The techniques can be selected via a number of algorithm parameters. We use a set of synthetic data to evaluate the column retrieval errors due to the algorithm alone, in order to select the set of retrieval parameters that gives the smallest errors. We compare the results of the optimum algorithm with those of methods used by other investigators. We also illustrate the application of the algorithm to real OMI data and show vertical column densities with and without tropospheric correction.

II. Ozone Monitoring Instrument

A. Instrument Characteristics

The OMI is one of four instruments onboard the EOS-Aura satellite designed to study the Earth's atmosphere [20]. The satellite was launched into a sun-synchronous orbit on July 15, 2004, with a local equator-crossing time of about 1345 at the ascending node. Aura is the third in a series of large Earth observing platforms to be flown by NASA, with international contributions. The OMI is an imaging spectrometer that measures solar light backscattered by the Earth's atmosphere and surface. The instrument consists of two spectrometers, one measuring the UV spectral range from 270 to 365 nm in two subranges (UV1: 270–314 nm, resolution: 0.42 nm, sampling: 0.32 nm; UV2: 306–380 nm, resolution: 0.45 nm, sampling: 0.15 nm), the other measuring the UV-visible spectrum from 350 to 500 nm (resolution: 0.63 nm; sampling: 0.21 nm). OMI uses a CCD array with one dimension resolving the spectral features and the other dimension allowing a 114° field of view, providing a 2600-km viewing swath transverse to the orbit track. Its nadir spatial resolution ranges from 13×24 to 24×48 km^2, depending on the instrument's operating mode. Spatial resolutions near the edges of the swath are lower by a factor of three along track and six across track. OMI provides daily global coverage of the sunlit, early-afternoon atmosphere. From the OMI-measured spectra, a number of important atmospheric trace gases are retrieved, including NO_2, ozone, formaldehyde, chlorine dioxide, and bromine monoxide, among others [21]. OMI data are also used to retrieve cloud fraction and cloud-top height [22]–[25], as well as aerosol optical depth and single-scattering albedo [26].

B. Differential Optical Absorption Spectroscopy

NO_2 slant columns are obtained from OMI spectra by the Differential Optical Absorption Spectroscopy (DOAS) technique [27], [28]. The DOAS technique has been used to retrieve trace gas concentrations from ground-based [4], [5], [29] and satellite-based [9] measurements. This method is well suited for the retrieval of NO_2, which has highly-structured absorption features on wavelength scales of less than 10 nm between 400 and 450 nm—a region where there is relatively little interference from other absorbers.

The traditional DOAS technique consists in a linear least-squares fit of the logarithm of the reflectance spectrum, over a chosen wavelength range, to a function of the form

$$\ln[R(\lambda)] = -\sum_i \sigma_i(\lambda) \cdot N_{s,i} - S(\lambda) \tag{1}$$

where $R(\lambda)$ is the ratio of the earth radiance spectrum $I(\lambda)$ to the incident solar irradiance spectrum $F(\lambda)$. Here, $\sigma_i(\lambda)$ is the reference absorption cross section spectrum of absorber i, and $S(\lambda)$ is a smooth, slowly-varying function that models the spectral "background" due to Rayleigh and Mie scattering and surface albedo. The free variables in the fitting procedure are $N_{\rm si}$, the slant column densities of species i, and the parameters of $S(\lambda)$. The wavelength range must be chosen so that the structures of the various σ_i are sufficiently different from each other within that range. Ideally, the σ_i should be nearly orthogonal to each other over the wavelength interval used; in practice, this is not always possible, and may lead to significant uncertainties in the least-squares solution.

In the operational NO_2 algorithm, a variation on the DOAS method is used. The approach accounts for the Ring effect, due to rotational Raman scattering, as well as the various absorbing species. In accordance with radiative transfer principles, the data have different functional dependences on the Ring and absorption spectra, so that (1) must be modified. The function fitted in the operational algorithm is

$$\begin{aligned} R &= \frac{I(\lambda)}{F(\lambda)} \\ &= P_n(\lambda) \exp\left[-\sigma_{NO2}(\lambda) N_{NO2} - \sigma_{O3}(\lambda) N_{O3}\right] \\ &\quad \times \left(1 + \sigma_{\rm Ring}(\lambda) C_{\rm Ring}\right). \end{aligned} \tag{2}$$

Here, P_n is an nth degree polynomial. The degree is an adjustable parameter in the program, but it is found that $n = 3$ is a suitable choice. The argument of the exponential function is the sum of the absorption terms, and the final factor accounts for the Ring effect. In this factor, $\sigma_{\rm Ring}$ is the Ring spectrum [30] and $C_{\rm Ring}$ is its fitted coefficient. Unlike the traditional DOAS approach, (2) requires the use of a nonlinear fit, since the Ring spectrum is effectively included as a source of photons, rather than as an additional absorber. Nonlinear methods must also be used if wavelength adjustments between the reference spectra and the data are required as part of the fitting procedure. However, in the present version of the algorithm, all wavelength corrections are applied prior to the spectral fit.

The wavelength range used is [405.0 nm, 465.0 nm]. We found that in this fitting window it is not necessary to include

the effects of absorption by water or oxygen dimers $(O_2 - O_2)$. Including them did not affect the fitted NO_2 slant columns. The reference spectra used for σ_{NO_2} and σ_{O_3} are those of Vandaele *et al.* [31] and Burrows *et al.* [32], respectively, and are convolved with the instrument's measured slit function. Because of the Doppler shift, it is generally necessary to interpolate the solar irradiance spectrum to the wavelength grid of the Earth radiance spectrum. In the OMI algorithm, a special interpolation method is used, based on knowledge of a highly over-sampled solar reference spectrum. Interpolation of the measured irradiance spectrum F from wavelength λ to $\lambda + \mathrm{d}\lambda$ is computed as $F(\lambda) = F(\lambda + \mathrm{d}\lambda)\,[F_o(\lambda + \mathrm{d}\lambda)/F_o(\lambda)]$, where F_o is the solar reference spectrum. This approach reduces interpolation errors related to the sampling rate and is an improvement over linear or spline methods. However, undersampling should not be significant for instruments like OMI with a large ratio between the slit function and sampling widths [28]. In OMI, this ratio is 3:1.

The best-fit values of N_{si} are the slant column densities of the absorbers. In the absence of scattering, these would be equivalent to the density integrated along the geometrical optical path, and the AMF would simply be the optical path length normalized to the vertical path length. However, because the photon path through the atmosphere is complex, and includes scattering by molecules and clouds as well as reflections off the terrain, the AMF is somewhat more complicated, and depends upon the NO_2 profile shape. Section III-B of this paper describes our development of an algorithm that determines the AMF appropriate to realistic NO_2 profiles.

III. Simulated Data

We employed a set of synthetic data, the output of chemical transport models, to develop and test the algorithm used to process OMI SCDs. Mixing ratio profiles from the models and ancillary geophysical data were used as input to radiative transfer calculations to calculate the SCDs. The mixing ratios were also integrated vertically to obtain the "true" *a priori* VCD values at each geographical location. The SCDs were used as inputs to the algorithm under development, to give retrieved VCDs that were compared directly with the true VCDs.

A. Atmospheric Models and Geophysical Data

Synthetic NO_2 profiles were generated by combining three-dimensional (3-D) stratospheric profiles from the GSFC CTM model [33], with 3-D tropospheric profiles calculated by the Harvard GEOS-CHEM model [34], [35]. The two sets of profiles were joined near 200 mbar. CTM (stratospheric) hourly NO_2 fields were obtained for 1998 and sampled on the 1345 local time (LT) meridian to simulate measurements at the OMI overpass. Due to Aura's orbital inclination, high latitudes are sampled at different LTs (earlier in the SH, later in the NH), but our algorithm study only used latitudes between $\pm 65^\circ$. Mean GEOS-CHEM (tropospheric) model results were available in the 0900 to 1200 LT interval for each day of the year from September 1996 to August 1997. Although this interval is earlier than the OMI overpass time, we took this to be an adequate representation of the statistical distribution of profiles that obtain at the actual overpass time. Combining model-generated fields for different years by the calendar date was dictated by the availability of model results. However, for the purposes of the development of the OMI algorithm, we only demand that the data *qualitatively* represent NO_2 fields at any given time of year. At worst, our procedure would include a small number of profiles that never actually occur in nature; these are not expected to significantly bias our evaluation of the algorithm.

Calculation of simulated slant columns required radiative transfer calculations based on relevant geophysical parameters. The ancillary data were obtained from various sources, as described below. These include cloud, temperature, and albedo for the grid cell in question. Cloud fractions (C) are estimated from 1998 TOMS reflectivities (R_{TOMS}) [36] using the following approximate relation to convert reflectivity to geometrical cloud fraction:

$$C = \frac{(R_{\mathrm{TOMS}} - R_g)}{(R_c - R_g)} \tag{3}$$

where R_g is the effective terrain reflectivity (including the effects of Rayleigh scattering) and R_c is a reflectivity for the cloud top, assigned a nominal value of 80%. The terrain reflectivities in this study are based on mean monthly data from the Earth Radiation Budget Experiment (ERBE) [37], which were measured in the 0.2 to 50 μm range. Cloud heights for the AMF calculation are mean monthly climatological values from the International Satellite Cloud Climatology Project (ISCCP) [38], [39]. Temperature profiles used were monthly mean values from NCEP [40]. No aerosols were included in the study due to the unavailability of a good aerosol model or data source. Aerosol effects, which can be large [19], will be incorporated into the AMF calculation in the future, as the OMI aerosol product is refined and validated.

B. Air Mass Factors

With a vertical density profile of NO_2 in hand, and a given optical geometry, we next turn to the calculation of the SCD that would be observed by the OMI instrument. Following a formalism similar to that of Palmer *et al.* [41], we can gain a great deal of computational efficiency by using the altitude-dependent *scattering weight* $m(z)$, which describes the scattering properties of the atmosphere, independent of the trace gas profile. If the NO_2 profiles are expressed in terms of number density $n(z)$ as a function of height z above the surface, then the slant column density S may be written

$$S = \int_z^\infty m(z')\alpha\,[T(z'), T_o]\,n(z')dz' \tag{4}$$

where z is the altitude of the lower boundary (terrain or cloud top) of the column seen from OMI. The factor $\alpha[T(z'), T_o]$ is a temperature correction related to the NO_2 absorption cross section and has a value close to unity at most altitudes. It accounts for the fact that the variable temperature T of observed NO_2 along the line of sight may differ from the fixed temperature T_0 used to obtain the SCD from the spectral fit. In general, $m(z')$ depends on the altitude z', the viewing geometry, and the albedo of the lower boundary of the observed column at altitude z. A

stratospheric slant column S_s may be defined using (4), with the lower limit z set equal to the altitude of the tropopause. Similarly, a tropospheric slant column S_t can be defined with the tropopause altitude used as the *upper* limit of integration. Values of $m(z')$ in the operational algorithm were calculated *a priori* using the radiative transfer code TOMRAD, used at the Goddard Space Flight Center (GSFC), Greenbelt, MD, for selected sets of values for solar zenith angle, viewing zenith angle, relative azimuth angle, surface albedo, and surface pressure. TOMRAD calculates solar backscattered radiances leaving the top of the atmosphere using the iterative scheme of Davé [42], [43], corrects for anisotropic scattering [44] and uses a pseudospherical approximation (approximation error $<1\%$, in the visible region of the spectrum, up to solar zenith angles of about 88°).

The scattering weight may also be used to produce total AMFs M, which can be used to convert observed or simulated slant column densities to vertical column densities V through the relation

$$M = \frac{S}{V} \tag{5}$$

where

$$V = \int_0^\infty n(z')dz'. \tag{6}$$

Equation (6) defines the total column density, but the respective stratospheric and tropospheric components V_s and V_t may be defined by selecting appropriate integration limits. The AMFs to be used in the retrieval algorithm (see Section IV) can be computed from (5) by using model output and radiative transfer calculations to obtain S and V. For the stratosphere, $M_s = S_s/V_s$, and, for the troposphere, $M_t = S_t/V_t$.

Until now we have assumed spatially homogeneous conditions with respect to cloud cover. In practice, two AMFs—clear and cloudy—are computed for each scene using the same NO_2 profile for both. An effective air mass factor is then computed as the radiance-weighted sum of the clear and cloudy AMFs. In Section IV, we use the symbols M_s' and M_t' to represent, respectively, the stratospheric and tropospheric effective AMFs for partly cloudy (spatially inhomogeneous) conditions. Radiance weights are calculated from the geometrical cloud fraction and the clear and cloudy reflectivities, with clouds treated as opaque Lambertian surfaces. For creation of the simulated data set, cloud-top pressures were based on the ISCCP climatology.

C. Using the Test Data Set

Synthetic slant columns were derived from vertical columns given by the models selected on a 2.5° × 2° longitude–latitude grid from 60° S to 60° N. These were then processed with the OMI algorithm, using different combinations of retrieval parameters. The *retrieval errors*, defined as the difference between the retrieved and original VCDs, were then examined in order to select a parameter combination that minimizes the retrieval error for the widest range of geophysical situations.

To summarize, our algorithm development process consisted of the following steps.

1) Development of the test data set:
   a) a grid of geographic locations and select time-of-year was established;
   b) 15 orbits of the Aura satellite, and derived set of OMI fields-of-view, were simulated;
   c) at each time and geographic location, stratospheric and tropospheric model profiles were combined into a profile $n(z)$;
   d) each $n(z)$ was integrated to give the true model VCD (V_{True});
   e) equation (4) was integrated with $m(z)$ values appropriate for albedo, cloud height and OMI viewing geometry, and a radiance-weighted sum was used to obtain the total slant column.
2) Evaluation of algorithm performance:
   a) a set of retrieval parameters for the OMI algorithm was chosen;
   b) slant columns were processed with the OMI NO_2 retrieval algorithm, described in Section IV, to get the retrieved NO_2 VCD (V_{Retr});
   c) the retrieval error was calculated at each grid point for both total and tropospheric VCDs.

IV. OMI Algorithm Description

The OMI NO_2 algorithm uses the OMI-measured earthshine radiances and solar irradiances to calculate slant column NO_2 densities, processed with an initial AMF to get initial estimates of the vertical column densities. A whole-day field of initial VCDs is used to develop smooth approximation of the global stratospheric field and identify small-scale regions that have significantly elevated NO_2, where the VCD is recalculated using an AMF for the tropospheric component. The algorithm contains a number of parameters that specify how the retrieval is accomplished. These parameters can be varied, and the retrieval errors for a set of parameter choices can be evaluated, as will be discussed in Section V.

Here, we summarize the steps in the NO_2 algorithm to convert slant to vertical columns. A description of the conversion follow, and is summarized in Fig. 1.

1) *Slant Column Densities (S)*—A spectral fit, using radiance and irradiance spectra, is performed, yielding NO_2 slant column densities.
2) *Initial Vertical Column Densities* (V_{init})—The slant column densities are divided by the stratospheric air mass factor.
3) *Stratosphere-Troposphere Separation (STS)*—Spatial filtering is applied to the geographic field of V_{init} to estimate the stratospheric and the tropospheric components of the vertical column density.
4) *Tropospheric Correction (TC)*—Where the tropospheric component exceeds a defined threshold, V_{init} is corrected for the tropospheric component, to give V; everywhere else, $V = V_{\mathrm{init}}$.

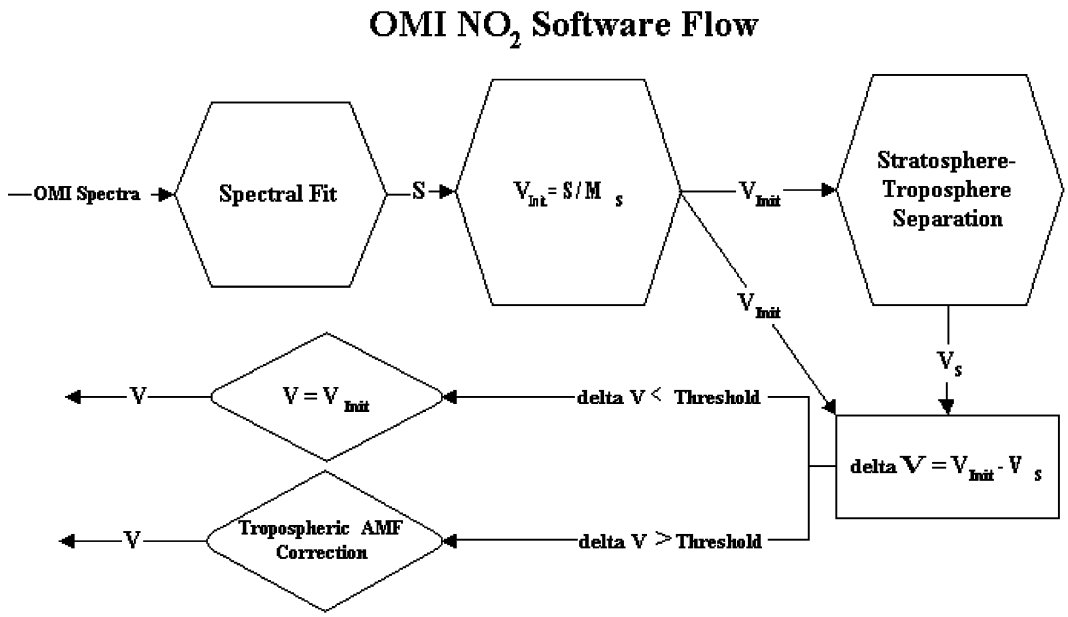


Fig. 1. Flow chart for the OMI NO_2 algorithm. The spectrum is fitted to give the SCD S, which is divided by stratospheric AMF M_S, to yield $V_{\rm init}$. The STS procedure computes the stratospheric VCD V_S, which is compared to $V_{\rm init}$ to determine the retrieved VCD V.

A. Initial Vertical Column Densities

Since most NO_2 is stratospheric, an initial AMF, based on a generic stratospheric profile, is used to estimate an initial vertical column density. The initial NO_2 vertical columns densities, $V_{\rm init}$, are found by dividing the *total* slant columns by the *stratospheric* air mass factor M_s, which is calculated by (5) with a model stratospheric NO_2 profile

$$V_{\rm init} = \frac{S}{M_s}. \tag{7}$$

In the OMI algorithm, a single, mean, stratospheric NO_2 profile is used in the calculation of M_s, since stratospheric AMFs are relatively insensitive to the structure of the NO_2 profile.

B. Stratosphere-Troposphere Separation (STS)

In principle, the vertical column density could be calculated by dividing the slant column density by an AMF based on the entire profile from ground to infinity. This would require accurate knowledge of the relative NO_2 distribution from ground to the upper stratosphere. The relative amount is difficult to estimate, since stratospheric and tropospheric NO_2 have different origins and vary independently in time and space. Therefore, the two are treated separately, using approximate profile shapes for each region. A major focus of this study was to explore the effects of using different methods to separate stratospheric and tropospheric NO_2 components.

Polluted regions are identified in the initial VCD field through spatial filtering. The method rests on two assumptions: 1) gradients in stratospheric NO_2 are much larger in the latitude direction than in the longitude direction, and 2) significant variation of tropospheric NO_2 occurs on smaller geographic scales than that of stratospheric NO_2. These same assumptions have been exploited in other NO_2 retrieval algorithms [10], [11], [13], [18].

We begin with a geographically gridded array of initial VCD values. The grid contains data from one day, which is the amount of time required by OMI for global coverage. To minimize tropospheric bias in the stratospheric field, we mask polluted areas in a two-step process. Initially, we use an *a priori* global mask to eliminate large areas with potentially high amounts of tropospheric NO_2. A variety of masking criteria are compared in Section V. Next, unmasked regions are used to calculate a running 10° boxcar average in the meridional direction. Planetary wave analysis is then applied in the zonal direction, to give a preliminary stratospheric field. However, this preliminary field can be affected by transient pollution events, so a second masking step is required. We identify and mask places where differences between the unmasked initial VCD data and the preliminary field are depart from the analysis by more than one standard deviation. The remaining data are then reanalyzed to obtain the final set of coefficients for the planetary waves. For the purposes of deriving accurate total and tropospheric VCDs, we have found that a wave-2 fit provides the best representation of the stratospheric field in most cases, as we will show.

C. Tropospheric Correction

The AMFs used to calculate the initial VCDs do not yield accurate VCDs in polluted regions. An air-mass factor adjustment is required to account for the difference in optical path through the tropospheric part of the profile in these regions. Air mass factors for profiles that peak near the boundary layer (tropospheric component) are generally smaller than high-altitude (stratospheric component) air mass factors. Thus, the adjustment usually corrects an underestimation of the total vertical column density. The tropospheric component of the vertical column density V_t and the corrected total vertical column V are given by (8) and (9), respectively

$$V_t = \frac{S - M_s' V_s}{M_t'} \tag{8}$$

$$V = V_s + V_t. \tag{9}$$

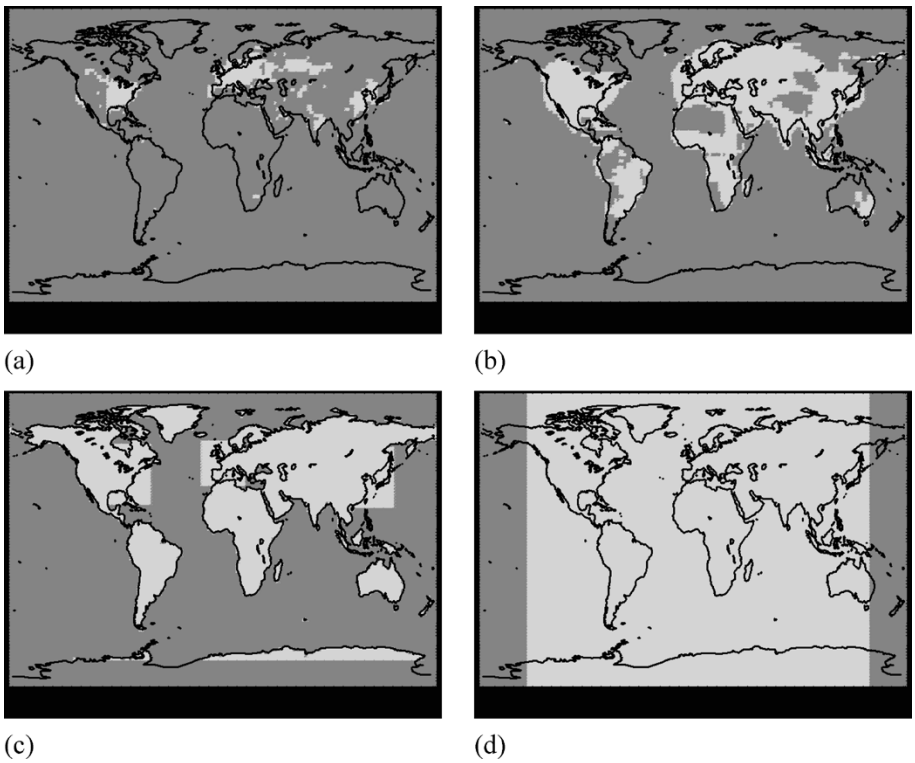


Fig. 2. Pollution masks: (a) M_1, (b) M_2, (c) M_3, and (d) M_4. The lighter areas are masked out prior to the wave analysis.

Here, S is the slant column density from the initial spectral fit, and V_s is the stratospheric component of the vertical column density obtained from the smoothing procedure. The partly cloudy air mass factors for the stratospheric and tropospheric parts of the profile are M_s' and M_t', respectively. They are constructed from a table of precalculated scattering weight functions.

In the OMI algorithm, the tropospheric correction is applied whenever value of the quantity $\Delta V = V_{\mathrm{init}} - V_s$ is greater than some threshold. Then the tropospheric vertical column density is calculated using (8), and used to obtain the total vertical column density according to (9). As we will show, setting the threshold to zero gave the best overall accuracy for the vertical column densities. Other criteria for applying the AMF correction, such as correcting negative as well as positive values of ΔV, were explored. These are discussed in Sections V and VI.

V. Algorithm Development

The algorithm described in the preceding section has a number of adjustable parameters. To optimize the accuracy of the retrieved NO_2 VCDs, we systematically explored the effects of parameter choices on the error statistics using the set of synthetic data. This section describes the adjustable parameters and their effects on the accuracy of the VCD retrievals.

A. Parameter Sets

There are four categories of retrieval parameters: 1) profiles; 2) masking; 3) stratospheric field estimate; and 4) threshold for correction of the AMF.

Profiles: The profiles used for the calculations of the AMFs were derived from the CTM and GEOS-CHEM models, as described in Section III-A. We examined three sets of profiles, which we designate as $\mathrm{P}_{\mathrm{daily}}$, $\mathrm{P}_{\mathrm{avg}}$, and $\mathrm{P}_{\mathrm{cmr}}$. $\mathrm{P}_{\mathrm{daily}}$ is the set of actual daily profiles that were used to generate the SCDs. This profile choice is, of course, not available for the operational algorithm, but we use it in our study as a "best case," against which we can measure the performance of the other choices of profile. $\mathrm{P}_{\mathrm{avg}}$ denotes the use of a single stratospheric profile and a geographically gridded set of tropospheric profiles. The stratospheric profile is an annual global mean of CTM profiles, and the tropospheric profiles are GEOS-CHEM annual means in $2.5^\circ \times 2^\circ$ longitude-latitude bins. The profile set $\mathrm{P}_{\mathrm{cmr}}$ employs the same single stratospheric profile used in $\mathrm{P}_{\mathrm{avg}}$, but assumes all tropospheric NO_2 exists in a layer below 1.5 km, with a constant mixing ratio.

Masks: Several choices were considered for the *a priori* pollution mask. The mask choices are designated M_1 through M_4. M_1 and M_2 were based on GEOS-CHEM model output. The four mask options are shown in Fig. 2. They are as follows.

- M_1 Mask all regions with annual mean tropospheric VCD $> 3 \times 10^{15}$ cm^{-2}.
- M_2 Mask all regions with annual mean tropospheric VCD $> 5 \times 10^{14}$ cm^{-2}.
- M_3 Mask all land areas and some coastal regions.
- M_4 Mask all areas except a narrow strip of longitudes in the central Pacific.

Stratospheric Field Estimate: We explored four parameterizations of the stratospheric field, all of which assume that the field can be expressed as a sum of zonal waves. These cases are denoted as W_0, W_1, W_2, and W_4, where the numeral indicates the maximum wavenumber used. Hence, W_0 models the

stratosphere as being zonally constant (one parameter to fit). W_4 models the stratosphere as a linear combination of sines and cosines of L, 2L, 3L, and 4L, where L is the longitude. The number of coefficients to be fit is, thus, 1 (W_0), 3 (W_1), 5 (W_2), or 9 (W_4).

Threshold for AMF Correction: As discussed in Section IV, we apply a tropospheric AMF correction when ΔV exceeds a predetermined threshold. The threshold was chosen from the following alternatives.

- $\mathrm{T}_{-\infty}$ Threshold of negative infinity (correct all for values of all ΔV, whether positive or negative).
- T_0 Threshold of zero (correct for only positive values of ΔV).
- $\mathrm{T}_{+\infty}$ Threshold of positive infinity (do not correct; trivial case).

B. Parameter Sets

To evaluate the STS/TC retrieval algorithm, we varied the M, W, and T parameters, fixing the set of profiles as $\mathrm{P}_{\mathrm{avg}}$. The profiles $\mathrm{P}_{\mathrm{avg}}$ are the ones currently used to calculate the AMFs in the operational OMI algorithm. We also compared results based on AMFs calculated with alternative profile sets, namely $\mathrm{P}_{\mathrm{daily}}$ and $\mathrm{P}_{\mathrm{cmr}}$, using the optimal parameter choices from the STS/TC algorithm. Results of profile comparisons are presented in Section V-D.

The STS/TC algorithm's accuracy was tested on the basis of the errors in retrieved vertical column density for each day. In all cases, we used simulated data fields for 24 days taken from January, April, July, and October. Statistics were examined regarding the magnitude, sign and geographical distribution of total VCD errors between 60° S and 60° N latitude. Although both clear and cloudy scenes were used to compute the stratospheric field, error statistics were based on regions with less than 25% cloud cover. This cloud fraction is approximately the upper limit at which modest levels of boundary-layer tropospheric NO_2 can be detected by the OMI algorithm [21].

We assume a nominal threshold of 0.2×10^{15} cm^{-2} for significance in the retrieval error. This threshold approximates the combined uncertainties in the spectral fit and the small-scale stratospheric variations. Boersma *et al.* [19] have found uncertainties of the same order of magnitude in their analysis of GOME data, and summarize similar results from other studies. The spectral fit errors are based on contributions from uncertainties in cross section shapes, estimates of instrument noise and errors in atmospheric temperature assumptions. For OMI, Boersma *et al.* [21] estimate that these factors contribute to an error of approximately 6% under unpolluted conditions, which corresponds to an absolute error slightly less than 0.2×10^{15} cm^{-2} for a typical VCD of 3×10^{15} cm^{-2}. Small-scale stratospheric variations also introduce a fundamental uncertainty in the retrieval, since such structures cannot be definitively assigned to the stratosphere or troposphere by the algorithm. Output from the CTM model shows variations about the smooth stratospheric field have standard deviations between 0.05×10^{15} and 0.2×10^{15} cm^{-2}, depending on latitude, with typical a standard deviation near 0.1×10^{15} cm^{-2}. Based on these considerations, the criteria used to evaluate the algorithm included the magnitude of the root mean-square (rms) error, the probability of errors larger than the significance threshold, and the relative probabilities of positive and negative errors. Only significant errors in both the retrieved total vertical column density and its tropospheric component were used in the evaluation. Consideration was given to cases with cloud fractions below the 25% upper limit.

C. Results and Error Comparisons

Results of the exploration of parameter space (M, W, T) are summarized in Tables I and II for total and tropospheric errors, respectively. We tabulate both the percentage of cases with errors greater than the significance threshold of 0.2×10^{15} cm^{-2}, and the rms error. Included are the 32 cases, covering the four values of M, four values of W, and two values of T. An additional case—the "no-correction" case—for $\mathrm{T}_{+\infty}$ is also listed (for this case, the M and W parameters are not applicable). We include only wave W_0 when applying mask M_4, since higher wavenumbers cannot be fitted to the narrow range of longitudes in the central Pacific associated with this mask. Based on the tabulated statistical parameters for the total and tropospheric VCD errors, we conclude that the following combination yields the best overall retrieval:

- Mask M2 (mean tropospheric VCD $> 5 \times 10^{14}$ cm^{-2}).
- Wave W2 (fit stratosphere with wave-2).
- Threshold T_0 (correct only positive ΔV).

When applied to our simulated data set, this case yields significant errors in only 8% of the total VCD retrievals and an rms error of 0.15×10^{15} cm^{-2}. The corresponding values for the tropospheric column errors are 25% and 0.25×10^{15} cm^{-2}, respectively. Although not shown in the table, for that case, the frequency of significant positive errors was approximately equal to that of significant negative errors. In most other cases, an appreciable sign bias was seen in the errors. In several cases, e.g., (M_4, W_0, $\mathrm{T}_{-\infty}$), the frequency of large errors is greater than the frequency when no correction is applied. However, the magnitudes of the rms errors are much greater in the no-correction case. The significance of the optimal parameter set is discussed further in the next section.

Fig. 3 shows total VCD error maps and additional statistical information for three of the sets of retrieval parameters. The three parameter sets are (M_4, W_0, $\mathrm{T}_{-\infty}$), (M_2, W_2, T_0), and $\mathrm{T}_{+\infty}$, and, in all cases, the profiles $\mathrm{P}_{\mathrm{avg}}$ are used as the basis for the AMFs. The first set describes the so-called "Pacific reference sector" (PRS) method, which assigns stratospheric column NO_2 at a given latitude a constant value equal to total column NO_2 in the central Pacific at that latitude. Corrections are then made for both positive and negative values of ΔV. Versions of the Pacific reference sector method have been used in numerous GOME studies [11], [13], [14]. The second parameter set is the optimal parameter combination discussed above, and the third corresponds to the "no-correction" case. The time-averaged VCD errors, mapped in Fig. 3, show the differences between retrieved V and the "true" VCD, V_{True}. In the figures, the color scale has been chosen so that all regions with retrieval errors less than the significance threshold appear uniformly shaded. Significant negative or significant positive errors are darker or lighter, respectively. The map for the no-correction case [Fig. 3(c)] shows that the simple assumption that the

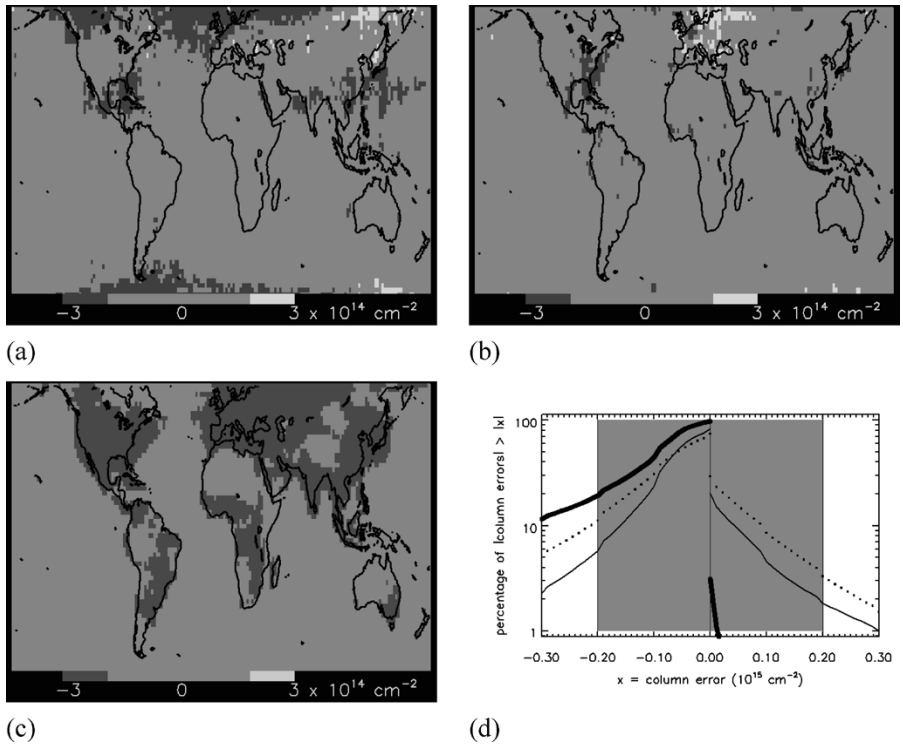


Fig. 3. Time-averaged total vertical column density retrieval errors for: (a) Pacific reference sector method, (b) OMI algorithm, and (c) no correction. In the maps, the lighter (darker) areas are regions of significant over- (under-) estimation of the VCD by the retrieval algorithm. Curves in (d) show the percentage of geographic area between 60° S and 60° N latitude for which the VCD error (not time averaged) exceeds (or is less than for negative errors) value on the x axis for (dotted) PRS method, (thin) OMI algorithm, and (thick) uncorrected. Errors inside the shaded region are presumed to be smaller than nonalgorithmic errors, such as measurement error and SCD retrieval error.

total VCD equals $V_{\rm init}$ significantly underestimates the vertical column density over a large fraction of land areas, particularly in polluted regions. Large areas of both over- and under-estimation of the total vertical column density are seen in the PRS map [Fig. 3(a)]. Significant positive and negative errors are also seen for the optimal case [Fig. 3(b)] but over smaller areas than in the other two cases.

A quantitative diagnostic of the performance of the retrieval algorithms is the probability that any VCD error is greater (less) than a given positive (negative) value, that is, the probability *distribution* function, shown in Fig. 3(d), for the cases used in the other three panels. These curves are based on the errors found for all geographic locations between 60° S and 60° N latitude with cloud fraction $<25\%$, and use all days of data, i.e., they were not calculated from the time-averaged data displayed in the maps. The frequencies are calculated with respect to the entire ensemble of cases, so the sum of the intersections of the left- and right-hand curves with the vertical line $x = 0$ is 100%; the fact that the negative curve meets this line at a higher point than the positive curve means that there is an overall negative bias. In particular, we see that the "no correction" method gives essentially no overestimations of the VCD (thick curves). We have shaded the assumed significance threshold of 0.2×10^{15} cm^{-2}, discussed above. The optimal OMI method ($\rm M_2, W_0 - W_2, T_0$) yields some overestimations, but is an improvement on the PRS method for both positive and negative errors. All three methods show some tendency to underestimate, rather than overestimate, the VCD. This is due mainly to an *over*estimation of the stratospheric field, due to the inclusion of small amounts of background tropospheric NO_2 in the un-masked regions. We examine this issue in Section VI.

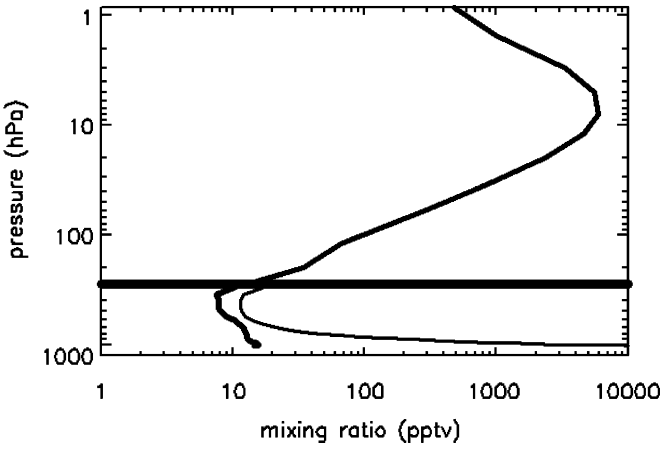


Fig. 4. Mean stratospheric profile from the CTM model (above the horizontal line), and two examples from the GEOS-CHEM model of tropospheric profiles (below horizontal line) for (thick) nonpolluted and (thin) highly polluted cases.

D. NO_2 Profiles

The results shown in the preceding section were based on AMFs calculated from the $\rm P_{avg}$ profiles, which are the ones used in the operational OMI algorithm. This set consists of annually averaged (and geographically averaged, in the case of the stratosphere) profiles, which are assumed to be approximately valid on any given day of any year. The approximation avoids the

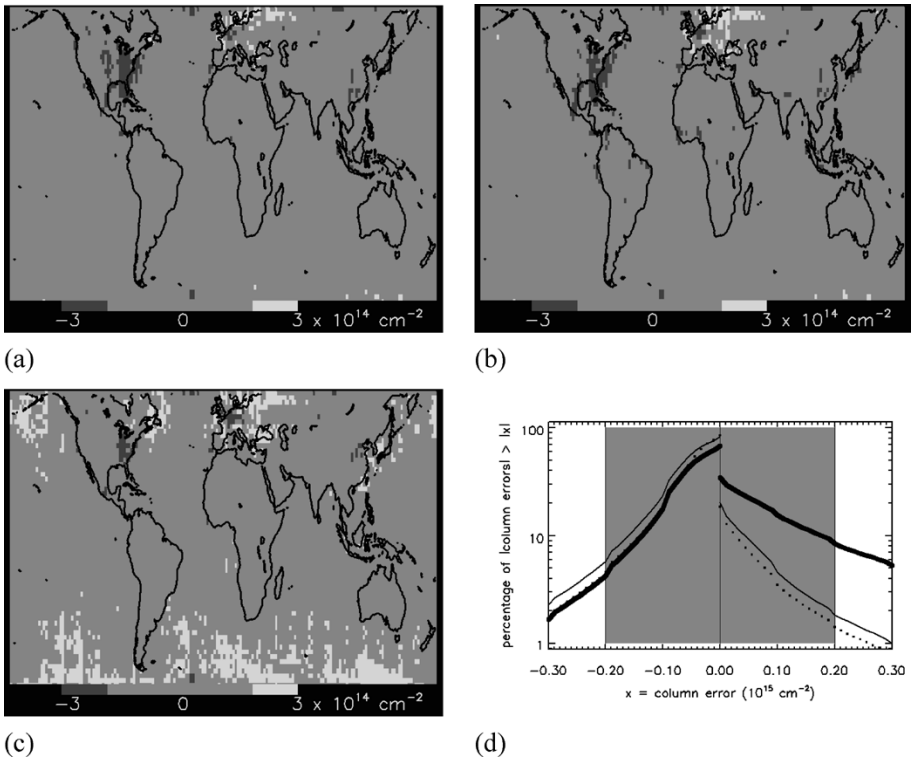


Fig. 5. Time-averaged total vertical column density retrieval errors from OMI algorithm but with AMFs calculated from: (a) daily profiles, (b) single average stratosphere and time-averaged geographically gridded troposphere, and (c) single average stratosphere and constant mixing-ratio troposphere. In the maps, (darker) the lighter areas are regions of significant over- (under-) estimation of the VCD by the retrieval algorithm. Curves in (d) show the percentage of geographic area between 60° S and 60° N latitude for which the VCD error (not time-averaged) exceeds (or is less than for negative errors) value on x axis for (dotted) $\mathrm{P_{daily}}$, (thin) $\mathrm{P_{avg}}$, and (thick) $\mathrm{P_{cmr}}$. Errors inside the shaded region are presumed to be smaller than nonalgorithmic errors, such as measurement error and SCD retrieval error.

need for daily model input to the retrieval algorithm. The stratospheric profile and two examples of tropospheric profiles—one from an unpolluted and the other from a polluted region—are shown in Fig. 4.

We examined the validity of the $\mathrm{P_{avg}}$ profile approximation for AMF calculations by comparing retrievals based on these profiles with retrievals based on two other profile sets, described in Section V-A). We tested these profile sets with the full set of M, W, and T parameters, but show here only the results of tests using the optimal set ($\mathrm{M_2}$, $\mathrm{W_2}$, $\mathrm{T_0}$). Maps of VCD errors and error statistics plots are shown in Fig. 5.

It can be seen from Fig. 5 that the total column retrieval accuracy attained using the $\mathrm{P_{avg}}$ profiles falls between that of the $\mathrm{P_{daily}}$ and $\mathrm{P_{cmr}}$ profiles, as expected. Notice that the degradation in accuracy is relatively small when the approximate $\mathrm{P_{avg}}$ profiles are substituted for the "true" $\mathrm{P_{daily}}$ profiles [cf. Fig. 5(a) and (b)]. However, retrieval accuracy is more substantially compromised, particularly in cases of VCD overestimation, when the simple assumption of a constant tropospheric mixing ratio, $\mathrm{P_{cmr}}$, is used [Fig. 5(c)]. We have also found that when the $\mathrm{P_{cmr}}$ profiles are used in combination with the PRS method, cases of significant underestimation of the VCD are nearly as probable as when no correction is applied.

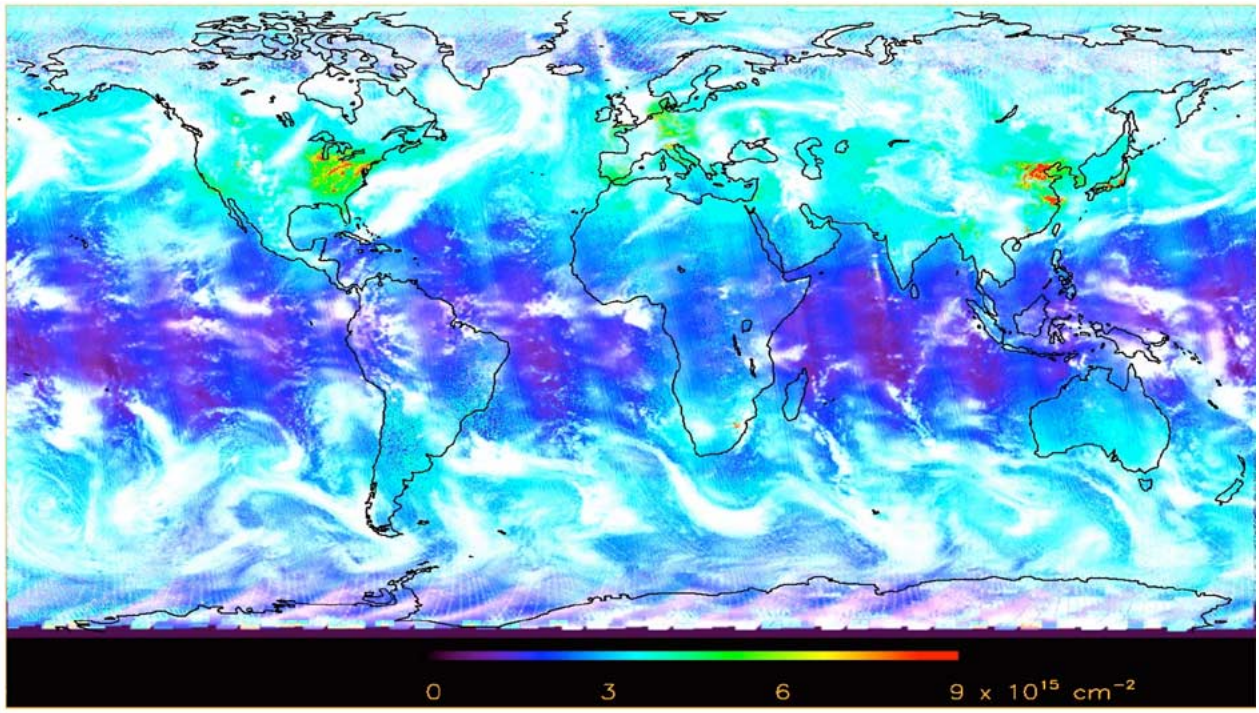
E. Application to OMI Data

The OMI STS/TC algorithm was applied to a day of OMI slant column data. Fig. 6 shows VCDs before and after processing with the OMI STS/TC algorithm. The data represent 15 orbits on April 15, 2005. Unlike the simulated data, the OMI measurements also contain noise, data gaps and other artifacts. Among these is an anomaly in the retrieved SCDs across the orbital track. The anomaly varies among the 60 cross-track pixels and causes "striping" in the retrieved NO_2 columns along each orbit. The problem is thought to be related to the performance of the CCD detector. We have removed most of this cross-track anomaly in a *post-hoc* fashion by subtracting a correction computed from the mean NO_2 column value at each of the 60 pixels.

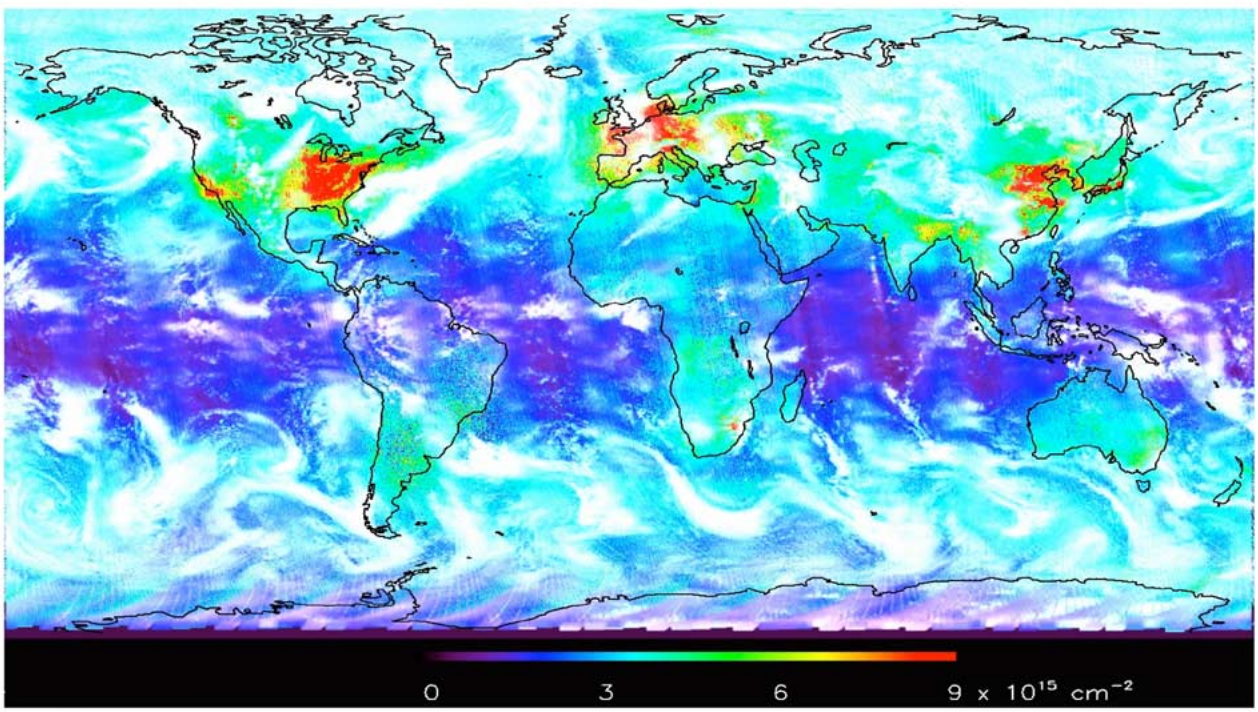
The images shown demonstrate the results of applying the STS/TC algorithm to the NO_2 initial VCDs. The effects are most evident in the enhanced column densities over the polluted regions of eastern North America, Europe, and eastern Asia. Tropospheric enhancements of factors of three or more can be seen in many areas. In the calculation, the tropospheric AMFs were computed from the viewing geometry, the albedos from the OMI data product, and the mean stratospheric and gridded average tropospheric profile set $\mathrm{P_{avg}}$. All retrieval parameters are the optimized values described in this paper, namely mask $\mathrm{M_2}$, wave-2 $(\mathrm{W_2})$, and threshold $\mathrm{T_0}$.

F. Discussion

In this study, the accuracy of the NO_2 retrieval has been assessed in terms of the NO_2 total vertical column density. For most investigations of the chemistry and dynamics of the atmosphere, individual values of the stratospheric and tropospheric components are required. The OMI algorithm separates these two parts of the total VCD to improve the total column



(a)



(b)

Fig. 6. NO_2 total vertical column densities from OMI on April 15, 2005. (a) Initial vertical column densities $V_{\rm init}$. (b) Vertical column densities corrected using the optimized OMI algorithm, as described in the text. All data have been corrected for cross-track anomaly (see text) and placed on a $1/4 \times 1/4$ deg^2 geographic grid. White areas are cloudy regions.

accuracy. We have described some of the considerations and assumptions used in the algorithm to estimate these components. In this section, we compare the OMI algorithm with methods used in other studies to infer stratospheric and tropospheric NO_2 from measurements and look at the selection of clear and cloudy scenes in generating the stratospheric NO_2 field. We also discuss the tropospheric AMF correction and its dependence on profile choice.

The preceding section describes optimization of the algorithm to determine a set of retrieval parameters (M2, W2, $\rm T_0$). As is evident from Table I, other parameter combinations yield similar-to-identical error statistics for the total column

TABLE I
ERROR STATISTICS p, σ, WHERE p IS THE PERCENTAGE OF CASES WITH *TOTAL* VCD RETRIEVAL ERRORS EXCEEDING THE SIGNIFICANCE THRESHOLD OF 0.2×10^{15} cm^{-2}, AND σ IS THE RMS ($\times 10^{13}$ cm^{-2}) OF THE *TOTAL* VCD ERRORS. STATISTICS ARE BASED ON DAYS FROM JANUARY, APRIL, JULY, OCTOBER WITH CLOUD COVER LESS THAN 25% AT ANY GIVEN LOCATION BETWEEN 60° S AND 60° N LATITUDE. STATISTICS FOR THE PARAMETER COMBINATIONS USED IN THE OPERATIONAL OMI ALGORITHM ARE IN BOLD

ΔV Correction Threshold	Fourier Wave	Mask type M_1	M_2	M_3	M_4
		% cases of significant errors , rms error (10^{13} cm^{-2})			
$T_{-\infty}$	W_0	18 , 20	14 , 18	15 , 18	14 , 20
	W_1	17 , 19	12 , 17	14 , 19	--
	W_2	17 , 19	11 , 17	15 , 22	--
	W_4	18 , 21	12 , 23	17 , 40	--
T_0	W_0	9 , 16	8 , 15	9 , 15	8 , 15
	W_1	10 , 16	8 , 15	9 , 16	--
	W_2	11 , 16	**8 , 15**	10 , 19	--
	W_4	12 , 18	9 , 18	12 , 25	--
$T_{+\infty}$ (no correction)		19 , 38			

TABLE II
ERROR STATISTICS p, σ, WHERE p IS THE PERCENTAGE OF CASES WITH *TROPOSPHERIC* VCD RETRIEVAL ERRORS EXCEEDING THE SIGNIFICANCE THRESHOLD OF 0.2×10^{15} cm^{-2}, AND σ IS THE RMS ($\times 10^{13}$ cm^{-2}) OF THE *TROPOSPHERIC* VCD ERRORS (SEE TABLE I FOR *TOTAL* VCD ERROR STATISTICS). STATISTICS FOR THE PARAMETER COMBINATIONS USED IN THE OPERATIONAL OMI ALGORITHM ARE IN BOLD

ΔV Correction Threshold	Fourier Wave	Mask type M_1	M_2	M_3	M_4
		% cases of significant errors , rms error (10^{13} cm^{-2})			
$T_{-\infty}$	W_0	45 , 33	42 , 32	43 , 33	45 , 35
	W_1	44 , 30	39 , 29	41 , 32	--
	W_2	43 , 29	36 , 29	40 , 41	--
	W_4	40 , 30	34 , 40	39 , 70	--
T_0	W_0	28, 27	30, 29	30 , 29	33, 29
	W_1	29 , 25	27, 25	29 , 28	--
	W_2	28 , 24	**25, 25**	28 , 36	--
	W_4	28 , 26	25, 32	28 , 47	--
$T_{+\infty}$ (no correction)		32 , 64			

retrieval—e.g., the simpler approach (M4, W0, T_0). While this result implies some flexibility in the choice of parameters for the total column retrieval, larger differences are seen in the tropospheric column retrieval (Table II), which was also considered in our study. We also found significantly larger errors for most combinations of the M and W parameters when the threshold $T_{-\infty}$ was used instead of T_0. The choice of the *a priori* mask M2 is based on a GEOS-CHEM model estimate of regions that, on average, contain significant tropospheric NO_2. The same model has generated the simulated data used in our study to evaluate the mask. While this data set is not independent, our use of a two-step process for eliminating tropospheric contamination of the computed stratospheric field somewhat reduces the algorithm's sensitivity to the choice of initial mask. The choices for all three retrieval parameters will be reevaluated during the OMI validation phase. Because we have based our algorithm's optimization on an ideal, noise-free case, the retrieval errors we report must be considered lower limits on the actual errors from OMI.

We have employed zonal analysis in the OMI algorithm to model the stratospheric field, and found wave-2 to be optimal. Another means of separating the stratosphere and troposphere is the image processing technique (IPT) used by Velders *et al.* [12], Leue *et al.* [10], and Wenig *et al.* [18]. In this method, continents and surrounding coastal waters are masked (using a mask similar to M3), and a two-dimensional (2-D) normalized convolution is applied to the NO_2 field to yield a smooth estimate of the stratospheric field. When applied to several days of GOME data, Wenig *et al.* [18] also use a "minimum filter," which selects the lowest clear-sky VCD value in the time period at a given location to generate the stratospheric NO_2 field. However, this approach is not applicable to OMI analysis, where the stratospheric field is calculated from a single day of data.

An issue in any method used to determine the stratospheric field from global NO_2 data is the treatment of clouds. Leue *et al.* [10] derive the stratosphere from areas with cloud fractions greater than 50%, while Velders *et al.* [12] and Wenig *et al.* [18] base their estimates on clear skies. Use of overcast regions can ensure that clouds mask low-level tropospheric NO_2, so that the stratospheric estimate will not be biased by boundary layer pollution. Such an approach requires that any tropospheric NO_2 above the clouds be carefully taken into account. Wenig *et al.* [8] have analyzed GOME data and noted NO_2 enhancements above clouds. Radiative transfer calculations based on their relatively simple NO_2 profile assumptions did not readily explain the enhancements. Thus, they excluded cloudy regions in their study and found that this reduced the mean stratospheric background and eliminated most cases of apparent negative tropospheric NO_2 columns. Martin *et al.* [13] show that low-level marine cloud decks typically increase tropospheric AMFs by 30% to 40% relative to AMFs over clear ocean scenes. Their stratospheric field is derived from data in the Pacific, where low clouds are frequent and can thus significantly enhance the effects of small amounts of tropospheric NO_2. The present version of the OMI algorithm uses both clear and cloudy pixels to generate the stratospheric field. If pixels from more than one orbit overlap, a weighted average is used, with greater weight assigned to the pixels having smaller cloud fractions. To assess the tropospheric retrieval accuracy in this study, we have only looked at cases with less than 25% cloud cover and have ignored aerosol contamination. Both of these effects have been shown to have a potentially large influence on the tropospheric AMF [19]. Instrument data along with ground and aircraft measurements are currently being used to improve the treatment of clouds and aerosols in the OMI algorithm.

Tropospheric NO_2 in areas used to derive the stratospheric field can create a bias in the retrieval. To correct for this potential bias, Martin *et al.* [13] subtract a GEOS-CHEM modeled estimate of the tropospheric NO_2 VCD—derived from daily NO_2 amounts over the Pacific—from their stratospheric field. This "tropospheric excess" is generally less than 0.2×10^{15} cm^{-2}. Failure to account for it in the application of our algorithm to the simulated data is, in part, responsible for the negative error bias evident in Fig. 3. However, we have chosen not to include it in the operational OMI algorithm in order to avoid adjusting our reported NO_2 columns by an additive amount derived from a model.

We have demonstrated that a tropospheric AMF correction threshold T_0 is optimal and found that thresholds of $\mathrm{T}_{-\infty}$ or $\mathrm{T}_{+\infty}$ yield less accurate retrievals. Additional tests showed no improvement in retrieval accuracy for any finite threshold setting T_ε (for example, ε could equal the 1σ variation of the actual stratospheric field about its smoothed value). It is clear that for the symmetric case of a purely stratospheric field containing random positive and negative variations, the use of threshold T_0 would lead to a positive bias in the total VCD retrieval. However, the tropospheric NO_2 in our simulated data set introduces an asymmetry in the VCD field. Use of the $\mathrm{T}_{-\infty}$ threshold would enhance the negative error bias (Fig. 3) and, as we have shown, increase the tropospheric retrieval error. We expect the same results to hold for OMI retrievals, even in the presence of noise.

Estimation of the tropospheric AMF, itself, depends mainly on clouds, surface albedo, and the NO_2 profile, the latter two being especially important over land areas [13]. Boersma *et al.* [19] have shown that uncertainties in cloud fraction and albedo are the two largest sources of error in estimating the tropospheric AMF, contributing relative errors of 20% to 30% and 20% to 50%, respectively, over moderately-to-heavily polluted regions. In contrast, NO_2 profile variability in these areas leads to AMF errors less than 15%. AMF errors from all sources over oceans were determined to be generally smaller than over land. The reported effects of profile uncertainties on AMF error are consistent with the findings of Martin *et al.* [13] and Heland *et al.* [15], who also estimate profile-related AMF errors $<15\%$ due to variability and/or uncertainties in the vertical distribution of boundary layer NO_2. We have estimated the effects of daily profile variability on the absolute NO_2 VCD error in our own comparison of retrievals based on $\mathrm{P}_{\mathrm{daily}}$ and $\mathrm{P}_{\mathrm{avg}}$ profiles and shown significant errors to be less prevalent than in the case of retrievals employing a constant mixing-ratio profile. Although we have not investigated albedo or cloud effects in this study, we have found that profile choice has a comparable—though generally smaller—effect on absolute retrieval error than choices regarding masking, the stratospheric field model and the AMF correction threshold.

VI. Conclusion

In this paper, we have described the algorithm that is used to retrieve NO_2 vertical column densities from the OMI instrument on the EOS-Aura satellite. The algorithm involves spectral fitting to determine NO_2 slant column densities, applying stratospheric AMFs to compute an initial VCD field, and then the using tropospheric AMFs to correct the VCDs in polluted regions. We have compared retrievals that use different methods for computing AMFs and identifying and correcting VCDs in polluted regions. These methods can be described in terms of a number of retrieval parameters. We tested the algorithm on simulated data sets from models in order to identify combinations of parameters that yield accurate retrievals of the total and tropospheric VCD. We have estimated minimum values for the retrieval errors expected in the processing of real OMI data.

An important consideration in computing AMFs is the assumed shape of the NO_2 profile. We have found that accurate AMFs can be obtained with a gridded yearly-averaged set of tropospheric NO_2 profiles and a single mean stratospheric profile. The stratosphere-troposphere separation is accomplished in a two-step process by masking regions with significant tropospheric NO_2 and then modeling stratospheric variation in the unmasked areas with a sum of sinusoidal terms. The optimal approach is to mask areas with mean annual tropospheric VCDs greater than 0.5×10^{15} cm^{-2}, and then fit the remainder with wave-2. This method has been shown to be an improvement on the Pacific reference sector method, which assumes a zonally invariant stratosphere calculated from VCDs in the central Pacific Ocean. A tropospheric AMF correction is applied to all areas where the initial NO_2 field exceeds the modeled stratospheric field.

An algorithm similar to that described in this study is currently being used to process NO_2 data from OMI. Although we have optimized the algorithm based on simulated data with cloud fractions less than 25%, the scenes actually observed often have larger cloud fractions. Since some of the effects of clouds on tropospheric AMFs are not yet well characterized, we currently report only the *visible* VCD as the official data product from the operational algorithm—i.e., we do not estimate amounts of NO_2 hidden beneath clouds. An example of the VCD data product from OMI has been presented in this paper. Analysis of the early results from OMI will be the basis for future algorithm modifications, including better treatment of clouds and aerosols. Additional comparisons of retrieval methods are planned for future studies.

Acknowledgment

The authors would like to thank R. Martin for generating the GEOS-CHEM tropospheric mixing ratios used in the AMF computations. They would also like to thank B. Polansky for stratospheric NO_2 computations with the Goddard CTM model.

References

[1] P. J. Crutzen, "The influence of nitrogen oxides on the atmospheric ozone content," *Q. J. R. Meteorol. Soc.*, vol. 96, pp. 320–325, 1970.

[2] D. Murphy, D. Fahey, M. Proffitt, S. Liu, C. Eubank, S. Kawa, and K. Kelly, "Reactive odd nitrogen and its correlation with ozone in the lower stratosphere and upper troposphere," *J. Geophys. Res.*, vol. 98, pp. 8751–8774, 1993.

[3] A. W. Brewer, C. T. McElroy, and J. B. Kerr, "Nitrogen dioxide concentrations in the atmosphere," *Nature*, vol. 246, pp. 129–133, 1973.

[4] J. F., Noxon, "Nitrogen dioxide in the stratosphere and troposphere measured by ground-based absorption spectroscopy," *Science*, vol. 189, pp. 547–549, 1975.

[5] J. F. Noxon, E. C. Whipple, Jr., and R. S. Hyde, "Stratospheric NO_2, 1. Observational method and behavior at mid-latitude," *J. Geophys. Res.*, vol. 84, pp. 5047–5065, 1979.

[6] J. F. Noxon, "Stratospheric NO_2, 2: Global behavior," *J. Geophys. Res.*, vol. 84, pp. 5067–5076, 1979.

[7] S. Solomon and R. R. Garcia, "On the distribution of long-lived tracers and chlorine species in the middle atmosphere," *J. Geophys. Res.*, vol. 89, pp. 11 633–11 644, 1984.

[8] M. Wenig, S. Kühl, S. Beirle, E. Bucsela, B. Jähne, U. Platt, J. Gleason, and T. Wagner, "Retrieval and analysis of stratospheric NO_2 from GOME," *J. Geophys. Res.*, vol. 109, no. D04315, 2004. DOI: 10.1029/2003JD003652.

[9] J. P. Burrows, M. Weber, M. Buchwitz, V. V. Rozanov, A. Ladstatter-Weissenmayer, A. Richter, R. DeBeek, R. Hoogen, K. Bramstedt, and K. U. Eichmann, "The Global Ozone Monitoring Experiment (GOME): Mission concept and first scientific results," *J. Atmos. Sci.*, vol. 56, pp. 151–175, 1999.

[10] C. Leue, M. Wenig, T. Wagner, O. Klimm, U. Platt, and B. Jaehne, "Quantitative analysis of NO_x emissions from global ozone monitoring experiment satellite image sequences," *J. Geophys. Res.*, vol. 106, pp. 5493–5505, 2001.

[11] A. Richter and J. P. Burrows, "Retrieval of tropospheric NO_2 from GOME measurements," *Adv. Space Res.*, vol. 29, pp. 1673–1683, 2002.

[12] G. J. Velders, M. C. Granier, R. W. Portmann, K. Pfeilsticker, M. Wenig, T. Wagner, U. Platt, A. Richter, and J. P. Burrows, "Global tropospheric NO_2 column distributions: Comparing 3-D model calculations with GOME measurements," *J. Geophys. Res.*, vol. 106, pp. 12,643–12 660, 2001.

[13] R. V. Martin, K. Chance, D. J. Jacob, T. P. Kurosu, R. J. D. Spurr, E. Bucsela, J. F. Gleason, P. I. Palmer, I. Bey, A. M. Fiore, Q. Li, R. M. Yantosca, and R. B. A. Koelemeijer, "An improved retrieval of tropospheric nitrogen dioxide from GOME," *J. Geophys. Res.*, vol. 107 (D20), no. 4437, 2002. DOI: 10.1029/2001JD001027.

[14] A. Lauer, M. Dameris, A. Richter, and J. P. Burrows, "Tropospheric NO_2 columns: A comparison between model and retrieved data from GOME measurements," *Atmos. Chem. Phys.*, vol. 2, pp. 67–78, 2002.

[15] J. Heland, H. Schlager, A. Richter, and J. P. Burrows, "First comparison of tropospheric NO_2 column densities retrieved from GOME measurements and *in situ* aircraft profile measurements," *Geophys. Res. Lett.*, vol. 29 (20), no. 1983, 2002. DOI: 10.1029/2002GL015528.

[16] H. Bovensmann, J. P. Burrows, M. Buchwitz, J. Frerick, V. V. Rozanov, K. V. Chance, and A. P. H. Goede, "SCIAMACHY: Mission objectives and measurement modes," *J. Atmos. Sci.*, vol. 56, pp. 127–150, 1999.

[17] R. V. Martin, C. E. Sioris, K. V. Chance, T. B. Ryerson, T. H. Bertram, P. J. Woolridge, R. C. Cohen, J. A. Neuman, A. Swanson, and F. M. Flocke, "Evaluation of space-based constraints on nitrogen oxide emissions with regional aircraft measurements over and downwind of eastern North America," *J. Geophys. Res.*, 2005, submitted for publication.

[18] M. Wenig, N. Spichtinger, A. Stohl, G. Held, S. Beirle, T. Wagner, B. Jähne, and U. Platt, "Intercontinental transport of a power plant plume of nitrogen oxides," *Atmos. Chem. Phys.*, vol. 3, pp. 387–393, 2003.

[19] K. F. Boersma, H. J. Eskes, and E. J. Brinksma, "Error analysis for tropospheric NO_2 retrieval from space," *J. Geophys. Res.*, vol. 109, no. D04311, 2004. DOI: 10.1029/2003JD003962.

[20] P. F. Levelt, E. Hilsenrath, G. W. Leppelmeier, G. B. J. van den Oord, P. K. Bhartia, J. Tamminen, J. F. de Haan, and J. P. Veefkind, "Science objectives of the Ozone Monitoring Instrument," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 5, pp. 1199–1208, May 2006.

[21] F. Boersma, E. Bucsela, E. Brinksma, and J. F. Gleason, OMI Algorithm Theoretical Basis Document, K. Chance, Ed., vol. 4, ch. 2, 2002.

[22] J. Joiner, P. K. Bhartia, R. P. Cebula, E. Hilsenrath, and R. D. McPeters, "Rotational-raman scattering (ring effect) in satellite backscatter ultraviolet measurements," *Appl. Opt.*, vol. 34, pp. 4513–4525, 1995.

[23] J. Joiner and P. K. Bhartia, "The determination of cloud pressures from rotational-Raman scattering in satellite backscatter ultraviolet measurements," *J. Geophys. Res.*, vol. 100, pp. 23 019–23 026, 1995.

[24] J. Joiner, A. Vasilkov, D. Flittner, E. Bucsela, and J. Gleason, "Retrieval of cloud pressure from rotational Raman scattering," OMI Algorithm Theoretical Basis Document, vol. 3, 2001.

[25] J. R. Acarreta, J. F. deHaan, and P. Stammes, "Cloud pressure retrieval using the $O_2 - O_2$ absorption band at 477 nm," *J. Geophys. Res.*, vol. 109, no. D05204, 2004. 10.1029/2003JD003915.

[26] O. Torres, R. Decae, P. Veefkind, and G. de Leeuw, "OMI aerosol retrieval algorithm," OMI Algorithm Theoretical Basis Document, vol. 3, 2001.

[27] U. Platt and J. Stutz, *Differential Optical Absorption Spectroscopy (DOAS), Principle and Applications*. Heidelberg, Germany: Springer Verlag, 2006.

[28] M. Wenig, B. Jähne, and U. Platt, "An operator notation as a new DOAS formalism," *Appl. Opt.*, vol. 44, pp. 3246–3253, 2005.

[29] G. Mount, R. Sanders, A. Schmeltekopf, and S. Solomon, "Visible spectroscopy at McMurdo station, Antarctica, 1. Overview and daily variation of NO_2 and O_3, austral spring, 1986," *J. Geophys. Res.*, vol. 92, pp. 8320–8328, 1987.

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

[31] A. C. Vandaele, C. Hermans, P. C. Simon, M. Carleer, R. Colin, S. Fally, M. F. Mérienne, A. Jenouvrier, and B. Coquart, "Measurements of the NO_2 absorption cross-section from 42 000 cm^{-1} to 10000 cm^{-1} (238–1000 nm) at 220 K and 294 K," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 59, pp. 171–184, 1998.

[32] J. P. Burrows, A. Richter, A. Dehn, B. Deters, S. Himmelmann, S. Voigt, and J. Orphal, "Atmospheric remote sensing reference data from GOME-2. Temperature-dependent absorption cross-sections of O_3 in the 231–794 nm range," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 61, pp. 509–517, 1999.

[33] A. R. Douglass, M. R. Schoeberl, R. B. Rood, and S. Pawson, "Evaluation of transport in the lower tropical stratosphere in a global chemistry and transport model," *J. Geophys. Res.*, vol. 108 (D9), no. 4529, 2003. 10.1029/2002JD002696.

[34] I. Bey, D. J. Jacob, R. M. Yantosca, J. A. Logan, B. D. Field, A. M. Fiore, Q. Li, H. Y. Liu, L. J. Mickley, and M. G. Schultz, "Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation," *J. Geophys. Res.*, vol. 106, pp. 23 073–23096, 2001.

[35] R. V. Martin, D. J. Jacob, J. A. Logan, I. Bey, R. M. Yantosca, A. C. Staudt, Q. Li, A. M. Fiore, B. N. Duncan, H. Liu, P. Ginoux, and V. Thouret, "Interpretation of TOMS observations of tropical tropospheric ozone with a global model and *in situ* observations," *J. Geophys. Res.*, vol. 107 (D18), no. 4351, 2002. 10.1029/2001JD001480.

[36] R. D. McPeters, P. K. Bhartia, A. J. Krueger, J. R. Herman, C. G. Wellemeyer, C. J. Seftor, G. Jaross, O. Torres, L. Moy, G. Labow, W. Byerly, S. L. Taylor, T. Swissler, and R. P. Cebula, Earth Probe Total Ozone Mapping Spectrometer (TOMS) Data Products Users' Guide, NASA/TP-1998-206895, Nov. 1998.

[37] B. R. Barkstrom, "The Earth Radiation Budget Experiment (ERBE)," *Bull. Amer. Meteorol. Soc.*, vol. 65, pp. 1170–1185, 1984.

[38] R. A. Schiffer and W. B. Rossow, "The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Programme," *Bull. Amer. Meteorol. Soc.*, vol. 64, pp. 779–784, 1983.

[39] W. B. Rossow and R. A. Schiffer, "ISCCP cloud data products," *Bull. Amer. Meteorol. Soc.*, vol. 72, pp. 2–20, 1991.

[40] R. M. Nagatani, A. J. Miller, K. W. Johnson, and M. E. Gelman, "An eight-year climatology of meteorological and SBUV ozone data," NOAA, Boulder, CO, NOAA Tech. Rep. 17, 333, 1990.

[41] P. I. Palmer, D. J. Jacob, K. Chance, R. V. Martin, R. J. D. Spurr, T. P. Kurosu, I. Bey, R. Yantosca, and A. Fiore, "Air-mass factor formulation for spectroscopic measurements from satellites," *J. Geophys. Res.*, vol. 106, pp. 14 539–14 550, 2001.

[42] J. V. Davé, "Meaning of successive iteration of the auxiliary equation in the theory of radiative transfer," *Astrophys. J.*, vol. 140, pp. 1292–1303, 1964.

[43] ——, "Multiple scattering in a nonhomogeneous, Rayleigh atmosphere," *J. Atmos. Sci.*, vol. 22, pp. 273–279, 1965.

[44] Z. Ahmad and P. K. Bhartia, "Effect of molecular anisotropy on the backscattered UV radiance," *Appl. Opt.*, vol. 34, pp. 8309–8314, 1995.

[45] J. F. Noxon, "Correction to 'Stratospheric NO_2, 1. Observational method and behavior' ," *J. Geophys. Res.*, vol. 85, pp. 4560–4561, 1980.

Eric J. Bucsela received the B.S. degree in physics from Emory University, Atlanta, GA, in 1983, and the M.S. degree in electrical science and the Ph.D. degree in physics from the University of Michigan, Ann Arbor, in 1986 and 1994, respectively.

He is currently an Associate Research Scientist at the GEST/University of Maryland Baltimore County (UMBC), Baltimore. His work has involved aeronomy of the dayglow using *in situ* sounding-rocket measurements. He completed a postdoctoral fellowship at the Naval Research Laboratory in 1997. He has worked with the NASA Goddard Space Flight Center, Greenbelt, MD, since 1999, through affiliations with Raytheon, SSAI, and, most recently, the UMBC. His work involves satellite data analysis to investigate physical and chemical properties of the stratosphere and troposphere. His current research is a collaborative effort with the Netherlands' KNMI to retrieve trace-gas amounts from the Ozone Monitoring Instrument, onboard the EOS-Aura satellite.

Edward A. Celarier received the B.S. degrees in chemistry and physics/astronomy from the University of Washington, Seattle, in 1981, and the Ph.D. degree in physical chemistry from the University of Toronto, Toronto, ON, Canada, in 1986.

He is currently a Scientist with SGT, Inc., Greenbelt, MD. He was a Postdoctoral Fellow with the Chemistry Department, Princeton University, Princeton, NJ, from 1986 to 1989, and went on to become an Assistant Professor of chemistry at Hampton University, Hampton, VA, from 1989 to 1995. He has been working on contract at the NASA Goddard Space Flight Center, Greenbelt, MD, for 11 years in the Atmospheric Chemistry and Dynamics Branch (Code 613.3) in the area of remote-sensing measurements of trace gases (ozone and nitrogen dioxide) using satellite instruments (TOMS, SBUV, SAGE, OMI), as well as radiative transfer in the atmosphere.

Mark O. Wenig received the Intermediate Diploma in physics from the University of Münster, Münster, Germany, in 1995, the Diploma in physics from the Interdisciplinary Center for Scientific Computing, University of Heidelberg, Heidelberg, Germany, in 1998, and the Ph.D. degree in physics from the Institute of Environmental Physics, University of Heidelberg, in 2001, for his work on satellite measurement of long-term global tropospheric trace gas distributions and source strengths.

Since 2005, he has been affiliated with the University of Maryland Baltimore County/GEST to work at the NASA Goddard Space Flight Center (GSFC), Greenbelt, MD, on new techniques to combine multidimensional data sets coming from satellite measurements to analyze the impact of trace gases and aerosols on air quality.

Dr. Wenig was awarded a National Research Council scholarship from 2002 to 2005 to work at the NASA GSFC, in the Atmospheric Chemistry and Dynamics Branch on atmospheric data assimilation.

James F. Gleason received the Ph.D. degree in chemistry from the University of Colorado, Boulder, in 1987.

He is currently a Senior Research Scientist with the NASA Goddard Space Flight Center, Greenbelt, MD. After a postdoctoral fellowship at Brookhaven National Laboratory, Upton, NY, he joined Goddard as an NRC Resident Research Associate in 1989. From 1993 to 1996, he was the EOS-Aura Project Scientist, and he currently serves as the Project Scientist for the NPOESS Preparatory Project (NPP) mission. He is the Principal Investigator for the NO_2 data product on the NASA OMI Science team.

J. Pepijn Veefkind is an expert in the field of satellite remote sensing of aerosols and trace gases. He has experience in developing and validating retrieval algorithms for satellite based radiometers and spectrometers. During his Ph.D. research, he developed different techniques for the retrieval of aerosol optical thickness from GOME, ATSR-2, and AVHRR data and participated in several international field experiments. Since 1999, he has been working with the Ozone Monitoring Instrument (OMI) PI Team, which is part of the Atmospheric Composition Division of Royal Netherlands Meteorological Institute (KNMI). He is responsible for the development and implementation of retrieval algorithms for the OMI ozone column, ozone profile, NO_2, aerosol, and cloud products. He is chairing the Algorithms Working Group of the International OMI Science Team. His research interests include retrieval techniques, global monitoring of aerosols and trace gases, intercomparison of satellite measurements from different sensors, radiative forcing of climate, and satellite detection of air pollution.

K. Folkert Boersma received the M.Sc. degree in physics and environmental physics from the Free University in Amsterdam, Amsterdam, The Netherlands, in 1998. For his M.S. thesis, he operated a lidar system to measure stratospheric ozone at the National Institute for Water and Atmospheric research (NIWA), Lauder, New Zealand. He received the Ph.D. degree in physics from the Technical University Eindhoven, Eindhoven, The Netherlands, in 2005, with a dissertation titled "Tropospheric nitrogen dioxide—retrieval, interpretation and modeling."

He is currently a Postdoctoral Researcher at Harvard University, Cambridge, MA. He was a Postdoctoral Researcher at the Royal Netherlands Meteorological Institute (KNMI)/Technical University Eindhoven, Eindhoven, The Netherlands. After a short stay at the Ministry of Foreign Affairs, he started as a Junior Scientist at KNMI, working on retrieval algorithms for the Ozone Monitoring Instrument together with NASA. Since 2002, he has been focusing on algorithm development and error and data analysis with the GOME and SCIAMACHY satellite instruments.

Ellen J. Brinksma studied astrophysics and received the Ph.D. degree in lidar studies of the middle atmospere from Vrije Universiteit, Amsterdam, The Netherlands, in 2001.

In January 2000, she joined the OMI Science Team at Royal Netherlands Meteorological Institute, where she works on satellite validation, satellite retrievals, and ground-based measurements, with a focus on ozone and nitrogen dioxide. She has published 15 peer-reviewed papers on atmospheric science topics.