Evaluation of Surface Shortwave Flux Estimates from GOES: Sensitivity to Sensor Calibration

MARGARET M. WONSICK, RACHEL T. PINKER, AND WEN MENG
Department of Atmospheric and Oceanic Science, University of Maryland, College Park, College Park, Maryland

LOUIS NGUYEN
Climate Sciences Branch, NASA Langley Research Center, Hampton, Virginia

(Manuscript received 1 June 2005, in final form 14 December 2005)

ABSTRACT

Parameters derived from satellite observations depend on the quality of the calibration method applied to the raw satellite radiance measurements. This study investigates the sensitivity of absolute reflectance, derived cloud cover, and estimated surface shortwave (SW) downward fluxes to two different calibration methods for the visible sensor aboard the eighth Geostationary Operational Environmental Satellite (GOES-8). The first method was developed at NOAA’s National Environmental Satellite, Data, and Information Service (NESDIS), and the second at the NASA Langley Research Center. Differences in visible reflectance ranged from ~0.5% to 3%. The average difference in monthly mean cloud amount was ~3%, and the average difference in monthly mean shortwave downward flux was 5 W m⁻². Differences in bias and rms of the SW fluxes when evaluated against ground station measurements were less than 3 W m⁻². Neither calibration method was shown to consistently outperform the other. This evaluation yields an estimate of the errors in fluxes that can be attributed to calibration.

1. Introduction

Information on surface radiative fluxes is needed in climate change research (Garrat et al. 1993; Wild et al. 1995), hydrological applications (Ohmura and Wild 2002; Sorooshian et al. 2002; Mitchell et al. 2004), mesoscale weather prediction (Yucel et al. 2002, 2003), and oceanic applications (Sui et al. 2003; Baumgartner and Anderson 1999). To obtain the spatial and temporal coverage desired for these applications, observations from geostationary satellites are routinely used as input to radiative flux inference schemes. For radiation balance studies, both IR and visible sensors are of interest. The IR sensors are calibrated on board the satellite, and some visible sensors, such as those on board the Earth Radiation Budget Experiment (ERBE) (Barkstrom et al. 1990) and the Tropical Rainfall Measuring Mission (TRMM) and Terra satellites (Lee et al. 1998), carry onboard calibration systems to maintain stable radiance measurements (Minnis et al. 2002). However, most visible sensors require vicarious calibrations. These sensors obtain a measure of visible reflectance in the form of raw counts, which need to be converted to a reflectance using calibration constants. The general equation for reflectance $R$ is

$$R = \frac{F_{sfc}}{\mu_0 F},$$

where $F_{sfc}$ is the upward flux at the surface, $\mu_0$ is the cosine of the solar zenith angle (SZA; the angle between a pixel’s zenith and the sun), and $F$ is the direct solar irradiance at the top of the atmosphere. The eighth Geostationary Operational Environmental Satellite (GOES-8), which was used in this study, does not have an onboard calibration system and must be calibrated vicariously.

The visible channel calibration procedure accounts for the following: the deep space signal, the degradation of the sensor based on its time since launch, the solar irradiance averaged over the spectral response function.
for the portion of the spectrum within the sensor’s visible channel, sun–earth distance, and solar zenith angle. The latter two parameters are calculated by standard methods based on date, time, and location, but space count and sensor degradation are empirically derived and may vary among different calibration methods. Likewise, the solar irradiance for the visible channel depends on the choice of the equivalent width (the bandwidth of a hypothetical perfect absorber that would absorb the same amount of energy as a given spectral band) for the visible channel, and may differ among calibration methods.

The derived parameters from satellite observations depend on the quality of the calibration, which generally is not perfect. It is therefore of interest to estimate the error limits of derived parameters due to uncertainties in available calibrations. The focus of this study is to investigate the sensitivity of surface shortwave (SW) downward fluxes to two different calibration methods for the visible sensor on board the GOES-8. The first method was developed at the National Oceanic and Atmospheric Administration’s (NOAA’s) National Environmental Satellite, Data, and Information Service (NESDIS) (Weinreb et al. 1997). The second method evolved from a study conducted at the National Aeronautics and Space Administration (NASA) Langley Research Center, in which data from research satellites carrying onboard calibration systems were used as reference sources to develop new calibration coefficients for GOES-8 (Minnis et al. 2002). In this paper, these two calibration methods will be referred to as NOAA and NASA, respectively.

The sensitivity of the surface SW downward fluxes to sensor calibration was evaluated by determining the impact on various parameters that are used as inputs in the radiation inference schemes. First, the absolute reflectances produced by the two calibrations were compared in order to gain an understanding of the impact on this most basic parameter. Since surface SW downward fluxes are highly dependent on cloud cover, the sensitivity of cloud detection to calibration was evaluated and the results were compared to those from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Platnick et al. 2003). Finally, the differences between the satellite-estimated SW downward fluxes obtained when applying the NOAA and NASA calibrations were validated using ground measurements from three sources: surface radiation (SURFRAD) stations (Augustine et al. 2000), the Arizona meteorological (AZMET) network (Brown 1989), and the Illinois State Water Survey network (Hollinger et al. 1994).

Section 2 will describe the data used in the study, and section 3 will explain the methodology. Results and conclusions are given in sections 4 and 5, respectively.

2. Data

2a. Satellite shortwave flux estimates

The Global Energy and Water Experiment/Surface Radiation Budget (GEWEX/SRB) model (Pinker et al. 2003) as implemented operationally at NOAA/NESDIS and used in the Land Data Assimilation System (LDAS) activity (Mitchell et al. 2004) was used to estimate satellite SW downward radiative fluxes for this study. The model was driven by GOES-8 observations in the region bounded by 25°–50°N and 70°–125°W. Visible data (0.67 μm) collected by the satellite at 1-km resolution was subsampled to 4 km during processing to match the resolution of other GOES-8 channels. The coupled cloud and snow detection algorithm developed at the University of Maryland (Li et al. 2006, manuscript submitted to J. Geophys. Res.; Pinker et al. 2006, manuscript submitted to J. Geophys. Res.) was used to discriminate between cloudy and clear skies. Final results were projected onto a 0.5° latitude–longitude grid. As will be shown in section 3, the difference between

---

**Table 1.** Postlaunch calibration coefficients used in NOAA and NASA GOES-8 calibration equations.

<table>
<thead>
<tr>
<th>Constant</th>
<th>NOAA</th>
<th>NASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_0$</td>
<td>0.6556</td>
<td>0.650</td>
</tr>
<tr>
<td>$\Delta g$</td>
<td>$1.107 \times 10^{-4}$</td>
<td>$1.341 \times 10^{-4}$</td>
</tr>
<tr>
<td>$C_0$</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>$E$</td>
<td>518.7</td>
<td>526.9</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** Comparison of NOAA and NASA gain trends from GOES-8 launch date in April 1994 through March 2003.
FIG. 2. (a) Instantaneous visible reflectance values (%) produced with NOAA calibration method for 1700 UTC 1 Aug 2000. (b) Difference between instantaneous visible reflectance values (%) calculated with the NASA and NOAA calibration methods. (c) Same time as in (b), but for cloud amount less than 25%. (d) As in (b), but for cloud amount greater than 60%.

FIG. 3. Monthly average cloud amount (%) at 1030 LST for August 2000 for (a) NOAA calibration, (b) NASA calibration, (c) MODIS, and (d) difference between NASA and NOAA monthly average cloud amount (%) at 1030 LST for August 2000.
the NOAA and NASA calibration methods increases with time since satellite launch. To analyze the impact of the calibration trend on the derived fluxes, this study covers the months of July and August 1998, August 2000, and November 2002.

b. Surface measurements of SW fluxes

Three sources of surface SW downward flux measurements were used to evaluate the satellite estimates. The first is the SURFRAD network, which had six stations available during the period of this study. More information on SURFRAD can be found at http://www.srrb.noaa.gov/surfrad/index.html. The second is the AZMET, which contains radiation measurements from 22 sites in Arizona. Information on AZMET and instruments used is given online at http://ag.arizona.edu/AZMET/. The third is the Illinois State Water Survey network, which has 19 sites in Illinois. Their data can be found at http://www.sws.uiuc.edu/.

3. Methodology

For each month of data, surface SW downward fluxes were computed using both the NOAA and NASA calibration methods. The basic calibration equation used for both methods is a variation of the general equation for reflectance [Eq. (1)], as applied to GOES-8 data:

\[ R = \frac{L}{E \mu_0 \delta}, \]

where \( R \) is reflectance (\%), \( L \) is the mean visible channel radiance (W m\(^{-2}\) sr\(^{-1}\) \( \mu \)m\(^{-1}\)), \( E \) is the solar irradiance for the GOES-8 visible channel (W m\(^{-2}\) sr\(^{-1}\) \( \mu \)m\(^{-1}\)), \( \mu_0 \) is the cosine of the solar zenith angle, and \( \delta \) is the sun–earth distance correction factor (astronomical units). Here \( L \) is determined from

\[ L = g(C - C_0), \]

where \( g \) is the gain (W m\(^{-2}\) sr\(^{-1}\) \( \mu \)m\(^{-1}\) count\(^{-1}\)), \( C \) is the 10-bit observed count, and \( C_0 \) is the 10-bit space count. The gain accounts for sensor degradation after launch and is calculated as

\[ g = g_0 + \Delta g DSL, \]

where \( g_0 \) is the gain on the GOES-8 launch date of 13 April 1994, \( \Delta g \) is the rate of change of gain, and DSL is the number of days since the launch date. The differ-
ence between the methods arises in the values used for $g_0$, $\Delta g$, $C_0$, and $E$ as listed in Table 1. As shown in Fig. 1, the difference in gain between the two methods increases linearly with time due to the dependence on DSL. During the early lifetime of the satellite, the impact of the gain differences is negligible, but its effect becomes more significant as the sensor ages. By August 2001, the gain difference between the two methods is 6%.

The NOAA method used a radiometrically stable reference site in the Sonoran Desert to monitor the GOES-8 sensor degradation and related the GOES-8 top-of-the-atmosphere albedo measurements to those from the calibrated Advanced Very High Resolution Radiometer (AVHRR) visible sensor to determine the rate of change of gain. Further details on how the values in Table 1 were chosen in the NOAA calibration are given in Weinreb et al. (1997) and Rao (2001). In the NASA approach, a technique was used to cross calibrate the GOES-8 sensor using the Visible Infrared Scanner (VIRS) instrument aboard the TRMM research satellite (Minnis et al. 2002). A 21-month sample
was used to perform a regression fit between GOES-8 data and the corresponding data from the VIRS data that were temporally and spatially matched. The absolute accuracy of the NASA calibration is dependent on the VIRS calibration. Minnis et al. (2002) evaluated the VIRS calibration by calibrating GOES-8 with other self-calibrated satellites and found that the VIRS visible calibration is within 2.75% of Terra MODIS and second Earth Resources Satellite (ERS-2) second Along Track Scanning Radiometer (ATSR-2) instruments. The resulting NASA GOES-8 calibration constants $g$ and $C_0$ are given in Table 1. The value of $E$ is set to 526.9 W m$^{-2}$ sr$^{-1} \mu$m$^{-1}$ as per Whitlock et al. (1994).

4. Results

a. Impact on absolute reflectance

The first step in evaluating the differences in calibration methods was to assess their impact on absolute reflectance. Figure 2a shows an instantaneous view of reflectance values calculated with the NOAA calibration method for 1700 UTC on 1 August 2000. Figure 2b shows the difference in absolute reflectance calculated by the NASA and NOAA methods for the same time. Overall, the differences ranged from 0.5% to 3%. Looking separately at the reflectance differences for mostly clear areas (Fig. 2c) and mostly cloudy areas (Fig. 2d), it is evident that the largest differences occur in cloudy areas where reflectance values are high. From Eq. (3), the two main factors that influence the difference in reflectance are the gain $g$ and the space count $C_0$. (Differences attributable to $g_0$ and $E$ are less than 1.3% and 0.008%, respectively.) As shown in Table 1, both the gain and space count are larger in the NASA calibration. The larger gain amplifies the reflectance, whereas the larger space count reduces it. The effect of the space count dominates only over dark scenes such as oceans, where the NASA reflectance is slightly lower than NOAA. The NASA calibration returns higher reflectances in all other cases due to the larger gain.

Table 2. Comparison of November 2002 validation statistics for the SURFRAD network with and without filter for anomalous points. Second and third rows show statistics after removing points outside $3\times$ rms and $2\times$ rms, respectively.

<table>
<thead>
<tr>
<th>SURFRAD network</th>
<th>Bias</th>
<th>Rms</th>
<th>$R$</th>
<th>$N$</th>
<th>% Eliminated</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOAA</td>
<td>6.65</td>
<td>4.27</td>
<td>0.88</td>
<td>1169</td>
<td>N/A</td>
</tr>
<tr>
<td>NASA</td>
<td>9.73</td>
<td>7.35</td>
<td>0.91</td>
<td>1151</td>
<td>1.5</td>
</tr>
<tr>
<td>$3\times$ rms</td>
<td>7.02</td>
<td>5.43</td>
<td>0.94</td>
<td>1100</td>
<td>6.0</td>
</tr>
<tr>
<td>$2\times$ rms</td>
<td>7.02</td>
<td>5.43</td>
<td>0.94</td>
<td>1099</td>
<td>6.0</td>
</tr>
</tbody>
</table>
b. Impact on cloud amount

Among the input parameters to the inference scheme for calculating surface SW fluxes from satellite-observed reflectances is information on cloud amounts. This requires the application of a cloud screening algorithm to differentiate clear sky from cloudy sky. A coupled cloud and snow detection algorithm (CCSDA) developed at the University of Maryland is used for this purpose (Li et al. 2006, manuscript submitted to J. Geophys. Res.). The CCSDA employs seven threshold tests using combinations of four channels of GOES-8 data. The tests are adapted from the Clouds for AVHRR (CLAVR) cloud detection algorithm for the AVHRR sensor onboard the NOAA polar-orbiting satellites (Stowe et al. 1999), and concepts from the MODIS cloud mask (Ackerman et al. 1998) that was developed for the Terra and Aqua Earth Observing System (EOS) satellite missions. The CLAVR and MODIS cloud thresholds are stratified by geography (land versus ocean) and day versus night. These thresholds work well under most conditions but have limitations during the snow season when the surface background may vary dramatically due to snowfall, snowmelt, and snow aging. The CCSDA creates an hourly background analysis using information from clear-sky pixels where available and filling in gaps with a Cressman (1959) analysis approach. Cloud thresholds are set as departures from the expected background values. The background analysis allows the cloud thresholds to vary at pixel level and evolve with time. The CCSDA also applies spectral tests to produce its own analysis of snow conditions that is updated hourly, so timely identification of snow in the background analysis greatly improves the algorithm’s skill in detecting clouds over snow.

To evaluate the impact of calibration on the derived cloud product, the monthly mean cloud amount for August 2000 was calculated using the reflectance data yielded by each calibration method. Figures 3a and 3b show the monthly mean cloud amount produced by the NOAA and NASA calibration methods, respectively. The corresponding monthly mean cloud amount produced from MODIS data at 1° resolution is given in Fig. 3c for comparison. [Note: MODIS data are only available at 1030 local standard time (LST) for the area of interest. The monthly mean cloud amounts for the NASA and NOAA calibrations were calculated by averaging the 1030 LST cloud amounts for every day of the month for each pixel.] Differences in cloud amount due to calibration method used were generally less than 3%, as shown in Fig. 3d. The largest cloud amount differences occurred in partly cloudy regions (cloud amount in the 50%–60% range), where the cloud screening algorithm is more sensitive to absolute reflectance values. Under overcast conditions, the reflectances well exceed the threshold for differentiating clear sky from clouds, so differences in reflectance have less impact on cloud amount.

c. Impact on surface SW downward flux

1) Absolute and normalized fluxes

Absolute differences in monthly mean surface SW downward fluxes were generally less than 8 W m⁻² and were inversely proportional to the actual flux, as shown in Figs. 4a and 4b. The NASA fluxes were lower than those of NOAA because the NASA calibration yielded higher reflectance values. This leads to higher cloud optical depths and lower transmission of SW energy through the atmosphere. The flux differences normalized by the average of the NOAA and NASA fluxes were less than 2.5% (see Fig. 4c).

2) Evaluation of fluxes against ground observations

Model-estimated fluxes were compared to measurements from three sets of ground station networks. The first is the SURFRAD network, with sites located at Bondville, Illinois; Fort Peck, Montana; Goodwin Creek, Mississippi; Boulder, Colorado; Desert Rock, Nevada; and Pennsylvania State University, University Park, Pennsylvania, as shown in Fig. 5. (Note: The Sioux Falls, South Dakota, site was added in 2003, after the period of this study.) These sites were chosen to represent diverse climatic regions of the United States. The second network is AZMET, which
has 22 sites within Arizona. The third is the Illinois State Water Survey network, with 19 sites dispersed throughout Illinois. The SW flux measurement from each ground station was compared to the satellite-estimated SW flux for the $0.5^\circ \times 0.5^\circ$ model grid box that contained the ground station, and the bias and root-mean-square (rms) error were calculated for each hour at each station. In the Arizona and Illinois networks, multiple ground stations that fell within the same grid box were evaluated separately. The hourly bias and rms values from all stations within a network were averaged over a month to produce the evaluation statistics shown in Figs. 6, 7, and 8. Each figure shows the evaluation results for both the NOAA and NASA calibrations. In respective figures, outliers beyond 3 sigma have been filtered out. The effect of the filter is shown in Tables 2, 3, and 4, which compare the statistics without filtering, with a 3-sigma filter, and with a 2-sigma filter. During the period used in the experiment, the maximum difference in hourly averaged bias and rms between the two calibration methods was 3 W m$^{-2}$. Neither of the calibrations consistently outperformed the other at all locations and all times.

5. Conclusions

Calibration of visible satellite data is the first step in deriving meteorological parameters from the raw data. This study gives an estimate of the uncertainties in derived parameters attributable to choice of calibration method. Using the NOAA and NASA calibration methods for GOES-8 visible data, it was found that

- differences in visible reflectance ranged from $-0.5\%$ to 0.6$\%$ for clear-sky conditions and 0$\%$–3$\%$ for cloudy conditions,
- average difference in monthly mean cloud amount was $\sim 3\%$,
- average difference in monthly mean surface SW downward flux was 5 W m$^{-2}$, with a maximum difference of 2.5$\%$ of the average of the NOAA and NASA monthly mean, and
- differences in bias and rms in validations against ground station measurements were less than 3 W m$^{-2}$.

Neither calibration method showed a distinct advantage over the other. These error estimates should be useful in quantifying the impact of choice of calibration method on applications of visible satellite data.

Acknowledgments. The work at the University of Maryland was supported by Grants NA030AR4310045 and NA06GPO404 from NOAA/OGP and NAG59916 from NASA to the University of Maryland. We thank NOAA for the SURFRAD data that were provided to the project at no cost. Thanks are due to Dr. S. Hollinger for arranging the prompt transfer of ground observations from the Illinois Water Survey and to the Arizona Meteorological network (AZMET) and to Drs. J. DeLuisi, J. Augustine, and C. Long for helpful discussions of the SURFRAD data. Some figures were drawn with GrADS, created by Brian Doty.

REFERENCES


Table 4. As in Table 2, but for the Illinois State Water Survey network.

<table>
<thead>
<tr>
<th>Illinois State Water Survey network</th>
<th>Bias</th>
<th>Rms</th>
<th>$R$</th>
<th>$N$</th>
<th>% Eliminated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NOAA</td>
<td>NASA</td>
<td>NOAA</td>
<td>NASA</td>
<td>NOAA</td>
</tr>
<tr>
<td>No filter</td>
<td>2.87</td>
<td>0.47</td>
<td>66.52</td>
<td>66.79</td>
<td>0.92</td>
</tr>
<tr>
<td>3× rms</td>
<td>3.28</td>
<td>0.38</td>
<td>60.94</td>
<td>60.54</td>
<td>0.94</td>
</tr>
<tr>
<td>2× rms</td>
<td>1.24</td>
<td>0.90</td>
<td>51.38</td>
<td>51.37</td>
<td>0.96</td>
</tr>
</tbody>
</table>

% Eliminated refers to the percentage of data points that fall within 2-sigma from the mean.


