Acceleration of multiple-scattering, hyperspectral radiative transfer calculations via low-streams interpolation

Christopher W. O'Dell[1]

[1]   A new generation of near-infrared instruments, such as the Greenhouse Gases Observing Satellite (GOSAT) and the ill-fated Orbiting Carbon Observatory (OCO), aims to measure concentrations of trace gases through the high-resolution measurements of spectra of reflected sunlight. Accurate retrievals of trace gas amounts from these data present a computational challenge for the underlying radiative transfer algorithms because these algorithms must simulate the angle-dependent radiances at high spectral resolution within near-infrared gas absorption bands in scattering atmospheres. Such calculations can be prohibitively expensive for certain applications. Building upon previous work by others, we describe a simple yet accurate technique to accelerate the radiative transfer of many monochromatic wavelengths in the visible and near infrared. It is tested over a large variety of profile types, including cases with high solar zenith angle, large amounts of atmospheric scattering, highly polarizing, non-Lambertian surfaces, and bands with overlapping gases. Speed gains of 1–2 orders of magnitude are demonstrated, with typical errors less than 0.1% for most atmospheres and less than 0.5% for even the most challenging cases.

1.   Introduction

[2] Trace gases such as carbon dioxide are now being measured from space using spectroscopic observations of reflected sunlight in the near infrared. The SCanning Imaging Absorption spectroMeter for Atmospheric CartograpHY (SCIAMACHY) has been measuring various trace gases such as CO_2, CO, CH_4, and O_3 since March 2002 aboard the ENVISAT satellite [*Bovensmann et al.*, 1999]. The Greenhouse gases Observing SATellite (GOSAT) was launched in January 2009 and carries the Thermal And Near-infrared Sensor for carbon Observation Fourier Transform Spectrometer (TANSO-FTS), designed to make global observations of near-surface CO_2 and CH_4 concentrations [*Hamazaki et al.*, 2005]. The Orbiting Carbon Observatory (OCO) was to make similar measurements of carbon dioxide [*Crisp et al.*, 2004] but experienced a launch failure in February 2009. The primary objective of these missions is to identify and quantify the sources, where these gases are emitted, and sinks, where they are removed from the atmosphere. This is a particularly challenging remote sensing measurement because the presence of surface sources and sinks must be inferred from the small spatial and temporal variations that they produce in the background distribution. Simulations of CO_2 fluxes with source-sink

inversion models indicate that estimates of the column-averaged CO_2 dry air mole fraction with accuracies of 1 ppm (out of the 388 ppm background) are needed on regional scales at monthly intervals to retrieve CO_2 sources and sinks on these scales [*Miller et al.*, 2007; *Chevallier et al.*, 2007; *Baker et al.*, 2008]. This requires retrieval models with accuracies of a fraction of one percent.

[3] For OCO, GOSAT, and other similar sensors, model simulations of spectra are generally required for parameter retrievals, data assimilation, and also simple visualization and sensitivity studies. However, in atmospheres with even a modest amount of scattering, multiple-scattering calculations are required to obtain accurate estimates of visible and near-infrared instrumental radiances. These simulations can be time-consuming as they often require tens of thousands of monochromatic radiative transfer (RT) calculations for simulating a single instrumental spectrum. Remote sensing retrieval calculations are even more demanding because they require not just radiances, but often the radiance Jacobians as well, which specify the derivatives of the radiances with respect to each parameter of interest (e.g., trace gas abundance, aerosol optical properties and distribution, surface properties) for each monochromatic spectral point.

[4] Many multiple scattering codes in use today rely on the adding-doubling method [e.g., *van de Hulst*, 1980; *Hansen*, 1971] or the discrete ordinate method [e.g., *Stamnes et al.*, 1988]. Both methods compute the radiation coming along a discrete set of zenith angles or streams, and interpolate between these angles to find the radiation at other angles. A code is said to be run with $2N_s$ streams when

[1]Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado, USA.
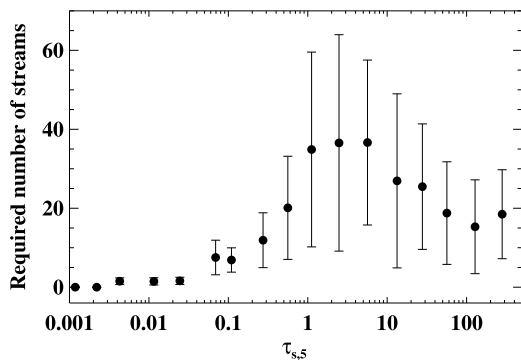
Figure 1. Number of streams required to achieve 0.1% accuracy in TOA intensity, versus a reference calculation with 256 streams. The independent variable $\tau_{s,5}$ is the column-integrated scattering optical depth above the point in the atmosphere where the cumulative gas absorption optical depth to the TOA equals 5.0. See text for details.

N_s is the number of streams in either hemisphere. In nadir viewing mode, the computational burden is proportional to N_s^3, because the codes largely rely on multiplication and inversion of $N_s \times N_s$ matrices. For off-nadir viewing, an additional loop over N_s is required due to the additional azimuthal Fourier modes required [see, e.g., *Liou*, 2002], leading to an approximate N_s^4 dependency. Therefore, using the smallest number of streams allowed for a given application is often highly desirable.

[5] The number of streams required to achieve a given accuracy varies with the amount and type of atmospheric scattering and absorption, as well as the surface properties. Generally speaking, the stronger the atmospheric scattering, the higher the number of streams required. As an example, Figure 1 shows the number of full-sphere streams required to achieve 0.1% accuracy in top-of-atmosphere (TOA) intensity for a wide variety of profiles and near-infrared wavelengths, versus a reference calculation with 256 streams. The profiles are a subset of the *Chevallier* [2001] set of 60-layer ECMWF profiles; see section 2 for details of the optical properties and radiative transfer. The solar zenith angle is variable, but the observation zenith angle is fixed at 0° The independent variable in Figure 1 is the integrated scattering optical depth from TOA to the level in the atmosphere where the cumulative gas absorption optical depth to the TOA equals 5.0; scattering below this point in the atmosphere has little effect on the TOA radiance, as virtually all radiation from below this point will be absorbed before reaching the TOA. The error bars represent ±1 standard deviation and represent the variability among different profiles. Note that these calculations were made with a doubling-adding code and employed both Delta-M scaling [*Wiscombe*, 1977] and the TMS (single-scattering) correction of *Nakajima and Tanaka* [1988]; these methods can substantially reduce the required number of streams to achieve a given accuracy for most cases. It is seen that for more significantly scattering profiles, a high N_s is often required to achieve this level of accuracy.

[6] To understand the implications of Figure 1, let us consider a typical optimal estimation approach to retrieving CO_2 concentration from OCO or GOSAT measurements

[e.g., *Bösch et al.*, 2006; *Oshchepkov et al.*, 2008; *Butz et al.*, 2009]. OCO and GOSAT measure roughly the same shortwave bands; Table 1 gives some details of the OCO bands. OCO and GOSAT were designed to yield CO_2 estimates with random and systematic biases no larger than ~0.3% [*Crisp et al.*, 2004; *Suto et al.*, 2008]. To obtain this, we suggest the corresponding radiative transfer errors to be in the neighborhood 0.1%. This is somewhat less than the expected instrument noise of OCO, and it was shown by *Hasekamp and Butz* [2008] that RT errors on the order of 0.1% lead to CO_2 retrieval errors of 0.05–0.20 ppm, which constitutes an acceptable 20% or less of the 0.3% total error budget.

[7] To fully resolve the spectral lines, on the order of 40,000 spectral points are required for the simulation of all three OCO bands. The standard OCO retrieval algorithm was to be run for scattering optical depth of 0.3 or less; the GOSAT team will process similarly clear profiles. From Figure 1 it is seen that roughly 16 streams would generally be needed to reach the desired 0.1% accuracy level for this amount of scattering. Performing a typical line-by-line approach thus would require tens of thousands of independent, 16-stream calculations.

[8] Many methods have been put forth to reduce the computational expense of the line-by-line approach described above. Most strive to reduce the number of monochromatic calculations needed, or the complexity of the multiple scattering calculation needed at each spectral grid point. Traditional correlated k distribution methods [e.g., *Goody et al.*, 1989; *Lacis and Oinas*, 1991; *Fu and Liou*, 1992] are an example of the first class of methods, but can lead to errors of tens of percent even in modestly scattering atmospheres [*Duan et al.*, 2005]. However, several techniques have been proposed to reduce the computational time for multiple scattering calculations in the visible and near IR [*West et al.*, 1990; *Min and Harrison*, 2004; *Natraj et al.*, 2005; *Duan et al.*, 2005; *Hasekamp and Butz*, 2008; *Boesche et al.*, 2009]. Most of these techniques replace the line-by-line calculations at each spectral wavelength with a smaller set of representative calculations, typically binned by absorption optical depth or something similar, and reconstruct the full spectrum from this small set of calculations. Some of the techniques also take into account the vertical structure of the absorption in some fashion in order to reduce errors. Several of the techniques achieve errors on the order of 1–2%, while some are substantially better. *Natraj et al.* [2005] (hereafter N05) reports errors of less than 0.3% in the O_2 A band, but only tests a single case with total aerosol optical depth (AOD) of 0.2. *Hasekamp and Butz* [2008] (hereafter H08) report errors of better than 0.13% in all three OCO bands for a similarly thin case (AOD = 0.3), and better than 0.5% for a more challenging

Table 1. Some Instrument Properties of the Orbiting Carbon Observatory

Band	Name	Spectral Range (cm^{-1})	ILS FWHM[a] (cm^{-1})
1	Oxygen A	12,950–13,200	0.63
2	WCO$_2$	6170–6290	0.27
3	SCO$_2$	4805–4900	0.21

[a]ILS FWHM, instrument line shape full-width at half-maximum.

high-cloud case. *Duan et al.* [2005] (hereafter D05) report errors better than 0.5% for all but the most challenging multilayered cloud systems. Most authors report results for nadir observations and modest solar zenith angles of 30°–40°.

[9] Of the studies described above, the work of H08, N05, and D05 achieve errors at or near the desired accuracy for OCO-type retrievals, but all have some drawbacks. H08 requires both radiances and Jacobians to be calculated, as it works using a first-order correction to the binned radiance involving the corresponding derivatives with respect to the atmospheric and surface optical properties. While the N05 technique does not require derivatives, it requires on the order of 1000 high-accuracy calculations in the oxygen A band alone, substantially more than the other approaches. The technique of D05 requires only a small number of bins and does not require radiance derivatives. They calculate the first order of scattering contribution to each spectral radiance exactly and model the multiple-scattering component as piecewise analytic functions in the absorption optical thickness and a variable related to the fraction of the absorption that occurs above scattering layers of cloud or aerosol. They apply their method for total intensity (polarization was not considered), and mostly consider simple one and two-layer cloud systems in their analysis; it is not clear if their analytic fits will generalize to situations with more complicated profiles of scattering or with polarizing atmospheric and/or surface properties.

[10] In this paper, we describe a simple, tunable method to accelerate visible and near-infrared multiple-scattering radiance calculations with a minimal loss in accuracy. This method, called Low-Streams Interpolation (LSI), can be used in conjunction with any scalar or vector multiple-scattering radiative transfer model, and does not require derivatives to be calculated. This method builds upon the work of several previous authors but differs in several important respects. It identifies a strong relationship between radiative transfer errors and two easy-to-calculate properties of the gas absorption, and utilizes this relationship to accelerate the RT calculations. The motivation for and description of the method are given in section 2. In section 3, results are shown depicting both the accuracy and computational efficiency of the method. Section 4 gives a summary of the results and compares and contrasts the technique to those of previous authors.

2. Methodology

2.1. Background and Reference Model

[11] High-resolution spectrometers are often sensitive to a linear combination of Stokes parameters [e.g., *Chandrasekhar*, 1960]; this is true in particular for both the GOSAT and OCO near-infrared sensors. This is typically expressed as

$$I_{meas} = m_{11}I + m_{12}Q + m_{13}U + m_{14}V, \qquad (1)$$

where I_{meas} is the instrument-measured radiance, and the m_{ij} coefficients are elements of the instrument Mueller matrix; these coefficients are a combination of instrument parameters and satellite viewing geometry and hence are a function of known quantities. Stokes parameter V is typi-

cally negligible in the atmosphere, and most instruments are insensitive to it. It will therefore not be considered further in this work.

[12] For the monochromatic RT computations described herein, the Stokes parameters are calculated as follows:

$$
\begin{aligned}
I &= I_{ss,tms} + I_{ms}, \\
Q &= Q_{ss} + Q_2, \qquad (2) \\
U &= U_{ss} + U_2,
\end{aligned}
$$

where ss stands for single scattering and ms stands for multiple scattering. $I_{ss,tms}$ is calculated not exactly, but rather using the TMS correction as defined by *Nakajima and Tanaka* [1988]. This is to cancel out errors due to phase function truncation that takes place in the multiple-scattering calculation. I_{ms} is the multiply scattered intensity calculated by a scalar code, the Successive Orders of Interaction (SOI) model [*Heidinger et al.*, 2006] adapted for shortwave use; it employs Delta-M scaling as well as a first-order correction to account for spherical geometry [*Spurr*, 2002]. Q_{ss} and U_{ss} are calculated using an exact first order of scattering calculation, and Q_2 and U_2 are the contributions to each from the second order of scattering. For relatively clear atmospheres (AOD < 0.3), *Natraj and Spurr* [2007] have shown that retaining only the second order of scattering components of Q and U is generally sufficient; we calculate Q_2 and U_2 using their exact method. This method is also quadrature-based and takes as input a given number of streams; in practice we use the same quadrature angles as in the calculation of I_{ms}. The results described below are expected to be general even though a full vector code has not been used.

[13] We begin with the hypothesis that errors incurred by RT calculations performed at a lowered accuracy, relative to a high-accuracy reference calculation, will be correlated with the optical properties of the profile in question. For the simulations presented here, we define the high-accuracy reference calculation with 24 streams and using optical properties calculated on the original set of atmospheric layers for all calculations required by equation (2). As noted by *Min and Harrison* [2004], reduced accuracy can result from (1) fewer streams in the scalar, multiple-scattering calculations; (2) reduced vertical resolution in the multiple-scattering calculations; and (3) complete neglect of the multiply scattered components of Stokes parameters Q and U, though in principle other parameters of the RT could be varied as well. For the work herein, a simple two-stream calculation for I_{ms} is employed. To reduce the number of layers for this calculation, we simply combine the optical properties (optical depth, single scattering albedo, and phase functions) of adjacent sets of N_{sum} layers together. Higher values of N_{sum} result in faster but less accurate calculations. For our calculations, N_{sum} was set to 3 for profiles with 60 atmospheric layers. The optical properties were combined for a set of layers exactly as they would be when combining optical properties for a set of scattering species: optical depths add, scattering optical depths add, and phase matrix elements weighted by scattering optical depth add. Table 2 provides a summary of the settings used in the high- and low-accuracy RT calculations.

Table 2. Typical Parameters of the Monochromatic Radiative Transfer Calculations

Calculation	Streams	N_{sum}	Q_2, U_2
High Accuracy	24	1	Yes
Low Accuracy	2	3	No

2.2. Study Cases

[14] Let us now examine the errors of the low-accuracy calculations relative to those of high accuracy, using the three OCO bands as a convenient test bed. We have chosen four diverse atmospheric profiles for case studies, taken from a standard set of ECMWF model profiles [*Chevallier*, 2001]. The profiles each have 60 vertical layers; the vertical cloud and aerosol distributions along with the water vapor path (WVP) for each are shown in Figure 2. Case 1 is a nearly clear case with a low, thin water cloud. Case 2 contains a high, thin cirrus cloud. Case 3 contains both water and ice cloud, and case 4 contains thick, overlapping water and ice clouds. Cases 1–3 also contain a low, thin aerosol layer of a standard continental type, while case 4 contains an oceanic aerosol layer near the surface; optical properties for the aerosols are drawn from *Shettle and Fenn* [1979]. Optical properties for water clouds were calculated from Mie theory assuming a Gamma particle size distribution. The optical properties of ice clouds with a given effective radius were interpolated using the model described by *Baum et al.* [2005a, 2005b]. The mean particle radius for ice and water cloud is generally different in different atmospheric layers. Water cloud, ice cloud, and aerosol all include a fully polarized treatment of the scattering phase function, using a standard formalism [*Siewert*, 1982; *de Rooij and van der Stap*, 1984]. Gas absorption cross sections are taken from HITRAN 2008 [*Rothman et al.*, 2009]. Rayleigh scattering, including polarization, follows the treatment of *Bodhaine et al.* [1999].

[15] Two viewing geometries and surface types are investigated. For nadir observations, the observation view angle is zero, and the surface is Lambertian with an albedo

that varies linearly from 0.09 to 0.11 over the width of each band. In glint mode, the observation direction points exactly at the solar glint spot on an ocean surface, which is described by a fully polarized Cox and Munk model [*Cox and Munk*, 1954], assuming an ocean wind speed of 10 m s^{-1}.

2.3. Relationships of Errors With Optical Properties

[16] Examples of errors in Stokes parameters I and Q for the low-accuracy calculations relative to the high-accuracy calculations in the oxygen A band are shown in Figure 3 for nadir observation mode with a solar zenith angle (SZA) of 40° The errors are plotted versus the column-integrated gas optical depth τ_g, which in this case is due only to oxygen, and are expressed in percent. While there is some scatter in the plots, it is clear that there is a strong but complicated relationship between the error and τ_g for each Stokes parameter.

[17] A simple RT scheme is therefore motivated by Figure 3. If one could calculate the line-by-line RT using only the low-accuracy calculations, and effectively sample the curves in Figure 3 with only a small number of high-accuracy calculations, one could reconstruct the error curve by interpolation. The resulting curve could then be used to correct the low-accuracy calculations. An alternative approach would be to somehow parameterize these error curves, but due to their apparent complexity this could prove challenging. A simple way to accomplish interpolation is to divide the τ_g axis into a small set of bins, and average together the optical properties of all spectral points falling in each bin. However, because the spectral bands are narrow, we simply average the gas absorption profiles, and take the scattering properties of the atmosphere and surface to be those at the band center.

[18] Unfortunately, after applying such a scheme one would still be left with the inherent scatter in the curves of Figure 3. Because the scattering optical properties change slowly within this narrow band, the scatter is likely due to spectroscopic features. Because the wings of individual absorption lines often overlap, a given τ_g can be a sum of the midwing contributions of two lines, or be due to the core
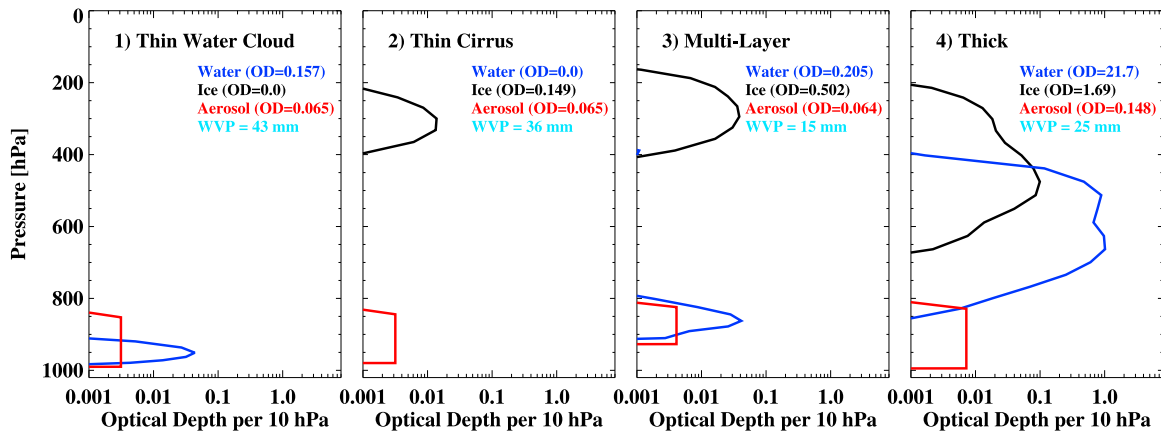

Figure 2. Vertical distribution of clouds and aerosol for the four ECMWF cases: case 1, thin water cloud; case 2, thin cirrus; case 3, multilayer; and case 4, thick. The cumulative optical depths at 0.755 μm for cloud water, cloud ice, and aerosol are shown in the legends.
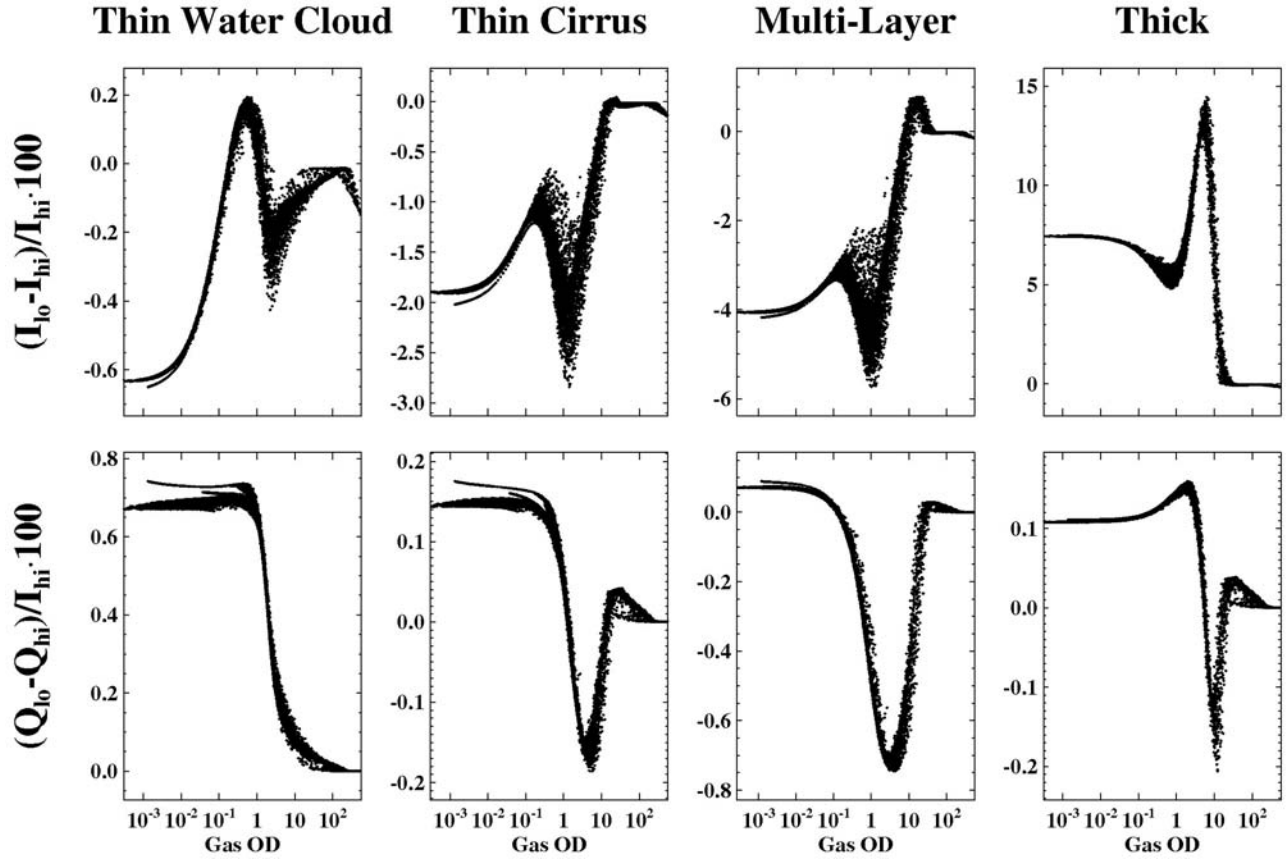
Figure 3. Monochromatic RT errors in Stokes parameters I and Q for the sample ECMWF cases in the oxygen A band. Errors in I are expressed in percent, while errors in Q are expressed as a percent of the continuum value of I. See text for details.

of a single line, or a large number of other similar variations. Thus, the same τ_g can result from a wide variety of profiles of gas absorption, some peaking lower in the atmosphere and some peaking higher in the atmosphere. It is the interplay of the particular shape of the τ_g profile with the location of the scattering layers that causes the scatter evident in Figure 3. For a more full account of this process, see *Duan et al.* [2005].

[19] To remove this dependence, we follow D05 and define a new variable τ'_g, which is the gas optical depth from the top-of-atmosphere down to the layer in the atmosphere where the scattering optical depth equals some critical value. For profiles with a scattering optical depth of less than two, the critical value is simply half the total scattering optical depth; for all other profiles it is unity. Note that this is the same as the variable k' of D05, except that the latter is the absorption optical depth due to all species (not just gases). Spectral points where most of the gas absorption occurs above the scattering layers will have $\tau'_g \approx \tau_g$; conversely, spectral points where most of the gas absorption occurs below the scattering layers will have τ'_g near zero.

[20] Again following D05, we consider the dependence of the multiple scattering error term on the ratio τ'_g/τ_g, termed ξ. Figure 4 shows the error in I for the thin cirrus case plotted as a function of $\xi^{1/2} \equiv \sqrt{\tau'_g/\tau_g}$, for three sets of narrow τ_g ranges as shown. Figure 4 top (bottom) plots show the results for the

oxygen A (SCO_2) band. Several important conclusions can be drawn from Figure 4. First, most points in most cases have an error that follows the quantity $\xi^{1/2}$ roughly linearly, except for the highest values of ξ where this often breaks down. This linear dependence on $\xi^{1/2}$ is thus far a purely empirical result. Second, only in the two middle plots does the error even depend on ξ; in the other plots there is no significant dependence. The physical reason for this is not completely clear. Third, as illustrated in Figure 4 (bottom), this relationship also holds in the strong CO_2 band, even though there can be absorption due to CO_2, water vapor, or both. Note that spectral points where greater than half of τ_g is due to water vapor have been colored blue in Figure 4. These points are seen to fall on the same curve as the "CO_2 points". Therefore, it is the vertical distribution of the gas absorption, not which particular gas is responsible, that is most important. Finally, these properties of the error versus $\xi^{1/2}$ generally hold for Stokes parameters Q and U as well (not shown).

2.4. LSI Formulation

[21] With these observations we can now fully formulate the LSI method, with the basic approach to be a two-dimensional interpolation of radiance errors in spectral bins of τ_g and $\xi^{1/2}$. The idea is to sample the error curves from Figures 3 and 4 by assigning each spectral point to a specific (τ_g, $\xi^{1/2}$) bin. Both high-accuracy and low-accuracy radiative transfer calculations will be performed on each binned
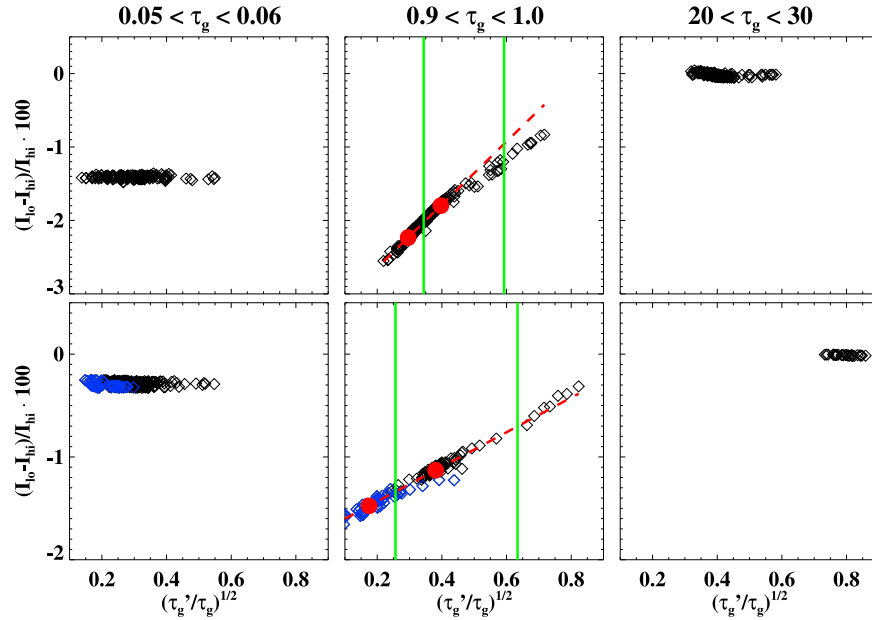
Figure 4. Monochromatic RT errors in stokes I for case 2 in the (top) oxygen A band and (bottom) SCO$_2$ band, plotted versus $\sqrt{\tau'_g/\tau_g}$, for SZA = 40° for nadir observations. Each plot corresponds to a set of points in a given range of τ_g as shown. Black diamonds represent spectral points dominated by oxygen (or carbon dioxide) absorption, while blue diamonds represent spectral points dominated by water vapor absorption. The green vertical lines in the middle plots represent 25% and 75% of the range of $\sqrt{\tau'_g/\tau_g}$ values in the plot. Red filled circles denote spectrally binned values used in the LSI calculations.

profile, in addition to low-accuracy RT calculations for each monochromatic spectral point. Because the low-accuracy calculations are extremely fast, and there are so few high-accuracy calculations, the total time for all RT calculations will be significantly faster than for the standard line-by-line approach. We begin with a description of how to construct the binned profiles of optical properties, and follow with a description of how to correct the monochromatic radiance errors using the high- and low-accuracy radiances from the binned profiles.

2.4.1. Constructing the Binned Profiles

[22] Bins spaced logarithmically (or nearly so) in τ_g provide a suitable binning for that dimension in order to effectively sample the error curves of Figure 3, though some points (particularly at low and high values of τ_g) prove to be nonessential. By examining error curves for hundreds of different types of profiles spanning a wide range of scattering properties, viewing geometry, and surface properties, we have arrived at sets of 18, 8, and 14 τ_g bins for the three OCO bands. The choices for the particular τ_g boundaries for the OCO bands are given in Table 3. It is important to note that the specific values of the τ_g bounds are unimportant; what matters most is that they are more closely spaced where the error curves vary fastest, and more sparsely spaced where the curves vary more slowly. Typically the curves have the most variation from τ_g of 0.1–10, though this varies slightly for the particular band and profile in question.

[23] For the other dimension, a simple two bin approach is found to be sufficient to correct most of the errors due to variation with $\xi^{1/2}$ as was illustrated in Figure 4. Furthermore, many of the τ_g bins at the low and high ends of the range need not be evaluated with two $\xi^{1/2}$ bins because they show virtually no error variation with $\xi^{1/2}$. $\xi^{1/2}$ subbins are used for τ_g bins 3 through 12 in the oxygen A band, bins 2 through 6 in the WCO$_2$ band, and bins 3 through 11 in the SCO$_2$ band. For a given τ_g bin to be subdivided into two $\xi^{1/2}$ subbins, the following procedure is used. Spectral points in the lowest 25% of the $\xi^{1/2}$ range for the τ_g bin in question are assigned to the lower $\xi^{1/2}$ subbin, and spectral points from 25% to 75% for the $\xi^{1/2}$ range are assigned to the upper $\xi^{1/2}$ subbin. Points above the 75% threshold are ignored for the purposes of the binning. This process leads to 28, 13, and 23 (τ_g, $\xi^{1/2}$) bins for the three OCO bands. The process of assigning spectral points to a particular $\xi^{1/2}$ subbin is illustrated in the middle plots of Figure 4. The green lines denote values of $\xi^{1/2}$ at 25% and 75% of the range of $\xi^{1/2}$ for the τ_g bin shown. The red filled circles represent the binned

Table 3. The τ_g Bin Boundaries for the Three Bands Considered

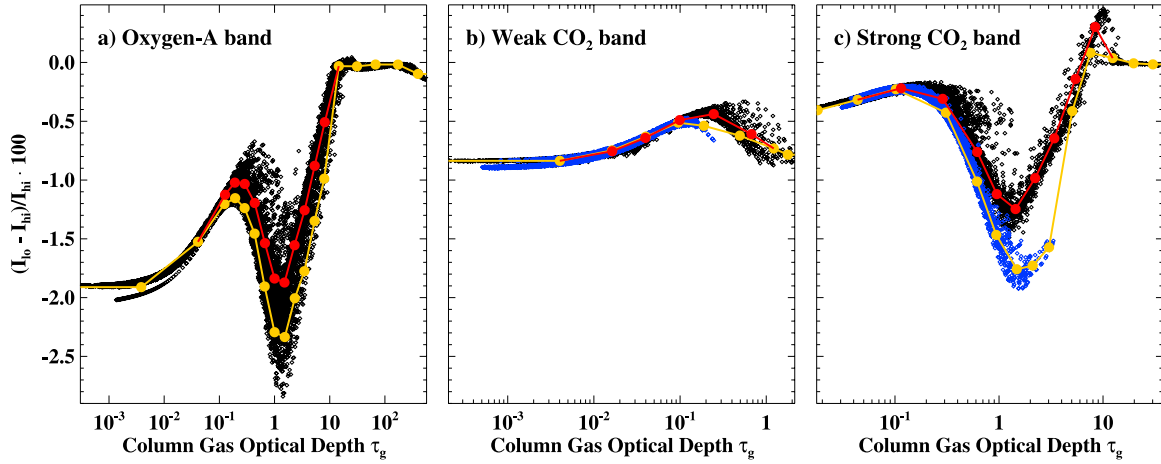| Band	Number of τ_g Bins	τ_g Bin Boundaries
Oxygen A	18	0, 0.005, 0.06, 0.15, 0.23, 0.35, 0.53, 0.8, 1.2, 1.9, 2.9, 4.3, 6.7, 10, 22, 46, 100, 300, 1000
WCO$_2$	8	0, 0.01, 0.025, 0.063, 0.16, 0.40, 1.0, 1.5, 1000
SCO$_2$	14	0, 0.02, 0.06, 0.17, 0.5, 0.77, 1.2, 1.8, 2.8, 4.3, 6.5, 10, 15, 25, 1000

Figure 5. Error plots for each OCO band, along with binned errors, plotted versus τ_g. Shown is case 2 for Stokes parameter I at a solar zenith angle of 40° for all three OCO bands. Small points represent the percent error in I for each spectral point; points with more than half their gas absorption due to water vapor have been colored in blue. Large filled circles represent the as-calculated values of the errors from each bin, in order to sample the error curve. Yellow (red) circles represent the binned calculations for the low (high) bins of $\xi^{1/2}$.

values (the specific averaging procedure is discussed below), and the dashed line represents a linear model of the error as a function of $\xi^{1/2}$. Clearly some points do not fall on this curve and thus estimates of their error values using this simple linear model will be slightly wrong, but the number of such points is generally small and it will be shown later that the resulting errors are also small.

[24] Once all the (τ_g, $\xi^{1/2}$) bin assignments have been made for each spectral point, the optical properties of each bin must be evaluated. This is done by simply averaging the gas absorption profiles together from all spectral points comprising the bin; all optical properties from scatterers, including extinction, single scattering albedo, and phase function are simply taken to be those at the band center. The surface properties are also taken from the band center. Therefore, the only difference between the various bins is in their profile of gas absorption. Effects due to varying values of the scattering properties across the band are considered later in section 2.5. The values of τ_g and $\xi^{1/2}$ for each binned profile are then directly evaluated according to definition.

[25] The errors for each spectral bin, calculated as described above, are illustrated in Figure 5 for the thin cirrus case. Figure 5 shows the relative errors in Stokes parameter I at a solar zenith angle of 40° in nadir observation mode for both the binned calculations as well as the individual spectral points. The curves for different profiles, Stokes parameters, and geometries behave in a qualitatively similar manner. Figure 5a shows the results for the oxygen A band, and can be seen to be identical to the corresponding plot in Figure 3, except now the binned results have been included. Clearly the binned results (large filled circles) are doing a good job of capturing the shape of the error curve as a function of τ_g. The yellow (red) circles denote the lower (upper) $\xi^{1/2}$ bins, representing points where the gas absorption occurs lower (higher) in the atmosphere. Figure 5a also illustrates that for very low and high values of τ_g a binning in $\xi^{1/2}$ is not performed (all points are simply assigned to the lower $\xi^{1/2}$ bin). Finally, note the bifurcation

of the errors on the left edge of Figure 5a, corresponding to continuum points. This is due to the different scattering and surface properties at the left and right sides of the band. The binned value ends up in between the two.

[26] Figures 5b and 5c show the same plot but for the other two OCO bands. Spectral points dominated by water vapor absorption are shown in blue, which have a different correlation of error with τ_g than do the "CO_2" points. The lower $\xi^{1/2}$ bins tend to track these points for the range of τ_g where they exist, implying that the correlation with $\xi^{1/2}$ naturally accounts for these effects as desired.

2.4.2. Radiance Correction Using Binned Results

[27] After the line-by-line and binned RT calculations have been performed, a two-dimensional bilinear interpolation is performed in (τ_g, $\xi^{1/2}$) for each Stokes parameter to determine the error, low accuracy relative to high accuracy, at each monochromatic spectral point; each spectral point is then corrected for this error accordingly. Mathematically we can express this procedure as follows. Let $S^{lo}(i, j)$ be the low-accuracy estimate of some Stokes parameter S (either I, Q, or U) for bin i in the τ_g dimension and bin j in the $\xi^{1/2}$ dimension, and similarly for $S^{hi}(i, j)$. We define the error in Stokes parameter S in the low-accuracy calculation relative to the high-accuracy calculation for an individual bin as

$$\varepsilon_S(i,j) = \frac{S^{lo}(i,j) - S^{hi}(i,j)}{I^{hi}(i,j)}. \tag{3}$$

[28] Once the $\varepsilon_S(i, j)$ for each bin is known, the corresponding quantity $\varepsilon_S(\lambda)$ for each monochromatic spectral point of wavelength λ can be calculated according to a bilinear interpolation in the τ_g and $\xi^{1/2}$ variables. However, because the space is unevenly sampled in τ_g and $\xi^{1/2}$ there is no single unique way to perform the interpolation. The details of the particular bilinear interpolation scheme adopted here are given in Appendix A. Finally, we compute the corrected monochromatic radiances from the
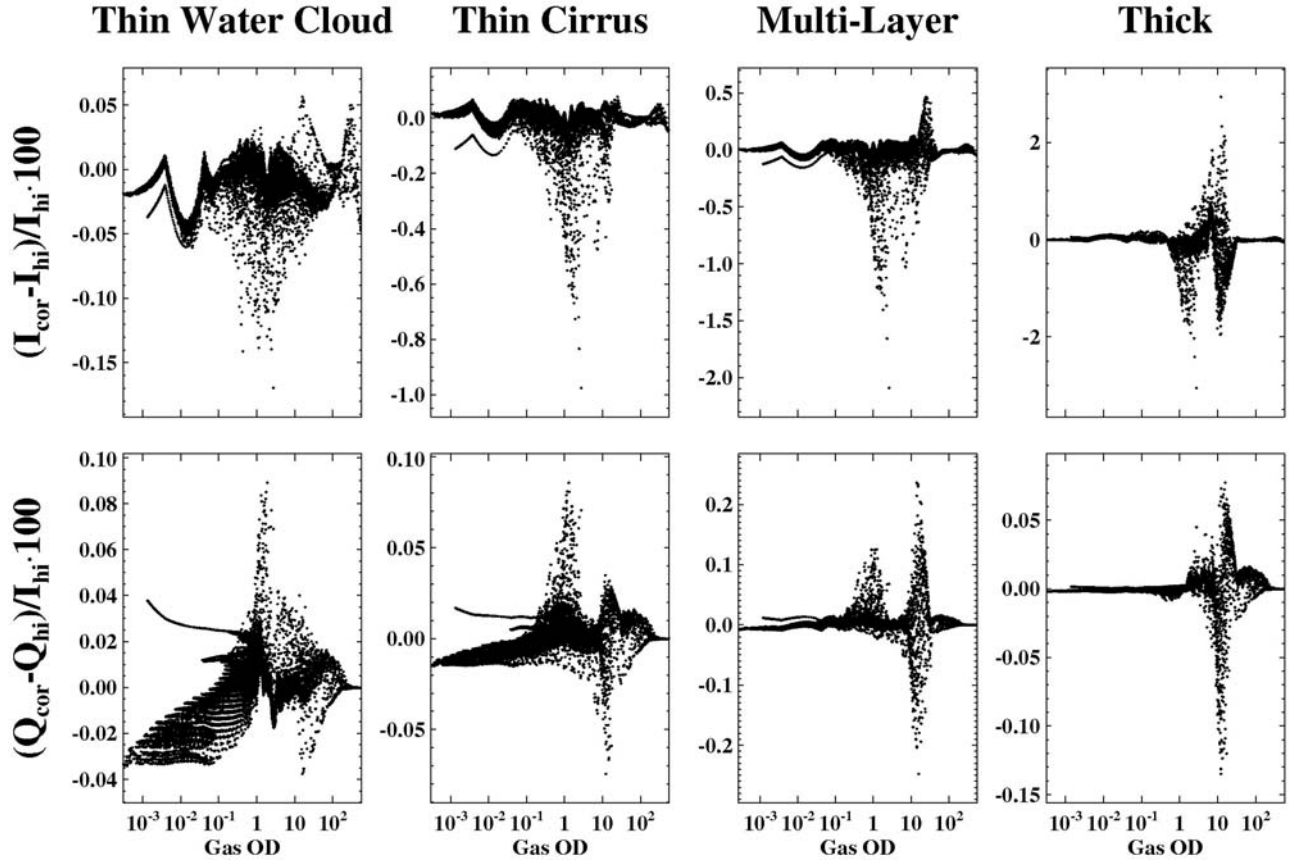
Figure 6. Same as Figure 3 but with errors corrected using the LSI technique. Note the dramatic difference in the vertical scales as compared with Figure 3.

low-accuracy monochromatic radiances using the correction terms $\varepsilon_S(\lambda)$ as follows:

$$I(\lambda) = \frac{I^{lo}(\lambda)}{1 + \varepsilon_I(\lambda)}, \tag{4a}$$

$$Q(\lambda) = Q^{lo}(\lambda) - \varepsilon_Q(\lambda)I(\lambda), \tag{4b}$$

$$U(\lambda) = U^{lo}(\lambda) - \varepsilon_U(\lambda)I(\lambda), \tag{4c}$$

where the quantity on the left-hand side of each equation denotes the corrected radiance. Note that the quantity $I(\lambda)$ used in equations (4b) and (4c) is the same quantity calculated by (4a). Figure 6 shows the results of this procedure, analogous to Figure 3 but now with the errors having been corrected using the LSI method. The LSI method removes the bias and much of the scatter, so the maximum errors are now tenths of percent for the thinner scattering cases, up to ~2% for the thicker cases. The effect of using a simple linear interpolation is apparent in some of the plots, but the error due to this effect is comparable to or smaller than the general magnitude of the other error sources. Finally, note that in this particular case the errors in I tend to dominate the errors in Q (or U); this is a feature of Lambertian surfaces and will not generally be the case for polarizing surfaces

such as the ocean surface. However, even over highly polarizing surfaces the LSI technique works effectively for all the Stokes parameters.

[29] The errors shown in Figure 6 are high spectral resolution errors on the original 0.01 cm^{-1} grid. Of greater importance are the errors in a spectrum as detected by a particular instrument. Consider again the case of OCO, which had full-width at half-maximum (FWHM) values in its instrument line shape (ILS) functions of 0.63, 0.27, and 0.21 cm^{-1} in its three bands. The measurement process is represented first by constructing the quantity I_{meas} from equation (1). Because both nadir and glint modes of OCO take place in the principle plane, equation (1) simplifies to

$$I_{meas} = \frac{I - Q}{2}. \tag{5}$$

When I_{meas} is convolved with the instrument ILS, the radiance errors are reduced to much lower levels, typically less than a few tenths of a percent as shown in Figure 7. The independent variable used is OCO pixel number, which is linearly related to wavelength, and each OCO band has been plotted in a different color. The thinner scattering profiles, cases 1 and 2, have errors on the order of a tenth of a percent, while the thicker cases (especially case 4) have higher errors. All errors are less than 0.5%, and the errors for each case are comparable for all three bands.
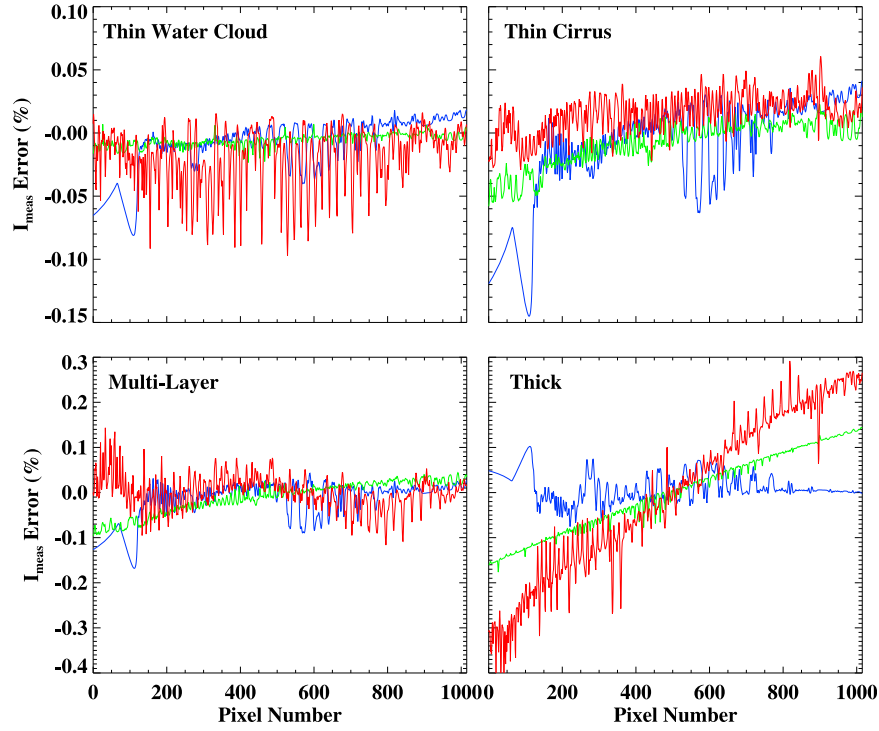
Figure 7. The percent errors in the instrument-measured spectra for the four ECMWF cases, at a solar zenith angle of 40° in nadir observation mode, for all three OCO bands. The oxygen A band is shown in blue, the WCO$_2$ band is in green, and the SCO$_2$ band is in red. Note the scale change from the top to bottom plots.

2.5. Correction for Spectral Variation in Scattering Properties

[30] A dominant feature of the errors from Figure 7 is the slope present in several of the curves; note in particular that the errors tend to cross zero in the middle of the band. This is a direct result of the fact that the binned radiances and their associated error terms have used the scattering optical and surface properties taken from the band center. Near the band edges this approximation may lead to significant errors, depending on how strongly the scattering properties change across the band in question.

[31] A simple first-order correction can reduce much of this error term. The technique is to create a single extra bin at the lowest value of τ_g with the atmospheric scattering and surface properties taken at one of the band edges, rather than at the band center as for all the other bins. Next, we make the assumption that the error due to this term behaves strictly linearly across the band. Knowing this error term at two points determines the line that passes through them.

[32] We take these two points to be the left-hand edge of the band and the center of the band, the latter of which has already been calculated for all bins. The error at the left edge of the band for the lowest τ_g bin is calculated as before. First, the profiles of optical properties for this bin are constructed. In this case, the gas absorption profile is the same as for the lowest τ_g bin, but the profile of scattering quantities (scattering optical depth, single scatter albedo, phase matrices, and surface properties) are taken from left edge of the band. Low- and high-accuracy radiances are calculated, and the error term is again calculated according to equation

(4). Then the error function $\varepsilon_S(\lambda)$ for arbitrary Stokes parameter S at any wavelength λ, calculated with equation (4), can be corrected according to

$$\widetilde{\varepsilon}_S(\lambda) = \varepsilon_{S,c}(\lambda) + \left(\varepsilon_{S,l}(1,1) - \varepsilon_{S,c}(1,1)\right)\frac{\lambda_c - \lambda}{\lambda_c - \lambda_l}, \qquad (6)$$

where $\widetilde{\varepsilon}_S(\lambda)$ represents the slope-corrected error, $\varepsilon_{S,c}(\lambda)$ is original error calculated from equation (4), and the subscripts l and c denote quantities evaluated with scattering properties taken from the left edge and center of the band, respectively. The (1, 1) notation simply means the first spectral bin in the τ_g and $\xi^{1/2}$ dimensions (note that for the lowest τ_g bin, there is always only a single bin in the $\xi^{1/2}$ dimension).

[33] Figure 8 shows the results of this process. For spectra where this effect is important, most notably the oxygen A band in case 2 and the CO$_2$ bands in case 4, the results are substantially better. Note that this method does not completely remove the effect, as we have only calculated the magnitude of the effect for a single bin in the continuum, and have explicitly assumed that the relative size of the effect is the same for all other spectral bins. A more accurate approach would be to repeat all binned calculations at two points across the band, instead of just for the lowest τ_g bin, but this would have the drawback of doubling the required number of spectral bins. The simple approach yields relatively low errors at the minor cost of a single additional binned calculation, and as such is used in all results presented below. With this modification, the total number of spectral bins in the LSI technique becomes 29, 14, and 24
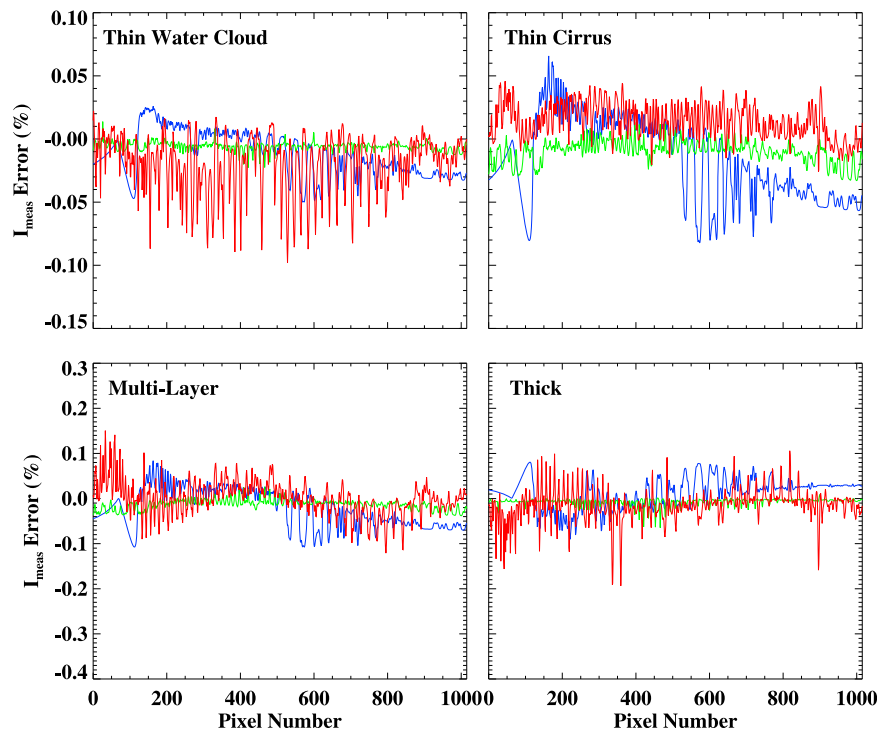
Figure 8. Same as Figure 7 but with the simple slope correction applied.

for the oxygen A, WCO_2, and SCO_2 bands, respectively. We note that *Boesche et al.* [2009] make a similar correction for their oxygen A work, but use 2–5 additional bins instead of just one as done here. *Hasekamp and Butz* [2008] make a correction for this effect using the derivatives of the radiances with respect to the scattering properties, that seems to work equally well.

3. Results

3.1. Accuracy of the LSI Technique: Study Cases

[34] Some basic error statistics for the LSI method as acting on the four sample ECMWF cases at several different solar zenith angles are shown in Table 4. As a reminder, the glint cases view the solar glint spot on an ocean surface described by a fully polarized model [*Cox and Munk*, 1954], while the nadir cases view a Lambertian surface with an albedo of ~0.1. The RMS errors are typically on the order of 0.05% or less, though they tend to be worse for glint than for nadir observations, and also are higher for higher SZA. All four ECMWF cases exhibit surprisingly similar errors, though the low water cloud case (case 1) has the lowest relative errors and the multilayer case (case 3) has the highest relative errors. Errors in the WCO_2 band tend to be substantially lower than in the other two bands. The worst RMS error is 0.42% which is the multilayer case in the oxygen A band for glint observations with a solar zenith angle of 70° Maximum errors are typically three times the RMS errors.

[35] In this technique as in most others, errors tend to increase at higher values of SZA. This is illustrated graphically in Figure 9, which shows the errors in the oxygen A and SCO_2 bands for case 2 for nadir observations as a function of SZA. The solid lines depict the RMS percent errors for the standard LSI technique wherein the low-accuracy calculations use two streams for the scalar, multiscattering calculation. The error is seen to rise steeply when SZA $\gtrsim 40°$, so much so that in the SCO_2 band, the error at 75° is close to 10 times larger than from 0 to 40°. For glint observations (not shown), the errors can be an even stronger function of SZA. The prime driver of this error term seems to be the number of streams in the scalar calculation. This is likely due to the fact that the diffuse radiance from a cloud or aerosol layer is more directional for higher SZA, and thus requires more streams to accurately resolve. The dashed curves show the corresponding results using four streams instead of two; the errors are seen to be somewhat smaller. This provides a potential mechanism to increase the accuracy of the result in such situations, at the expense of slower computations.

3.2. Accuracy of the LSI Technique: Full-Orbit Ensemble

[36] In order to further validate the accuracy of the LSI technique, we have also tested it on a large ensemble of profiles. The profiles were generated using the OCO Orbit Simulator [*O'Brien et al.*, 2009], which simulates a single dayside orbit of an OCO-like instrument in the 1:30 PM Sun synchronous Earth Observing System Afternoon Constellation (EOS A-train). For complete details of this tool, see *O'Brien et al.* [2009]. The orbit was for 8 September 2006; the instrument was assumed to be in nadir observation mode and only overland cases were analyzed. Profiles of cloud water and ice density were drawn from actual CloudSat observations [*Austin et al.*, 2009]. ECMWF model forecasts were used for temperature and water vapor profiles, inter-

Table 4. RMS Percent Errors in OCO-Like Radiances Due To LSI Method

View Mode	Case	Oxygen A Band			WCO$_2$ Band			SCO$_2$ Band		
		10°	40°	70°	10°	40°	70°	10°	40°	70°
Nadir	1	0.012	0.022	0.080	0.003	0.007	0.032	0.010	0.028	0.118
Nadir	2	0.020	0.037	0.078	0.006	0.013	0.034	0.015	0.020	0.121
Nadir	3	0.024	0.047	0.068	0.008	0.017	0.016	0.054	0.037	0.090
Nadir	4	0.046	0.035	0.028	0.007	0.010	0.014	0.025	0.037	0.040
Glint	1	0.008	0.022	0.050	0.001	0.008	0.029	0.006	0.042	0.139
Glint	2	0.013	0.035	0.188	0.003	0.016	0.053	0.015	0.020	0.306
Glint	3	0.022	0.079	0.420	0.007	0.037	0.137	0.036	0.056	0.108
Glint	4	0.051	0.041	0.029	0.006	0.015	0.012	0.033	0.056	0.032

polated to the CloudSat vertical grid. Concurrent Cloud-Aerosol Lidar and Infrared Pathfinder Satellite (CALIPSO) observations were used to find the vertical distribution of aerosol, when detected. Because CALIPSO cannot easily infer the type or size distribution of the underlying aerosol, the aerosol type was chosen to be a continental or oceanic type, based on which better matched the CALIPSO observations. Profiles of carbon dioxide concentration were taken from the Parameterized Chemistry Transport Model (PCTM) [*Kawa et al.*, 2004]. The land surface bidirectional reflectance distribution function (BRDF) was fully polarized and combines the MODIS MCD43B1 BRDF product (A. H. Strahler and J. P. Muller, MODIS BRDF/albedo product: Algorithm theoretical basis document version 5.0, 1999, http://modis.gsfc.nasa.gov/data/atbd/atbd_mod09.pdf) with a polarized BRDF model based on POLDER observations (F.-M. Bréon, personal communication, 2008); note that the resulting BRDF also varied linearly within an OCO band. The fractional polarization of the surface ranged from negligible to tens of percent. Due to the large number of CloudSat layers (typically greater than 100), the N_{sum} parameter was set at 5 instead of 3. Also, the number of streams for the high-accuracy calculation was 20 instead of 24. All other RT parameters were unchanged.

[37] A total of 709 profiles were processed from this single orbit. They had a variety of cloud and aerosol distributions, solar zenith angles and surface properties, and thus provide a fairly rigorous test of the LSI method. The results are summarized in Figure 10, which shows histograms of the RMS errors for each individual spectrum, separated by band. The median RMS errors are 0.022%, 0.007%, and 0.012%, while 95% of the RMS (maximum) errors are below 0.07% (0.18%), 0.03% (0.04%), and 0.09% (0.25%) for the three OCO bands. The few cases with errors higher than this generally have large amounts of scattering and/or occur at a very high SZA (>70°). These results confirm that the LSI technique has extremely promising error characteristics, even for a difficult array of profiles.

3.3. Computational Savings

[38] The computational savings of any spectral binning technique are difficult to quantify as they generally depend on the desired accuracy level, the number of vertical layers, and the details of the particular monochromatic RT code used. The following speed gains for the LSI technique are relative to the chosen reference model of section 2.1, with 24 streams and 60 vertical layers. All three OCO bands are calculated, and the times include some overhead (such as the calculation of optical properties and the convolution with

the ILS), but in general most of the time is spent in the RT calculations. For nadir observations, the average computational time was 670 seconds for the reference calculation versus 15 seconds using the LSI technique, a speed gain of about 45. For glint calculations, the corresponding times were 6300 seconds for the reference calculation versus 30 seconds for the LSI calculation, a factor of 210 faster. The speed-up factor for glint mode can be expected to apply to virtually all off-nadir calculations, which require an additional loop over azimuthal Fourier mode.

[39] The above computational savings must be placed in the context of the reference model adopted here. For instance, if the reference calculation had used 48 streams instead of 24, the speedup factors would be an order of magnitude larger or more. A more useful timing feature for the LSI technique is the time spent in the low-accuracy calculations versus the high-accuracy calculations. For nadir observations, roughly 60% of the time is spent on the low-accuracy calculations and 40% on the high-accuracy calculations. The low-accuracy calculations involve essentially two calculations, a vector first order of scattering calculation and a scalar two-stream multiple scattering calculation; the latter dominates the overall low-accuracy calculation time. Potential speed improvements could be realized by using a dedicated two-stream code, versus a standard multiple-scattering code such as the one used in this work [*Heidinger*
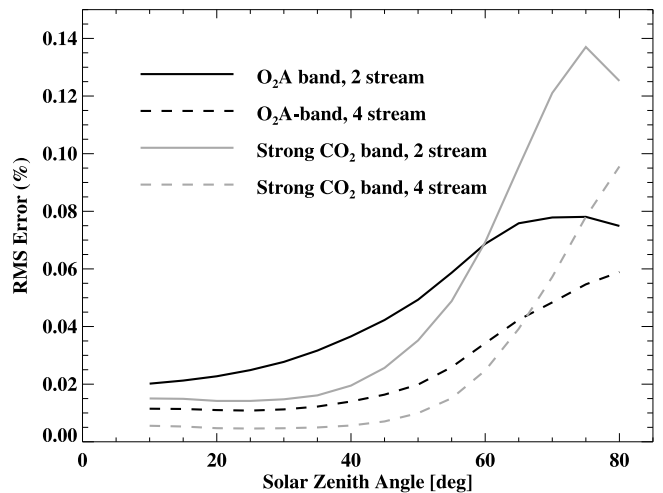

Figure 9. RMS errors in the oxygen A band (black) and SCO$_2$ band (gray) for case 2, nadir mode, as a function of SZA. The solid lines depict the nominal two-stream results, while the dashed lines show four-stream results.
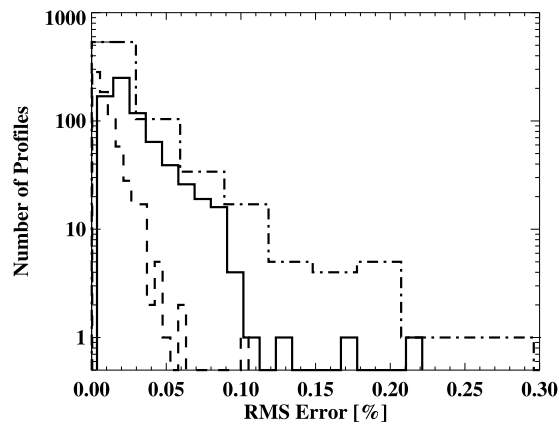
Figure 10. Histograms of RMS errors for each OCO band, over a total of 709 profiles spanning a wide range of cloud and aerosol distributions, solar zenith angle, and surface properties. Depicted are oxygen A band errors (solid line), WCO_2 band errors (dashed line), and SCO_2 band errors (dash-dotted line).

et al., 2006]. However, for off-nadir viewing it is the high-accuracy binned calculations that dominate the overall time. This may represent most simulations (especially for GOSAT, which observes the local nadir in only one out of every 5 observations), and thus speed gains here can only be realized either by reducing the number of spectral bins or reducing the complexity of the high-accuracy calculations, for example through a lower number of streams.

[40] It is illuminating to place the LSI technique in the context of the other acceleration techniques available in the literature mentioned previously. What parameters are important in defining the overall acceleration factor of a given technique? These can be broken down into roughly three categories: the number of high-accuracy RT calculations, the use of line-by-line single-scattering calculations, and the use of low-accuracy, line-by-line multiple-scattering intensity calculations (and the associated number of streams and layers). These parameters are shown in Table 5 for four selected techniques in addition to LSI. The number of spectral bins varies somewhat for each technique in the oxygen A band, but is on the order of 30 for four of the techniques; it is some 30 times higher for N05, though N05 explicitly states that this could probably be reduced with minimal loss in accuracy. Only LSI and the work of H08 explicitly consider the CO_2 bands, which contain gases (in particular water vapor) with different vertical profiles and hence can represent a challenge for spectral binning techniques. H08 requires 400 bins for the SCO_2 band, while LSI requires only 24 to achieve similar accuracy to the oxygen A band results. The former treats water vapor and CO_2 separately and uses a 20×20 grid in the τ_g of each gas, while LSI uses 24 bins spanning a useful range of τ_g and $\xi^{1/2}$ values. This illustrates the effectiveness of the $\xi^{1/2}$ variable to capture errors due to differences in the vertical gas absorption profiles in each τ_g bin. It should be noted that while the work of D05 did not explicitly consider the CO_2 bands, our results here indicate that their technique would likely work well for those bands, employing a similar number of bins as the LSI technique.

[41] The various methods considered in Table 5 also differ in which line-by-line calculations they require. Three of the techniques need an exact single-scattering calculation for each monochromatic wavelength, but as stated previously this is a fast computation and does not often substantially contribute to the overall calculation time. Only two of the techniques, LSI and that of N05 require a line-by-line, scalar multiple scattering calculation. Both use a fast two-stream calculation, and the LSI speeds this calculation up further by reducing the number of vertical layers via its N_{sum} parameter as described in section 2.1. As stated previously, for all off-nadir simulations the two-stream calculations do not greatly impact the overall calculation time, and for nadir calculations the two-stream time is comparable to that of the binned calculations. The conclusion is that it can be beneficial to use line-by-line single-scattering and two-stream, scalar calculations to enhance accuracy over the binned calculations alone, as there is typically not a great computational penalty associated with them. Also, this implies that the computational savings of a given technique relative to others can be assessed directly through the number of high-accuracy binned calculations alone.

4. Discussion

[42] In this work we have presented a fast, accurate method to accelerate the radiative transfer calculations required to simulate the spectra of gas absorption bands in the near-infrared, under conditions where scattering processes are important. Standard line-by-line methods are replaced by a series of several calculations: fast, low-accuracy line-by-line calculations that utilize both single-scattering and scalar, two-stream computations, as well as a small number of low- and high-accuracy calculations on spectrally binned profiles. The accuracy of this technique varies depending on the details of the atmospheric profile in question, and in particular depends on the vertical distribution of scattering, the vertical distribution of gas absorption relative to scattering, as well as the solar zenith angle. In general, the RMS accuracy is on the order of 0.05% in the ILS-convolved radiances for low to moderate solar zenith angles, and is slightly worse at high solar zenith angles; however, this is most likely a feature of any such approximation technique which relies on spectral binning. In the three OCO bands, RMS errors for the LSI technique are less than 0.1% for most profiles, and less than 0.5% for the most challenging profiles coupled with a high solar zenith angle. These error statistics were arrived at both through case studies as well as a large variety of profiles generated through a simulation of a hypothetical OCO orbit.

Table 5. Required Calculations of Selected Acceleration Methods

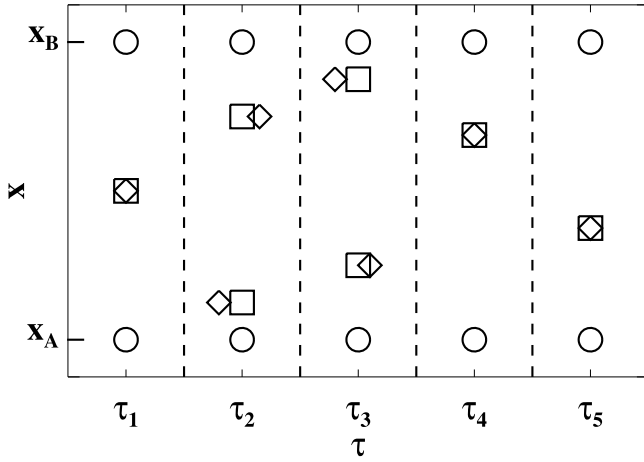| Method	Number of Bins Oxygen A (SCO_2)	Single Scattering (Line By Line)	Multiple Scattering (Line By Line)	Jacobians Required
D05	28 (?)	Yes	No	No
N05	945 (?)	Yes	Two-stream	No
H08	21 (400)	No	No	Yes
Boesche et al. [2009]	37 (?)	No	No	No
LSI	29 (24)	Yes	Two-stream	No

Figure A1. Illustration of error grid regularization necessary for the bilinear interpolation of the radiance error. Illustration is for a purely hypothetical band. The output of the RT code yields errors at irregularly spaced (τ, x) points, as depicted by the diamonds. Note that we have defined $\tau \equiv \tau_g$ and $x \equiv \xi^{1/2}$ for this discussion. The original τ bin boundaries are represented by the dashed vertical lines. In a first step, all τ values are transformed to the mean τ for a given τ bin (squares). In the second step, all x values are transformed to two constant values, x_A and x_B (circles).

[43] In the context of trace gas retrievals from sensors such as OCO or GOSAT, this level of error is somewhat less than the instrument noise and do not greatly impact the retrieval errors [*Hasekamp and Butz*, 2008]. Furthermore, because the algorithm is primarily based on a simple bilinear interpolation, linearization of the algorithm to form radiance derivatives is straightforward, allowing the approach to be used in a standard optimal estimation retrieval framework [*Rodgers*, 2000] or for data assimilation. After implementing the scheme with linearization we have found that although the Jacobians are not quite as accurate as the radiances, retrieval products are not noticeably affected by these slight errors.

[44] This work has been presented in the context of previously published acceleration techniques available in the literature. The LSI technique builds upon the strengths of some of these techniques, such as utilizing the enhanced accuracy of a fast, monochromatic two-stream calculation as done by N05, binning in optical properties similar to those used by D05, and including a lower number of vertical layers in the multiple-scattering calculation as proposed by *Min and Harrison* [2004]. It is appealing in that it can be used with most scalar or vector radiative transfer codes, and does not require Jacobians which can be a benefit in certain situations, though neither of these features are unique.

[45] Indeed, the most valuable and unique contribution of the LSI method is in its recognition that radiance errors due to a variety of simplifying approximations can be represented in a piecewise linear way as function of two easy-to-calculate variables. Further, it has been demonstrated to achieve high accuracy over an unprecedented variety of situations, including bands with overlapping gases, large numbers of vertical layers, profiles with large amounts of atmospheric scattering, and scenes with highly polarizing surfaces

described by nontrivial BRDFs. The LSI technique is found to be comparable in accuracy to the most accurate of these previous techniques, and comparable in computational efficiency to the most efficient such techniques. A Fortran-90 code implementing the LSI technique is available from the author upon request.

Appendix A

[46] The bilinear interpolation discussed in section 2.4 is nontrivial primarily because the error function ε_S ($S = I$, Q, or U) is unevenly sampled in its two independent variables τ_g and $\xi^{1/2}$. In the following discussion let $\tau_g \to \tau$, $\xi^{1/2} \to x$ and $\varepsilon_S \to \varepsilon$. For a given τ bin, there can be one or two x subbins. For the case of two subbins, each individual subbin will have its own value of τ and x. In order to simplify the bilinear interpolation, we seek to first put our sampled functional values on a more regular grid. Specifically, we want to end up with samples on $\tau = \tau_i$, $i = 1.M$ and $x = (x_A, x_B)$. For simplicity we take $x_A = 0$ and $x_B = 1$ (this bounds all physically allowed values of x). This is shown pictorially in Figure A1 for a hypothetical band with five τ bins, two of which have two x subbins.

[47] This grid regularization takes place in two steps. First, we construct the mean error as a function of only τ. For any τ bin i with two distinct x subbins ($j = 1, 2$), we take

$$\varepsilon(\tau_i) = \frac{N(i,1)\varepsilon(i,1) + N(i,2)\varepsilon(i,2)}{N(i,1) + N(i,2)}, \quad (A1)$$

where $N(i, j)$ is the number of spectral points that fall within bin (i, j).

[48] Consider those τ bins containing two x subbins ($j = 1, 2$). The values of τ in each subbin, $\tau_{i,1}$ and $\tau_{i,2}$, are close but not typically identical. Thus, we first wish to "center" these values on a common value of τ. To account for the variation of ε over this small range of τ variation while holding x constant, we use simple linear interpolation in $\ln\tau$ without extrapolation (outside the range of available τ points, we assume the function remains constant at the value of the last available point). After this "centering adjustment", we are left with the error function on a two-dimensional grid of (τ, x) pairs (squares in Figure A1) such that there is a single one-dimensional grid of τ points (τ_i, $i = 1.M$), and for each i there is one or two values of x.

[49] Recall that for fixed τ, we assume a purely linear dependence of the error on x. Thus, in the second step we find ε for each τ_i bin at the predefined points x_A and x_B (circles in Figures A1). For τ bins with no x dependence, we simply take

$$\varepsilon(\tau_i, x_A) = \varepsilon(\tau_i, x_B) = \varepsilon(\tau_i) \quad (A2)$$

for a τ bin i with two x subbins, having mean x values of $x_{i,1}$ and $x_{i,2}$, the desired function values are constructed using a linear fit in x,

$$\varepsilon(\tau_i, x_A) = \frac{\varepsilon(\tau_i, x_{i,1}) - \varepsilon(\tau_i, x_{i,2})}{x_{i,2} - x_{i,1}}(x_A - x_{i,1}) + \varepsilon(\tau_i, x_{i,1}), \quad (A3)$$

and similarly for $\varepsilon(\tau_i, x_B)$.

[50] We now have the error function ε evaluated on a rectangular grid of (τ, x) as depicted by the circles in Figure A1.

To evaluate this function at an arbitrary (τ, x) point, we use standard bilinear interpolation as follows. Find i' such that $\tau_{i'} \leq \tau < \tau_{i'+1}$. If $\tau < \tau_1$ we take $i' = 1$, and if $\tau \geq \tau_M$ we take $i' = M - 1$. Define variables f_τ and f_x as

$$f_\tau = \frac{\ln \tau - \ln \tau_{i'}}{\ln \tau_{i'+1} - \ln \tau_{i'}} \tag{A4}$$

$$f_x = \frac{x - x_A}{x_B - x_A}. \tag{A5}$$

We also enforce the condition that $0 < f_\tau < 1$; this ensures that beyond the τ boundaries, we assume that the function ε remains constant in τ. Next we define linear weights for each of the four bounding points,

$$
\begin{aligned}
w_{i',A} &= (1-f_\tau)(1-f_x), \\
w_{i',B} &= (1-f_\tau)f_x, \\
w_{i'+1,A} &= f_\tau(1-f_x), \\
w_{i'+1,B} &= f_\tau f_x,
\end{aligned}
\tag{A6}
$$

The $\varepsilon(\tau, x)$ at an arbitrary point is then simply given by

$$\varepsilon(\tau, x) = \sum_{i=i'}^{i'+1} \sum_{j=A}^{B} w_{i,j} \varepsilon(\tau_i, x_j). \tag{A7}$$

References

Austin, R. T., A. J. Heymsfield, and G. L. Stephens (2009), Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature, *J. Geophys. Res.*, *114*, D00A23, doi:10.1029/2008JD010049.

Baker, D. F., H. Bösch, S. C. Doney, and D. S. Schimel (2008), Carbon source/sink information provided by column CO_2 measurements from the Orbiting Carbon Observatory, *Atmos. Chem. Phys. Discuss.*, *8*, 20,051–20,112.

Baum, B. A., A. J. Heymsfield, P. Yang, and S. T. Bedka (2005a), Bulk scattering properties for the remote sensing of ice clouds. Part I: Microphysical data and models, *J. Appl. Meteorol.*, *44*, 1885–1895.

Baum, B. A., P. Yang, A. J. Heymsfield, S. Platnick, M. D. King, Y.-X. Hu, and S. T. Bedka (2005b), Bulk scattering properties for the remote sensing of ice clouds. Part II: Narrowband models, *J. Appl. Meteorol.*, *44*, 1896–1911.

Bodhaine, B. A., N. B. Wood, E. G. Dutton, and J. R. Slusser (1999), On Rayleigh optical depth calculations, *J. Atmos. Oceanic Technol.*, *16*, 1854–1861.

Boesche, E., P. Stammes, and R. Bennartz (2009), Aerosol influence on polarization and intensity in near-infrared O_2 and CO_2 absorption bands observed from space, *J. Quant. Spectrosc. Radiat. Transfer*, *110*, 223–239.

Bösch, H., et al. (2006), Space-based near-infrared CO_2 measurements: Testing the Orbiting Carbon Observatory retrieval algorithm and validation concept using SCIAMACHY observations over Park Falls, Wisconsin, *J. Geophys. Res.*, *111*, D23302, doi:10.1029/2006JD007080.

Bovensmann, H., J. P. Burrows, M. Buchwitz, J. Frerick, S. Noël, V. V. Rozanov, K. V. Chance, and A. P. H. Goede (1999), SCIAMACHY: Mission objectives and measurement modes, *J. Atmos. Sci.*, *56*, 127–150.

Butz, A., O. P. Hasekamp, C. Frankenburg, and I. Aben (2009), Retrievals of Atmospheric CO_2 from simulated space-borne measurements of back-scattered near-infrared sunlight: accounting for aerosol effects, *Appl. Opt.*, *48*, 3322–3336.

Chandrasekhar, S. (1960), *Radiative Transfer*, Dover, Mineola, N. Y.

Chevallier, F. (2001), Sampled databases of 60-level atmospheric profiles from the ECMWF analyses, *SAF Programme Res. Rep. 4*, EUMETSAT, Darmstadt, Germany.

Chevallier, F., F.-M. Bréon, and P. J. Rayner (2007), Contribution of the Orbiting Carbon Observatory to the estimation of CO_2 sources and sinks: Theoretical study in a variational data assimilation framework, *J. Geophys. Res.*, *112*, D09307, doi:10.1029/2006JD007375.

Cox, C., and W. H. Munk (1954), Statistics of the sea surface derived from Sun glitter, *J. Mar. Res.*, *13*, 198–227.

Crisp, D., et al. (2004), The Orbiting Carbon Observatory (OCO) mission, *Adv. Space Res.*, *34*, 700–709.

de Rooij, W. A., and C. C. A. H. van der Stap (1984), Expansion of Mie scattering matrices in generalized spherical functions, *Astron. Astrophys.*, *131*, 237–248.

Duan, M., Q. Min, and J. Li (2005), A fast radiative transfer model for simulating high-resolution absorption bands, *J. Geophys. Res.*, *110*, D15201, doi:10.1029/2004JD005590.

Fu, Q., and K. N. Liou (1992), On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres, *J. Atmos. Sci.*, *49*, 2139–2156.

Goody, R., R. West, L. Chen, and D. Crisp (1989), The correlated-k method for radiation calculations in nonhomogeneous atmospheres, *J. Quant. Spectrosc. Radiat. Transfer*, *43*, 191–199.

Hamazaki, T., Y. Kaneko, A. Kuze, and K. Kondo (2005), Fourier transform spectrometer for Greenhouse Gases Observing Satellite (GOSAT), *Proc. SPIE Int. Soc. Opt. Eng.*, *5659*, 73–80.

Hansen, J. E. (1971), Multiple scattering of polarized light in planetary atmospheres. Part I. The doubling method, *J. Atmos. Sci.*, *28*, 120–125.

Hasekamp, O. P., and A. Butz (2008), Efficient calculation of intensity and polarization spectra in vertically inhomogeneous scattering and absorbing atmospheres, *J. Geophys. Res.*, *113*, D20309, doi:10.1029/2008JD010379.

Heidinger, A. K., C. O'Dell, R. Bennartz, and T. Greenwald (2006), The successive-order-of-interaction radiative transfer model. Part I: Model development, *J. Appl. Meteorol. Climatol.*, *45*, 1388–1402.

Kawa, S. R., D. J. Erickson, S. Pawson, and Z. Zhu (2004), Global CO_2 transport simulations using meteorological data from the NASA data assimilation system, *J. Geophys. Res.*, *109*, D18312, doi:10.1029/2004JD004554.

Lacis, A. A., and V. Oinas (1991), A description of the correlated-k distribution method for modelling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres, *J. Geophys. Res.*, *96*, 9027–9064.

Liou, K. N. (2002), *An Introduction to Atmospheric Radiation*, Academic, San Diego, Calif.

Miller, C. E., et al. (2007), Precision requirements for space-based X_{CO_2} data, *J. Geophys. Res.*, *112*, D10314, doi:10.1029/2006JD007659.

Min, Q., and L. C. Harrison (2004), Retrieval of atmospheric optical depth profiles from downward-looking high-resolution O_2 A-band measurements: Optically thin conditions, *J. Atmos. Sci.*, *61*, 2469–2477.

Nakajima, T., and M. Tanaka (1988), Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation, *J. Quant. Spectrosc. Radiat. Transfer*, *40*, 51–69.

Natraj, V., and R. J. D. Spurr (2007), A fast linearized pseudo-spherical two orders of scattering model to account for polarization in vertically inhomogeneous scattering absorbing media, *J. Quant. Spectrosc. Radiat. Transfer*, *107*, 263–293.

Natraj, V., X. Jiang, R.-L. Shia, X. Huang, J. S. Margolis, and Y. L. Yung (2005), Application of principal component analysis to high spectral resolution radiative transfer: A case study of the O_2 A band, *J. Quant. Spectrosc. Radiat. Transfer*, *95*, 539–556.

O'Brien, D., I. Polonsky, C. W. O'Dell and A. Carheden (2009), Orbiting Carbon Observatory (OCO) algorithm theoretical basis document: The OCO simulator, technical report, 48 pp., Coop. Inst. for Res. in the Atmos., Colo. State Univ., Fort Collins.

Oshchepkov, S., A. Bril, and T. Yokota (2008), PPDF-based method to account for atmospheric light scattering in observations of carbon dioxide from space, *J. Geophys. Res.*, *113*, D23210, doi:10.1029/2008JD010061.

Rodgers, C. D. (2000), *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Sci., Hackensack, N. J.

Rothman, L. S., et al. (2009), The HITRAN 2008 molecular spectroscopic database, *J. Quant. Spectrosc. Radiat. Transfer*, *110*, 533–572.

Shettle, E. P., and R. W. Fenn (1979), Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties, *AFGL-TR-79-0214*, Air Force Geophys. Lab., Hanscom Air Force Base, Mass.

Siewert, C. E. (1982), On the phase matrix basic to the scattering of polarized light, *Astron. Astrophys.*, *109*, 195–200.

Spurr, R. (2002), Simultaneous derivation of intensities and weighting functions in a general pseudo-spherical discrete ordinate radiative transfer treatment, *J. Quant. Spectrosc. Radiat. Transfer*, *75*, 129–175.

Stamnes, K., S.-C. Tsay, K. Jayaweera, and W. Wiscombe (1988), Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, *Appl. Opt.*, *27*, 2502–2509.

Suto, H., K. Takahiro, J. Yoshida, J. Ishida, A. Kuze, M. Nakajima, and T. Hamazaki (2008), The pre-launch performance test and calibration results of Thermal And Near-infrared Sensor for carbon Observation (TANSO) on GOSAT, *Proc. SPIE Int. Soc. Opt. Eng.*, *6297*, doi:10.1117/12.677113.

van de Hulst, H. C. (1980), *Multiple Light Scattering. Tables, Formulas, and Applications*, vol. 1, Academic, San Diego, Calif.

West, R., D. Crisp, and L. Chen (1990), Mapping transformations for broadband atmospheric radiation calculations, *J. Quant. Spectrosc. Radiat. Transfer*, *43*, 191–199.

Wiscombe, W. (1977), The delta-M method: Rapid yet accurate radiative flux calculations for strongly asymmetric phase functions, *J. Atmos. Sci.*, *36*, 1408–1422.

C. W. O'Dell, Department of Atmospheric Science, Colorado State University, 1371 Campus Delivery, Fort Collins, CO 80523, USA. (odell@atmos.colostate.edu)