The Spherical Harmonics Discrete Ordinate Method for Three-Dimensional Atmospheric Radiative Transfer

K. FRANKLIN EVANS

Program in Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado

(Manuscript received 10 July 1996, in final form 30 June 1997)

ABSTRACT

A new algorithm for modeling radiative transfer in inhomogeneous three-dimensional media is described. The spherical harmonics discrete ordinate method uses a spherical harmonic angular representation to reduce memory use and time computing the source function. The radiative transfer equation is integrated along discrete ordinates through a spatial grid to model the streaming of radiation. An adaptive grid approach, which places additional points where they are most needed to improve accuracy, is implemented. The solution method is a type of successive order of scattering approach or Picard iteration. The model computes accurate radiances or fluxes in either the shortwave or longwave regions, even for highly peaked phase functions. Broadband radiative transfer is computed efficiently with a \( k \) distribution. The results of validation tests and examples illustrating the efficiency and accuracy of the algorithm are shown for simple geometries and realistic simulated clouds.

1. Introduction

Observations (e.g., Harrison et al. 1990) have confirmed that clouds have a large impact on the radiative energy flows in the atmosphere. Besides the problem of predicting the distribution of cloud properties, one difficulty in modeling radiative transfer in clouds has been the ubiquitous inhomogeneity of clouds. Almost no cloud fields on Earth are horizontally uniform, which is the assumption of plane-parallel models that are the mainstay of atmospheric radiative transfer. Modeling studies have shown that inhomogeneity effects are significant even in overcast clouds (Cahalan et al. 1994) and potentially large in broken cloud fields (e.g., Welch and Wielicki 1984; Barker and Davies 1992). Thus there is a need for 3D radiative transfer models to serve as numerical tools for understanding these inhomogeneity effects, parameterizing their effects in climate models, and correcting for their effects in remote sensing inversions.

The atmospheric science community has been developing 3D radiative transfer models for over two decades. The reason that this area of modeling is still the subject of ongoing research is because it is so computationally demanding. Solving the radiative transfer equation in a 3D medium is a 5D (three space, two angle) boundary value problem. There are still needs for computationally efficient, accurate, and flexible 3D models.

The first and most common type of radiative transfer model used to study 3D cloud effects has been the Monte Carlo method (e.g., Marchuk et al. 1980). This method, which can be thought of as simulating photon paths in the medium, has a reputation as being rather slow for results with good accuracy. The alternatives to Monte Carlo methods are those methods (such as the one described here) that explicitly represent the radiance field in the computational domain. This class of model is the standard for plane-parallel transfer. Gabriel et al. (1993) contains an overview of some of these types of 3D methods. Monte Carlo methods are generally more efficient than explicit methods when relatively few quantities need to be computed. For example, to compute the domain average reflected and transmitted solar flux, a Monte Carlo calculation with 100,000 photons will often suffice, which would be faster than an explicit radiative transfer method for a 3D medium. On the other hand, if many quantities, such as an image of upwelling radiances from the domain top or the 3D distribution of heating rate, are needed, then an explicit method can be much faster than Monte Carlo methods. For explicit methods, since the whole radiance field is computed, any desired radiative quantity may be computed at little extra cost. For Monte Carlo methods, either many more photons or successive calculations are required for more output quantities. Monte Carlo methods will always use much less memory. The exact cross-over point in computational speed between Monte Carlo and explicit methods depends on the particular problem being solved and the models used.

Corresponding author address: Frank Evans, Campus Box 311, University of Colorado, Boulder, CO. 80309.
E-mail: evans@nit.colorado.edu

© 1998 American Meteorological Society
There are many hopelessly inefficient ways to compute 3D radiative transfer. Focusing on the general, computationally efficient methods, we find that they all represent the spatial variation of the radiance field with a grid, rather than spectrally. There are methods with discrete ordinate representation of the angular aspects of the radiance field, such as the discrete ordinate method (DOM) or \( S_n \) method (e.g., Gerstl and Zardecki 1985; Sanchez et al. 1994) or others (Kuo et al. 1996). Other methods use a spectral approach for the angular part of the field, such as the spherical harmonic spatial grid (SHSG) method (Evans 1993). The method described here is an outgrowth of SHSG, combining aspects of discrete ordinate methods, using a new solution method, and implemented in 3D.

The spherical harmonics discrete ordinate method (SHDOM) uses both spherical harmonics and discrete ordinates to represent the radiance field during different parts of the solution algorithm. The spherical harmonics are used to compute the source function including the scattering integral. The discrete ordinates are used to integrate the radiative transfer equation spatially. One innovation is that the discrete grid, which represents the spatial variation of the fields, is adaptive, so that extra resolution is provided where needed. The solution method is simply iterating between the source function and radiance field, which amounts to a successive order of scattering approach [or a Picard iteration (Kuo et al. 1996)].

The model can perform 1D, 2D, or 3D unpolarized radiative transfer with general medium properties (extinction, phase function, etc., varying arbitrarily). The angular and spatial resolution may be specified, which allows the trade-off between calculation speed and accuracy to be explored. The modeled transfer may be monochromatic or broadband (with a distribution), and the source of radiation may be collimated solar and/or thermal emission. The lower boundary surface may have spatially varying albedo or temperature and a general bidirectional reflectance function can be specified. The horizontal boundary conditions may be either periodic or open. The model output consists of radiances, hemispheric flux, net flux, or net flux convergence at desired locations. Standard methods are used to compute accurate upwelling solar reflection with highly peaked phase functions. The model is freely available to the scientific community.

This paper gives an overview of the SHDOM algorithm (section 2), shows test results to validate the model’s correctness (section 3), and illustrates the performance and how best to use the model (section 4). The details of the algorithm are described in the appendix.

2. SHDOM algorithm

a. Representation of the radiation field

The primary novel aspect of SHDOM is to combine aspects of the spherical harmonic and discrete ordinate representations of the radiation field. As with the SHSG method (Evans 1993), the radiance field is represented on a discrete grid and the spherical harmonic representation is used to compute the scattering integral quickly. SHDOM, however, solves the integral form of the radiative transfer equation, computing radiance from the source function along discrete ordinates throughout the grid. This more physical simulation of radiation streaming along discrete ordinates was found to be more accurate and efficient than the purely spherical harmonic approach of SHSG, especially for nonscattering portions of the domain (e.g., clear air). SHDOM is generally superior in performance to SHSG, which only computes 2D transfer and is not implemented for general use by the scientific community.

If the discrete ordinate representation is best for representing the streaming of radiation, why use spherical harmonics at all? One reason is that the scattering integral in the radiative transfer equation may be more efficiently computed using spherical harmonics. As long as scattering depends only on the scattering angle, the scattering integral source function in spherical harmonic space reduces to simple multiplication:

\[
J_{\text{scat}}(\chi) = \frac{\omega_X}{2l + 1} I_{\text{in}}, \tag{1}
\]

where \( I_{\text{in}} \) are the radiance spherical harmonic coefficients, \( \omega_X \) are the Legendre phase function coefficients, and \( \omega \) is the single-scattering albedo. Hence the scattering integral may be computed very quickly (of order \( N \) operations). Of course, transforms between the spherical harmonic space discrete ordinate representations are required to take advantage of both. The partial separation of the azimuthal and zenith angle parts of the transforms results in the forward and backward transforms together taking approximately \( 9N^{3/2} \) operations (see appendix). This compares with \( 2N^3 \) operations for a purely discrete ordinate representation, where \( N \) is the number of discrete ordinates.

Rather than storing the radiance field, SHDOM stores the source function as a spherical harmonic series at each grid point. Solving for the source function is equivalent to solving for the radiance field, because radiance may be derived by integrating the source function in the radiative transfer equation. Another advantage of the spherical harmonic representation is that less storage is required as compared with a purely discrete ordinate approach. An adaptive spherical harmonic truncation is implemented so that the order of the series varies with grid point. This new approach can save considerable computer memory when parts of the domain have a source function that is zero (e.g., no scattering for solar problems) or smooth (thermal emission, scattering inside optically thick media), for which only a few terms are required.

The source function and radiance are defined on discrete grid points. This representation is far more efficient
than a spectral approach (see, e.g., Evans 1993), because the streaming of radiation is inherently local. In addition, an adaptive discrete grid is implemented. This approach, which is believed to be original for radiative transfer models, allows extra spatial resolution to be supplied only where needed. In the computational fluid dynamics literature (e.g., Melton et al. 1995) this approach is called adaptive mesh refinement with an unstructured Cartesian grid. A regular “base grid” is defined at the start, but as the iterative solution method proceeds, grid cells may be split in half to achieve higher spatial resolution. The criterion for dividing a grid cell is based on the change in source function across a cell. Extra cells may be split to make a “smoother” adaptive grid, so there are not large discontinuities in the grid. A tree data structure is used to keep track of the grid cells (volumes) and grid points so that neighboring cells can be located. The following sections illustrate the adaptive grid, and details are in the appendix.

The medium properties (extinction, single-scattering albedo, Legendre series of the phase function, and temperature) may be specified at every grid point of the “property grid.” For 3D media with complex phase functions, however, just the specification of the phase function may overwhelm available memory. Therefore a limited size table of phase functions may be input (e.g., as a function of effective radius) with each grid point having an index to specify which phase function.

b. Solution method

The solution method is a Picard iteration, or \( \Lambda \) iteration in the astrophysical literature (Stenholm et al. 1991). Each iteration consists of four steps in which

1) the spherical harmonics source function is transformed to discrete ordinates,
2) the source function is integrated to obtain the radiance field,
3) the radiance field is transformed to spherical harmonics, and
4) the source function is computed from the radiance field in spherical harmonics.

A flowchart illustrating the major components of the algorithm is shown in Fig. 1.

The second step is to use the integral form of the radiative transfer equation,

\[
I(s) = \exp \left[ - \int_0^s k(s') \ ds' \right] I(0) \\
+ \int_0^s \exp \left[ - \int_0^s k(t) \ dt \right] J(s')k(s') \ ds', \tag{2}
\]

to calculate the radiance \( I \) along discrete ordinates from the source function \( J \). The extinction \( k \) and the product of the extinction and source function \( (Jk) \) are assumed to vary linearly with distance \( s \) across a grid cell. The source function is integrated backward along a discrete ordinate from each grid point to a grid cell face that has known radiances at its bounding grid points; usually this is across just one grid cell. The radiance is interpolated between the surrounding grid points using bilinear interpolation to give the initial radiance for the integration. The extinction and source function product \( (Jk) \) are also bilinearly interpolated to the entrance point.

The integration in (2) is approximated for \( k \) and \( Jk \) linear in \( s \) by a formula that is accurate for small optical path and for constant extinction for all optical paths (see appendix). The accuracy of this spatial integration scheme is limited by using only quantities from one cell. To some extent the adaptive grid improves upon this (see section 4a). Kuo et al. (1996) uses an 11th-order spatial integration scheme involving many cells, but the high accuracy is not realized because the angular resolution limits the overall radiative transfer accuracy.
There are two aspects to how the spatial resolution affects the accuracy of the solution. The first is the source function integration across each grid cell. The adaptive grid cell generation is designed for this aspect. The second is the interpolation of the source function and initial radiance from the face grid points. In regions where the source function is large this aspect will also be taken care of by the adaptive grid, but where the source function is small (such as clear sky) only the resolution of the base grid is available to limit the radiance interpolation error. It should be noted that the error from the spatial integration is systematic, for example, tending to produce extra reflected flux. This is unlike a Monte Carlo method, for which averaging over pixels will reduce the error in domain-averaged quantities.

The SHDOM model allows either periodic or open horizontal boundary conditions. With the periodic boundary condition, discrete ordinate rays exiting one side of the domain wrap around to the opposite side. With the open boundary condition there is no reflection from the boundaries. In this case the incident radiation at the boundaries is determined by plane-parallel calculations for the boundary columns. The reflection from the lower boundary may be either Lambertian or a general bidirectional reflectance factor (BRF). The BRF model of Rahman et al. (1993) and Fresnel reflection are included, and other models are straightforward to implement. The radiance at the top boundary for downward discrete ordinates is set (usually to zero), and the radiance at the bottom boundary for upward discrete ordinates is related to the reflected downwelling radiance and any emission. The lower surface properties may be variable, in which case the reflection function parameters and temperature are specified on a regular grid. The surface properties are bilinearly interpolated to the adaptive grid points at the bottom of the medium.

The Picard iterations are a type of successive order solution, and as such are slow to converge for optically thick, conservative scattering media. To speed convergence a sequence acceleration method based on geometrical convergence is performed every other iteration. The accelerated source function is an extrapolation of the change

\[ J^{(n)} = J^{(n)} + a(J^{(n)} - J^{(n-1)}) \]  

where \( a \) is determined from the geometric convergence properties. The solution iterations are stopped when the solution criterion, which is the normalized rms difference between successive source function fields, is below a specified value.

The radiance and source function are initialized before the solution iterations with an Eddington radiative transfer solution on independent columns of the base grid. A spectrally broadband integration may be performed using a correlated \( k \)-distribution approach (e.g., Fu and Liou 1992). In this case there is a loop over successive monochromatic radiative transfer calculations in the \( k \) distribution (see Fig. 1), starting with the most absorbing \( k \). The SHDOM algorithm is very efficient at \( k \) distributions because strong gaseous absorption causes the iterations to converge in just a few iterations, and the previous \( k \) is used to initialize the next. It is assumed that the absorbing gases making up the \( k \) distribution are stratified and have no horizontal variability.

The desired radiometric quantities are summed with appropriate weights over the \( k \) distribution (if there is one). Hemispheric fluxes are computed from the discrete ordinate radiances during the solution process. Net fluxes (in \( x \), \( y \), and \( z \)) and the mean radiance are simply related to the lowest-order spherical harmonic radiance terms. The source function for a specified direction may be output. The radiance at specified angles and locations is computed by integrating the source function through the medium. For solar problems with the delta-\( M \) method, the TMS method of Nakajima and Tanaka (1988) is used to compute the source function. This method replaces the scaled, truncated Legendre phase function expansion for the singly scattered solar radiation by the full, unscaled phase function expansion. The multiply scattered contribution still comes from the truncated phase function. The TMS method is quite accurate for the backscattering directions encountered in satellite remote sensing, but not as accurate for directions in the solar aureole region.

An “independent pixel” (IP) mode of operation may be chosen. This mode calculates the radiative transfer solution on separate columns (1D), or on separate planes (2D) in either direction (\( XZ \) or \( YZ \)). For the 2D mode of operation in a three-dimensional medium the plane should be aligned with the solar azimuth, and this could be called the “independent scan” (IS) method. The independent pixel mode is implemented by simply changing the neighbor cell pointers, so that the discrete ordinate rays treat the columns (IP) or planes (IS) as separate periodic domains.

3. Validation testing

It is necessary to validate a new model to assure that it is relatively error free and does indeed solve the radiative transfer equation. Furthermore, it is important to characterize the accuracy and running time for a given angular and spatial resolution. The approach taken here is to show several validation tests that will also demonstrate the performance of the algorithm. Three tests of solar and thermal radiative transfer are shown: 1) fluxes in plane-parallel geometries, 2) radiances and fluxes from a 3D Gaussian field, and 3) fluxes from a 2D fractal field.

a. Independent pixel test

A good place to start validating a 3D model is with the well-understood monochromatic plane-parallel
Two solar transfer situations and one thermal transfer situation are run. Both have a Henyey–Greenstein phase function with asymmetry parameter of 0.85, and no surface reflection. For the solar cases the solar zenith angles are 15° and 60°, the scattering is conservative, the delta-M method is used, and there are $N_m = 16$ discrete zenith angles. For the thermal cases the wavelength is 10 μm, the single-scattering albedo is 0.1, the surface and medium bottom temperature is 300 K, the medium top temperature is 250 K, and $N_m = 32$. A plane-parallel doubling–adding model run with 32 streams computes the outgoing fluxes for comparison (Evans and Stephens 1991). Plane-parallel validation tests including Lambertian surface reflection have also been performed.

Figure 2 illustrates the albedo accuracy and the adaptive grid for the solar 1D transfer. The accuracy is relatively constant except for small jumps caused by the addition of adaptive grid points. As the cell splitting accuracy parameter is lowered by about a factor of 3, the rms albedo error is reduced by a factor of 3. The adaptive gridpoint depiction in the figure shows how the number of grid points necessary for a given accuracy increases with optical depth. The grid points are preferentially added to the top of the domain because that is where the source function is changing most rapidly from the incident solar beam. The grid spacing for optical depth of 20 is four times smaller at the medium top than near the center.

Table 1 lists the number of grid points and the rms error in upwelling and downwelling flux for several independent pixel cases. The high angular resolution implies that the accuracy is limited by the vertical resolution. For the high sun case ($\theta_s = 15^\circ$) the absolute flux accuracy is just about the same as the splitting accuracy parameter, but for the low sun case the flux accuracy is much better than the cell splitting parameter. For the thermal case the flux error is also much smaller than the splitting accuracy parameter. These examples illustrates that while the accuracy will tend to increase with reduction of the cell splitting parameter, there is not a simple relationship between it and the desired accuracy.
b. 3D Gaussian test

The next two series of tests rely on validation by a 3D backward Monte Carlo radiative transfer code. The code is based on the algorithm of O'Brien (1992) including the use of quasi-random numbers (the Halton sequence) and acceleration of the order of scattering series. The medium properties vary trilinearly between the grid points, which is computationally expensive, but matches how the medium is defined in SHDOM.

The second test series is 3D solar and thermal transfer in a Gaussian extinction field. The domain size is $2 \times 2 \times 1$ ($X-Y-Z$) and the extinction varies as $k = 4.63 \exp\left[-4(x-1)^2 - 4(y-1)^2 - 16(z-\frac{1}{2})^2\right]$ on a $20 \times 20 \times 11$ grid. The peak optical depth is 2, which is chosen to be small in order for the Monte Carlo code to give reasonably accurate results. The solar radiative transfer is for a wavelength of 1.65 $\mu$m and has a solar zenith angle of 45°. The thermal transfer test is at 10.7 $\mu$m and has a cloud temperature from 260 to 265 K above a 295–K surface. In both cases the optical properties are computed with Mie scattering theory for liquid cloud droplets with an effective radius of 10 $\mu$m and an effective variance of 0.1. The surface is nonreflective and there is no absorption other than the cloud droplets. For each output location the backward Monte Carlo code was run with $10^6$ photons and 26 scatterings for the solar case and $10^5$ photons and 10 scatterings for the thermal case. The adaptive grid is not used for SHDOM, because the optical depth across the the base grid cells is already quite small.

The zenith upwelling radiance and upwelling and downwelling flux at locations with a horizontal spacing of 0.2 are compared. The difference between Monte Carlo and SHDOM is expressed as an rms difference normalized by the mean of the quantity (e.g., flux) over the 60 comparison points. Figure 3 shows the upwelling flux and flux difference between SHDOM and Monte Carlo for the solar case. The peak albedo is only about 0.06 because of the low optical depth and forward-scattering phase function. The flux errors for this high angular resolution case are of the order of 1% and show a complex pattern. This is a result of flux at a point depending on the radiance from all upwelling angles and hence many locations, including the effects of the periodic boundaries. In contrast, the error pattern for radiance is very similar to the Gaussian optical depth image.

The Gaussian comparison results are also shown in Table 2, which lists the SHDOM accuracy for cases with various angular and spatial resolution. One test to validate the SHDOM algorithm is that as the angular and spatial resolutions increase, the error should generally decrease. The first part of the table shows how the error decreases with angular resolution. In order to have a few percent accuracy, moderate angular resolution ($N_\mu \geq 8$) is required. This is another illustration of how for 3D radiative transfer “two-stream” approaches do not produce adequate local fluxes. As the angular resolution is increased to ($N_\mu = 16$), the error for fluxes in the solar case does not keep decreasing. This is due to the need for more spatial resolution to further improve the accuracy. The lower part of the table shows how the error changes with the spatial resolution. For the $N_\mu = \ldots$
8 cases the error does not decrease much, because angular resolution is the limitation. For \( N_x = 16 \) there is a significant increase in accuracy for the higher spatial resolutions (except for the solar radiance). Based on the change in Monte Carlo results with the number of photons, even with \( 10^6 \) photons per location, the solar transfer comparison is probably affected by the Monte Carlo noise. The computer time (CPU) in minutes for the solar case is also listed. This may be compared with the 8455 min taken by the backward Monte Carlo code for the two fluxes and one radiance at the 60 locations. This is a rather unfair comparison for the Monte Carlo method, as this implementation is not optimized for speed (e.g., it performs path integrations through a trilinearly varying extinction in each grid).

c. 2D fractal test

The last validation test is for outgoing fluxes from a 2D fractal field described in Evans (1993). The same test configurations for the \( N_x = 64, N_y = 17 \) medium is used here, but SHSG, SHDOM, and backward Monte Carlo are compared. The medium has a mean optical depth of 2.1 with a logarithmic standard deviation of 1.0, and a Henyey–Greenstein phase function with \( g = 0.7 \). The delta-\( M \) scaling method is not used. The Monte Carlo results are rerun with \( 3 \times 10^8 \) photons and 26 orders of scattering for the solar case and \( 10^8 \) photons and 10 orders of scattering for the thermal case. The upwelling and downwelling flux are compared at 16 locations.

Table 3 lists the normalized rms flux difference from
Monte Carlo for SHDOM and SHSG. SHSG and SHDOM have equivalent angular resolution and the same spatial grid, though the adaptive grid is used for SHDOM. For SHDOM with no adaptive grid points, SHSG is substantially more accurate. This illustrates the advantage of SHSG’s second-order finite differencing. The adaptive grid is able to improve the accuracy of SHDOM substantially. Although each iteration of SHDOM takes more computation than SHSG, many fewer are needed, so for this test SHDOM is 2–8 times faster (see CPU times in Table 3).

4. Performance testing and examples

It is important to assess a new radiative transfer model’s computational efficiency and performance characteristics, both in order to demonstrate its usefulness and to understand the best way to operate the model. This section illustrates how the adaptive grid affects the accuracy of the results and shows examples of the capabilities of SHDOM.

a. Adaptive grid accuracy

Two tests show the accuracy convergence of SHDOM with the grid resolution (both base grid and adaptive grid) for 2D solar transfer, illustrating important points about the advantages and limitations of the adaptive grid. The first is a realistic cloudy medium obtained from cloud liquid water content output from a large eddy simulation (LES) of stratocumulus performed by Moeng (Moeng et al. 1996). For this test a 2D slice is selected from the 3D cloud LWC field, and the vertical grid is subsampled to produce a $64 \times 21$ grid with resolution $55 \text{ m}$ by $50 \text{ m}$. The cloud droplet size distribution is obtained by assuming a constant droplet concentration of $50 \text{ cm}^{-3}$ and a gamma distribution with an effective variance of 0.1. The optical properties are computed with Mie theory at a wavelength of 1.65 $\mu\text{m}$.

Figure 4a displays the extinction field for the LES test. The peak extinction is about 100 $\text{km}^{-1}$, and the highest optical depth is 11.8. The SHDOM property file has a tabulated phase function format, so phase functions are defined at discrete effective radii with 0.5-$\mu\text{m}$ spacing. There is no absorption or scattering by gases. The solar angle is $45^\circ$ (on the left), and the surface albedo is 0.06. Figure 4b shows the adaptive grid cell structure. The adaptive grid points are primarily at cloud top where the extinction is highest and where the direct solar beam first interacts with the cloud. This is where the source function is changing most rapidly.

The accuracy test is done by comparing SHDOM results at different resolutions to a very high resolution reference case. The reference case has $N_x = 32 \times N_z = 64$ discrete ordinates and a base grid of $N_x = 512 \times N_z = 161$ points with no additional adaptive grid points. Obtaining spatial resolution with only the base grid is contrasted with gaining resolution from only the adaptive grid. The upwelling radiance from the medium at seven angles ($\theta = 0^\circ, 15^\circ, 30^\circ, 45^\circ$ for $\phi = 0, 180^\circ$) at 55-m spacing is compared. In addition the flux and mean radiance at points on the input grid ($N_z = 64 \times N_x = 21$) are compared. Table 4 lists the rms accuracy and CPU time for two angular resolutions and four spatial resolutions. For the base grid cases the resolution is increased by halving one of the grid spacings ($X$ or $Z$) at each step, while for the adaptive grid case it is done by decreasing the splitting accuracy parameter a factor of 3.
The radiance error starts out at about 8% and reaches 1% for the highest resolution case. The CPU time (for an HP 715/75) is basically proportional to the product of the number of discrete ordinates and the number of grid points. It starts out quite modest at 0.3 min, but grows to be substantial. The characteristics of how the accuracy varies with the number of grid points is best seen in Fig. 4c, which shows the upwelling radiance accuracy. The radiance accuracy for the even grid saturates at 3% for \( N_x = 8 \) discrete ordinates, but drops to 1% for \( N_x = 16 \) ordinates as the spatial resolution increases. This illustrates the principle that the angular and spatial resolution must be advanced in tandem to optimally achieve a specified accuracy. The adaptive grid accuracy saturates above the even grid value (for \( N_x = 16 \)). Decreasing the cell splitting accuracy does not keep on improving the spatial resolution. The adaptive grid shows its benefits before the saturation occurs, with the radiance error dropping from 8.7% to 1.5% (for \( N_x = 16 \)) when the splitting accuracy is 0.03 (for solar flux of 1). At this sweet point the accuracy obtained for a certain number of grid points with the adaptive grid is substantially better than with the even grid. The number of adaptive grid points at this “optimal” point is problem dependent, but generally the number of additional points would not be much greater than the number of base grid points.

A similar convergence test is shown in Fig. 5 for a 2D extinction field derived from visible and thermal infrared channels in a Landsat cumulus scene. The optical depth is obtained from the visible channel and the extinction varies vertically according to adiabatic LWC. The cloud-top height is derived from the infrared channel and the cloud base from cloud shadows. The grid spacing is 100 m over the 5 km by 3 km domain. A single phase function for 10-\( \mu m \) effective radius droplets at a wavelength of 0.63 \( \mu m \) is used. The Lambertian surface albedo is 0.2 except the center 2 km, where it is 0.4. The solar zenith angle is 45° and the clear sky has no extinction. The adaptive grid method works extremely well in this case, reducing the radiance error from 10% to 1.6% with only 40% more adaptive grid cells than the base grid (\( N_x = 16 \) angles and splitting accuracy of 0.03). The equivalent accuracy using an even grid requires more than three times the number of grid points. The other radiometric quantities behave in a similar manner. The even grid accuracy must eventually surpass that of the adaptive grid because spatial resolution is required in the clear-sky regions for the correct streaming of radiation.

### Examples

To illustrate some of the capabilities of the SHDOM model, three examples are presented. The first is for 2D solar transfer in a uniform medium, but where the surface albedo has a discontinuity. The optical properties are for 0.5-\( \mu m \) sunlight in a haze represented by water droplets with an effective radius of 0.5 \( \mu m \). SHDOM operates on a 3D medium where the optical depth increases along one axis, so that in effect 21 separate 2D cases are run at one time. Figure 6 shows the nadir reflectance trace for four optical depths. The contrast of the surface albedo boundary decreases as the optical
FIG. 5. (a) The 2D Landsat cloud extinction. (b) The adaptive grid for cell splitting accuracy of 0.03. (c) The rms accuracy relative to the high resolution reference case for seven upwelling radiances at 100-m spacing.

FIG. 6. The 2D reflectance profile across a surface albedo discontinuity for four different optical depths of visible haze.

depth increases, but there is also a horizontal smoothing effect due to multiple scattering of the sunlight.

The second example demonstrates the ability of SHDOM to compute broadband radiative transfer quickly in multidimensional media. The LES cloud slice discussed in section 4a is used here but with the full vertical resolution. The correlated $k$-distribution method, specifically the parameterization of Fu and Liou (1992), is the basis of the broadband integration of the gaseous line absorption. This $k$ distribution (Fu and Liou 1992) has six bands from 0.2 to 4.0 μm (centers at 0.45, 1.0, 1.6, 2.2, 3.0, 3.75 μm). A total of 54 monochromatic transfer calculations (54 $k$’s) are required across the six bands. A composite sounding appropriate for marine stratocumulus is used to compute the volume absorption coefficients for the $k$-distribution input to SHDOM.

Along with the 25-m spaced cloud levels from 400 to 850 m, two levels below and nine above (to 30 km) are added for accurate solar absorption calculation. Since the desired output is fluxes at the 0.4- and 0.9-km levels, fine base grid resolution is required only between those levels, because there is little scattering outside with which the inhomogeneous radiation field can interact. The cloud droplet scattering is computed with Mie theory using the band center wavelength and solar-weighted band-averaged index of refraction. Molecular Rayleigh scattering (Fu and Liou 1992) and an ocean albedo of 0.06 are included. SHDOM is run with $N_m = 8$ and $N_f = 16$ discrete ordinates, which gives adequate accuracy for flux and heating rates. The cell splitting accuracy for the six bands is (10, 10, 5, 2, 2, 2), decreasing with the solar flux (the parameter is an absolute measure, e.g., W m$^{-2}$). The largest adaptive
grid has about 3300 points, compared with 1950 in the base grid.

Figure 7 shows the outgoing visible and near IR 2D broadband fluxes, as might be measured by aircraft radiometers along a flight track in the solar plane. The downwelling fluxes below the cloud have much more variability than the upwelling fluxes above the cloud. This is due to the highly directional nature of much of the downwelling radiation from the direct beam and low orders of scattering. The net flux convergence computed from the net fluxes above and below the cloud averages almost zero for the visible band, but has large negative excursions due to cloud side reflection. SHDOM can be used to explore schemes that use the visible net fluxes to correct the inferred infrared solar absorption for 3D effects (Ackerman and Cox 1981).

The number of iterations for the broadband calculation was 361, which is less than seven iterations per monochromatic calculation. The accurate 2D broadband solar calculation took 12.2 min (on an HP 715/75 computer). The successive order type iterations are very efficient in a $k$-distribution calculation. The highly absorbing $k$'s complete in one iteration because little solar flux reaches the cloud, moderate $k$ converge rapidly because of the low single scattering albedo, and low absorbing $k$'s benefit from the initialization by the previous field. SHDOM without initializing by the previous $k$ took 504 iterations and 14.4 min. The largest benefit of the previous $k$ initialization was for the visible band, for which the number of iterations decreased from 130 to 76.

The last example compares the use of periodic and open boundary conditions for broadband longwave radiative transfer. The 2D medium contains a cloud from 1.0- to 2.0-km altitude. The liquid water content is Gaussian in $X$ with full width half maximum of 1.0 km and linear in $Z$, with a peak of 0.2 g m$^{-2}$. The broadband transfer is performed using Fu and Liou (1992) longwave $k$ distribution of 12 bands (67 total $k$'s) for a standard midlatitude atmosphere. The cloud-top temperature is 280 K, cloud base is 285 K, and the black surface is 295 K. The cloud optical properties are determined from Mie calculations for each band assuming a droplet concentration of 200 cm$^{-3}$. The grid spacing for the cloud layer is 0.1 km, and there are five additional levels below and seven above (to 10 km).

The SHDOM output shown is the broadband flux at the surface, as might be observed by a pyranometer as these Gaussian cloud streets advect over. Figure 8 shows the downwelling longwave flux for three boundary condition cases. The flux is significantly higher when assuming a 3-km-wide periodic domain, as compared to a 12-km-wide domain (for which case open and periodic boundaries have the same fluxes). This is due to the flux contributions from the periodic array of clouds at 3 km spacing. For the 12-km domain the periodic clouds do not contribute significantly. The flux trace for the 3-km domain, open boundary condition case agrees very well with the 12-km domain result. For $N_{\mu} = 16 \times N_{\phi} = 32$ discrete ordinates and a cell splitting accuracy of 1.0, the broadband calculation took 16.8 min (203 iterations) for the 3-km open domain and 62.1 min for the 12-km periodic domain. This illustrates that the open boundary
condition may be used to substantially reduce the domain size and hence computation for isolated cloud situations.

5. Summary and conclusions

SHDOM is an algorithm and computer program used to model general three-dimensional atmospheric radiative transfer. Unpolarized monochromatic or broadband transfer may be computed with either or both solar and thermal radiation. Any combination of radiances, fluxes, or net flux convergence (heating rate) may be output.

The radiative transfer source function (from which any radiative quantity may be derived) is computed on a discrete spatial grid with the angular distribution represented in a spherical harmonic series. The spherical harmonic representation saves memory because with the adaptive series truncation only the required number of spherical harmonic coefficients are stored at each grid point. It is also faster to compute the source function using spherical harmonics than with discrete ordinates. On the other hand, the streaming of radiation is best modeled using the integral form of the radiative transfer equation along discrete ordinates.

The iterative solution method 1) transforms the source function to discrete ordinates, 2) integrates the source function along discrete ordinates to compute the radiance field, 3) transforms the discrete ordinate radiance to spherical harmonics, and 4) computes the source function from the radiance field in spherical harmonics. This is a successive order of scattering solution method, except that the source function is initialized with a 1D Eddington solution. A sequence acceleration method is implemented to speed convergence in optically thick, scattering media.

An adaptive grid is implemented to add grid points where they are most needed during the solution iterations. The criterion for splitting a cell in half is based on the difference in source function across the cell. The adaptive grid controls the error in the source function integration, but not the interpolation error of the cell entering radiance and source function. To deal with the interpolation error, adequate resolution is needed in the initial regular “base” grid. The angular resolution is determined by the number of discrete ordinates (which also determines the number of spherical harmonic terms).

Good radiance accuracy for solar problems with highly peaked phase functions is assured by using the delta-M scaling method along with the untruncated phase function single-scattering solution (Nakajima and Tanaka 1988). Broadband transfer is accomplished with a sum over a k distribution of gaseous absorption lines. This is very efficient in SHDOM because the highly absorbing k’s converge quickly. Independent pixel (1D) and independent scan (2D) modes are available for testing these approximations to 3D transfer.

Verification is shown for solar and thermal radiative transfer in three situations: 1) independent columns in a linearly increasing optical depth field, 2) a 3D Gaussian extinction field with a Mie phase function, and 3) a 2D fractal extinction field. Validation is provided with doubling–adding and backward Monte Carlo models. Tests of 2D solar radiative transfer in stratocumulus and broken cumulus cloud fields illustrate how the error decreases with increasing angular and spatial resolution. The spatial and angular resolution must be increased together to achieve desired accuracy with minimum computer resources. The spatial resolution is increased using the regular base grid or the adaptive grid accuracy parameter. Increasing the resolution with the adaptive grid alone decreases the error rapidly at first, but then much more slowly with the number of grid cells, while using the base grid alone decreases the error more steadily. This shows how the base grid and the adaptive grid must be used in conjunction for optimal results.

The capabilities of SHDOM, such as handling spatially varying surface properties and open and periodic boundary conditions, are illustrated. Broadband shortwave and longwave radiative transfer examples are shown. The broadband solar transfer is computed for a 2D stratocumulus cloud in a 64 × 30 gridpoint domain. A correlated k distribution with six bands and 54 k’s is used, which takes only 12 min to run on a workstation computer. This implies that SHDOM is the first explicit radiative transfer model efficient enough to perform broadband 3D atmospheric radiative transfer for significant sized domains. When many radiative quantities are desired from a 3D transfer simulation, SHDOM should be much faster than Monte Carlo models.

The SHDOM model, a k-distribution program, and a program to compute cloud optical properties, all written in FORTRAN, are being distributed. Example input files and UNIX scripts are included. The distribution and associated information are available on the Internet World Wide Web at URL http://nit.colorado.edu.

Potential applications of SHDOM include the following.

- Comparing with other multidimensional models to determine the situations for which each is most suited.
- Comparing 1D, 2D, and 3D approximations for inhomogeneous radiative transfer.
- Studying effects of cloud inhomogeneities on satellite remote sensing of cloud properties and inferred fluxes.
- Exploring heterogeneity effects on aircraft flux observations of clouds and developing corrections for cloud variability.
- Investigating whether 3D radiative transfer can cause real or apparent “anomalous” solar absorption in clouds.
- Determining effects of spatial variations in surface albedo on aerosol and cloud retrievals and radiative forcings.
- Improving the theory for how radiation flows in realistic cloud fields.
• Developing radiative parameterizations for climate models and cloud resolving models.

The SHDOM model should have the efficiency and flexibility to be a significant tool to investigate the above issues and many others in the field of atmospheric radiation.

Acknowledgments. The author thanks Chin-Ho Moeng for providing the LES model cloud liquid water output and Quiang Fu for the correlated k-distribution code. The efforts of Lin Chambers in testing versions of SHDOM are greatly appreciated. Warren Wiscombe and the editor Eric Smith are thanked for suggestions that greatly improved the readability of the paper. The idea for SHDOM grew out of a visit to Ted Smith and Jeff Haferman at the University of Iowa. Financial support for this research was provided by NSF Grant ATM-9421733.

APPENDIX

SHDOM Algorithm Details

a. Spherical harmonic transforms

The SHDOM algorithm computes the radiative transfer equation source function in spherical harmonics space \( J_{lm} \), which is related to the source function in azimuthal directions by

\[
J(\mu, \phi) = \sum_{lm} Y_{lm}(\mu, \phi) J_{lm},
\]

where \( Y_{lm} \) are orthonormal real-valued spherical harmonic functions. In the course of solving the radiative transfer equation, SHDOM transforms between the discrete ordinate and spherical harmonic representations. The set of discrete ordinates implemented are a reduced Gaussian grid. There are \( N_\mu \) Gaussian quadrature cosine zenith angles, \( \mu_j \), and \( N_\phi \) evenly spaced azimuth angles, \( \phi_k \). The discrete ordinate set is reduced by having fewer azimuth angles at larger \( |\mu| \) (near the poles), so \( N_\mu \) depends on \( J \). For 2D radiative transfer with symmetry around \( \phi = 0 \), the \( \phi \) range from 0 to \( \pi \), while in the more general case they range from 0 to 2\( \pi \) with a maximum number of \( \phi \) of \( N_\phi \). The reduced Gaussian grid of discrete ordinates has about 70% of the number of angles of a regular Gaussian grid \( (N_\mu \times N_\phi) \).

The spherical harmonic representation has the meridional index \( l \) between 0 and \( L \) and the Fourier azimuthal mode \( m \) from 0 to \( M \). If \( M = L \) then the spherical harmonic truncation is triangular, which has equal angular resolution in all directions. The spherical harmonic truncation is set by the number of discrete ordinates according to \( L = N_\mu - 1 \) and \( M = N_\phi/2 - 1 \), which assures that the spherical harmonics are orthogonal when their product is integrated. For 2D transfer only cosine azimuth modes are needed, while for 3D both cosine and sine modes are needed. The Fourier modes are implemented by having \( m < 0 \) refer to sine modes and \( m > 0 \) refer to cosine modes. For \( M = L \) the number of spherical harmonic terms is \( N_{lm} = (L + 1)^2 \) for both cosine and sine modes and \( N_{lm} = (L + 1)(L/2 + 1) \) for only cosine modes.

The source function \( (J_{lm}) \) is transformed to discrete ordinates by

\[
J_{\mu} = \sum_{m=-M}^{M} u(m \phi) \sum_{j=|m|}^{N_\mu} \Lambda_{jm}(\mu_j) J_{lm},
\]

where \( \Lambda_{jm}(\mu) \) are the normalized associated Legendre functions, and \( u(m \phi) = \cos(m \phi) \) for \( m \geq 0 \) and \( u(m \phi) = \sin(m \phi) \) for \( m < 0 \). For 2D situations with symmetry in azimuth only the non-negative \( m \) modes are used. The transform coefficients \( \Lambda_{jm}(\mu) \) are precomputed and stored. The functions \( \Lambda_{jm}(\mu) \) are computed as a set (all desired \( l \) and \( m \)) for a particular \( \mu \), using upward \( l \) recursion with normalization. The discrete ordinate radiance \( I_{\mu} \) is derived from the source function \( J_{\mu} \) by integrating the radiative transfer equation. The discrete ordinate radiance at each grid point is transformed to spherical harmonic space according to

\[
I_{lm} = \sum_{j=1}^{N_\mu} w_j \Lambda_{jm}(\mu_j) \sum_{k=1}^{N_\phi} w_{jk} u(m \phi_k) J_{lm},
\]

where \( w_j \) are the Gauss-Legendre quadrature weights and \( w_{jk} \) are the azimuthal integration weights normalized appropriately.

The explicit form of the transforms illustrates how the azimuthal and zenith angle parts partially separate, leading to a substantial decrease in the number of operations. For more than about 12 azimuthal angles an FFT is used for the azimuthal Fourier transform (the FFT does not have to be a power of 2). If there are \( N \) discrete ordinates then the number of floating point operations for both transforms together is approximately \( 9N^{3/2} \), and less when using the azimuthal FFT (asymptotically \( \sim 3N^{3/2} \)). The source function computation has only of order \( N \) operations.

The computation of the source function is very fast in spherical harmonic space because the scattering integral reduces to a simple weighting by the Legendre phase function coefficients \( \chi_l \) for each \( l \). The Legendre coefficients are defined by

\[
P(\cos \Theta) = \sum_{j=0}^{N_\mu} \chi_j \mathcal{P}_j(\cos \Theta),
\]

where \( \mathcal{P}_j \) are Legendre polynomials and \( P(\cos \Theta) \) is the phase function at the scattering angle \( \Theta \). At each grid point the source function is computed from the radiance according to

\[
J_{lm} = \frac{\omega \chi_l}{2l + 1} I_{lm} + S_{lm},
\]

where \( \omega \) is the single scattering albedo, and \( S_{lm} \) is the solar pseudosource and/or thermal source. The thermal
source is isotropic and so involves just the first spherical harmonic term:

\[ S_{\text{in}} = (1 - \omega)B(T)(4\pi)^{1/2}\delta
\]

where \( B(T) \) is the Planck function, either monochromatic or band integrated as appropriate. The solar pseudosource of diffuse radiation is

\[ S_{\text{in}} = \frac{F_0}{\mu_0} e^{-d} Y_0(\mu_0, \phi_0) \frac{\omega X_0}{2l + 1}, \]

where \( F_0 \) is the solar flux on a horizontal plane, \( \mu_0 \) is the cosine of the solar zenith angle, \( \phi_0 \) is the solar azimuth angle, and \( \tau \) is the optical path from the grid point to the upper domain boundary along the solar direction. Computing the source function in the truncated spherical harmonics space acts to smooth the discrete ordinates solution, thereby reducing ray effects, which are commonly a problem with the S-N method (Gerstl and Zardecki 1985; Chai et al. 1993).

The angular resolution in the model, which is governed by the spherical harmonic truncation level, is locally adaptive. The number of terms in the expansion of the source function is based on which terms are above a specified threshold. This flexible storage scheme is implemented by having a pointer to the memory location containing the beginning of each grid point’s spherical harmonic expansion. The radiance field also has an adaptive spherical harmonic truncation, but this is taken to be a few more terms than the source function truncation so that the angular resolution may increase during the solution iterations. The primary benefit of the adaptive truncation is a savings in memory storage.

**b. Adaptive grid**

The accuracy of the radiative transfer solution is governed in part by the resolution of the spatial grid at which the source function and radiance are defined. The accuracy of the source function integration depends on how well the source function and extinction variability are represented with the grid. If the optical path across a grid cell (in all directions) is small compared to one, then the integration across that cell will be accurate. This criterion, however, is too conservative for the center of optically thick media where the source function changes slowly. Thus for optically thick or highly variable media it is useful to have the spatial resolution vary across the medium.

The base grid does not need to be the same resolution as the input optical property grid, though it often is. The base grid is evenly spaced horizontally but may be arbitrarily spaced vertically. The input properties (extinction, single-scattering albedo, phase function, and temperature) are trilinearly interpolated from the input property grid to the base grid (actually the product of the extinction and single-scattering albedo, etc., are interpolated). If a \( k \) distribution is being used, then the vertical gaseous absorption profile is interpolated and the volume extinction for the particular \( k \) is added to the optical properties of the medium. If the delta-\( M \) scaling method (Wiscombe 1977) is specified, then the optical properties are scaled before their use. For solar problems the direct beam flux is computed for each grid point by tracing the path to the sun through the input property grid. The exact formula for the integral of a trilinearly interpolated extinction field is used for the optical path across each grid cell.

The adaptive grid evolves from the base grid by subdividing or “splitting” cells where more resolution is judged to be needed. The data structure for the adaptive grid is a tree structure with a root starting at each base grid cell. At each node a cell has pointers to its parent cell, perhaps two child cells, six neighboring cells, and to the eight grid points that the cell borders. The neighbor pointers are required in order to trace rays through the medium. There is an indication of whether the neighboring cell has been subdivided so that there are several bordering cells, or that there is no neighboring cell in that direction (a domain boundary). Only the end nodes of the tree (the highest resolution) are used in the radiative transfer computation. A fair amount of complexity is required to use the cell structure during ray tracing and to update the structure during cell splitting.

The criterion for splitting cells is based on how much the source function times extinction changes across a cell. A cell may be split in half in either of the three Cartesian directions, depending on whether any of them exceed the splitting criterion. The source function depends on angle, of course, but a single criterion is required, so it uses the average difference over all angles:

\[ |\Delta J| = \left[ \frac{1}{k} \int \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^{1} [k_s J^{(1)}(\mu, \phi) - k_e J^{(1)}(\mu, \phi)]^2 \, du \, d\phi \right]^{1/2}, \]

where \( J^{(1)} \) and \( k_s \) are the source functions and extinction, respectively, at two grid points bounding a grid cell. To determine how the source function difference might affect the radiance, the source function difference is multiplied by a function of the optical depth across a grid cell. Thus the cell-splitting criterion is

\[ C = |\Delta J| [1 - e^{-\tau}], \]

where the average optical path across the grid cell is \( \tau = \bar{k}d \), \( \bar{k} \) is the average extinction, and \( d \) is the distance between the two grid points in question. The criterion is averaged over the four edges of a cell (in 3D) that cross a potential cell splitting plane.

All end-node cells are tested for cell splitting in each direction \((X, Y, Z)\). The cells are sorted by the maximum (over \( X, Y, Z \)) of the cell-splitting criterion. Those cells
with the highest criterion above a certain value are split first, which assures that the most appropriate cells are divided when memory to create new cells is exceeded. After these cells are split, the adaptive grid is examined to find more cells to divide to make the grid “smoother.” A cell is divided if it has smaller cells on opposite sides or a neighboring cell has a grid spacing more than three times finer (Karman 1995). The new cells that were formed by splitting may themselves be divided during one solution iteration. Thus a cell that needs to be split in two directions (say X and Z) will not have to wait for the next iteration. As the solution iterations proceed, the cell splitting accuracy is gradually lowered to the desired final cell-splitting accuracy, and so more grid points are added at each iteration during this process. This procedure is computationally faster than achieving the full resolution in one iteration.

After each cell is divided, up to four new grid points bordering the cell are added (if not existing already). The medium properties are trilinearly interpolated from the property grid, using the same procedure as for the base grid. The spherical harmonic representation of the radiance is linearly interpolated between the two old points of the edge, and from this the source function for the new point is computed.

Tests have shown that the addition of adaptive grid points does not guarantee improved accuracy. The finite grid solution method apparently is sensitive to the abrupt variations in grid density that are implicit with an adaptive grid. This criterion for dividing cells has been found to be useful in many cases, but will not always be the optimal method of cell splitting.

c. Integration along discrete ordinates

The SHDOM algorithm solves for the source function \( J \) of the radiative transfer equation. The differential form of the radiative transfer equation for the radiance \( I(s) \) in the direction of a ray at a distance \( s \) along the ray is

\[
\frac{dI(s)}{ds} = -k(I - J),
\]

(A11)

where \( k \) is the volume extinction coefficient. The integral solution for the radiance, \( I(s) \), at distance \( s \), in terms of the initial radiance \( I(0) \) and the extinction and source function along the path is given in 2.

The radiative transfer equation is integrated for each of the discrete angles assuming that the source function field is fixed during the integration. The transform from spherical harmonics to discrete angles produces the source function \( J(x, \mu, \phi) \) at the gridpoint locations \( x \), for angles \( (\mu, \phi) \). The source function is integrated backward from each grid point to a grid cell face that has valid (known) radiances at its bounding grid points. The radiance is interpolated between the bounding grid points using bilinear interpolation to give the initial radiance for the integration. Usually the ray must be traced back across just one grid cell before getting a known radiance. To reduce the error from interpolating the radiance at the grid cell face, the ray may be traced backward until the transmission falls below some minimum specified value. The backward ray tracing, as contrasted with forward tracing, is necessary for implementation of the adaptive grid, because it guarantees that a radiance value will be computed for every grid point.

The radiance exiting a grid cell depends on the source function and extinction variation across the grid cell and the initial radiance on the other side of the cell. To find an efficient yet accurate approximation to the grid cell path integral, the variation of the extinction and source function within the grid cell are approximated from the values at the entering \( (k_e, J_e) \) and exiting \( (k_i, J_i) \) points. The extinction is assumed to vary linearly with distance across the grid cell from \( s' = 0 \) at the entering location to \( s' = s \) at the exiting location:

\[
k(s') = k_i + (k_i - k_e)\frac{s'}{s}.
\]

(A12)

With the assumption of linear variation of extinction, the optical path across the grid cell is

\[
\tau = \int_0^s k(s') \, ds' = \frac{1}{2}(k_i + k_e)s.
\]

(A13)

Because the product of the extinction and the source function appears in the integral form of the radiative transfer equation, this product, rather than the source function, is assumed to vary linearly across the grid cell, according to

\[
J(s')k(s') = J_0k_0 + (J_1k_1 - J_0k_0)\frac{s'}{s}.
\]

(A14)

The entering and exiting values of the extinction \( k \) and extinction/source function product \( Jk \) are computed using bilinear interpolation of the four gridpoint values of the faces pierced by the ray.

There is no simple and efficiently computed form of (A2) for the case where both the extinction and extinction/source function product vary linearly across the cell. If the extinction is constant and the source function is linear, then the exiting radiance is

\[
I(s) = e^{-\tau}I(0) + (1 - e^{-\tau}) \times \left[ J_0 + (J_0 - J_1)\left(\frac{1}{\tau} - \frac{e^{-\tau}}{1 - e^{-\tau}}\right)\right].
\]

(A15)

The solution for linear extinction and linear extinction/source function product can be expanded to first order in the path distance \( s \), resulting in

\[
I(s) = e^{-\tau}I(0) + (1 - e^{-\tau}) \times \left[ J_0k_0 + J_1k_1 + \frac{k_0J_1k_1 - k_1J_0k_0}{6(k_0 + k_1)} s\right].
\]

(A16)

This expansion solution is not accurate for large optical
path across a grid cell \((\tau)\), and in fact is unstable in the successive order iterations. Therefore, the expression for the path integral across a grid cell uses the expansion for linear extinction solution \((A16)\) for \(\tau \leq 2\), and for \(\tau > 2\) a combination of the expansion and the constant extinction solution \((A15)\):

\[
I(s) = e^{-\tau}I(0) + (1 - e^{-\tau}) \times \frac{1}{k_0 + k_1} \left[ I_{p0} + J_{p1} + \left( k_0 J_{p0} - k_1 J_{p1} \right) \right] \\
\quad \quad \quad \quad \times \frac{2}{k_0 + k_1} \left( 1 - 2 + \frac{2e^{-\tau}}{1 - e^{-\tau}} \right).
\]

(A17)

This has the correct small optical path solution and is correct for constant extinction or constant source function for all optical paths. For the integration formula to have the correct asymptotic behavior for \(\tau \to \infty\) when \(k_0 \neq k_1\), the cell entering extinction is adjusted toward the exiting extinction by \(k'_0 = k_0 + 4(k_0 - k_1)/\tau\). The cell integration formulas are somewhat complicated. A simpler integration scheme used in the DOM (Sanchez et al. 1994) was also tried; the 25% faster execution did not justify the considerably lower accuracy.

The order in which grid points are selected for the discrete ordinate radiance integration is chosen to minimize the number of cells that are traced back until a known radiance is found. Thus, for downwelling ordinates, grid points on the bottom of the top row of cells are done first. The particular ordering for this “sweeping” through the grid depends on the octant of the discrete ordinate. The order of the sweeping is precomputed for a given grid to speed the integration.

d. Boundary conditions

There is no reflection from the top of the domain, though isotropic incident radiance may be specified. The upwelling radiation from the bottom of the domain is the sum of reflected and emitted radiation. For a specular surface \((e.g., Fresnel)\), the integration cannot be used because there may not be an incident discrete ordinate corresponding to the upwelling direction. In this case interpolation (based on cubed inverse distance) between incident discrete ordinates is used. The solar direction must be along a discrete ordinate direction, so that the reflected direct beam is not lost. For a Lambertian surface, \(\rho\) is simply the albedo, though separate routines are used for improved efficiency. The parameters specifying the reflectance function at each boundary grid point are bilinearly interpolated from those in the input surface file. The general surface reflection routines are designed to simplify adding other reflection types.

The horizontal boundary conditions may be specified (independently for \(X\) and \(Y\) directions) as periodic or open. The periodic boundaries are implemented by having the boundary cells point to neighbors on the opposite side. As a ray is traced back from cell to cell, it will then automatically wrap around the domain. For simplicity in coding, the upper boundaries in \(X\) and \(Y\) have an extra plane of grid points that duplicates the lower boundary grid points. The open horizontal boundary conditions are implemented by letting the lower and upper boundary cells behave as “independent scans.” This is done by having the cells along the boundary point to themselves rather than their interior neighbors.

By starting the cell integration “sweeping” process with the boundary cells, the boundary planes become independent of the interior cells, and the radiance at their grid points serves as the boundary radiance for the interior points. For 2D \((X-Z)\) transfer this means that the upper and lower \(X\) boundary columns have plane-parallel radiance values.

e. Solution method

The solution method consists of iterations (section 2d) in which the spherical harmonics source function is transformed to discrete ordinates, which is is integrated to obtain the discrete ordinates radiance, which is then transformed back to spherical harmonics. The iterations are stopped when the solution criterion, which is the normalized rms difference between successive source function fields, is below a specified value. To save memory, phases 1–3 of the iterations are combined by doing all three phases inside the loop over the discrete ordinate zenith angles. Therefore, the discrete ordinate version of the source function/radiance must be stored only for \(N_a\) ordinates for each grid point.

Since the solution method is like an order of scattering approach, convergence is slower for optically thicker or more scattering media. The sequence acceleration method \((A3)\), which speeds convergence for these situations, is based on the observed geometric convergence of the solution. Because the source function difference changes its pattern each iteration, a 1D geometric acceleration approach does not work well. Instead a 2D approach is taken. Imagine three noncolinear points in a plane corresponding to the source function vectors \((\mathbf{J})\) for three successive iterations, so \(\mathbf{u}^{(n-1)} = \mathbf{J}^{(n-1)} - \mathbf{J}^{(n-2)}\) and \(\mathbf{u}^{(0)} = \mathbf{J}^{(0)} - \mathbf{J}^{(n-1)}\) are two adjacent vectors. The
The maximum memory needed for every grid point. The source function difference array (radiance, source function, and source function difference) to not decrease monotonically. situations, though it sometimes causes the solution criterion to not decrease monotonically.

If the acceleration method is implemented, then the primary memory requirements are for three large arrays (radiance, source function, and source function difference) containing the spherical harmonic expansions at every grid point. The source function difference array is not needed if the acceleration method is not used. The maximum memory needed for $N_p = 8$, $N_g = 16$ angular resolution is 266 $N_{ps}$, where $N_{ps}$ is the number of grid points.

f. Computing radiometric quantities

Hemispheric fluxes are computed at every grid point during the solution process. The discrete ordinate radiances are summed according to

$$F^z(x_i) = \sum_{j=1}^{N_{z}/2} w_j \sum_{k=1}^{N_g} \hat{\mathbf{w}}_k |\mu| I_{\phi k}, \quad (A21)$$

where the zenith angle sum ($j$) is over either the upward or downward hemisphere ordinates.

The rest of the output quantities are computed from the spherical harmonic representation of the radiance or source function. One output mode, useful for studying radiative energy flows, lists gridpoint values of the mean radiance and net flux vector, which are

$$\bar{I}_0 = \frac{F_0 e^{-r_i}}{4\pi \mu_0}, \quad (A24)$$

$$F_{x0} = \frac{F_0}{\mu_0} e^{-r_i} \sin \theta_0 \cos \phi_0 \quad F_{y0} = \frac{F_0}{\mu_0} e^{-r_i} \sin \theta_0 \sin \phi_0 \quad F_{z0} = -F_0 e^{-r_i}. \quad (A25)$$

The net flux convergence, which is related to the heating rate, is computed from the volume absorption coefficient $k(1 - \omega)$, the mean intensity, and the thermal and/or solar source, using

$$-\nabla \cdot \mathbf{F}_{\text{net}} = k(1 - \omega) \left[ 4\pi \bar{I}_0 + \frac{F_0}{\mu_0} e^{-r_i} - 4\pi B(T) \right]. \quad (A26)$$

The radiance at specified angles and locations is computed by integrating the source function through the medium. The spherical harmonic representation of the source function is transformed to the desired viewing direction ($\mu$, $\phi$). For solar problems with the delta-M method, the TMS method of Nakajima and Tanaka (1988) is used to compute the source function. This method replaces the scaled, truncated Legendre phase function expansion for the singly scattered solar radiation by the full, unscaled phase function expansion. The multiply scattered contribution still comes from the truncated phase function. The source function at the specified direction is then

$$J(\mu, \phi) = \sum_{lm} \left[ J_{lm} Y_{lm}(\mu, \phi) - \frac{F_0}{\mu_0} e^{-r_i} \right] \frac{\omega_i \chi_i}{2l+1} Y_{lm}(\mu_0, \phi_0) Y_{lm}(\mu, \phi) + \frac{F_0}{\mu_0} e^{-r_i} \sum_{\phi} \frac{\omega}{4\pi} \left\langle \beta_i \cos \Theta \right\rangle_1 \chi_{\beta_i}(\cos \Theta), \quad (A27)$$

where the primed quantities are delta-M scaled with truncation fraction $f$ and $\beta_i(\cos \Theta)$ is the $i$th Legendre polynomial evaluated at the scattering angle between ($\mu_0$, $\phi_0$) and ($\mu$, $\phi$). The last term of (A27) is summed over all the coefficients in the input Legendre expansion of the phase function. The TMS source function is integrated with the delta-M scaled extinction.

The radiance at the viewing direction is obtained using the integral form of the radiative transfer equation (2) with the above source function. The path through each grid cell is divided into segments with optical path less than 0.1, the extinction and extinction/source function product is trilinearly interpolated, and (16) used to compute the radiance contribution from each segment. Radiance values at many locations are calculated for a single viewing angle using one source function field.

When using a $k$ distribution, the output radiometric quantities are summed over the distribution using the weights associated with each $k$. 

ratio of the lengths of the two vectors is the convergence rate $r$, and the angle $\psi$ between them is a measure of how much the pattern in $\mathbf{u}$ is changing:

$$r = \frac{|\mathbf{u}^{(n)}|}{|\mathbf{u}^{(n-1)}|}, \quad \cos \psi = \frac{\mathbf{u}^{(n-1)} \cdot \mathbf{u}^{(n)}}{|\mathbf{u}^{(n-1)}||\mathbf{u}^{(n)}|}. \quad (A19)$$

Assuming the sequence converges geometrically, the maximum distance of the geometric spiral along the $\mathbf{u}^{(n-1)}$ vector is used for the acceleration distance:

$$a = \frac{1 - r \cos \psi + r^{1+\pi/2(\psi)}}{1 + r^2 - 2r \cos \psi} - 1. \quad (A20)$$

The acceleration extrapolation is only done every other iteration so there are valid $\mathbf{u}^{(n-1)}$ and $\mathbf{u}^{(n)}$. This solution sequence acceleration method substantially reduces the number of iterations required for conservative scattering situations, though it sometimes causes the solution criterion to not decrease monotonically.

The acceleration extrapolation is only done every other iteration so there are valid $\mathbf{u}^{(n-1)}$ and $\mathbf{u}^{(n)}$. This solution sequence acceleration method substantially reduces the number of iterations required for conservative scattering situations, though it sometimes causes the solution criterion to not decrease monotonically.
The complete source function, grid structure, and other items needed to compute radiometric quantities may be output to a save file. Later, the save file may be input to the SHDOM model to compute more radiometric quantities or to perform more iterations to reach a higher accuracy.

REFERENCES


