Satellite remote sensing methods for estimating clear Sky shortwave Top of atmosphere fluxes used for aerosol studies over the global oceans

Sundar A. Christopher *

Department of Atmospheric Science, The University of Alabama in Huntsville, AL 35899, USA
The University of Alabama in Huntsville, 320 Sparkman Drive, Huntsville, AL 35899, USA

A R T I C L E   I N F O

Article history:
Received 22 April 2011
Received in revised form 9 June 2011
Accepted 11 June 2011
Available online 5 August 2011

Keywords:
Aerosol
Climate forcing
CERES
MODIS
Solar radiation

A B S T R A C T

The difference between the top of atmosphere shortwave clear sky (cloud and aerosol free, SWCLR) and aerosol sky radiative fluxes is known as direct radiative effect (DRE) for all aerosols or Direct Climate Forcing (DCF) for anthropogenic aerosols. There are several methods for calculating SWCLR including satellite-based methods and radiative transfer approaches. Since uncertainties in SWCLR can propagate into errors in DRE or DCF, we assess the SWCLR estimates over the global oceans using three approaches and quantify the differences among these methods both as a function of space and season. Our results indicate that the more commonly used intercept ($73.4 \pm 3.6$) and radiative transfer methods ($74.7 \pm 4.0$ Wm$^{-2}$) are in close agreement to within $\pm 1.3$ Wm$^{-2}$. Values of SWCLR are provided as a function of space and season that can be used by other studies that require such values or as a source of validation. We further recommend that research studies report the methods and assumptions used to estimate SWCLR to facilitate easier intercomparisons among methods.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Tropospheric aerosols are usually defined as solid or liquid particles suspended in air that can be produced from both natural and anthropogenic sources. Examples of natural sources include sea salt aerosols over the global oceans and mineral dust from arid deserts while anthropogenic sources include smoke from agricultural burning and fossil fuel combustion. Although sub-groupings of anthropogenic aerosols do exist, they are in general characterized as sulfate (SU), black carbon (BC), and particulate organic matter (POM). Aerosol particles are usually categorized based on size including nucleation mode ($0.001-0.1$ μm diameter), accumulation mode ($0.1-1$ μm diameter), and coarse mode ($>1$ μm diameter). Much of the anthropogenic aerosols are found in the accumulation mode, whereas mechanically produced aerosols such as dust and sea salt are predominantly coarse mode aerosols. Considering that nearly 70% of the earth is ocean, it is not surprising that the total median sea salt aerosol source strength is the largest ($6000$ Tg year$^{-1}$) when compared to dust ($1600$ Tg year$^{-1}$) and BC, POM, and SU put together ($300$ Tg year$^{-1}$). However, in terms of median mass loading, dust has the highest mass ($20$ Tg) when compared to sea salt ($6$ Tg) with SU, BC, and POM totaling $4$ Tg (CCSP 2009).

Aerosols have a wide range of impacts from affecting visibility to absorbing/scattering sunlight (direct radiative effect), acting as cloud condensation nuclei, and modifying cloud hydrological processes (indirect radiative effect). Although the lifetimes of tropospheric aerosols are less than a week, the ubiquitous nature of these sources and source strengths are important for various applications including climate and air quality. Moreover, their strong regional impacts are based on their absorptive and scattering properties, as well as the myriads of chemical compositions making the study of aerosols both challenging and interesting. A comprehensive review of aerosols and their climate effects from both an observational and a modeling perspective is presented in the CCSP report and by Yu et al., (2006 and references therein), Both DRE and DCF quantify the change in shortwave radiation (0.2-4.5 μm) with and without aerosols at the top of atmosphere (TOA) (usually at 20 km ASL). The TOA is especially useful since global measurements of reflected solar radiation are available from a routine basis from satellites over several decades (Anderson et al., 2005). In general, aerosols tend to reduce the amount of solar radiation reaching the surface thereby cooling the surface. In contrast, the effect of the absorption and emission of anthropogenic CO$_2$ is to warm the earth's surface. Therefore, the competing cooling of the aerosol effect compared to the warming CO$_2$ effect has been a topic of much debate (Andreae et al., 2005).

The definition of CO$_2$ forcing is well established in the literature and represents the change in TOA radiative fluxes from current day CO$_2$ conditions to pre-industrial values. However, complications arise when discussing aerosol forcing since definitions among methods vary (Bellouin et al., 2008). Modeling studies define aerosol forcing (due to anthropogenic aerosols only) as the change in radiative fluxes at the TOA from current day aerosol concentrations to pre-industrial values.

* Tel: +1 256 961 7872; fax: +1 256 961 7755
E-mail address: sundar@nsstc.uah.edu.
Since it is difficult to pin down pre-industrial aerosol concentrations spatially and seasonally, satellite-based studies define aerosol effects as the change in radiative fluxes between current day aerosol conditions and conditions when aerosols are not present. Bellouin et al. (2008) examines this definition-problem and discuss the utility of using current day natural aerosol distribution as pre-industrial aerosols. They note the complexity of this issue since most satellite-based studies are only adept at assessing aerosols in cloud-free conditions and open questions such as accuracy of emission sources and processes in numerical simulations remain. Assessing aerosol impacts in the presence of clouds is still an unsolved issue, since neither satellite-based studies nor modeling simulations adequately characterize the relative positions of aerosols and clouds and their associated radiative properties. While differences between modeling and satellite-based studies have been reduced, they have yet to be fully and completely reconciled due to varying definitions of aerosol types (Bellouin et al., 2008) and aerosol absorption (Myhre et al., 2009).

Many uncertainties exist when calculating DRE using observational methods and include sensor calibration, sampling resolution, cloud clearing quality, accuracy of radiance to flux conversions, among many others (Christopher et al., 2006) which are not the focus of the paper. They are addressed in Yu et al. (2006). Another important uncertainty that requires further analysis is that pertaining to the cloud and aerosol free backgrounds (that we label clear sky flux) used to calculate DRE. Recall that DRE is defined as the difference between clear and aerosol sky fluxes. Thus, any uncertainty in clear sky fluxes will propagate through to the final DRE calculations although compensating effects are possible. To address this issue, we take a closer look at the shortwave clear sky fluxes (SWCLR) that are used to compute aerosol forcing in satellite-based studies. Satellite-based studies define DRE or DCF as SWCLR-SWAER where SWCLR is clear sky (no clouds or aerosols) TOA flux and SWAER is aerosol sky flux. Note that the effects of aerosols on thermal radiation are not considered in this paper and depending upon particle size and absorption characteristics, their effects could be significant.

Several methods have been used to create SWCLR from CERES-based observations. The first uses a regression technique whereby top of atmosphere clear sky fluxes are assumed to have a linear correlation to AOD (while also being a function of atmospheric conditions, satellite viewing geometry, and surface conditions). The intercept to the regression equation where AOD = 0 is then defined as SWCLR. In some cases, the regression method is not applied and the minimum flux value within a region of cloud-free data is assumed to be the clear sky flux (Christopher & Huang, 2002). However, this method breaks down in regions where cloud and/or aerosol cover are consistently high. Another method often used by the research community is the radiative transfer (RT) model method. Using this technique, clear sky fluxes are calculated from radiative transfer models using atmospheric conditions, surface albedo, and viewing geometry inputs for a pristine, clear sky conditions. At least one study (Kato et al., 2002) has examined the differences in these methods (regression and RT) using CERES and VIRS data from the TRMM satellite. For the region bounded by ±40° latitude, they found a total uncertainty of 1.2 Wm⁻² between these methods with the RT method producing albedos 3% to 4% higher than the regression method, resulting in higher clear sky SW fluxes.

However, the combination of MODIS and CERES on Terra and Aqua is better suited to assess cloud free fluxes. The VIRS had limited spectral channels and the TRMM only provided data between -40°N-40°S. Better angular models for surface and aerosols (Zhang et al., 2005) have been developed as well. Furthermore, the AOD retrievals from MODIS are considered a much newer generation of retrievals compared to the two channels methods employed for the VIRS. We examine three methods of obtaining SWCLR over the global oceans and quantify these values as a function of season and region to determine which methods are best suited for studies of DRE and/or DCF, while noting the advantages and disadvantages of each.

2. Methods

We use the CERES Single Satellite Footprint (CERES-SSF, FM3, Edition 2C) product from the Aqua satellite between May 2006 and August 2007 (Wielicki et al., 1996). This CERES contains point spread function weighted MODIS collection 5 aerosol optical depth within each CERES footprint. The nadir spatial resolution is 20 km for CERES and the MODIS aerosol product is 10 km at nadir. For aerosols, the CERES converts the TOA measured radiances to fluxes based on theoretical radiative transfer calculations and TRMM data (Loeb et al., 2005). To improve upon this approach, the CERES total upward shortwave radiances between 0.2 and 4.5 μm (SDS-35) are converted to SW fluxes using angular dependence models that were specifically derived for aerosols from Terra (Zhang et al., 2005). This product has been used extensively for studying aerosol forcing (e.g. Christopher et al., 2006) over the global oceans and therefore used in this study. Note that the anthropogenic fraction of the total AOD must be first obtained before calculating DCF estimates, (e.g. Kaufman et al., 2005). Over the global oceans, the cloud-free DCF is about 4 times smaller (−1.4 Wm⁻²) than the total DRE (−5.5 Wm⁻²) (Christopher et al., 2006). However, we are primarily concerned with the total aerosol DRE in this research and the additional uncertainties associated with separating the anthropogenic component of AOD are discussed in other literature (e.g. Christopher et al., 2006).

The CERES Fixed Swath Width (FSW) data product (Aqua, FM3, Edition 2C) at 1 × 1 degree resolution contains the monthly gridded FSW product that were from hourly single satellite swath fluxes. We use the upward SW flux for pristine conditions (no-clouds and no-aerosols) parameter (SDS-46) in the FSW product that was calculated from the instantaneous Clouds and Radiative Swath (CRS) using a 4-stream radiative transfer model (Fu & Liou, 1993; Rutan et al., 2006). The 4-stream model uses atmospheric profiles (temperature and water vapor) from Goddard Earth Observing System Data Assimilation System, version 4 GEOS-4, ozone profiles from National Center for Environmental prediction (NCEP), and spectral surface albedo based on wind speed, chlorophyll concentrations, and sea foam to estimate TOA flux for those conditions in the wavelength band between 0.2 and 4.5 μm. After the first model pass, TOA results are compared with CERES TOA fluxes (that use empirical ADM’s from above) and the model is adjusted by changing surface and atmospheric conditions to produce a better match. The resulting instantaneous SWCLR values have an estimated error of ±5 Wm⁻². It should be noted that while the model-based SW flux values are derived directly from the RT model, the final values are adjusted to fit actual observations that do use the ADM. Thus, this step does improve the overall agreement between the two products. We make use of the FSW-based SWCLR estimates over the same spatial and temporal domains as our observational based SWCLR estimates to examine the differences between aerosol-free fluxes obtained from these two methods. While there are several approaches for simulating the clear sky values over the global oceans, we use the FSW product merely as an example.

Our goal is to obtain the seasonal and spatial distribution of SWCLR both spatially and seasonally from the methods mentioned above. Fig. 1 shows the framework for the problem. In this figure, SWCLR below the aerosol layer labeled SWCLR-A is the desired value. However, when aerosols are present above the ocean surface, CERES can only obtain reflected solar radiation information from the aerosol layer. Therefore, approximations are necessary to obtain SWCLR. Perhaps the most commonly used method to obtain is to examine the relationship between AOD and the shortwave flux for all cloud-free pixels. Then the regression line is extrapolated back to zero AOD and the ordinate value is approximated as SWCLR (Christopher et al., 2006).
The inset in Fig. 1 shows this SWCLR value as SWCLR-1 by the square symbol. However, care must be taken to ensure that the scatter in the data is not large and the relationship is indeed linear and statistically significant. At higher AODs closer to the source region this linearity may break down and therefore it is important to establish this relationship in the linear regime usually for AOD<0.5. Most studies do not report the AOD range for which these regression lines are established; and therefore, we provide this as one recommendation of this paper. Obviously zero AOD conditions and the SWCLR obtained for these conditions are merely approximations and this denotes the darkest surface conditions and thereby the lowest possible SWCLR values.

The second method (SWCLR-B) approximates SWCLR by calculating the lowest or second lowest shortwave flux value (to account for possible cloud shadows and/or noise) over a one-month period and encompassing a certain grid size (2×2° degree, in this case) The lowest shortwave flux is defined as corresponding to clear sky, aerosol free conditions. This method assumes that these conditions do not change over the temporal window that is selected. This is a better approximation if the atmospheric conditions over the pixel do not change within the temporal window, since it is the same pixel above which the aerosol layer was present on another day. Christopher and Zhang (2002) used this method to obtain SWCLR values. Another less common method is to use pixels that are cloud and aerosol free in proximity to the pixel of interest (SWCLR-B). However, in major aerosol regimes and cloudy conditions, the nearest cloud and aerosol free pixel may be several tens of kilometers away and therefore is not a good approximation. Radiative transfer methods that require SWCLR for aerosol climate forcing calculations assume surface and atmospheric conditions based on climatology or observations and then calculate SWCLR values. This is also an approximation since the accuracy of these values depends upon how well the surface and atmosphere can be characterized at discrete wavelengths before integrating it across the entire solar spectrum. Recall that the CERES-FSW product also computes the SWCLR values based on a four-stream radiative transfer model (Fu & Liou, 1993). These models require knowledge of the state of the atmosphere and the surface to compute SWCLR and various assumptions are invoked in these calculations resulting in instantaneous uncertainties of up to ± 5 Wm−2.

3. Results

We compare three methods used for calculating SWCLR, 1) the intercept method (INT) where AOD is regressed against CERES shortwave flux, 2) the second minimum shortwave flux value method (MIN) over a 30-day period, and 3) SWCLR from radiative transfer method (RT) from the CERES-FSW product to quantify the SWCLR as a function of space and season. The global ocean-only average and standard deviation SWCLR for each method is 73.4±3.6, 78.2±5.1, and 74.7±4.0 Wm−2 respectively. The minimum pixel SWCLR is both higher and noisier than the other two methods employed. Fig. 2 shows the scatter plot between the other two methods, INT and RT, for 13 regions over the global oceans. The four seasons are shown in different colors and the regression line corresponding to the annual means is also shown in black. Each panel also contains the annual mean SWCLR values for both the INT and the RT methods within a particular region. The reader is referred to Christopher and Jones (2007) who reported the aerosol radiative efficiencies for the same 13 regions shown here. Fig. 2 and Table 1 (that contains statistics for all 3 methods) reveal several features that are of interest. The SWCLR changes as a function of space and season due to seasonal variations in solar zenith angle and in surface conditions (e.g. wind speed). Sample size in each region varies according to the amount of ocean surface covered in an area and the degree of cloud cover present. As a result, sample size is generally smaller in the Northern Hemisphere where larger landmasses exist. In the Southern Hemisphere, a greater proportion of the Earth’s surface is covered by water, increasing the sample size. Above 30°N, most SWCLR values (from both methods) are clustered between 65 and 75 Wm−2. Further south, the range of values increases to between 70 and 80 Wm−2. The seasonal means for SWCLR vary within a range of 5–10 Wm−2 with the SON season almost always having the lowest values in the Northern Hemisphere.

In the Southern Hemisphere, a greater seasonal dependence is apparent. Between −30°S and 0°S, SWCLR values change up to 20 Wm−2 as a function of season. The lowest values occur during DJF (southern hemisphere winter) with the highest values occurring in DJF (southern hemisphere summer). Since land interaction is much less of a factor in many of these regions south of the equator, the natural seasonal variability in SWCLR is more apparent. Scatter between INT and RT methods is generally less then ± 5 Wm−2, with correlations in excess of 0.9. Overall mean values between these two methods only differ by approximately ± 1.5 Wm−2 for all regions and all seasons, with INT method fluxes generally being slightly lower. The close agreement between INT and RT methods provides confidence in SWCLR estimates, since these methods are essentially independent of each other. The INT largely uses only observed satellite data (although AOT retrievals require radiative transfer calculation) and the RT uses four stream radiative transfer calculations. As indicated in each panel, the INT produces smaller SWCLR values when compared to the RT for the annual mean since the INT extrapolates the values back to zero AOD values, thereby providing a low end to the SWCLR estimate. The RT method on the other hand reflects higher values due to enhanced atmospheric scattering.

The MIN method produces consistently higher values compared to the other two methods for each season and in the overall mean (Table 1). As seen in the inset of Fig. 1, the MIN will always produce higher values than the INT method since it calculates SWCLR from non-zero AOD conditions. The MIN method is also susceptible to cloud or aerosol contamination that is often difficult to reconcile. Since the MIN method values are consistently larger than the INT and RT methods (near 5 Wm−2 in some regions) and since time and space averaging is responsible for these uncertainties, we conclude the MIN method is not adequate for estimating broadband SWCLR estimates. If we did not include the SWCLR values from the MIN method then the difference between the INT and the RT methods are often less than ± 1.5 Wm−2 for all seasons, well below the uncertainty of either method. This indicates an excellent agreement between the four stream radiative transfer calculations in the FSW product and the INT method. The annual mean SWCLR value is about 74 Wm−2 averaged between these two methods.

![Fig. 1. Schematic of how top of atmosphere aerosol free shortwave fluxes are calculated. SWCLR-A is the required shortwave flux but SWCLR-B is used as an approximation. SWCLR-1 is the value used from the intercept method and SWCLR-2 is the value used from a temporal composite. See text for further details.](image-url)
4. Conclusions

Satellite-based methods for studying aerosol direct radiative effects and climate forcing require a robust estimate of aerosol free backgrounds. Using 1 year of MODIS Aqua data, we assess the differences between and among three commonly used methods and conclude that using the minimum shortwave flux over a certain time period for a certain grid produces SWCLR estimates that are noisy with the average aerosol-free SW flux value over the ocean is at least 3 Wm$^{-2}$ greater than either of the two other methods. As a result, using this method in aerosol radiative effect calculations would result in a significant underestimation of DRE. However, the intercept method where the MODIS AOD is regressed against the CERES SW flux from the SSF product is space and time (the INT method) compared with

Table 1
Seasonal and regional statistics for three commonly used methods (intercept, INT; minimum pixel, MIN, and model-based, RT) for shortwave top of atmosphere cloud free regions over the ocean.

<table>
<thead>
<tr>
<th>Region</th>
<th>DJF</th>
<th>MAM</th>
<th>JJA</th>
<th>SON</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INT</td>
<td>MIN</td>
<td>RT</td>
<td>INT</td>
<td>MIN</td>
</tr>
<tr>
<td>1</td>
<td>71.9</td>
<td>77.8</td>
<td>72.6</td>
<td>72.1</td>
<td>79.4</td>
</tr>
<tr>
<td>2</td>
<td>71.7</td>
<td>78.2</td>
<td>72.4</td>
<td>72.4</td>
<td>78.0</td>
</tr>
<tr>
<td>3</td>
<td>72.1</td>
<td>77.5</td>
<td>72.9</td>
<td>72.8</td>
<td>77.4</td>
</tr>
<tr>
<td>4</td>
<td>72.0</td>
<td>81.7</td>
<td>73.4</td>
<td>72.5</td>
<td>82.4</td>
</tr>
<tr>
<td>5</td>
<td>74.5</td>
<td>78.0</td>
<td>75.5</td>
<td>76.2</td>
<td>80.1</td>
</tr>
<tr>
<td>6</td>
<td>74.4</td>
<td>81.1</td>
<td>75.7</td>
<td>76.4</td>
<td>82.7</td>
</tr>
<tr>
<td>7</td>
<td>75.0</td>
<td>83.7</td>
<td>76.8</td>
<td>76.6</td>
<td>83.9</td>
</tr>
<tr>
<td>8</td>
<td>75.0</td>
<td>80.2</td>
<td>76.0</td>
<td>76.9</td>
<td>82.9</td>
</tr>
<tr>
<td>9</td>
<td>79.3</td>
<td>82.9</td>
<td>80.9</td>
<td>72.9</td>
<td>75.4</td>
</tr>
<tr>
<td>10</td>
<td>79.9</td>
<td>83.5</td>
<td>81.2</td>
<td>72.9</td>
<td>75.6</td>
</tr>
<tr>
<td>11</td>
<td>79.7</td>
<td>82.8</td>
<td>81.1</td>
<td>72.4</td>
<td>75.3</td>
</tr>
<tr>
<td>12</td>
<td>80.1</td>
<td>83.7</td>
<td>81.6</td>
<td>71.8</td>
<td>74.1</td>
</tr>
<tr>
<td>13</td>
<td>76.1</td>
<td>80.6</td>
<td>77.2</td>
<td>70.4</td>
<td>74.4</td>
</tr>
<tr>
<td>Mean</td>
<td>75.5</td>
<td>80.9</td>
<td>76.7</td>
<td>73.6</td>
<td>78.6</td>
</tr>
</tbody>
</table>

Fig. 2. Clear-sky shortwave flux from the radiative transfer method (RT) versus the clear sky SW flux from CERES (INT) over global oceans for 13 separate geographical regions. Different colors represent data from one of four seasons (blue = DJF, green = MAM, purple = JJA, and red = SON. The linear regression line between INT and RT for all data points is also shown. Large dots on each panel correspond to the seasonal mean values of both data sets. Values in the lower right-hand corner of each panel indicate the all-year mean values for INT and RT clear-sky fluxes respectively.
values from four stream radiative transfer calculations available from the CERES-FSW product are in excellent agreement (within $\pm 1.3 \text{ Wm}^{-2}$), an agreement similar to that reported by Kato et al. (2002).

Both INT and RT methods produce consistent results over all regions of the globe and for most aerosol types. The advantage of the FSW product is that it can give better spatial sampling in regions where not enough aerosol retrievals are made to create the regression-based SWCLR. This can occur in predominately cloudy regions, areas where aerosol concentrations are consistently small, and where extreme solar zenith angles ($>60^\circ$) prevents proper SW flux retrievals to be made. We therefore conclude that either of these methods can be used to obtain SWCLR for studying top of atmosphere aerosol impacts, with the RT method having the advantage in data sparse regions.

Acknowledgements

This research is supported by NASA’s Radiation Sciences, Interdisciplinary Sciences, an EOS grant, and ACMAP programs. The CERES SSF data that contains the merged MODIS and CERES was obtained through the NASA Langley Distributed Active Archive Systems.

References