



# Top-of-the-atmosphere shortwave flux estimation from satellite observations: an empirical neural network approach applied with data from the A-train constellation

Pawan Gupta<sup>1,2</sup>, Joanna Joiner<sup>2</sup>, Alexander Vasilkov<sup>3,2</sup>, and Pawan K. Bhartia<sup>2</sup>

<sup>1</sup>Universities Space Research Association, Greenbelt, MD, USA

<sup>2</sup>NASA Goddard Space Flight Center, Greenbelt, MD, USA

<sup>3</sup>Science Systems and Applications, Inc., Greenbelt, MD, USA

Correspondence to: Pawan Gupta (pawan.gupta@nasa.gov)

Received: 31 December 2015 – Published in Atmos. Meas. Tech. Discuss.: 1 February 2016

Revised: 29 May 2016 – Accepted: 7 June 2016 – Published: 7 July 2016

**Abstract.** Estimates of top-of-the-atmosphere (TOA) radiative flux are essential for the understanding of Earth's energy budget and climate system. Clouds, aerosols, water vapor, and ozone (O<sub>3</sub>) are among the most important atmospheric agents impacting the Earth's shortwave (SW) radiation budget. There are several sensors in orbit that provide independent information related to these parameters. Having coincident information from these sensors is important for understanding their potential contributions. The A-train constellation of satellites provides a unique opportunity to analyze data from several of these sensors. In this paper, retrievals of cloud/aerosol parameters and total column ozone (TCO) from the Aura Ozone Monitoring Instrument (OMI) have been collocated with the Aqua Clouds and Earth's Radiant Energy System (CERES) estimates of total reflected TOA outgoing SW flux (SWF). We use these data to develop a variety of neural networks that estimate TOA SWF globally over ocean and land using only OMI data and other ancillary information as inputs and CERES TOA SWF as the output for training purposes. OMI-estimated TOA SWF from the trained neural networks reproduces independent CERES data with high fidelity. The global mean daily TOA SWF calculated from OMI is consistently within  $\pm 1\%$  of CERES throughout the year 2007. Application of our neural network method to other sensors that provide similar retrieved parameters, both past and future, can produce similar estimates TOA SWF. For example, the well-calibrated Total Ozone Mapping Spectrometer (TOMS) series could provide estimates of TOA SWF dating back to late 1978.

## 1 Introduction

The Earth's energy budget constrains the general circulation of the atmosphere and determines the climate of the Earth–atmosphere system; it is therefore also an indicator of possible climate changes (Hatzianastassiou, et al., 2004). There is a long history of attempts to estimate Earth's albedo and energy budget (Dines, 1917; Hartmann et al., 1986). With the advent of the satellite remote sensing era, it became possible to directly measure the albedo of the Earth. Subsequently, the shortwave (SW) energy balance at the top of the atmosphere (TOA) and the role of clouds, aerosols, and trace gases has been studied using satellite measurements (Ramanathan et al., 1989; Yu et al., 2006; Bellouin et al., 2005; Loeb et al., 2005; Patadia et al., 2008; Joiner et al., 2009).

The Earth Radiation Budget Experiment (ERBE) was launched in October 1984 by the space shuttle Challenger and provided long- and shortwave radiation parameter measurements. TOA SW radiative parameter estimates from ERBE (Barkstrom, 1984; Barkstrom and Smith, 1986) showed that clouds approximately double the albedo of Earth to an all-sky value of 0.3 from an estimated clear-sky value of 0.15 (Ramanathan et al., 1989; Harrison et al., 1990). The next generation of broadband instruments, the Cloud and the Earth's Radiant Energy System (CERES), draws heavily on ERBE heritage. Since its first launch in 1997 on board the NASA Tropical Rainfall Measurement Mission (TRMM), CERES has provided continuous observations that can be used to understand the role of clouds and the energy cycle

in global climate change (Wielicki et al., 1995; Loeb et al., 2012).

Continuous and coincident measurements of radiative fluxes and atmospheric components facilitate research studies to estimate and understand the role of different atmospheric components on the planetary energy balance. Although CERES provides state-of-the-art estimates of TOA radiative fluxes, it was not designed to make measurements of individual atmospheric components that impact those fluxes. Several studies have utilized aerosol and cloud information from high spatial resolution MODerate resolution Imaging Spectroradiometer (MODIS) measurements to quantify their impact on TOA fluxes (Yu et al., 2006; Pata-dia et al., 2008; Zhang et al., 2005b; Loeb et al., 2005; Oreopoulos et al., 2009). Several attempts have also been made to convert narrowband radiances into broadband fluxes using regression or more sophisticated statistical approaches (Chevallier et al., 1998; Hu et al., 2002; Domenech and Wehr, 2011; Vázquez-Navarro et al., 2013).

The Ozone Monitoring Instrument (OMI), flying on NASA's Aura satellite since 2004, provides information about components important for the Earth's SW radiation budget, including the effective cloud/aerosol fraction (Stammes et al., 2008; Joiner and Vasilkov, 2006) and total column ozone (TCO) (Veefkind et al., 2006; McPeters et al., 2008; Kroon et al., 2008). OMI-retrieved parameters can be utilized to understand their role in the Earth's SW energy budget.

Modeling the spatial and temporal distribution of the TOA shortwave flux (SWF) requires a description of the components that control the transfer of solar radiation within the Earth–atmosphere system. When required parameters are missing or incomplete, a statistical approach is an alternative for estimation of TOA SWF. Here, we develop an artificial neural network model (NNM) to estimate total reflected TOA outgoing SWF. Artificial NNs are algorithms that simulate biological NNs by learning and pattern recognition (Bishop, 1995). NNs have been used by many scientific disciplines, including Earth science, to identify patterns and extract trends in imprecise and complicated nonlinear data (e.g., Lee et al., 1990; Gupta and Christopher, 2009). In radiation studies, NNs have been used to estimate TOA and surface SWF based on radiative transfer calculations with or without data from satellites (e.g., Krasnopolsky et al., 2008, 2010; Take-naka et al., 2011; Vázquez-Navarro et al., 2012; Jiang et al., 2014). CERES TOA flux algorithms have also used NNs to generate angular distribution models (ADMs) in the absence of sufficient high-resolution imager information for reliable scene identification (Loukachine and Loeb, 2003, 2004).

In this study, we utilize OMI cloud and ozone products along with other ancillary data to estimate TOA SWF. We develop NNs that take OMI-derived quantities as inputs and provide CERES-equivalent TOA SWF as the output. In essence, the trained NNs have learned or incorporated all of the complexity that is essential to the CERES algorithms

(ADMs, scene identification, etc.) and are then able to predict TOA SWF directly and efficiently based on a limited number of retrieved products from OMI or sensors that provide similar data. The trained NN models are optimized to run with data sets from OMI or similar sensors and can be applied generally to different seasons and years. For example, the NN-based models we develop here can be applied to similar measurements from the Total Ozone Mapping Spectrometer (TOMS) instruments. One objective of this study is to assess how well TOA SWF can be estimated using OMI-derived cloud and ozone products with NNs when nearly coincident CERES data are used for training. The developed NNs can then be applied to other data sets with similar products and accuracy. Alternatively, the general training approach could be applied with similar data sets such as with MODIS cloud and ozone products.

The strength of a NN approach is that it is highly efficient and, if well-trained, should be precise and accurate for this type of problem. NNs may be used to examine the sensitivity of TOA SWF to various input data sets but do not themselves provide specific insight into the physical mechanisms behind those sensitivities. NNs will of course only be as good as the data that are used for training. In addition, they may not perform well for unusual conditions that are not present in the training data set. Therefore, NNs cannot replace dedicated TOA SWF estimates from instruments like CERES.

The paper is organized as follows: Sect. 2 describes the various satellite data sets utilized in the study. Section 3 discusses the development of NN models including the selection of input parameters. Section 4 evaluates our NN estimation of TOA SWF using independent CERES data over ocean and land. Section 5 summarizes the results and discusses future work.

## 2 Satellite data sets and coincident sampling

Under clear-sky conditions, TOA SWF is affected by the Earth's surface properties, atmospheric absorbers such as water vapor, ozone, and aerosols, and scattering by air molecules and particulates. Over ocean, surface properties can be characterized by ocean color and roughness of the ocean surface. Under cloudy sky conditions, cloud optical properties such as the cloud optical thickness, geometrical cloud fraction, effective radius, and phase function affect TOA SWF. In clear and cloudy skies, the solar zenith angle (SZA) and Sun–Earth distance (SED) impact the TOA SWF.

In this work, we make use of data sets mainly from two passive sensors in A-train constellation of satellites that fly within 15 min of each other: (1) Aura OMI with an equatorial crossing time of  $\sim 13:45 \pm 15$  min LT (local time) and (2) Aqua CERES with an equatorial crossing time of  $\sim 13:30$  LT. For testing and optimization of the NNMs, we primarily use 2007 data over global oceans. Starting around 2008, OMI experienced an anomaly presumably due to mate-

rial outside the sensor that adversely affects the quality of the level 1B and level 2 data products in a portion of its 60 rows across the swath. Our study focuses on data in 2007 that are not significantly affected by these anomalies.

## 2.1 CERES

The first CERES instrument flew on the TRMM satellite, launched in November 1997, and provided data until 2000. Five CERES instruments are currently operating: two on NASA's Terra satellite (FM1 and FM2), two on NASA's Aqua satellite (FM3 and FM4), and one on the Suomi National Polar-orbiting Partnership (NPP) satellite (FM5). These CERES instruments provide radiometric measurements of the Earth's atmosphere from three broadband channels: (1) a shortwave channel to measure reflected sunlight (0.3–5  $\mu\text{m}$ ), (2) a long-wave channel to measure Earth-emitted thermal radiation in the window region (8–12  $\mu\text{m}$ ), and (3) a total channel to measure radiation from 0.3 to 200  $\mu\text{m}$ .

CERES radiances are converted to TOA fluxes using ADMs. The CERES science team has an extensive database of ADMs for clear- and cloudy-sky over both land and ocean (Loeb et al., 2005). The ADMs heavily depend upon the observed scene type and are sensitive to surface characteristics as well as cloud and aerosol optical properties (Loeb et al., 2003; Zhang et al., 2005a; Patadia et al., 2011). The ADMs over ocean are dependent upon wind speed and aerosol optical thickness along with sun-satellite geometry (Zhang et al., 2005a).

The Aqua spacecraft carries two identical CERES instruments: one operates in a cross-track scan mode (FM3) and the other in a biaxial scan mode (FM4). Measurements from the biaxial scan mode were used to develop the ADMs; this provided considerable improvement over the previous generation of instruments, including the ERBE (Loeb et al., 2003, 2007).

This study uses the Single Scanner Footprint (SSF, Edition 3A) TOA SWF obtained from the Aqua CERES FM3. The SSF product is an instantaneous merge of CERES parameters with coincident cloud and aerosol parameters derived from the Aqua MODIS (Loeb et al., 2003) at the footprint level (i.e., not daily averages). The high-resolution ( $1 \times 1 \text{ km}^2$  at nadir) MODIS imager data are used to characterize the clear and cloudy portions of the larger CERES pixel ( $20 \times 20 \text{ km}^2$  at nadir). We also used broadband surface albedo over land derived using measurements from MODIS and CERES to characterize the land surface in the NN trained specifically over land.

## 2.2 OMI

OMI provides hyper-spectral measurements of Earth-backscattered sunlight from UV to visible wavelengths ( $\sim 270\text{--}500 \text{ nm}$ ) with a spectral resolution of the order

of 0.5 nm (Levelt et al., 2006). Its spatial resolution is  $13 \times 24 \text{ km}^2$  at nadir with a swath width of about 2600 km. Cloud, aerosol, and TCO products from OMI are used in this study. Specifically, the cloud–aerosol optical centroid pressure (OCP), effective cloud fraction (ECF;  $f_c$ ), Lambertian equivalent reflectivity (LER) at 354.1 nm, SZA, relative azimuth angle (RAA), and viewing zenith angle (VZA) are obtained from the OMI cloud products as detailed below, and aerosol index (AI) and TCO are obtained from the OMITOMS total ozone product (OMTO3, version 8.5, collection 3) (McPeters et al., 2008).

Cloud–aerosol OCP, also known as effective cloud pressure, is a measure of the reflectance-weighted pressure reached by incoming solar photons (Joiner et al., 2012). It is distinct from the cloud-top pressure (CTP). While CTP is the more important parameter needed for TOA long-wave flux, OCP is more related to atmospheric absorption in the shortwave. OCP is derived from OMI observations using two different methods (Stammes et al., 2008): (1) filling-in of solar Fraunhofer lines from rotational Raman scattering in the UV (the OMCLDRR product) (Joiner and Bhartia, 1995; Joiner et al., 2004) and (2) collision-induced oxygen absorption ( $\text{O}_2\text{--O}_2$ ) at 477 nm (the OMCLDO2 product) (Stammes et al., 2008; Acarreta et al., 2004). Unless otherwise specified, we use the  $f_c$  and OCP from OMCLDRR product here.

OMI cloud and trace-gas algorithms use a simplified mixed Lambertian cloud model to estimate observed radiances  $I_m$ . In this scenario, a pixel is modeled as having components from clear and cloudy sub-pixels weighted using an ECF  $f_c$ ; i.e.,

$$I_m = I_g (1 - f_c) + I_c f_c, \quad (1)$$

where  $I_g$  and  $I_c$  are the radiances computed in the Rayleigh atmosphere for Lambertian surfaces corresponding to the clear and cloudy portions of the scene, respectively;  $f_c$  is defined as the fraction of the Lambertian cloud covering the pixel and is related to both the geometric cloud fraction and cloud optical thickness. It contains information similar to the LER of the scene (related to cloud and surface reflectivities). However, because it attempts to account for variations in the Earth's surface reflectivity, it is a more spectrally invariant quantity and therefore potentially more highly correlated with TOA SWF.

Formally, the ECF is wavelength dependent because it is defined by spectral quantities (Stammes et al., 2008). We conducted a simulation experiment to evaluate the wavelength dependence of  $f_c$ . In this experiment, we simulate observed TOA radiances as a weighted sum of the clear-sky and cloudy radiances; i.e.,

$$I_m = I_g (1 - f_g) + I_c f_g \quad (2)$$

where  $I_c$  is the cloudy radiance computed with a plane-parallel cloud model that depends on cloud optical thickness, and  $f_g$  is the geometrical cloud fraction. In our simulation, clouds have a vertically uniform distribution of the

extinction coefficient and phase function. We use a cloud-top height of 5 km and a cloud layer thickness of 1 km. The assumed cloud optical depth (COD) of 20 is spectrally independent within the 320–1400 nm wavelength range. The spectrally independent optical thickness is a good approximation for clouds with sufficiently large particles (Deirmendjian, 1969). We neglect gaseous absorption in the specified spectral range. Three models of cloud phase function are used: (1) ice crystals with an effective diameter of 60  $\mu\text{m}$  (Baum et al., 2014), (2) Deirmendjian's C1 model of a water cloud having droplets with an effective diameter of 12  $\mu\text{m}$  (Deirmendjian, 1969), and (3) the Henyey–Greenstein model (e.g., van de Hulst and Irvine, 1963) with an asymmetry parameter of 0.85. We use a simplified model of the spectral ground reflectance:  $R_g = 0.05$  at  $\lambda < 700$  nm and  $R_g = 0.2$  at  $\lambda > 700$  nm. We then calculate  $f_c$  by inverting Eq. (1) assuming a Lambertian cloud with a reflectivity of 0.8; this is commonly used for trace-gas algorithms (Stammes et al., 2008).

If  $f_c$  does not vary much with wavelength, then it should be highly correlated with TOA SWF and thus a good predictor in a statistical model of TOA SWF. We investigate the spectral dependence of  $f_c$  in Fig. 1 where we performed calculations using  $f_g = 0.5$ ,  $\text{SZA} = 45^\circ$ , and at nadir for the three phase functions assuming a cloud single-scattering albedo of unity. It can be seen that  $f_c$  is nearly invariant with wavelength over a wide spectral range; it changes by only a few percent even for the steep change in the ground reflectance (simulating the so-called red edge) at 700 nm. COD of less than 20 at the same geometrical cloud fraction leads to a lower value of the ECF. However, the spectral variations of ECF remain similar to that shown in Fig. 1. The weak spectral dependence of ECF is explained by the fact that the Lambertian cloud model with cloud reflectivity of 80 % effectively accounts for Rayleigh scattering in partially cloudy scenes as it has been shown by Ahmad et al. (2004). This result holds when other input parameters in our simulation are varied. We will show the implications of the spectral invariance of  $f_c$  below.

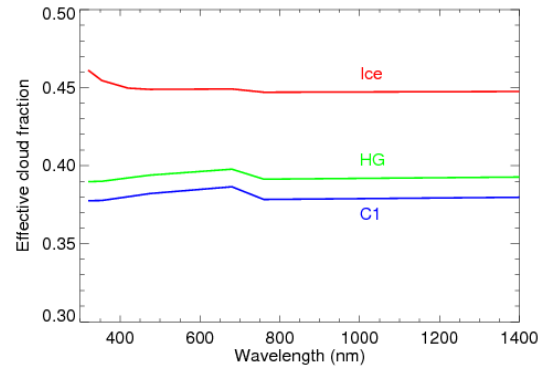
We have used the following modified cloud fraction parameter,  $f_{c\_mod}$ , as a predictor to estimate TOA SWF:

$$f_{c\_mod} = f_c \times \cos(\text{SZA}) \times \left(1/\text{SED}^2\right), \quad (3)$$

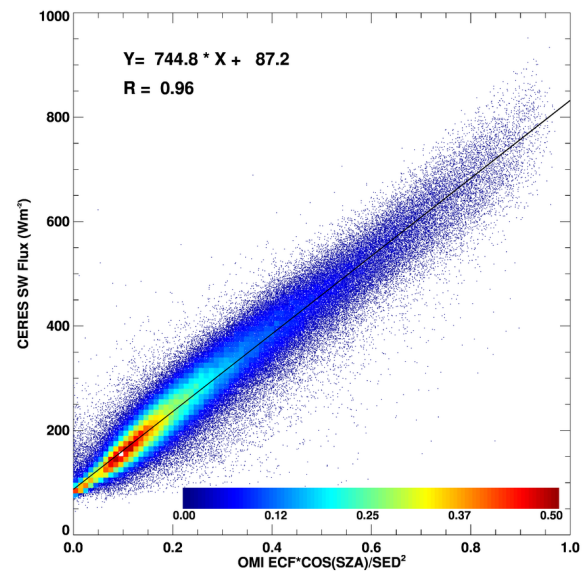
where SED is the Sun–Earth distance. The modification accounts for variation in the incoming solar irradiance with respect to SZA and SED. Figure 2 shows a 2-D histogram for a month of collocated CERES TOA SWF and OMI  $f_{c\_mod}$  (collocation criteria described in detail below) that demonstrates the near-linear relationship between the two parameters. It also indicates that the single parameter  $f_{c\_mod}$  captures much of the variability in TOA SWF.

### 2.3 Ancillary data

In addition to OMI data, a SeaWiFs-derived chlorophyll (Chl) concentration climatology is used as an input predic-



**Figure 1.** The spectral dependence of the effective cloud fraction ( $f_c$ ) for land,  $f_g = 0.5$ ,  $\text{SZA} = 45^\circ$ , observation at nadir. Red: ice cloud phase function; green: the Henyey–Greenstein (HG) function; blue: C1 cloud model.



**Figure 2.** Two-dimensional histogram of effective cloud fraction (ECF or  $f_c$ ) normalized with respect to incoming solar irradiance (i.e.,  $f_{c\_mod} = \text{ECF} \times \cos(\text{SZA}) / \text{SED}^2$ ) vs. CERES TOA shortwave (SW) flux over ocean for January 2007 showing a highly linear relationship. The color bar shows the sample density (normalized, in terms of fractional amount, in percent) for each bin in the collocated data set; when only a single point is in a bin, it is shown as a dark blue dot.  $R$  is the linear correlation coefficient. The solid black line is a linear fit to the data.

tor when  $f_c < 1$ . The precipitable water (PW) and 2 m surface wind speed (wind) are also used as predictors; these are provided in the CERES SSF data set and are taken from the Goddard Earth Observing System (GEOS) 4 reanalysis (Bloom et al., 2005).

## 2.4 Coincident sampling of OMI and CERES

Because the sizes of the OMI ( $13 \text{ km} \times 24 \text{ km}$ ) and CERES ( $20 \text{ km}^2$ ) pixels are similar at nadir, we perform a simple spatial collocation by finding the closest CERES pixel corresponding to each OMI pixel. OMI and CERES collocated pixels are only included in our training and validation samples when the distance between centers of OMI and CERES pixels is less than 20 km. We do not include pixels with viewing zenith angles  $> 60^\circ$ . At these angles, OMI and CERES pixels become significantly larger ( $\sim 150 \text{ km}$  for OMI and  $\sim 200 \text{ km}$  for CERES in the cross track direction) and therefore may contain many different scene types. We also mask OMI pixels with  $\text{AI} > 1$  to avoid heavy absorbing aerosol loaded scenes where the  $f_c$  and OCP are known to contain errors (Vasilkov et al., 2008). This does remove not non-absorbing aerosols. Non-absorbing aerosols are essentially treated as clouds within the OMI cloud data sets. We examined the frequency distribution of the distance between OMI and CERES pixels of all the collocated data sets and found that most of the collocated data, 98 and 60 % have distances less than 20 and 10 km, respectively. The quality-controlled collocated data are then averaged on equal latitude and longitude grids of  $1^\circ \times 1^\circ$  (unless otherwise specified) for training, testing, and validation of the neural networks.

The NN inputs (predominantly retrievals from OMI measurements) are used to train the NN to match the output (CERES-derived TOA SWF). Once the network is trained, input data sets can be used to calculate TOA SWF with characteristics similar to the CERES product. Therefore, the NN-produced TOA SWF will be referred as OMI estimated SWF throughout the paper.

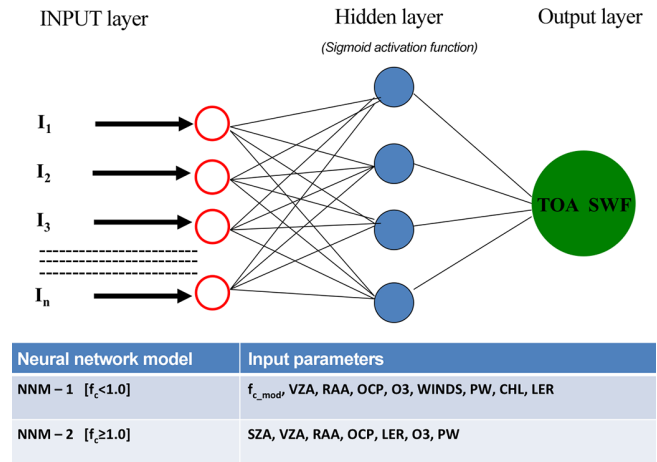
## 3 Artificial neural network model

### 3.1 General NN architecture and training approach

The general neural network architecture has three layers of neurons: an input layer, a hidden layer, and an output layer with standard multi-layer network architecture. We use the same number of neurons in the hidden layer as in the input layer as this produced generally good results. The input layer has an identity activation function; all other layers are connected by sigmoid activation functions (Eq. 4).

$$y(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

The network normalizes both input and output data sets with a unique linear mapping for each input and output parameter. Figure 3 provides an example of a schematic of the network used in our study. Here we used two different NNMs: one with nine nodes or parameters (NNM1) and a second with seven nodes (NNM2) in the input layer. Both of these models have one node (TOA SWF) in the output layer. Figure 3 also



**Figure 3.** A schematic of the neural network model used for estimation of TOA SW flux with OMI UV retrieved parameters. The table in the bottom lists all the input parameters corresponding to two NN models used.

lists the input parameters corresponding to the NNM1 and NNM2 models. The NNM1 model is optimized for ocean cases where the OMI  $f_c < 1.0$ , whereas NNM2 is optimized for cases where  $f_c = 1$  (saturated cases).

NNMs require optimized training to produce accurate outputs. Here we use a standard back propagation training algorithm (Hertz et al., 1991), where inputs are iteratively sent to the neural network. In back propagation, the hidden layer weights associated with each input parameter are modified through the training process that minimizes errors between the targets and outputs (Bishop, 1995; Gardner and Dorling, 1998). After each iteration, the error is propagated backward through the network and weights are modified to bring the actual response of the network closer to the desired output in a statistical sense. The function minimized during the training is a sum of squared errors of each output for each training pattern. Once the network is trained, it can be evaluated using independent data (i.e., not used in the training data set).

### 3.2 Impact of different input parameters

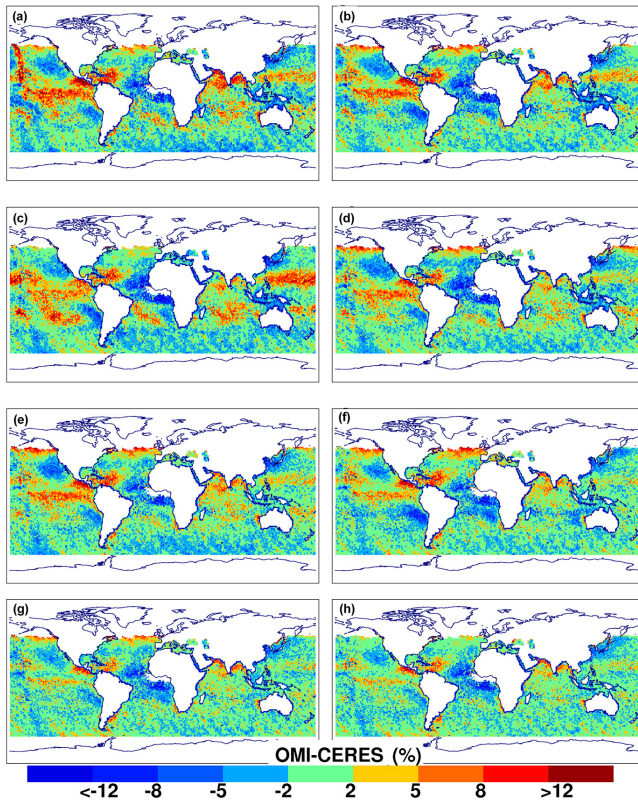
Here we examine the impact of using various input parameters on the derived neural networks. This exercise is performed using data with  $f_c < 1$  with 1 month of the data over ocean (January 2007). Table 1 presents the performance of eight different NNMs, denoted models “a” through “h”, with various input parameters listed in Table 1 and described in more detail in Sect. 3.2.1–3.2.2.

#### 3.2.1 Inclusion of OMI UV-derived parameters

In model a, we have combined the effects of SZA, SED, and  $f_c$  into a single input parameter called  $f_{c\_mod}$ , which was defined in Sect. 2.2. Use of this modified input parameter alone explains about 94 % ( $R = 0.97$ ) of the variabil-

**Table 1.** Statistical analysis of the input parameter selection exercise. The correlation coefficient ( $R$ ), slope, bias, and standard deviation (SD) of OMI–CERES TOA SW flux for eight different NN models are presented. These numbers correspond to daily intercomparison between OMI and CERES TOA SW flux. Data from January 2007 are used for this exercise.

Model	Parameters	$R$	Slope	Bias	SD ( $\text{Wm}^{-2}$ )
a	$f_{\text{c\_mod}}$	0.971	0.941	0.051	37.1
b	$f_{\text{c\_mod}}$ , VZA, RAA	0.979	0.959	0.000	31.4
c	LER, SZA, VZA, RAA	0.979	0.959	−0.030	31.1
d	$f_{\text{c\_mod}}$ , VZA, RAA, $\text{O}_3$	0.980	0.960	−0.009	30.9
e	$f_{\text{c\_mod}}$ , VZA, RAA, OCP	0.981	0.962	−0.004	30.0
f	$f_{\text{c\_mod}}$ , VZA, RAA, $\text{O}_3$ , OCP	0.981	0.963	0.002	29.9
g	$f_{\text{c\_mod}}$ , VZA, RAA, $\text{O}_3$ , OCP, PW, Wind	0.982	0.964	0.002	29.2
h	$f_{\text{c\_mod}}$ , VZA, RAA, $\text{O}_3$ , OCP, PW, Wind, Chl, LER	0.983	0.967	−0.010	28.3



**Figure 4.** Monthly mean (January 2007) maps of OMI minus CERES TOA SW flux (percent) for eight different NN models. The letters on the map corresponds to model number in Table 1.

ity in TOA SWF. As we add other parameters in models b through h, we observe small improvements in the OMI-estimated TOA SWF. Figure 4 shows the spatial distribution of monthly mean OMI–CERES SWF differences for these models.

The TOA SWF is estimated from a measured radiance and therefore the observational geometry factors in. The addition of satellite-viewing geometry parameters (VZA, RAA) to model a provides improvements in areas of high biases

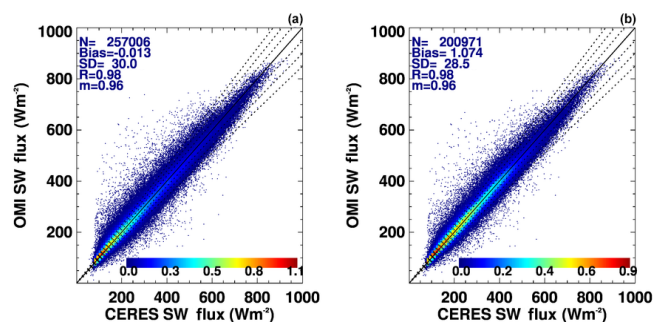
and reduces the standard deviation from 37.1 to 31.4  $\text{Wm}^{-2}$ . Model c tests the ability of the 354 nm reflectivity (LER) to predict TOA SWF in place of  $f_{\text{c\_mod}}$ . Although the statistical parameters in Table 1 corresponding to models b and c are very similar, we note spatial differences in the OMI–CERES TOA SWF in Fig. 4. Further analysis reveals that the  $f_{\text{c}}$ -based model b provides more accurate flux estimation as compared with the LER-based model for a larger range of fluxes.

The inclusion of TCO (model d) as an input parameter positively impacts TOA SWF estimation as shown in Fig. 4d; the high positive biases in the tropical Pacific and Indian oceans and in the region near 60° N have been reduced. The percentage of monthly mean OMI–CERES data that fall within  $\pm 8\%$  increases from 91 % in model c to 94 % in model d.

Model e adds cloud OCP to the input parameters included in model b. OCP also improves SWF estimates; the regions where improvement occurs are different from those improved by using TCO. Model f shows that when TCO and OCP are used together as input parameters, there is further improvement in SWF estimation. Although the global statistics in Table 1 do not clearly reflect this improvement, Fig. 4f shows that inclusion of OCP and TCO reduces biases in many regions, most prominently in the tropics. The percentage of total OMI samples (monthly mean) within  $\pm 8\%$  of CERES increases from 92 % in model b to 95 % in model f. Apart from these parameters, we also evaluated the inclusion of AI as an input parameter (not shown here). We found that overall it does not significantly improve the results; however it does provide some improvement in regions with positive AI values.

### 3.2.2 Addition of meteorological and other ancillary data

The impact of surface winds and total column water vapor (model g in Fig. 4g) is more prominent in the tropics than in other regions. Inclusion of Chl and LERs in model h removes some of the notable low biases in TOA SWF near the coast of northern China, the Caspian Sea, and the Black Sea. Further-



**Figure 5.** Two-dimensional histograms (similar to Fig. 2) of daily mean CERES- and OMI-derived (NN) TOA SW flux for January 2007: (a) the NN is trained on and applied to data from January 2007 and compared with CERES data from that month and (b) the NN is trained using data from January 2007 and applied with input data from January 2006 and compared with CERES data from January 2006. The color bar shows sample density (in fractional (percent) amount) of the collocated data set (see text for more details). The three dotted lines are 5, 10, and 15 % error envelope lines.

more, model h corrects for negative biases in areas with high TOA SWFs, most likely due to the inclusion of LER. The model “h” produces 89 % (99 %) of OMI-estimated monthly mean TOA SWFs within  $\pm 5$  % ( $\pm 12$  %) of CERES and is the best of the eight models. In all of the results, some striped patterns appear in the difference maps. These stripes are also seen in difference maps for July and will be discussed below in Sect. 4.3.

### 3.3 Consistency over time

We next examine the performance of the NN model h with respect to different input samples. We first examine the robustness of the NN for detection of interannual variability. In this exercise, we trained the NN with data from the first 15 days of January 2007 as above and applied it to input data from the entire months of January 2007 (Fig. 5a) and January 2006 (Fig. 5b). Figure 5a–b present 2-D histograms similar to that in Fig. 2 but here compare the TOA SWF from the NN with that from CERES for January of the 2 different years over ocean. The colors represent the 2-D histogram (or density) of coincident pairs using a bin size of  $10 \text{ Wm}^{-2}$ . The solid black 1 to 1 line is shown with three dotted lines on both sides that represent envelopes of  $\pm 5$ ,  $\pm 10$ , and  $\pm 15$  % OMI–CERES differences. Figure 5a–b show that the NN performance is consistent between years. Although the number of samples in January 2006 and 2007 is a bit different, the NN model produces similar statistics.

A NN trained on 1 particular month of data is not guaranteed to perform well for a different month. In the next test, we used the NN model trained on January 2007 data with input data from July 2007. Results for July (using CERES as the benchmark) were degraded as compared with application to

January data. This is due presumably to changes in observing conditions between the 2 different months (changes in solar angles).

We then trained the same NN (identical input parameters) using a subset of data from July 2007 and applied it to data from the entire month of July 2007. Results (again using CERES as the benchmark) were of similar quality to those where the NN was trained and applied to January. This exercise suggests that we may need to use different models for different months or expand our training data set for application to different months.

We next use data from the 1st day of each month of 2007 for training and data from the 16th day of each month of 2007 for evaluation. The comparisons with CERES using the training and validation data are consistent as shown in Fig. 6. The almost identical values of statistical parameters for training and validation data demonstrate that the neural network has been well trained. For example, there is a high degree of linear correlation ( $R = 0.98$  or  $R^2 = 0.96$ ) and slopes close to 1 (0.96) in both training and validation comparisons. The mean bias in both training and validation data sets is close to zero, whereas the global standard deviation remains stable and close to  $30 \text{ Wm}^{-2}$  in the two independent model runs. Further analysis shows that 83, 70, and 43 % of TOA SWF estimated from OMI (training and validation data combine) lie within the 15, 10, and 5 % of the CERES TOA SWF, respectively.

Further evaluation of the entire year reveals that this NN (model h) is appropriate for all months. Therefore this model will be used for subsequent analyses in this study. Creating more networks as a function of scene type or for different latitude belts or even for different months/seasons may improve results in certain regions. However, based on our results, we simplified the approach by minimizing the number of networks.

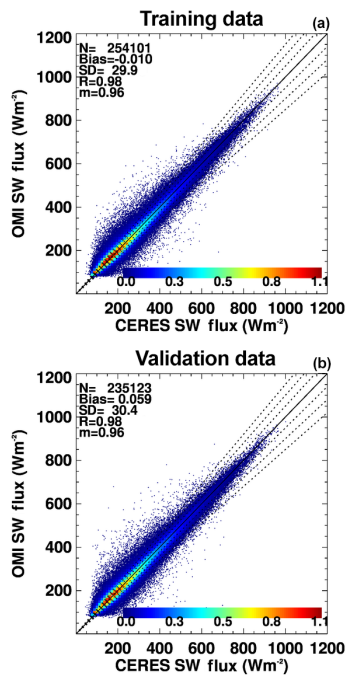
### 3.4 Case of $f_c = 1$

About 1–2 % of total coincident data correspond to  $f_c = 1$ , typical of overcast conditions with optically thick clouds. These cases were modeled using a simpler NN model with inputs of LER, SZA, VZA, RAA, OCP,  $O_3$ , and PW; the surface-related parameters (surface wind speed and Chl content) do not produce a significant impact for  $ECF = 1$  and have therefore been removed. Subsequent all-sky results shown in Fig. 6 use combined outputs from the two separate models for  $f_c < 1$  and  $f_c = 1$ .

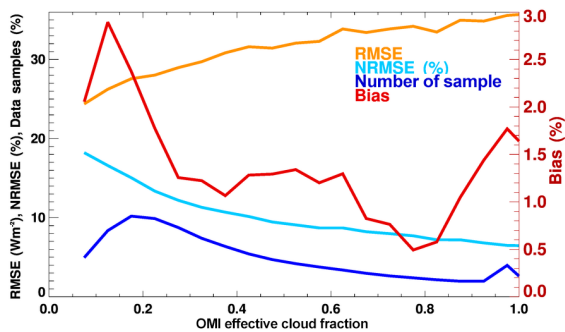
## 4 Results and discussion

### 4.1 Bias and root mean square error (RMSE) as a function of ECF

Here onward, all the results presented in Figs. 7–12 and discussed are produced using model h, which is the most op-

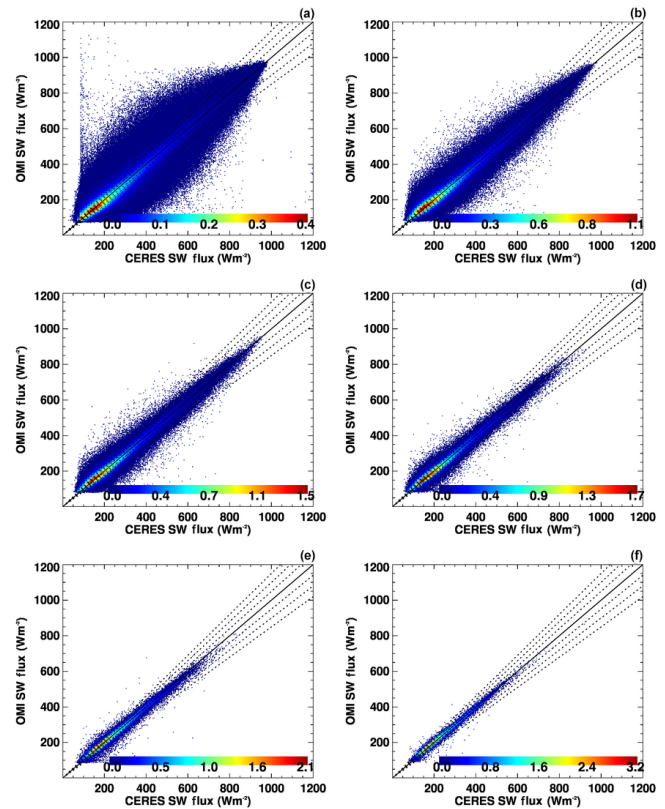


**Figure 6.** Two-dimensional histogram similar to Fig. 5 but showing training (top) and validation (bottom) results from the combined all-sky NN models (input parameters listed in Fig. 3, model “h” for  $f_c < 1$  and as in Fig. 3 for  $f_c > 1$ ) as final selected models for estimation of TOA SW flux. The data from two NN models have been combined in these plots. These are instantaneous flux values averaged over  $1^\circ \times 1^\circ$  grid boxes. Data from the 1st day of each month of 2007 are used for training and data from the 16th day of each month of 2007 are used for validation. The three dotted lines are 5, 10, and 15 % error envelope lines.



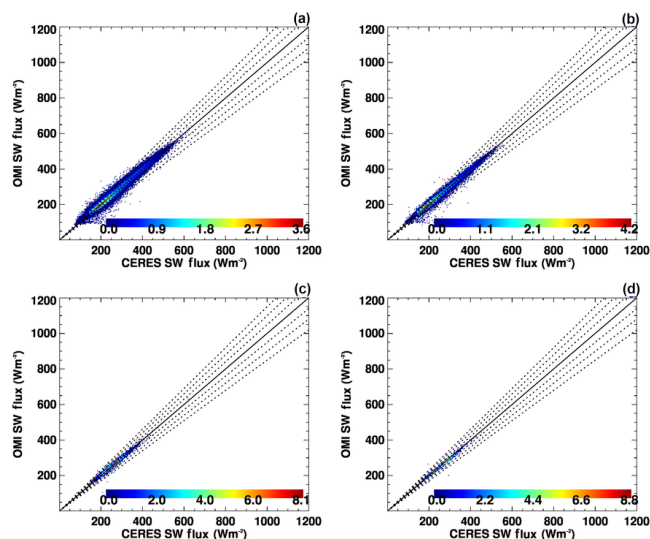
**Figure 7.** Root mean square errors (RMSE), normalized RMSE (NRMSE in percent), data samples (percent), and bias (percent) in training and validation data sets (same model and data as used in Fig. 6) as a function of effective cloud fraction for the data presented in Fig. 4. Model “h” is used to produce these results.

timized model. Figure 7 presents the RMSE, RMSE normalized by CERES flux (NRMSE in percent), data sample (percent), and bias (percent) for 5 % ECF bins. This analysis includes both training and validation data as pre-



**Figure 8.** Two-dimensional histograms (similar to Fig. 5) of the daily OMI and CERES TOA SW flux averaged over different spatial grid sizes for July 2007: (a) at OMI’s native pixel resolution, (b)  $0.5^\circ \times 0.5^\circ$ , (c)  $1^\circ \times 1^\circ$ , (d)  $2^\circ \times 2^\circ$ , (e)  $5^\circ \times 5^\circ$ , and (f)  $10^\circ \times 10^\circ$ . The corresponding statistical parameters are listed in Table 2. Model “h” is used to produce these results. The three dotted lines are 5, 10, and 15 % error envelope lines.

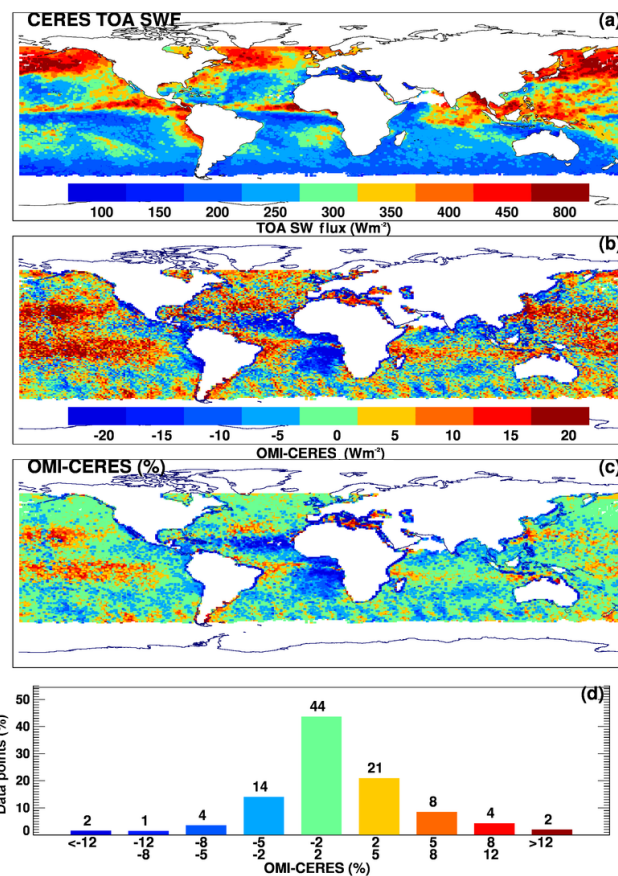
sented in Fig. 6. The RMSE varies between about 24 and  $35 \text{ Wm}^{-2}$  and continuously increases with cloud fraction (and observed TOA SWF). The NRMSE, in contrast, continuously decreases with ECF from about 18 % for 5 % ECF to  $\sim 6$  % for overcast conditions. The bias represents the mean error (in percent) for each ECF bin. The mean global bias shows more variability than RMSE and is highest (2.9 %) for about 10 % ECF. The bias decreases sharply from 2.9 % at  $f_c = 0.1$  to about 1.2 % at  $f_c = 0.4$ . The bias remains low ( $< 1.2$  %) for  $f_c > 0.4$  (usually associated with frontal or deep convective clouds). The higher biases for lower  $f_c$  (usually associated with thin cirrus and broken clouds) are likely related to higher noise and uncertainties in OMI cloud parameters. For example, Joiner et al. (2012) showed that cloud OCP errors increase with decreasing  $f_c$ . The biases may also be related to absorbing aerosol in the scene, particularly when it overlies clouds. This will be illustrated in more detail below as we show spatial variations in OMI–CERES differences.



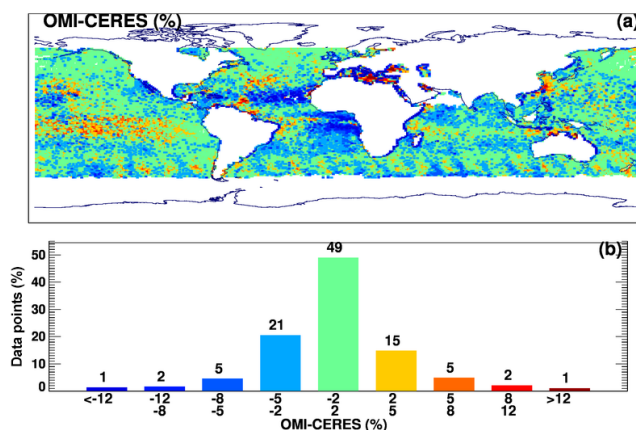
**Figure 9.** Similar to Fig. 8 but for monthly mean data (July 2007) OMI and CERES TOA SW flux averaged over different spatial grid sizes: (a)  $0.5^\circ \times 0.5^\circ$ ; (b)  $1^\circ \times 1^\circ$ ; (c)  $2^\circ \times 2^\circ$ ; (d)  $5^\circ \times 5^\circ$ . The corresponding statistical parameters are listed in Table 2. The three dotted lines are 5, 10, and 15 % error envelope lines.

## 4.2 Effects of spatial and temporal averaging

In order to evaluate the NN performance at different spatial and temporal scales similar to those used by the climate community, we use data from July 2007. Figure 8 presents a comparison of daily CERES and OMI TOA SWF over ocean for six spatial scales: the OMI native pixel ( $13 \times 24 \text{ km}^2$  at nadir) and  $0.5$ ,  $1$ ,  $2$ ,  $5$ , and  $10^\circ$  gridded spatial resolutions. Statistical parameters for these comparisons are reported in Table 2. As expected, the pixel level data are much noisier than the gridded data due to collocation noise in partly cloudy cases. This collocation noise averages out over larger spatial and temporal scales. Regardless of the noise, the slope (0.96) is still close to 1, and the linear correlation coefficient is 0.96 with a standard deviation of  $47.7 \text{ Wm}^{-2}$ . Below  $300 \text{ Wm}^{-2}$ , where the sample density is highest, the NN slightly underestimates the CERES SWF. The mean bias of the OMI-estimated SWF with respect to CERES is  $-1.4 \text{ Wm}^{-2}$ . This bias may be due to a combination of effects including uncertainties in the input parameters as well as the limitations of the NN model itself. For example, we have excluded pixels with a clear signature of absorbing aerosols (OMI-derived  $\text{AI} > 1$ ) where OMI ECFs and pressures may be in error in both the training and validation data. However, in some regions where smoke and dust overlaying clouds is common (e.g., western coast of Africa), pixels with erroneous cloud fractions due to small amounts of absorbing aerosol may be present in both the training and validation data. This may produce errors in the NN model and will be examined in more detail below.



**Figure 10.** (a) Monthly mean (July 2007) spatial distribution of TOA SW flux from CERES; (b) OMI–CERES (in  $\text{Wm}^{-2}$ ); (c) OMI–CERES (percent); (d) histogram of OMI–CERES (percent). The colors in (c) correspond to histogram colors in (d).



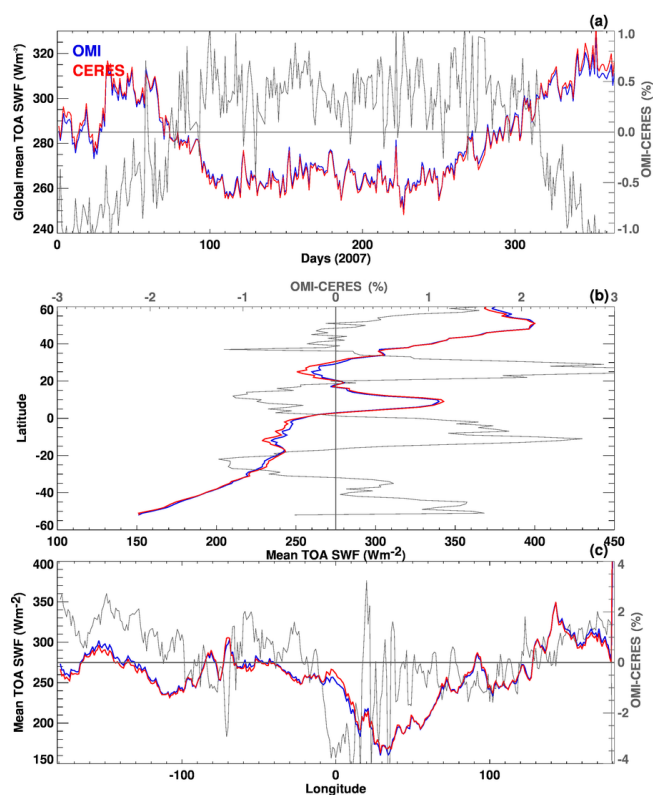
**Figure 11.** Similar to Fig. 10c and d but with a NN trained using data from the OMI cloud  $\text{O}_2\text{--O}_2$  product for July 2007.

For the daily data, as the spatial averaging scales increase from  $0.5$  to  $10^\circ$ , the OMI-estimated SWF becomes almost identical to CERES; the correlation coefficient in-

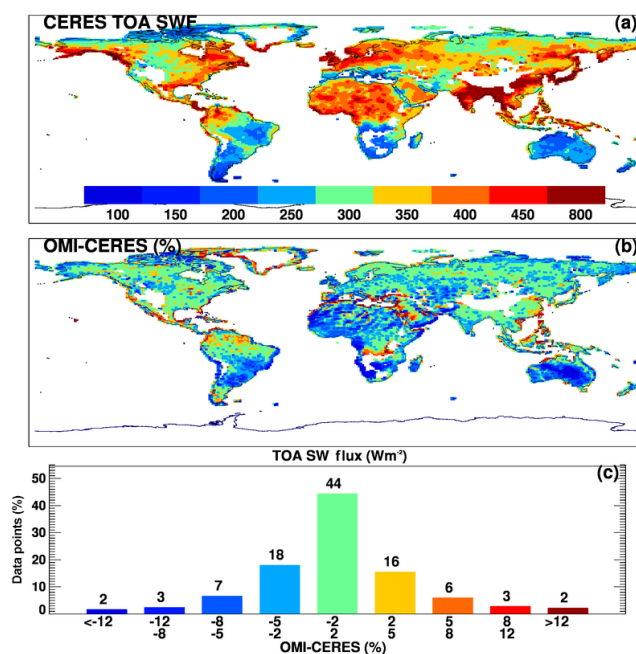
**Table 2.** Statistical parameters corresponding daily and monthly intercomparisons of pixel and gridded TOA SW flux data from CERES and OMI.

	<i>N</i>	<i>R</i>	<i>M</i>	<i>I</i>	Bias	SD	EE5 %	EE10 %	EE15 %
Pixel	8 109 323	0.96	0.96	10.5	−1.4	47.7	30	53	69
Daily									
0.5°	1 512 726	0.97	0.96	11.2	0.94	34.4	37	62	77
1°	529 679	0.98	0.96	11.2	1.0	27.9	43	69	83
2°	168 181	0.98	0.96	11.0	0.33	23.7	50	76	87
5°	35 454	0.99	0.96	9.2	−1.8	20.3	60	84	92
10°	10 834	0.99	0.97	7.0	−0.0	14.6	69	90	97
Monthly									
0.5°	108 620	0.99	0.98	6.1	1.5	12.9	74	93	97
1°	28 849	0.99	0.98	6.9	1.2	11.4	79	94	98
2°	7642	0.99	0.96	9.8	0.25	6.6	94	99	100
5°	1325	0.99	0.96	9.9	−1.8	7.0	95	99	100

Note: *N* is number of pairs, *R* is the correlation coefficient, *M* the slope, *I* is the intercept, bias is the mean of (OMI−CERES in  $\text{Wm}^{-2}$ ), SD is standard deviation of (OMI−CERES) in  $\text{Wm}^{-2}$ , EE is the error envelope for 5, 10, and 15 % errors. All flux values have units of  $\text{Wm}^{-2}$ .



**Figure 12.** (a) The daily global mean time series of TOA SW fluxes from the OMI NN and CERES and OMI−CERES (percent) on the secondary y axis. (b) Mean TOA SWF from CERES and OMI averaged along each latitude belt and OMI−CERES (percent) on the secondary x axis; (c) same as (b) but along longitude belts. The data used in (b) and (c) are from July 2007.



**Figure 13.** Similar to Fig. 10a, c, and d except over land for July 2007.

creases from 0.97 to 0.99, and the slope increases from 0.96 to 0.97. The percentage of OMI data that fall within  $\pm 5\%$  of CERES increases from 37 % for  $0.5^\circ$  to 69 % for  $10^\circ$  grids. About 87 % percent of OMI-estimated  $2^\circ$  gridded daily mean TOA SWFs are within 15 % of CERES data.

Figure 9a–d show 2-D histograms of monthly mean gridded data over ocean at  $0.5^\circ$ ,  $1^\circ$ ,  $2^\circ$ , and  $5^\circ$  spatial resolutions, respectively. The monthly intercomparisons of OMI

and CERES SWF show excellent agreement at all spatial resolutions with correlation coefficients of 0.99 and slopes of 0.98 (Table 2). The global mean biases vary between  $-1.8$  and  $0.25 \text{ Wm}^{-2}$ . The standard deviations vary between  $6.6$  and  $12.9 \text{ Wm}^{-2}$  for the different spatial resolutions. Of monthly mean  $1^\circ$  gridded OMI estimated TOA SWFs, 97 % are within 15 % of those derived from CERES and 93 % are within 10 %.

#### 4.3 Spatial distribution of TOA SWFs over ocean

Figure 10 presents the spatial distribution of  $1^\circ$  gridded monthly mean (July 2007) TOA SWF from CERES (Fig. 10a) and the difference with the OMI in  $\text{Wm}^{-2}$  (Fig. 10b) and percent difference (Fig. 10c). There are subtle differences between the NN and CERES estimates of TOA SWF as shown in Fig. 10b and c. The OMI minus CERES histograms (Fig. 10d) show that for 44 % (79 %) samples, NN fluxes are within  $\pm 2$  ( $\pm 5$ ) % of CERES fluxes. About 9 % of the samples have biases of  $\pm 8$  % or more. Overall, the Northern Hemisphere shows better agreement than the Southern Hemisphere during July (boreal summer). This could be due to larger errors in the OMI cloud products at higher SZAs in the Southern Hemisphere. The low biases on the western coast of Africa may be due to the presence of absorbing aerosols, particularly when they occur over clouds. The striped pattern in the Southern Hemisphere (latitudes  $> 40^\circ \text{ S}$ ) is mainly associated with high VZAs in conjunction with high SZAs that occur on one side of the OMI swath.

Figure 11 similarly shows differences between CERES and OMI TOA SWF over ocean derived using  $f_c$  and OCP from the OMI  $\text{O}_2\text{--O}_2$  product in place of the OMI RRS product. Because there are slight differences in the two cloud products, we retrained the network with OMI  $\text{O}_2\text{--O}_2$  cloud parameters to be consistent. The use of the  $\text{O}_2\text{--O}_2$   $f_c$  and OCP improves the accuracy of the estimated TOA SWF. The regions of improvement include the western coasts of South America and southern Africa and some parts of the Indian Ocean. The  $\text{O}_2\text{--O}_2$  cloud product, which uses visible wavelengths, is less affected by absorbing aerosol; this may explain the improvement in these areas where absorbing aerosol, especially over clouds, is common. However, negative biases remain over large regions off the western coast of Africa.

Figure 12a shows a time series of daily global mean values of TOA SWF over ocean from OMI and CERES for 2007. Both instruments show almost identical daily variations with differences within  $\pm 1$  %. Figure 12b and c provide monthly averaged (July 2007) CERES and OMI zonal and meridional means of TOA SWF. The OMI-derived TOA SWF is able to reproduce the variability shown in the CERES data well.

#### 4.4 Spatial distribution of TOA SWFs over land

We developed a similar land-only NN model that utilizes most of the input parameters from our ocean NN (e.g., OMI RRS cloud parameters). The only change is that for surface characterization we use a monthly climatology of surface broadband albedo in place of the Chl concentration and surface wind speed. The albedo product is derived using a combination of CERES and MODIS observations at  $1^\circ$  spatial resolution (Rutan et al., 2009). The land NN model is trained and validated over land only and it follows a similar approach as was used over ocean but with a slightly different set of input parameters. We have not performed extended analyses and testing over land as we have over ocean. Here, we demonstrate that the same methodology with different input parameters can reproduce CERES TOA SWF over land with similar accuracy and precision as was obtained over ocean.

Figure 13 shows results from the OMI-derived CERES-trained NN that produces TOA SWF over land. Statistical comparison with CERES over land provides results similar to those over ocean. The NN performs well over Asia and parts of Europe and the Americas. The OMI-based NN tends to underestimate TOA SWF over the high albedo desert areas of Northern Africa, Australia, and also over some regions of South America. Note that the large differences that occur in coastal regions may be due to imperfect collocations.

### 5 Summary and conclusions

We have developed a neural network approach to estimate TOA SWF based primarily on UV parameters retrieved with the Aura OMI and Aqua CERES-derived TOA SWF used for training. One year of data from OMI and CERES has been used to train/validate/analyze several separate neural networks for different conditions, which together provide estimation of TOA SWF under all-sky conditions. The most important input parameters are ECF and sun-satellite geometry. TCO and cloud optical centroid pressure from OMI, as well as surface-related parameters, provide secondary positive impacts.

Independent validation at different spatial and temporal scales shows that the OMI NN-based approach reproduces CERES-derived TOA SWF with high fidelity. Correlation coefficients for all comparison are  $> 0.95$ , and slopes are close to unity. A high percentage of OMI-estimated monthly mean TOA SWF at  $0.5^\circ$  spatial resolution over global oceans (97 %) falls within 15 % of CERES. The global mean bias in pixel level data of about  $-1.4 \text{ Wm}^{-2}$  over oceans with respect to CERES is likely due in part to errors in OMI cloud parameters that occur in the presence of absorbing aerosols.

We plan to apply our derived neural networks to long-term, well-calibrated UV measurements from TOMS. The TOMS series provides a long-term data record dating back to late 1978 (about half a decade before the first ERBE launch)

with a few small gaps between that time and the first CERES launch. We should be able to apply NN models derived with CERES/OMI to TOMS, provided that the input parameters are either available and compatible or can be estimated independently. For example, in place of actual cloud OCPs that are available from OMI, but not from TOMS, we could use a cloud OCP climatology that was developed from OMI data for use in the TOMS total ozone algorithm. The lower spatial resolution of TOMS is not expected to present any difficulties. This approach can also be extended to the future geostationary missions that provide the relevant input data, such as the NASA Earth Ventures Tropospheric Emissions: Monitoring of Pollution (TEMPO), the Korean Geostationary Environmental Monitoring Spectrometer (GEMS), and the European Space Agency (ESA) Sentinel 4 (Al-Saadi et al., 2015). Finally, we may apply the NN training and evaluation approach to data from CERES and the nadir mapper on the Ozone Mapping Profiling Suite (OMPS) that provides information similar to OMI. Both instruments fly on the Suomi NPP satellite. This may reduce collocation noise and small biases that result from the time difference between OMI and CERES measurements. The final NN models developed in this study (e.g., NNM-1 and NNM-2 in Fig. 3) along with the instructions on how to use them have been provided in the Supplement.

**The Supplement related to this article is available online at doi:10.5194/amt-9-2813-2016-supplement.**

**Acknowledgements.** This material is based upon work supported by the National Aeronautics and Space Administration (NASA) issued through the Science Mission Directorate (SMD) for the Aura Science Team managed by Kenneth Jucks and Richard Eckman. We thank the CERES, OMI, MODIS, and GEOS-DAS data processing teams for providing the data used for this study. We would also like to thank Norman Loeb and Arlindo da Silva for useful discussion and comments during the preparation of the paper.

Edited by: V. Sofieva

## References

- Acarreta, J. R., De Haan, J. F., and Stammes, P.: Cloud pressure retrieval using the O<sub>2</sub>-O<sub>2</sub> absorption band at 477 nm, *J. Geophys. Res.*, 109, D05204, doi:10.1029/2003JD003915, 2004.
- Ahmad, Z., Bhartia, P. K., and Krotkov, N.: Spectral properties of backscattered UV radiation in cloudy atmospheres, *J. Geophys. Res.*, 109, D01201, doi:10.1029/2003JD003395, 2004.
- Al-Saadi, J., Carmichael, G., Crawford, J., Emmons, L., Song, C.-K., Chang, L.-S., Lee, G., Kim, J., Park, R.: NASA Contributions to KORUS-AQ: An International Cooperative Air Quality Field Study in Korea, NASA White Paper available at: <https://goo.gl/VhssdX> (last access: 3 May 2016), 2015.
- Barkstrom, B. R.: The Earth radiation budget experiment (ERBE), *B. Am. Meteorol. Soc.*, 65, 1110–1185, 1984.
- Barkstrom, B. R. and Smith, G. L.: The Earth radiation budget experiment: Science and implementation, *Rev. Geophys.*, 24, 379–390, 1986.
- Baum, B. A., Yang, P., Heymsfield, A. J., Bansemer, A., Merrelli, A., Schmitt, C., and Wang, C.: Ice cloud bulk single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100  $\mu$ m, *J. Quant. Spectrosc. Ra.*, 146, 123–139, doi:10.1016/j.jqsrt.2014.02.029, 2014.
- Bellouin, B., Boucher, O., Haywood, J., and Reddy, M. S.: Global estimates of aerosol direct radiative forcing from satellite measurements, *Nature*, 438, 1138–1140, doi:10.1038/nature04348, 2005.
- Bishop, M.: Neural networks for pattern recognition, Oxford University Press, Inc., New York, ISBN-13: 978-0198538646, 1995.
- Bloom, S., da Silva, A., and Dee, D.: Technical Report Series on Global Modeling and Data Assimilation, edited by: Suarez, M. J., Documentation and Validation of the Goddard Earth Observing System (GEOS) Data Assimilation System–Version 4, NASA GSFC NASA/TM – 2005–104606, 26, available at: <http://gmao.gsfc.nasa.gov/systems/geos4/Bloom.pdf> (last access: 24 June 2016), 2005.
- Chevallier, F., Cheruy, F., Scott, N. A., and Chedin, A.: A neural network approach for a fast and accurate computation of a long-wave radiative budget, *J. Appl. Meteorol.*, 37, 1385–1397, 1998.
- Deirmendjian, D.: Electromagnetic scattering on spherical polydispersions, Elsevier Sci., New York, 290 pp, 1969.
- Dines, W. H.: The heat balance of the atmosphere, *Q. J. Roy. Meteor. Soc.*, 43, 151–158, 1917.
- Domenech, C. and Wehr, T.: Use of Artificial Neural Networks to Retrieve TOA SW Radiative Fluxes for the EarthCARE Mission, *IEEE Trans. Geosci. Remote S.*, 49, 1839–1849, 2011.
- Gardner, M. W. and Dorling, S. R.: Artificial neural networks: A review of applications in the atmospheric sciences, *Atmos. Environ.*, 32, 2627–2636, doi:10.1016/S1352-2310(97)00447-0, 1998.
- Gupta, P. and Christopher, S. A.: Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: 2. A neural network approach, *J. Geophys. Res.*, 114, D20205, doi:10.1029/2008JD011497, 2009.
- Harrison, E. F., Minnis, P., Barkstrom, B. R., Ramanathan, V., Cess, R. D., and Gibson, G. G.: Seasonal variation of cloud radiative forcing derived from the Earth Radiation Budget Experiment, *J. Geophys. Res.*, 95, 18687–18703, 1990.
- Hartmann, D. L., Ramanathan, V., Berroir, A., and Hunt, G. E.: Earth radiation budget data and climate research, *Rev. Geophys.*, 24, 439–468, 1986.
- Hatzianastassiou, N., Fotiadis, A., Matsoukas, Ch., Pavlakis, K., Drakakis, E., Hatzidimitriou, D., and Vardavas, I.: Long-term global distribution of earth's shortwave radiation budget at the top of atmosphere, *Atmos. Chem. Phys.*, 4, 1217–1235, doi:10.5194/acp-4-1217-2004, 2004.
- Hertz, J. A., Krogh, A. S., and Palmer, A.: Introduction to the Theory of Neural Computation, Addison-Wesley, Redwood City, Calif., 352 pp., ISBN-13: 9780201515602, 1991.
- Hu, Y., Zhang, H., Wielicki, B., and Stackhouse, P.: A neural network MODIS-CERES narrowband to broadband conversion,

- IEEE Geoscience and Remote Sensing Symposium, 6, 3227–32296, doi:10.1109/IGARSS.2002.1027138, 2002.
- Jiang, B., Zhang, Y., Liang, S., Zhang, X., and Xiou, Z.: Surface daytime net radiation estimate using artificial neural networks, *Remote Sens.*, 6, 11031–11050, doi:10.3390/rs6111031, 2014.
- Joiner, J. and Bhartia, P. K.: The determination of cloud pressures from rotational-Raman scattering in satellite backscatter ultraviolet measurements, *J. Geophys. Res.*, 100, 23019–23026, 1995.
- Joiner, J. and Vasilkov, A. P.: First results from the OMI Rotational Raman Scattering Cloud Pressure Algorithm, *IEEE Trans. Geosci. Remote S.*, 44, 1272–1282, 2006.
- Joiner, J., Vasilkov, A. P., Flittner, D. E., Gleason, J. F., and Bhartia, P. K.: Retrieval of cloud chlorophyll content using Raman scattering in GOME spectra, *J. Geophys. Res.*, 109, D01109, doi:10.1029/2003JD003698, 2004.
- Joiner, J., Schoeberl, M. R., Vasilkov, A. P., Oreopoulos, L., Platnick, S., Livesey, N. J., and Levelt, P. F.: Accurate satellite-derived estimates of the tropospheric ozone impact on the global radiation budget, *Atmos. Chem. Phys.*, 9, 4447–4465, doi:10.5194/acp-9-4447-2009, 2009.
- Joiner, J., Vasilkov, A. P., Gupta, P., Bhartia, P. K., Veefkind, P., Sneep, M., de Haan, J., Polonsky, I., and Spurr, R.: Fast simulators for satellite cloud optical centroid pressure retrievals; evaluation of OMI cloud retrievals, *Atmos. Meas. Tech.*, 5, 529–545, doi:10.5194/amt-5-529-2012, 2012.
- Krasnopolsky, V. M., Fox-Rabinovitz, and Belochitski, A. A.: Decadal climate simulations using accurate and fast neural network emulation of full, longwave and shortwave, radiation, *Month. Weather Rev.*, 136, 3683–3695, 2008.
- Krasnopolsky, V. M., Fox-Rabinovitz, M. S., Hou, Y. T., Lord, S. J., Belochitski, A. A.: Accurate and fast neural network emulations of model radiation for the NCEP coupled climate forecast system: climate simulations and seasonal predictions, *Month. Weather Rev.*, 138, 1822–1842, 2010.
- Kroon, M., Veefkind, J. P., Sneep, M., McPeters, R. D., Bhartia, P. K., and Levelt, P. F.: Comparing OMI-TOMS and OMI-DOAS total ozone column data, *J. Geophys. Res.*, 113, D16S28, doi:10.1029/2007JD008798, 2008.
- Lee, J., Weger, R. C., Sengupta, S. K., and Welch, R. M.: A neural network approach to cloud classification, *IEEE Trans. Geosci. Remote S.*, 28, 846–855, 1990.
- Levelt, P. F., van den Oord, G. H. J., Dobber, M. R., Mälikki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O. V., and Saari, H.: The Ozone Monitoring Instrument, *IEEE Trans. Geosci. Remote S.*, 44, 1093–1101, 2006.
- Loeb, N. G. and Manalo-Smith, N.: Top-of-Atmosphere direct radiative effect of aerosols over global oceans from merged CERES and MODIS observations, *J. Climate*, 18, 3506–3526, 2005.
- Loeb, N. G., Manalo-Smith, N., Kato, S., Miller, W. F., Gupta, S. K., Minnis, P., and Wielicki, B. A.: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Tropical Rainfall Measuring Mission Satellite, Part I: Methodology, *J. Appl. Meteorol.*, 42, 240–265, 2003.
- Loeb, N. G., Kato, S., Loukachine, K., and Manalo-Smith, N.: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Terra Satellite. Part I: Methodology, *J. Atmos. Ocean. Tech.*, 22, 338–351, doi:10.1175/JTECH1712.1, 2005.
- Loeb, N. G., Kato, S., Loukachine, K., Manalo-Smith, N., and Doelling, D. R.: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Terra satellite. Part II: Validation, *J. Atmos. Ocean. Tech.*, 24, 564–584, doi:10.1175/JTECH1983.1, 2007.
- Loeb, N. G., Kato, S., Su, W., Wong, T., Rose, F. G., Doelling, D. R., Norris, J. R., and Huang, X.: Advances in understanding top-of-atmosphere radiation variability from satellite observations, *Surv. Geophys.*, 33, 359–385, doi:10.1007/s10712-012-9175-1, 2012.
- Loukachine, K. and Loeb, N. G.: Application of an artificial neural network simulation for top-of-atmosphere radiative flux estimation from CERES, *J. Atmos. Oceanic Technol.*, 20, 1749–1757, 2003.
- Loukachine, K. and Loeb, N. G.: Top-of-atmosphere flux retrievals from CERES using artificial neural networks, *J. Remote Sens. Environ.*, 93, 381–390, 2004.
- McPeters, R. D., Kroon, M., Labow, G. J., Brinksma, E., Balis, D., Petropavlovskikh, I., Veefkind, J. P., Bhartia, P. K., and Levelt, P. F.: Validation of the Aura Ozone Monitoring Instrument Total Column Ozone Product, *J. Geophys. Res.*, 113, D15S14, doi:10.1029/2007JD008802, 2008.
- Oreopoulos, L., Platnick, S., Hong, G., Yang, P., and Cahalan, R. F.: The shortwave radiative forcing bias of liquid and ice clouds from MODIS observations, *Atmos. Chem. Phys.*, 9, 5865–5875, doi:10.5194/acp-9-5865-2009, 2009.
- Patadia, F., Gupta, P., and Christopher, S. A.: First observational estimates of global clear sky shortwave aerosol direct radiative effect over land, *Geophys. Res. Lett.*, 35, L04810, doi:10.1029/2007GL032314, 2008.
- Patadia, F., Christopher, S. A., and Zhang, J.: Development of empirical angular distribution models for smoke aerosols: Methods, *J. Geophys. Res.*, 116, 1984–2012, 2011.
- Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., and Hartmann, D.: Cloud radiative forcing and climate: Results from the Earth Radiation Budget Experiment, *Science*, 243, 57–63, 1989b.
- Rutan, D., Rose, F., Roman, M., Manalo-Smith, N., Schaaf, C., and Charlock, T.: Development and assessment of broadband surface albedo from Clouds and the Earth's Radiant Energy System Clouds and Radiation Swath data product, *J. Geophys. Res.*, 114, D08125, doi:10.1029/2008JD010669, 2009.
- Stammes, P., Sneep, M., de Haan, J. F., Veefkind, J. P., Wang, P., and Levelt, P. F.: Effective cloud fractions from the Ozone Monitoring Instrument: theoretical framework and validation, *J. Geophys. Res.*, 113, D16S38, doi:10.1029/2007JD008820, 2008.
- Takenaka, H., Nakajima, T. Y., Higurashi, A., Higuchi, A., Takamura, T., Pinker, R. T., and Nakajima, T.: Estimation of solar radiation using a neural network based on radiative transfer, *J. Geophys. Res.*, 116, D08215, doi:10.1029/2009JD013337, 2011.
- van de Hulst, H. C. and Irvine, W. M.: General report on radiation transfer in planets: Scattering in model planetary atmospheres, *Mem. Soc. R. Sci. Liege*, 7, 78–98, 1963.
- Vasilkov, A. P., Joiner, J., Spurr, R., Bhartia, P. K., Levelt, P. F., and Stephens, G.: Evaluation of the OMI cloud pressures derived from rotational Raman scattering by comparisons with other satellite data and radiative transfer simulations, *J. Geophys. Res.*, 113, D15S19, doi:10.1029/2007JD008689, 2008.

- Vázquez-Navarro, M., Mayer, B., and Mannstein, H.: A fast method for the retrieval of integrated longwave and shortwave top-of-atmosphere upwelling irradiances from MSG/SEVIRI (RRUMS), *Atmos. Meas. Tech.*, 6, 2627–2640, doi:10.5194/amt-6-2627-2013, 2013.
- Veefkind, J. P., de Haan, J. F., Brinksma, E. J., Kroon, M., and Levelt, P. F.: Total ozone from the ozone monitoring instrument (OMI) using the DOAS technique, *IEEE Trans. Geosci. Remote S.*, 44, 1239–1244, 2006.
- Wielicki, B. A., Harrison, E. F., Cess, R. D., King, M. D., and Randall, D. A.: Mission to planet Earth: Role of clouds and radiation in climate, *Bull. Amer. Meteorol. Soc.*, 76, 2125–2153, 1995.
- Yu, H., Kaufman, Y. J., Chin, M., Feingold, G., Remer, L. A., Anderson, T. L., Balkanski, Y., Bellouin, N., Boucher, O., Christopher, S., DeCola, P., Kahn, R., Koch, D., Loeb, N., Reddy, M. S., Schulz, M., Takemura, T., and Zhou, M.: A review of measurement-based assessments of the aerosol direct radiative effect and forcing, *Atmos. Chem. Phys.*, 6, 613–666, doi:10.5194/acp-6-613-2006, 2006.
- Zhang, J., Christopher, S. A., Remer, L. A., and Kaufman, Y. J.: Shortwave aerosol radiative forcing over cloud-free oceans from Terra. I: Angular models for aerosols, *J. Geophys. Res.*, 110, D10S23, doi:10.1029/2004JD005008, 2005a.
- Zhang, J., Christopher, S. A., Remer, L. A., and Kaufman, Y. J.: Shortwave aerosol radiative forcing over cloud-free oceans from Terra. II: Seasonal and global distributions, *J. Geophys. Res.*, 110, D10S24, doi:10.1029/2004JD005009, 2005b.