A 3D cloud-construction algorithm for the EarthCARE satellite mission

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1002/qj.824

Publisher: Royal Meteorological Society

www.reading.ac.uk/centaur

CentAUR

RMetS

Royal Meteorological Society

A 3D cloud-construction algorithm for the EarthCARE satellite mission

H. W. Barker,[a][*][†] M. P. Jerg,[a] T. Wehr,[b] S. Kato,[c] D. P. Donovan[d] and R. J. Hogan[e]

[a]*Environment Canada, Toronto, ON, Canada*
[b]*European Space Agency, Noordwijk, the Netherlands*
[c]*NASA–Langley Research Center, Hampton, VA, USA*
[d]*KNMI, De Bilt, the Netherlands*
[e]*University of Reading, Reading, UK*

[*]Correspondence to: H. W. Barker, Environment Canada, Cloud Physics and Severe Weather Research Section (ARMP), 4905 Dufferin St., Toronto, ON, Canada M3H 5T4. E-mail: howard.barker@ec.gc.ca
[†]The contribution of H. W. Barker to this article was prepared as part of his official duties as a Canadian government employee

This article presents and assesses an algorithm that constructs 3D distributions of cloud from passive satellite imagery and collocated 2D nadir profiles of cloud properties inferred synergistically from lidar, cloud radar and imager data. It effectively widens the active–passive retrieved cross-section (RXS) of cloud properties, thereby enabling computation of radiative fluxes and radiances that can be compared with measured values in an attempt to perform radiative closure experiments that aim to assess the RXS. For this introductory study, A-train data were used to verify the scene-construction algorithm and only 1D radiative transfer calculations were performed.

The construction algorithm fills off-RXS *recipient* pixels by computing sums of squared differences (a cost function F) between their spectral radiances and those of potential *donor* pixels/columns on the RXS. Of the RXS pixels with F lower than a certain value, the one with the smallest Euclidean distance to the recipient pixel is designated as the donor, and its retrieved cloud properties and other attributes such as 1D radiative heating rates are consigned to the recipient. It is shown that both the RXS itself and Moderate Resolution Imaging Spectroradiometer (MODIS) imagery can be reconstructed extremely well using just visible and thermal infrared channels. Suitable donors usually lie within 10 km of the recipient. RXSs and their associated radiative heating profiles are reconstructed best for extensive planar clouds and less reliably for broken convective clouds.

Domain-average 1D broadband radiative fluxes at the top of the atmosphere (TOA) for $(21\,\text{km})^2$ domains constructed from MODIS, CloudSat and Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) data agree well with coincidental values derived from Clouds and the Earth's Radiant Energy System (CERES) radiances: differences between modelled and measured reflected shortwave fluxes are within $\pm 10\,\text{W m}^{-2}$ for $\sim$35% of the several hundred domains constructed for eight orbits. Correspondingly, for outgoing longwave radiation $\sim$65% are within $\pm 10\,\text{W m}^{-2}$. Copyright © 2011 Royal Meteorological Society and Crown in the right of Canada.

1. Introduction

In late 1993, at a World Climate Research Programme (WCRP) workshop, retrieval of cloud properties from satellite data was recognized as essential for the development and assessment of cloud and associated radiation parametrizations as used by numerical weather prediction and climate models. At that time, however, data available for doing such retrievals came entirely from passive instruments and were demonstrably bound to yield just vertically integrated or 'effective' values of cloud properties via inversion of 1D solutions of the radiative transfer equation. The key recommendation made at the workshop (WCRP, 1994) was that

> ... a combination of active [lidar and radar] systems flown on satellites, [should] make significant progress on determining the four-dimensional distribution of cloud optical properties and the relationships between these properties and cloud liquid water, ice mass, and water vapor.

The seriousness of this statement was underscored by the fact that clouds had by then been singled out as important components of Earth's climate system and responsible for much uncertainty in estimates of climate sensitivity (Ramanathan, 1987; Cess *et al.*, 1992). Eventually, the cloud-profiling radar (CPR) ended up on CloudSat (Stephens *et al.*, 2002) and lidar on Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO; Winker *et al.*, 2007). These satellites, together with others, most notably Aqua with its passive Moderate Resolution Imaging Spectroradiometer (MODIS) imager and Clouds and the Earth's Radiant Energy System (CERES) broadband radiometer (Wielicki *et al.*, 1996), make up the so-called A-train of formation-flying satellites.

In the wake of WCRP-1994, plans for active–passive sensors on satellites now dominate the cloud–aerosol remote sensing subfield of atmospheric science. Most notable are the Earth, Clouds, Aerosols and Radiation Explorer (EarthCARE) Mission and the Aerosol–Cloud–Ecosystems (ACE) experiment. EarthCARE, a collaboration between the European Space Agency (ESA) and the Japan Aerospace Exploration Agency (JAXA), is the third ESA Explorer Core Mission (ESA, 2001). It is expected to be launched in 2015 into a near-sun-synchronous orbit with an equator-crossing time (ascending) between 1345 and 1400 LST at a mean altitude of $\sim$400 km. Observations will be made from a single satellite equipped with a 355 nm high-spectral-resolution lidar (HSRL), a 94 GHz Dopplerized CPR, a seven-channel Multi-Spectral Imager (MSI) and a broadband radiometer (BBR) with three along-track views. The active instruments will be near-nadir pointing while the passive instruments will view narrow swaths about the nadir. Retrievals of cloud and aerosol property profiles will be obtained through synergistic combination of data. The NASA Decadal Survey-based ACE is proposed to fly around 2020 (see http://dsm.gsfc.nasa.gov/ace/index.html).

The ultimate goal of EarthCARE is to retrieve cloud and aerosol properties to sufficient accuracy that, when used to initialize multilayer radiative transfer models, estimated top-of-atmosphere (TOA) radiative fluxes will be, *on average*, within ± 10 W m^{-2} of fluxes inferred from BBR radiances for each (10×10) km^2 to (10×50) km^2 column of atmosphere centred along the nadir active–passive retrieved cross-section (RXS) (ESA, 2001). Verification of this goal requires carrying out a continuous radiative closure experiment aimed at assessing the quality of retrievals using BBR data that will not be used in the retrieval process.

Such an experiment requires that broadband radiative transfer calculations be performed on the retrieved profiles, but the RXS is $\lesssim$1 km wide and BBR footprints are usually much larger (e.g. CERES at 20 km and the NASA Earth Radiation Budget Experiment (ERBE) Earth Radiation Budget Satellite (ERBS) at 40 km). While EarthCARE's BBR footprint is $\sim$1 km wide, before data from it can become usable it is likely that they will have to be integrated to at least a 5 km wide swath due to the nature of the BBR's point-spread function and signal-to-noise ratio. Hence, to perform proper multilayer/multidimensional radiation calculations requires 3D domains of cloud, aerosol and surface properties around the RXS that cover at least the BBR swath.

The focus of this study is development and assessment of a 3D cloud-construction algorithm that uses as input RXS data and conventional passive imagery. The construction algorithm is a more refined version of an earlier procedure that was implimented within an EarthCARE simulator environment (Donovan *et al.*, 2008). Merged data from CloudSat, CALIPSO, MODIS and CERES (Kato *et al.*, 2010) are used here to perform assessments. For this introductory study it suffices to limit radiation calculations to 1D independent column approximations (ICAs). In a complimentary study, 3D transfer models are applied and recommended for use by EarthCARE in what would be the first *operational* use of 3D radiative transfer models in the atmospheric sciences.

The following section discusses the generality of the problem at hand and presents the cloud-construction algorithm. This is followed by descriptions of the merged A-train dataset and the 1D radiative transfer models used to compute TOA fluxes. The fifth section documents results of several tests. The final section offers a summary and recommendations for future work.

2. 3D cloud-field construction algorithm

2.1. General remarks

Similarly to the A-train, EarthCARE will synergistically retrieve, from active and passive sensor data, a near-continuous 2D profile of cloud and aerosol properties along the satellite ground-track, referred to hereinafter as the RXS. Independent verification of the RXS is uncertain, as is computation of accurate radiative flux profiles for the RXS. The plan presented here advocates addressing these problems simultaneously via *construction* of 3D atmosphere–surface domains around the RXS. Once constructed, the 3D scene will serve as input to either 1D radiative transfer (RT) models in independent column approximation (ICA) mode (Stephens *et al.*, 1991) or 3D RT models (see several chapters in Marshak and Davis, 2005a). Either model would produce TOA fluxes or radiances as well as flux profiles for the constructed domain or subsets of it. At this stage, the focus is on the construction of clouds as they are the most difficult aspect of the problem.

Before explaining the algorithm, here is a list of datasets that are assumed to be available:

(A) a series of profiles (i.e. the nadir RXS) of cloud and aerosol properties retrieved from either active instruments alone or in synergy with data from a passive multispectral imager (MSI);

(B) MSI data at wavelengths typical of conventional imagers at resolutions ideally less than the coarsest active instrument and extending in the across-track direction on both sides of the RXS and wider than the BBR's swath;

(C) profiles of pressure, temperature, humidity and ozone along the RXS;

(D) spectral optical properties of Earth's surface at the resolution of the passive imager.

All these datasets are, or will be, available within the A-train and EarthCARE, and presumably ACE, experiments.

The scene construction methodology advocated here is a match-and-substitute algorithm. That is, MSI radiances for an off-nadir, *recipient*, pixel are compared with corresponding values for a range of pixels along the RXS; close matches are identified as potential *donors* with the closest to the recipient pixel being designated as the proxy to literally stand in for the recipient. This is repeated until all pixels in the required 3D domain are filled.

The essence of the proposed 3D cloud-construction algorithm rests on the following hypothesis:

> If two cloudy pixels in *close proximity* to one another, having (almost) identical temperature and moisture profiles as well as surface optical properties, exhibit arbitrarily small differences for an arbitrarily large number of spectral TOA radiances, differences in cloud-property profiles associated with the pixels are arbitrarily small as well.

In practice, the number of MSI channels is often not large; EarthCARE will have just seven. This, coupled with the fact that the best available estimate of the vertical structure of cloud is the narrow active–passive RXS (hence an active–passive mission's raison d'être), the operational version of the hypothesis is:

> If MSI radiances for an off-nadir pixel and a pixel along the RXS are sufficiently similar, the column associated with the RXS pixel can be considered to be a proxy for the off-nadir column, thereby leading to the construction of a 3D domain of cloud (and aerosol and surface too).

The following subsections explain the algorithm and discuss some issues surrounding its usage, including the application of 1D RT models to constructed domains.

2.2. Identification of donor columns

Let all pixels/columns along the RXS be at positions $(i, 0)$ where $i = 1$ is the most southerly point. As the satellite tracks north and then south back to the most southerly point, i increases. The intention is to broaden the RXS by filling off-nadir columns at (i, j) for all i and $j \in [-J, -1] \cup [1, J]$. To fill a column at (i, j) that has MSI radiances $r_k(i, j)$, where k is spectral interval, with an RXS column that has similar

radiances, begin by computing

$$F(i, j; m) = \sum_{k=1}^{K} w_k \left[\frac{r_k(i,j) - r_k(m, 0)}{r_k(i,j)} \right]^2$$
$$: m \in [i - m_1, i + m_2], \qquad (1)$$

where the summation is over K spectral radiances. Weights w_k can vary but were all set to unity, as no obvious benefits were observed otherwise. Only pixels along the RXS that satisfy the following conditions and are fewer than m_1 pixels behind and m_2 pixels in front of $(i, 0)$ are considered as potential donors:

(1) surface types at $(m, 0)$ and (i, j) must be the same;

(2) depending on availability of suitable columns, the column at $(m, 0)$ should ideally have sufficiently small uncertainties associated with key retrieved variables, such as cloud mask, water content, particle phase and particle size;

(3) absolute value of the difference between the cosine of solar zenith angles μ_0 at $(m, 0)$ and (i, j) must be $\leq \Delta\mu_0$;

(4) absolute value of the difference between solar azimuth angles relative to the satellite's ground track at $(m, 0)$ and (i, j) must be $\leq \Delta\varphi_0$.

For the first condition, only open water and land were considered in this study. Refinements that address land-surface types, including snow and ice, are straightforward to include. Moreover, this condition could be relaxed if clouds are optically very thick. The second condition is difficult to describe in detail at this stage. While active–passive synergistic algorithms for ice clouds, with rigorous treatment of errors, have been applied to A-Train data (Delanoë and Hogan, 2010), exploration of the second condition will need to wait until this is extended to liquid clouds and precipitation. The last two conditions ensure similar solar illumination.

Upon computing $F(i, j; m)$, one could require that the donor pixel be defined simply by

$$\underset{m \in [i - m_1, i + m_2]}{\arg\min} F(i, j; m), \qquad (2)$$

with conditions (3) and (4) listed above setting m_1 and m_2. If, however, m_1 and m_2 are large, this could identify a terribly inappropriate profile, the spectral radiances of which just happened to resemble the recipient's. In fact, using $\Delta\mu_0 = 0.01$ and $\Delta\varphi_0 = 10°$ leads to total search lengths $(m_1 + m_2 + 1)$ in excess of 1000 km near the peak values of μ_0.

While Eq. (1) is a commonly used cost function, the fact that each term in square brackets is a squared difference of a variable separated by regular increments implies that on average $F(i, j; m)$ should resemble the second-order structure function, defined as

$$S_2(\Delta x; r_k) = \left\langle [r_k(x) - r_k(x + \Delta x)]^2 \right\rangle, \qquad (3)$$

where x is position, Δx is displacement and angular brackets indicate averaging over all x in a defined range. It is well known that, for cloud-related variables,

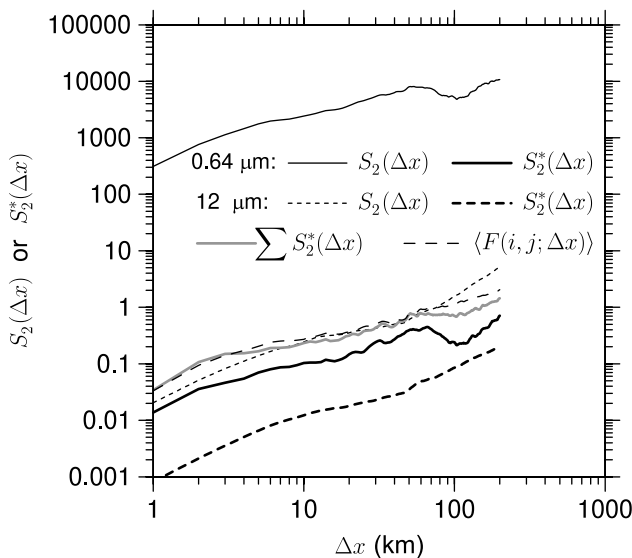$$S_2(\Delta x; r_k) \sim \Delta x^{\zeta(2)} \qquad (4)$$

Figure 1. Second-order structure functions $S_2(\Delta x)$ for 0.64 and 12 μm radiances for a 400 km stretch of the RXS for an orbit on 25 January 2007 (CloudSat granule 3960). $S_2^*(\Delta x)$ are related to $S_2(\Delta x)$ and defined in Eq. (5). The curve labelled $\sum S_2^*(\Delta x)$ is as $S_2^*(\Delta x)$ but each term in the ensemble is a sum of contributions from 0.64 and 12 μm radiances[see Eq. (6)]. The curve labelled $\langle F(i, j; \Delta x) \rangle$ is the ensemble average of the cost function defined in Eq. (1), where m has been replaced by Δx.

is followed fairly well where $\zeta(2) \in [0, 2]$ but is typically near 0.5 for Δx often up to 10 km or more (Monin and Yaglom, 1975; Nastrom and Gage, 1985; Davis *et al.*, 1997; Marshak and Davis, 2005b). This indicates that cloud-related variables tend to persist.

Equation (3) can be made to resemble Eq. (1) by normalizing differences as

$$S_2^*(\Delta x; r_k) = \left\langle \left[\frac{r_k(x) - r_k(x + \Delta x)}{r_k(x)} \right]^2 \right\rangle : \; r_k(x) > 0. \tag{5}$$

In general, S_2^* and S_2 scale almost identically with Δx. Figure 1 shows a typical example. It also shows that

$$\left\langle \sum_{k=1}^{K} \left[\frac{r_k(x) - r_k(x + \Delta x)}{r_k(x)} \right]^2 \right\rangle \simeq \langle F(i, j; \Delta x) \rangle$$
$$\sim \Delta x^{\zeta(2)}, \tag{6}$$

where $\langle F(i, j; \Delta x) \rangle$ implies averaging over all i in a defined range with Δx being the distance between donor and recipient. Therefore, based on the behaviour of cloud variables from structure function analyses, it is desirable that donor and recipient pixels have small $F(i, j; m)$, as noted in Eq. (2), and be as close as possible, thereby tapping into the spatial persistence of cloud properties and raising the chance that the donor's profiles resemble the recipient's. A simple way to effect this is to begin by ordering $F(i, j; m)$ from smallest to largest to form the set

$$\{\min[F(i, j; m)], \dots, \max[F(i, j; m)]\}$$
$$= \{\mathcal{F}(i, j; 1), \dots,$$
$$\mathcal{F}(i, j; m_1 + m_2 + 1)\} : \mathcal{F}(i, j; n) \leqslant \mathcal{F}(i, j; n + 1). \tag{7}$$

Defining distance between a potential donor at $(m, 0)$ and the recipient at (i, j) as

$$D(i, j; m) = \Delta L \sqrt{(i - m)^2 + j^2}, \tag{8}$$

where ΔL is imager resolution, let $D(i, j; m)$ and m go passively along with the ordering of $F(i, j; m)$ so that $\{\mathcal{F}(i, j; n)\}$ has an associated co-ordered set of distances $\{\mathcal{D}(i, j; n)\}$. One then solves for the index m^* as

$$\underset{m^* \in [1, (m_1 + m_2 + 1)f]}{\arg\min} \{\mathcal{D}(i, j; m^*)\} : f \in (0, 1), \tag{9}$$

which means, 'find the index m^* that corresponds to the smallest distance between recipient and those pixels that constitute the smallest $100f\%$ of $F(i, j; m)$'. For this study $f = 0.03$, meaning that just the smallest 3% of the $(m_1 + m_2 + 1)$ values of $F(i, j; m)$ are considered. Knowing m^* leads directly to m and thus the donor column. Finally, the profiles of retrieved cloud properties at $(m, 0)$ and all other properties are simply replicated at (i, j).

Although one could resort to a cloud mask derived from MSI imagery and seek to fill only those off-nadir pixels identified as cloudy, there is no need to, as the algorithm simply searches radiances; whatever is in the donor column at $(m, 0)$ is delegated to recipient at (i, j). Figure 2 shows a schematic of this process, which is applied until the desired 3D scene is constructed. In practice this would be until an entire orbit's RXS is converted into a 3D band. Obviously, RXS columns $(j = 0)$ identify themselves, so for them there is no need to apply the algorithm; the RXS forms the centre of the constructed domain.

In general, search lengths backward and forward along the RXS, m_1 and m_2, need not be equal. For this study, however, $m_1 = m_2 = 200 (= 200 \text{ km}/\Delta L)$. As will be seen later, this is excessive as most distances between donor and recipient are less than ~30 km, which ensures that for the overwhelming majority of cases solar illumination and weather conditions are very similar for donor and recipient. Moreover, searching 200 km in both directions with $f = 0.03$ in Eq. (9) yields 12 samples from which to select the donor.[†]

To a great extent, the width of the constructed domain is determined by the problem at hand. Here it is limited by the resolution of the BBR or possibly several cross-track BBR pixels. It must be borne in mind, however, that when radiative closure experiments are of concern, the larger the constructed scene, the more one will be assessing the construction algorithm and the less the RXS. For a very small BBR footprint that matches the RXS, radiative quantities from a 3D model would be averaged only along the RXS and the constructed domain would act as a *buffer zone* to facilitate horizontal transport of radiation through the cross-track sides of the RXS. The width of suitable buffer

---

[†]A potential donor pixel that is far removed from the recipient could be an excellent match for the recipient and, taken to the extreme, one could envisage a global and indeed mission-wide database of RXS profiles where, for example, a donor in the south Pacific in January 2006 gets identified as the proxy for a recipient in the central Atlantic in July 2008. This, however, can be terribly risky without very reliable ancillary information, such as temperature and moisture profiles, and a vast array of passive channels. Thus, it is logical to give the closest potential donors highest priority.
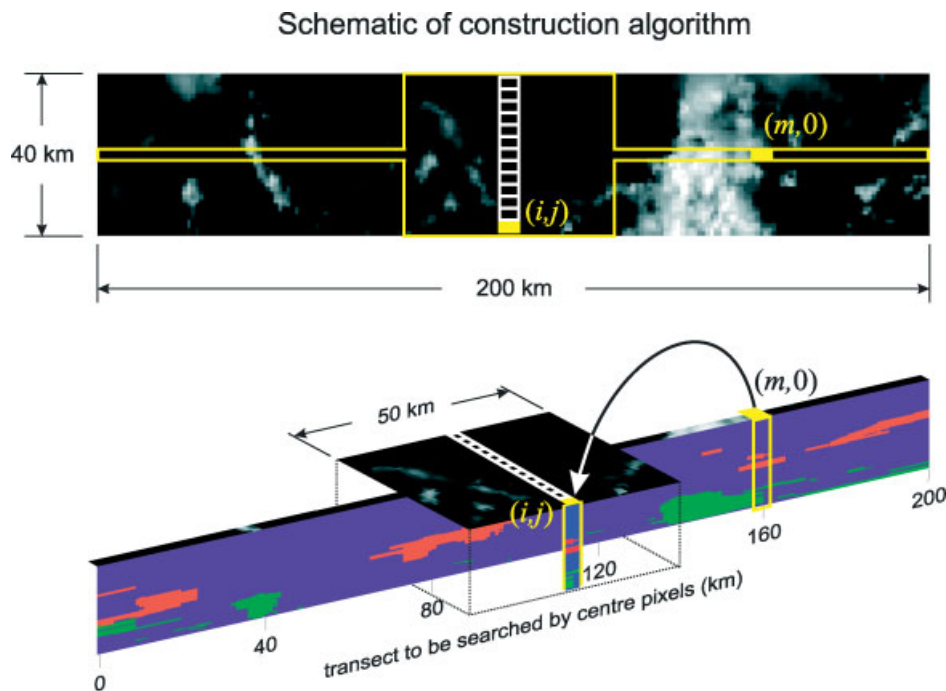
Schematic of construction algorithm

Figure 2. The top panel shows a stretch of MODIS visible imagery. The thin yellow-bordered line spanning it marks the position of the RXS. The objective is to fill the central volume marked by the large yellow-bordered region with cloud properties drawn from the RXS. For example, the column associated with the pixel at $(m, 0)$ has been designated as the proxy for the pixel at (i, j). The lower panel attempts to show that the cloud-radiation attributes associated with $(m, 0)$ are donated to (i, j). The algorithm is applied until all pixels in the central region are filled by donor RXS columns.
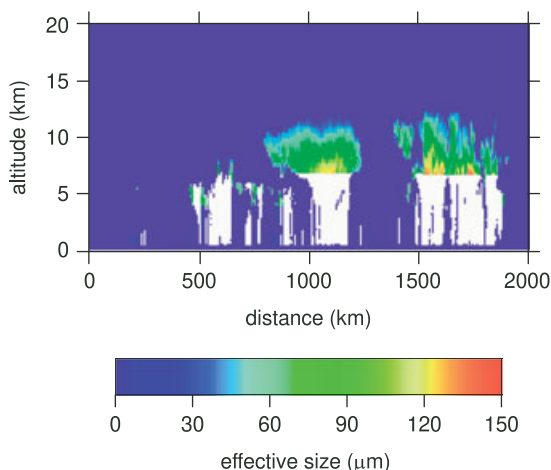

Figure 3. Typical example, from the equatorial Indian Ocean, of CloudSat retrieval through the Tropics. Blanked areas are where the retrieval algorithm reported only the strong likelihood of cloud but failed to report water content and effective particle size. As most failures occur well into clouds, CALIPSO's lidar is no help either.

zones could depend on the estimated geometric properties of cloud. When 1D radiative transfer is employed, as in this study, and the BBR footprint is relatively large (20 km for CERES), the constructed scene must cover at least the BBR's footprint. The ability of the construction algorithm to operate at distances progressively further away from the RXS is discussed in section 4.3.

2.3. 1D radiative transfer for constructed domains

As with CloudSat (L'Ecuyer, 2007), shortwave (SW) and longwave (LW) 1D radiative transfer algorithms, either narrow-band or broadband, can be applied to

each RXS column. Hence, during domain construction, calculated radiative heating rate and flux profiles and possibly radiances at $(m, 0)$ are treated just as any other variable: they are transplanted into off-RXS recipients. Once the constructed domain is complete, radiative quantities can be averaged to produce subdomain average ICA estimates.

As the focus of this study was the construction algorithm and not detailed radiative transfer, radiative transfer was limited to 1D models used in the Canadian Centre for Climate Modelling and Analysis global circulation model (GCM). SW fluxes were computed by a two-stream approximation, while for the LW an emissivity-type approach with corrections for scattering was used (Li, 2002). Gaseous transmittances (H_2O, CO_2 and O_3 for SW; H_2O, CO_2, O_3, CH_4, N_2O, $CFCl_3$ and CF_2Cl_2 for LW) were computed using the correlated k-distribution method with 31 quadrature points in cumulative probability space for SW and 46 for LW (Scinocca *et al.*, 2008). Optical properties for liquid droplets (Dobbie *et al.*, 1999; Lindner and Li, 2000) and ice crystals (Fu, 1996; Fu *et al.*, 1998) were resolved into four SW bands and nine LW bands. Aerosols were neglected. The SW component of this code participated in Barker *et al.*'s (2003) model intercomparison and performed well relative to benchmarks. Both SW and LW codes were featured in Oreopoulos and Mlawer (2010).

Surface albedos were set to values inferred from MODIS data as reported in the NASA–Langley merged dataset (Rutan *et al.*, 2009). For the SW model's 0.2–0.7 μm band, values obtained from 0.469 μm radiances were used. For the three near-infrared bands, values for 1.24 μm were used. For the LW, emissivity for all surfaces was set to 1.

3. Merged A-train data

A merged A-train dataset has been compiled at NASA–Langley in which CloudSat and CALIPSO profiles were mapped to 1 km resolution, thereby associating each column with a $\Delta L = 1$ km MODIS pixel (Kato *et al.*, 2010). It also includes 20 MODIS pixels either side of the CloudSat–CALIPSO RXS, making for 41 km wide MODIS imagery.

Figure 3 shows that CloudSat and CALIPSO retrieval algorithms often fail for optically thick clouds, which in all likelihood contain some large hydrometeors, but together they provide a cloud mask and reliable indication that clouds are likely to be present. Because these failures can blank out significant fractions of an orbit, it was decided to work with an alternate merged dataset that utilizes CloudSat's and CALIPSO's cloud masks in conjunction with a CERES-like approach for ascribing cloud phase, water content and effective particle size to the cloud mask along the RXS (Kato *et al.*, 2005).

First, a merged cloud mask was created by interpolating CloudSat's and CALIPSO's cloud masks on to CERES's arbitrary vertical grid (Kato *et al.*, 2005). This was deemed to be simpler than interpolating temperature, humidity and ozone (from the Goddard Earth Observing System–Data Assimilation System (GEOS-4); see Bloom *et al.*, 2005) and one or the other of CloudSat's or CALIPSO's cloud masks to either CALIPSO's or CloudSat's grid. If a CloudSat or CALIPSO cell overlaps a CERES layer, the cell is designated to be filled with cloud.

For each RXS MODIS pixel, effective visible optical depth $\tau_{\rm eff}$, effective particle size $r_{\rm eff}$ and particle phase were retrieved based on inversion of a 1D radiative transfer model (Minnis *et al.*, 2010a). If CloudSat's columnar classification (Sassen and Wang, 2008) is cirrus, altostratus, altocumulus or deep convection, those cells in the column identified as cloud with temperatures < 273 K are classed as ice and $r_{\rm eff}$ is set uniformly to the value inferred from MODIS data (Minnis *et al.*, 2008, 2010a, 2010b). Warmer cells in the column are set to liquid with $r_{\rm eff} = 10\,\mu$m. When the columnar classification is stratus, stratocumulus, cumulus or nimbostratus, the column is set to liquid with $r_{\rm eff}$ set uniformly to the value inferred from MODIS data.

Based on the cloud optical-property parametrization used for radiation calculations in this study (see section 2.3) and $r_{\rm eff}$ as defined above, values of cloud water content are assigned to each cell identified as cloud in the merged cloud mask field such that total optical depth equals $\tau_{\rm eff}$. Though not a perfect solution, especially when the fields are intended to be used by 3D solvers (Barker and Liu, 1995), one should bear in mind that at this stage exact replication of CERES TOA fluxes is not of paramount concern. Hence, lack of detail regarding the setting of important variables such as ice-crystal optical properties, phase-dependent extinction profiles and surface spectral characteristics was accepted up front, knowing that substantial disagreements with CERES estimates are likely for certain conditions; mostly tortuous convective cloud systems with little more than cloud masks from CloudSat and CALIPSO and rough 1D estimates from MODIS.

TOA fluxes are CERES's standard single scanner footprint (SSF) product for 20 km diameter pixels that coincide with the RXS. They were derived from angular distribution models and documented extensively by Loeb *et al.* (2005, 2006).

4. Results

This section consists of five separate subsections. In the first four, merged A-train data were used to assess specific aspects of the construction algorithm. In the fifth subsection, TOA fluxes were derived from constructed scenes and compared with corresponding CERES values.

4.1. *Spectral channel selection and degradation*

Most passive imagers found on satellites have several channels across visible and thermal wavelengths. The MODIS imager has 36 channels, while EarthCARE's will have just seven. One could feasibly use all available channels to evaluate Eq. (1), but the information content in some channels is far from independent and without a detailed weighting system it could even be detrimental to include too many channels. In this section, Eq. (1) was applied to the Sun-up side of an A-train orbit (19 April 2007 –CloudSat granule 5184) using all combinations of four distinct MODIS channels: 0.62–0.67 µm, 2.105–2.155 µm, 8.4–8.7 µm and 11.77–12.27 µm. These channels are similar to four in EarthCARE's MSI. This serves the following purposes: (1) to help select bands to use in Eq. (1), (2) to demonstrate algorithm performance for Sun-up and Sun-down conditions (when the first two channels are neglected the algorithm is essentially operating at night) and (3) to demonstrate algorithm performance in the event of channel failure.

The construction algorithm was applied between 60°S and 60°N using different combinations of MODIS channels in Eq. (1) as indicated on the ordinates of the plots in Figure 4. For domains measuring 21×40 km^2, mean radiances, cloud τ and cloud-top altitude h were computed for actual and reconstructed imagery. Least-squares linear regressions were performed for actual versus reconstructed values for individual channels or fields as indicated along the abscissa. Figure 4 shows coefficients of determination (R^2) for the fits ($R^2 = 1$ being the perfect reconstruction for all 262 domains). Each plot corresponds to domains with differing extents of total cloudiness, as listed in the titles. Those channels numbered 5–8 were *not* used in Eq. (1). Their actual and reconstructed fields were regressed to show how well channels not used in Eq. (1) were reconstructed.

Note that the upper four entries on the main diagonal in all plots have $R^2 \simeq 1$. This is because only that particular channel was used in the reconstruction. The fact that the 8.55 µm and 12.02 µm channels share substantial amounts of information is indicated by the fact that the R^2 values that involve them track each other closely. This indicates fortuitous redundancy, for if one of the channels fails the other can be used in its place. Similar situations exist for other channels not shown, for instance the 0.469, 0.555 and 0.645 µm channels.

Regarding the plots that apply to cloudy conditions, it is clear that in order to obtain good results for both τ and h a visible and a thermal channel are needed. The explanation for this almost goes without saying: the visible homes in on τ while the thermal responds greatly to cloud-top temperature or h. For example, note the slippage in performance for very
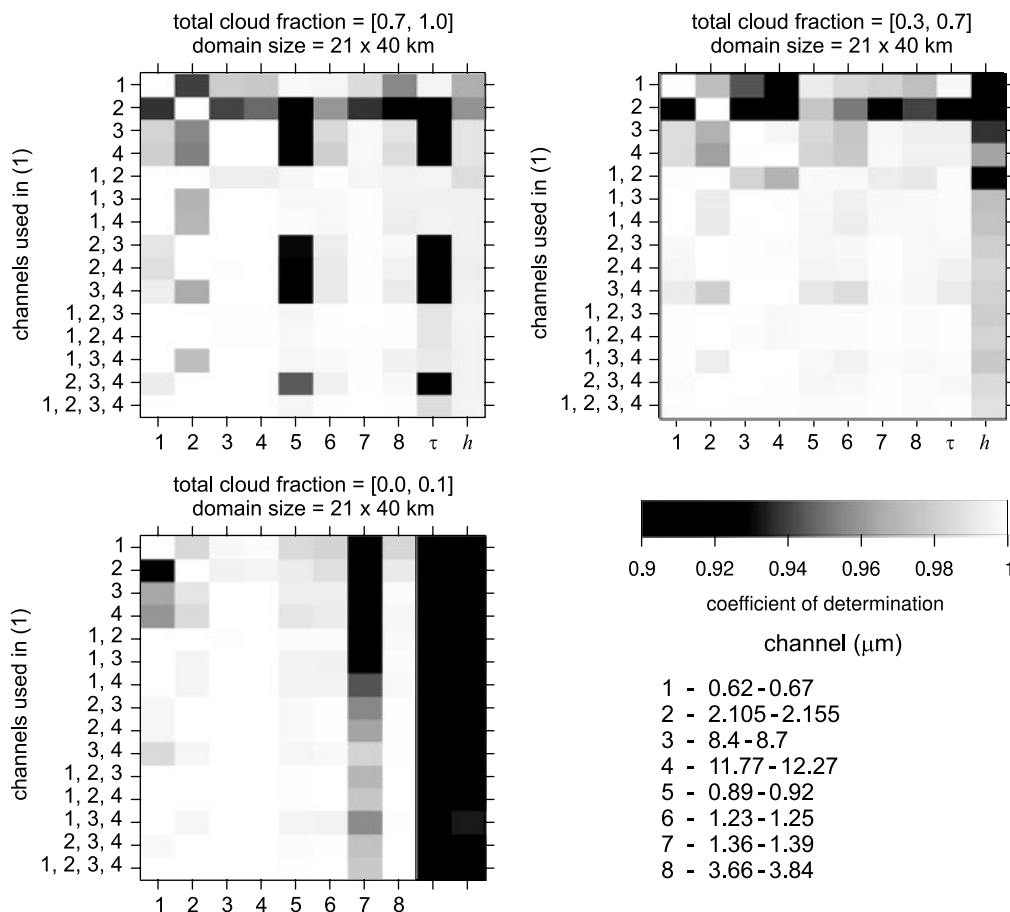
Figure 4. MODIS channels are listed on the lower right. The construction algorithm was applied to [60°S, 60°N] for an orbit on 19 April 2007 using various combinations of MODIS channels in Eq. (1) as listed on the ordinate of the plots. For domains measuring $21 \times 40 \text{ km}^2$, mean radiances were computed for actual and reconstructed MODIS imagery. The same was done for cloud optical depth τ and cloud-top altitude h. Least-squares linear regressions were then performed for actual versus reconstructed values for channel or field as indicated along the abscissa. The resulting coefficients of determination (R^2) for the fits are plotted. The upper left plot is for predominantly cloudy domains, the upper right for domains with intermediate cloud fractions and the lower left for predominantly cloudless domains. τ and h are not reported for the mostly cloudless domains. As an example for mostly cloudy domains, when channels 1 and 4 were used in Eq. (1), channel 1 was reconstructed almost perfectly with $R^2 \simeq 1$, but reconstruction of channel 2 [not used in Eq. (1)] was relatively poor, although R^2 was still large at 0.97. Reconstructions of τ and h were excellent as well, with $R^2 = 0.99$.

cloudy domains for τ when the 0.645 μm channel is left out of Eq. (1).

The lower left plot is for domains with either no cloud or very small total cloud fractions and so τ and h were not reported. While most channels are reconstructed well for most combinations used in Eq. (1), the 1.375 μm channel, used often for cirrus detection and lacking from EarthCARE's MSI, has frequent problems, although R^2 is still greater than 0.95, which is very good by most standards, when two or more channels are used in the reconstruction process. It is not clear why this is but it is likely related to spectral surface albedo. Hereinafter, unless mentioned otherwise, scenes were constructed using the four channels mentioned at the start of this section.

4.2. *Distance between donor and recipient columns*

The purpose of this section is to provide information regarding distance $D(i, j; m)$ between an off-nadir recipient at (i, j) and the RXS donor pixel at $(m, 0)$, where $D(i, j; m)$ is defined in Eq. (8).

When cost function values $F(i, j; m)$ are *not* ordered from smallest to largest but rather donors are found by applying Eq. (2) with m_1 and m_2 set by solar geometry and satellite location, there is the potential for large $D(i, j; m)$, especially

when clouds are 'broken' due to the necessity of having to conduct searches across cloudless gaps (analogous to the 'Levy flights' of photons, e.g. Davis and Marshak, 2004).

Along the top of Figure 5 are 41 km wide MODIS images for two important channels in the cost function (0.645 and 12.02 μm) for ~500 km across the southern Indian Ocean. The RXS is along the centre of the swaths. The scatter plot shows latitude θ_{lat} of recipient pixels plotted against θ_{lat} of donor pixels when only Eq. (2) was used to find donors. The locations indicated by A and B show tight clusters around the $D(i, j; m) = 0$ line. These clusters tend to correspond to the interior of fairly large clouds. In essence, a match may be found for points off the RXS by simply going over to the RXS and searching for just a few kilometres at most. Conversely, pixels near C were close to the edge of a cloud, where conditions were transitioning. For these cases the chance of finding an appropriate match on the local RXS ($\theta_{lat} \simeq 47.3°$) was fairly small. Indeed many proxies were found near C′, some 200 km away ($\theta_{lat} \simeq 45.5°$) at the other edge of the cloud (system). By the same token, many off-nadir pixels at C′ had matches back near C, hence, the psuedosymmetry seen in the plot. Note that many off-nadir pixels at C′ also found matches further along at C″ ($\theta_{lat} \simeq 43.5°$). When the minimum distance condition in Eq. (9) was applied, the large distances seen here vanished and the corresponding
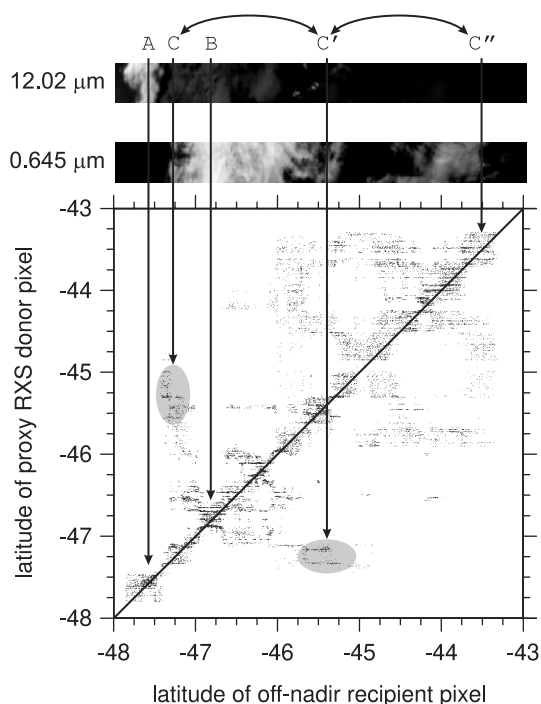
Figure 5. Images show swaths of MODIS imagery measuring 41 km wide and 5° long in the tracking direction from the south-central Indian Ocean on 25 January 2007. The construction algorithm was applied to this domain with donor pixels found by selecting the closest one from the ordered set of smallest cost-function values. The scatter plot shows ordered pairs of latitudes for off-nadir recipient pixels and their RXS proxy for all cloudy pixels ±20 km from the RXS.

scatter plot became much less interesting-looking, with most points clustered close to the $D(i, j; m) = 0$ line.

In recognition of the fact that the smallest possible distance between a recipient at (i, j) and a donor is $|j|$, we redefine donor–recipient distance as

$$D'(i, j; m) = \Delta L \max\{1, |i - m|\} \quad (10)$$

rather than using Eq. (8). Figure 6 shows the median of $D'(i, j; m)$ for the Sun-up portion of the 25 January 2007 orbit when constructed domains were 40 km long and either 11 or 41 km wide, and construction was done by ordering cost-function values and finding the shortest distance. For the most part, median values of D' are less than 10 km with the wider domains having, as expected, systematically slightly larger median D'.

4.3. Reconstruction of the RXS

This algorithm is purposely called a *construction algorithm* as opposed to a *reconstruction algorithm*, because it *constructs* a digital field where one never existed. For experiments reported in this section, however, attempts were made to *reconstruct* digital fields, namely the RXS and some of its attributes.

Verification of the construction algorithm using A-train data is not straightforward, for the answer is unknown. Nevertheless, one can go a certain distance by attempting to reconstruct the RXS itself. In so doing, one attempts to fill an RXS column at $(i, 0)$ by searching the RXS and applying Eq. (9) to potential donor pixels in $[i - m_1, i - n] \cup [i + n, i + m_2]$, which bars the first $\pm n$ pixels next to i, hence defining a *dead zone* in the search
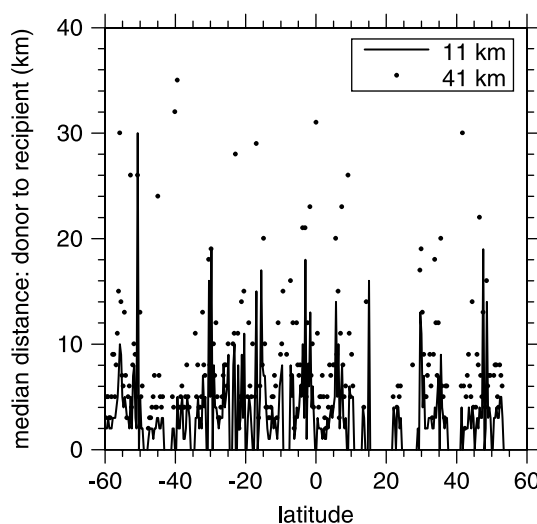

Figure 6. Median of distance between off-nadir recipient pixels and their RXS donors as a function of latitude for domains measuring 40 km long and either 11 or 41 km wide for the same orbit as shown in Figure 5. Distance is defined by Eq. (10).

process. This test, as illustrated in Figure 7, is meant to mimic filling off-RXS columns that are $\pm n$ pixels away from the RXS. For example, when $n = 5$ searching for a proxy column begins five pixels away, just as for off-RXS pixels at $(i, \pm 5)$.

Figure 8 shows attempts to reconstruct a 400 km-long stretch of RXS. The upper image in all three columns is the actual RXS merged cloud mask. This is an especially demanding case, as it involves fairly dense multilayer clouds. Over much of this domain passive-only retrievals would yield very little, if any, useful information about cloud vertical structure. Lower images are reconstructions for discrete values of $n = 1$, 5, 10 and 20. The leftmost column shows results for the full construction algorithm (sorting on F and selecting the closest) using the four spectral channels listed in Section 4.1. By $n = 5$, which corresponds to the outer edge of an 11 km wide domain similar to those to be computed for EarthCARE, it is clear that some error is creeping in; nevertheless, a significant amount of detail is captured. For the most part, even out at $n = 20$, multilayers of clouds have been replicated well. The region seeming to have the greatest difficulty is between 100 and 200 km along the horizontal. This is where layer clouds were transitioning or dying out entirely and is reminiscent of results in Figure 5 where large donor–recipient distances were required near cloud edges.

The centre column of plots in Figure 8 shows results for the same four channels, but donors were identified as the minimum cost function within ±200 km along the RXS (less the dead zone). Immediately, at $n = 1$, performance degraded relative to the full algorithm. Obviously a significant amount of noise emerged as a result of finding good radiance matches at great distances. This demonstrates the importance of minimizing donor–recipient distances and the algorithm's banking on what is already known from studies into horizontal fluctuations of cloud (Marshak and Davis, 2005b). By $n = 20$ the amount of noise in the reconstructed scene would likely spell trouble for 3D radiative transfer simulations.

Images in the rightmost column of Figure 8 are for the full algorithm using just the 12 μm channel. This can be thought of as representing performance either at night-time or under total failure of shortwave channels. Without the visible

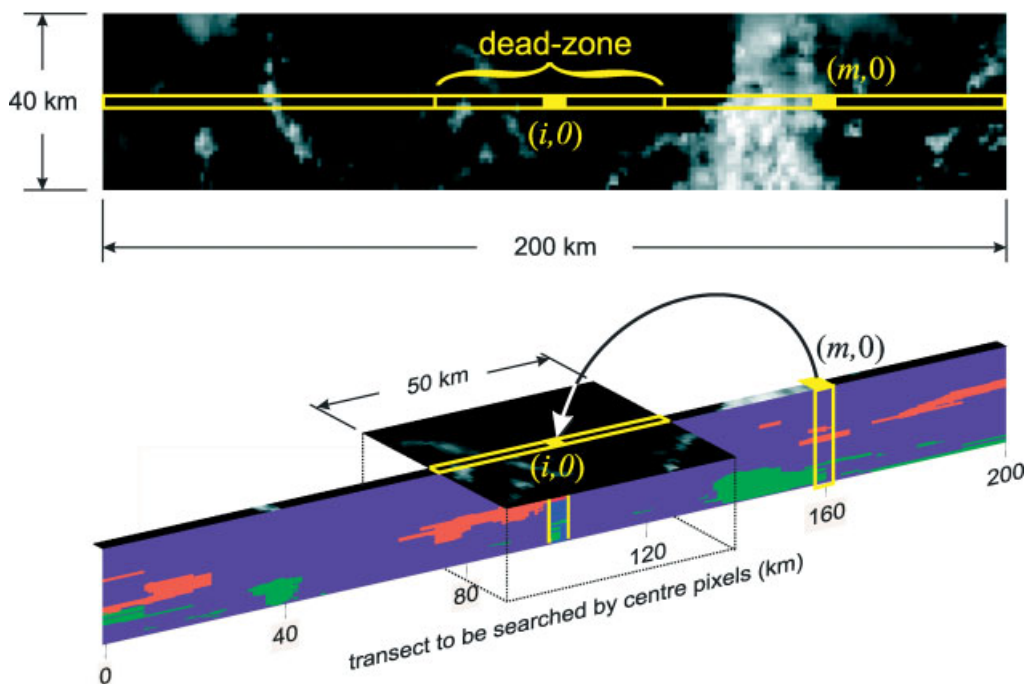## Schematic of construction algorithm applied to RXS

Figure 7. As in Figure 2 but this illustrates the situation when an attempt is made to reconstruct the RXS only. The pixel at $(i, 0)$ is to be filled by applying the algorithm to RXS pixels outside the *dead zone*, shown here as 25 km either side of $(i, 0)$. The pixel at $(m, 0)$ is the donor. This is repeated for all pixels along the RXS.
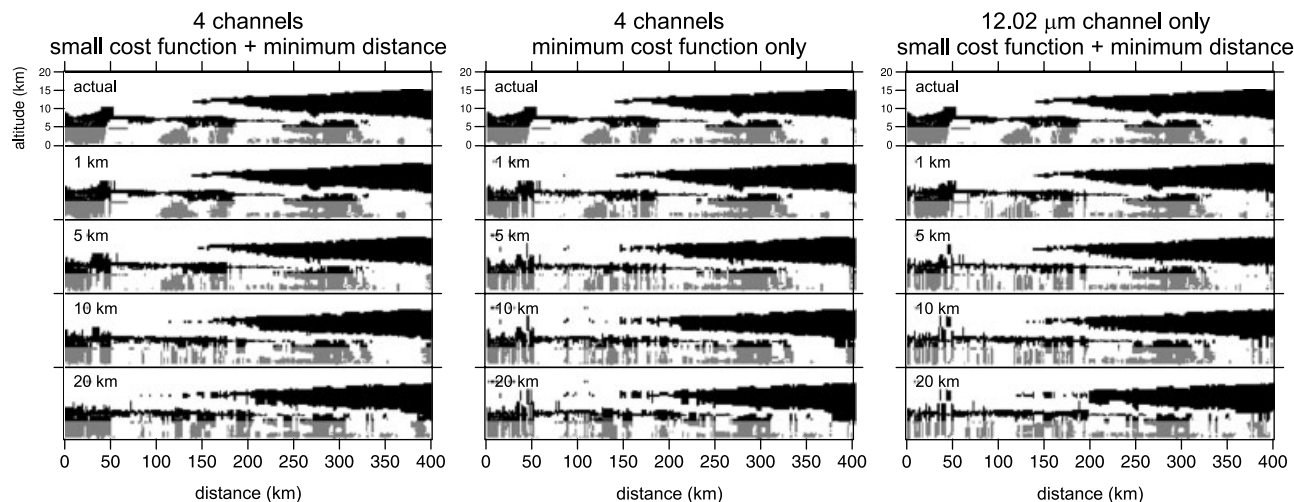

Figure 8. The topmost images are merged cloud masks (1 km horizontal resolution) for a stretch of tropical RXS along an orbit on 19 April 2007. Black and grey indicate ice and liquid, respectively. The sequences of lower images represent the corresponding masks produced by the construction algorithm for various dead-zone lengths as listed. The leftmost column of plots corresponds to results for the full algorithm using four spectral channels, the centre column is for the simple algorithm that does not try to minimize distances between donors and recipients, also using four channels, and the rightmost column is again for the full algorithm but this time using just the 12 μm channel.

constraint guiding optical depth, which is related closely to cloud geometric thickness, the lower reaches of cloud were reconstructed poorly. The upper portions of cloud, however, were reconstructed decidedly well. This demonstrates the control exerted by the thermal channel via its relation to cloud-top altitude (temperature). Nevertheless, it is still encouraging to note that this reduced application of the algorithm reproduced the multilayeredness of these clouds fairly well even out to $n = 20$.

Figure 9 shows mean values of several variables accumulated out to n as functions of n for the field shown in Figure 8. Results for accumulated fields of width

$2n + 1$ are shown, because the algorithm is intended to produce full 3D domains and not single rows. For these accumulations, averaging included the original RXS as it is included in constructed fields, yet gave double weight to the reconstructed lines so as to represent scene construction on both sides of the RXS. Results are shown for the three renditions of the algorithm used to produce the masks shown in Figure 8.

Generally, the full algorithm using four channels performed best. For small domains up to ~20 km wide ($n = 10$), performance of the full algorithm using the thermal channel was comparable to the simple algorithm (no
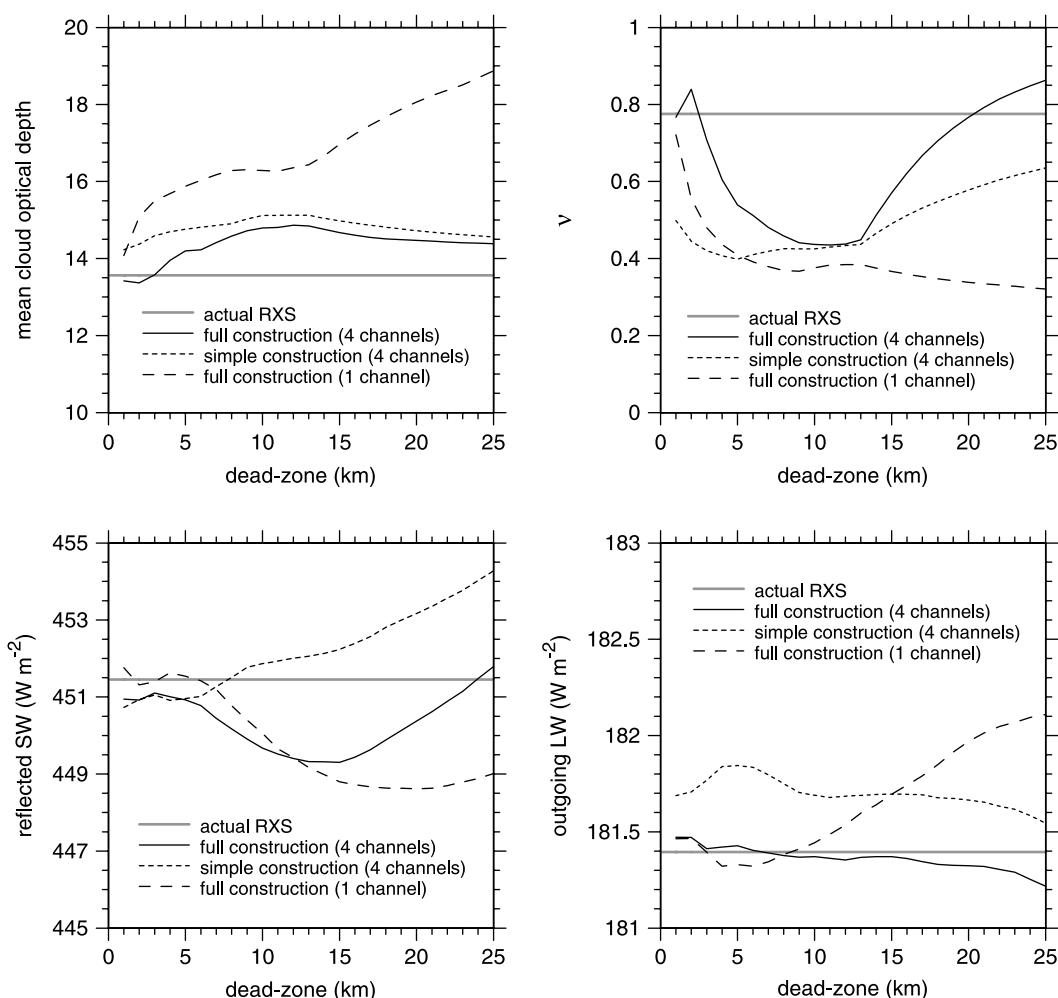
Figure 9. Grey lines indicate the values of mean cloud optical depth $\langle \tau \rangle$, $\nu = (\langle \tau \rangle / \sigma_\tau)^2$, where σ_τ is the standard deviation of τ, mean reflected shortwave (SW) flux at TOA and mean outgoing longwave (LW) flux at TOA for the actual RXS shown in the uppermost panels of Figure 8. Solid black lines indicate values for reconstructed accumulated cross-sections as functions of dead-zone length for the full algorithm using four channels; dotted curves are for the simple algorithm using the same four channels and dashed lines are for the full algorithm using only the 12 μm channel.

sorting of F) with four channels. Note that while mean cloud optical depths $\langle \tau \rangle$ were reconstructed quite well, errors in median values (not shown) were almost negligible. Likewise, the upper-right plot in Figure 9 shows $\nu = (\langle \tau \rangle / \sigma_\tau)^2$ where σ_τ is the standard deviation of τ. Corresponding maximum-likelihood estimates (not shown), which are less subject to infrequent, and often inconsequential, outliers, exhibit excellent agreement with the actual value. The most encouraging aspect of these plots is that TOA fluxes, computed by applying the 1D models to the RXS and its reconstructed counterparts, for this complex field are never in error by more than $\sim 2 \, \text{W m}^{-2}$ for reflected SW and $\sim 1 \, \text{W m}^{-2}$ for outgoing LW.

Figure 10 shows some domain-average profiles for the actual RXS shown in Figure 8 as well as for several values of n. These results are for the full algorithm using four channels. For clouds higher than 10 km, layer cloud fraction A_c, mean cloud water contents and ν for water content were reproduced extremely well for all n. This shows why outgoing longwave radiation (OLR) was captured so well. The largest errors in A_c occur for $n = 20$ at altitudes near 5 km. Mean water contents were reconstructed very well with the exception of $n = 5$ and 10 for clouds between 2 and 5 km high. Associated values of ν were also very small and indicate that these errors were the result of having selected a small

number of columns with anomalously large water contents. On the other hand, for $n = 20$ almost all reconstructed clouds below 10 km lacked sufficient horizontal variability, as indicated by values of ν being twice as large as they should be, but mean water contents were fine. This was due to too many occurrences of a single RXS column being used multiple times.

Figure 11 shows heating-rate profiles for the leftmost column of domains in Figure 8: full algorithm using four channels. Both SW and LW heating-rate patterns for the reconstructed fields, out to at least $n = 10$ (~ 20 km wide domain), agreed nicely with those for the actual field. While plots of accumulated mean heating-rate profiles were made out to $n = 20$, they were not worth showing, as differences relative to those for the actual field were almost imperceptible.

As mentioned, the field used here as an example was a difficult case of multilayer tropical cloud. Most 400 km sections of cloud do not exhibit as much intricacy as this one. Numerous other examples were examined and almost all reconstructions performed equally well or better than the one shown here. In general, the more extensive and planar the clouds and the fewer the number of definite layers, the better the reconstruction.
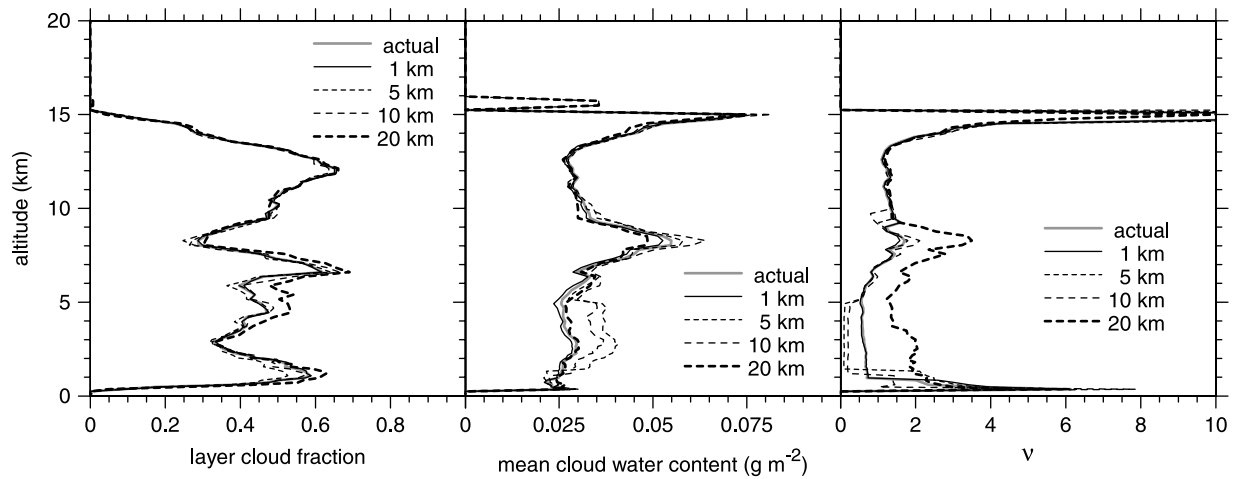
Figure 10. Domain-average profiles of layer cloud fraction, cloud water content $\langle w \rangle$ and $\nu = (\langle w \rangle / \sigma_w)^2$, where σ_w is the standard deviation of w, for the RXS shown in Figure 8 as well as for reconstructed cross-sections corresponding to various dead-zone lengths as listed.
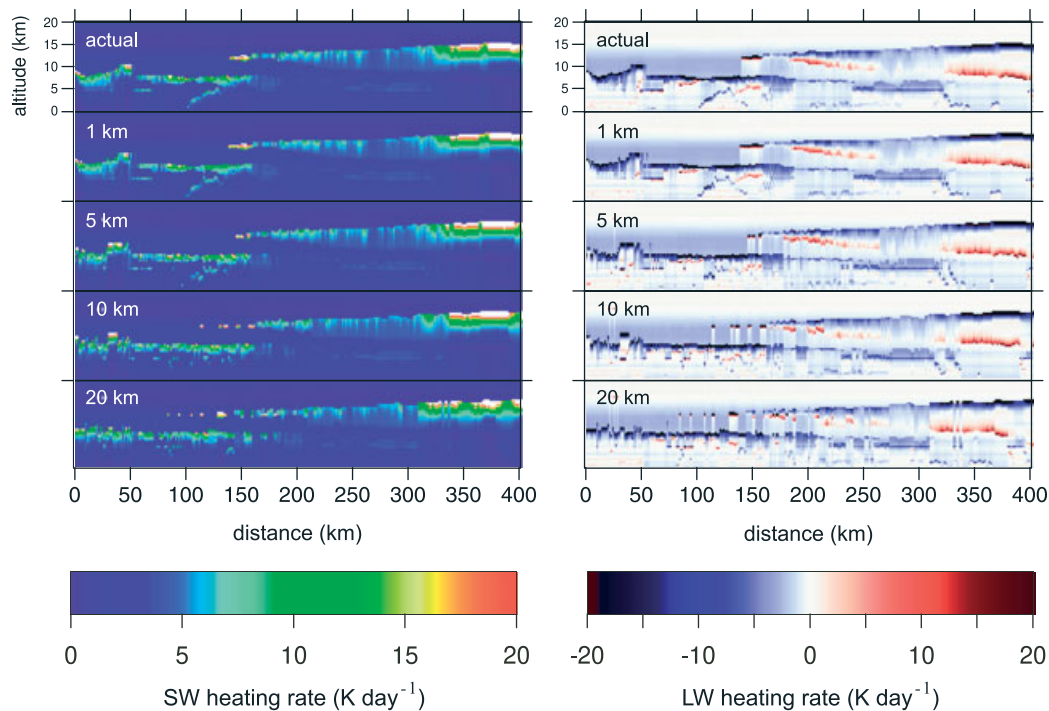

Figure 11. Broadband SW and LW radiative heating cross-sections for the RXS (upper panel) and its reconstructions for several dead-zone values as listed (see Figure 8). White patches on the SW panels are values greater than 20 K day^{-1} while black patches on the LW panels are beyond ± 20 K day^{-1}.

4.4. Image reconstruction

An additional way to demonstrate the credibility of the construction algorithm is to show what its reconstructed images look like. Figure 12 shows two examples of very disparate cloud fields. With the exception perhaps of the stratiform 0.645 μm reconstruction, which is almost a perfect reproduction, one can see that the reconstructed fields appear to be reproductions but with fewer grey-scale bits. This is basically the case, as they were fabricated from just small samples of potential donors. Also, one can see areas where an extreme off-nadir event, such as that two-thirds of the way down on the left side of the 0.645 μm image for (1.57°N, 93.82°E), failed to find a satisfactory match from the available RXS pixels. Nevertheless, its 12.02 μm counterpart was very satisfactory.

Figure 13 shows the square root of the ensemble average second-order structure functions, as defined in Eq. (3), made in the tracking direction, for the images shown in Figure 12. It also shows results for the 3.79 μm channel, which was not used in Eq. (1). The obvious difference between the actual and reconstructed images is the addition of a minor amount of noise at scales smaller than $\sim$5 km for 3.79 and 12.02 μm. This is evident from Figure 12. Differences for the 0.645 μm images, however, were almost nil.

Several key cloud properties are inferred directly from MODIS radiances (Minnis *et al.*, 2008). Hence, as an extension to imagery reconstruction, it is instructive to see how well the algorithm can reconstruct MODIS-inferred cloud properties. Recall that whatever is associated with the donor RXS pixel, including MODIS cloud properties, is transplanted to the off-nadir recipient.
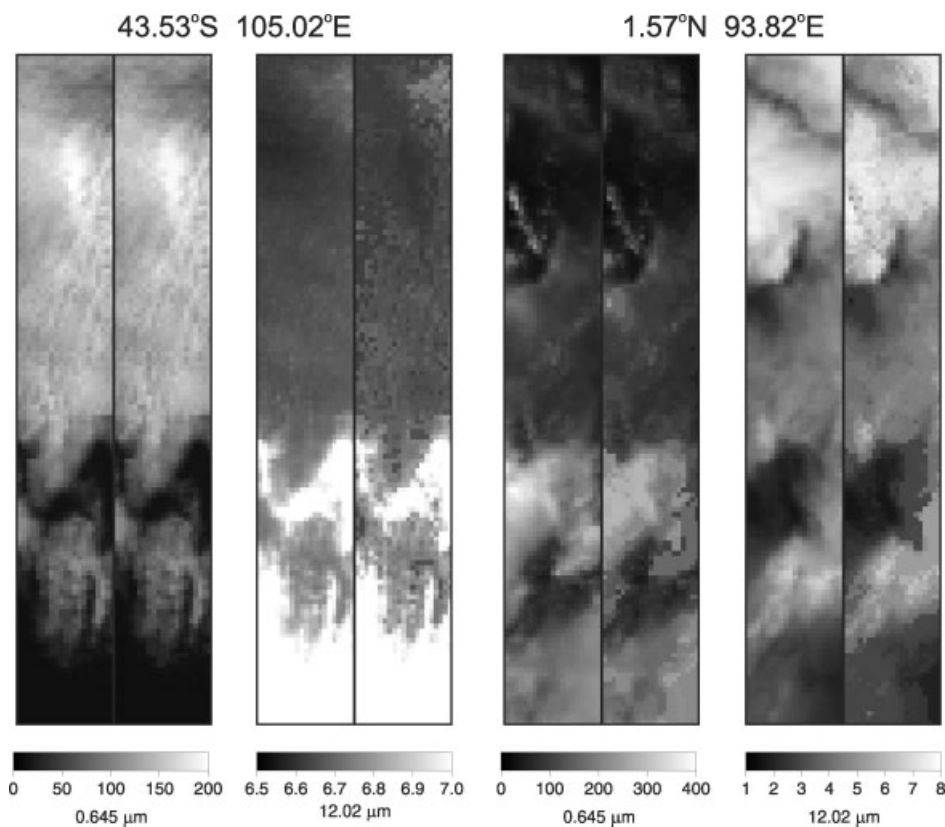
Figure 12. For each pair of images, the left is the actual MODIS data while the right is its reconstruction based on Eq. (1) using channels 0.62–0.67 μm, 2.105–2.155 μm, 8.4–8.7 μm and 11.77–12.27 μm. Both examples measure 21 km wide by 135 km long and units are W m^{-2} sr^{-1} μm^{-1}.
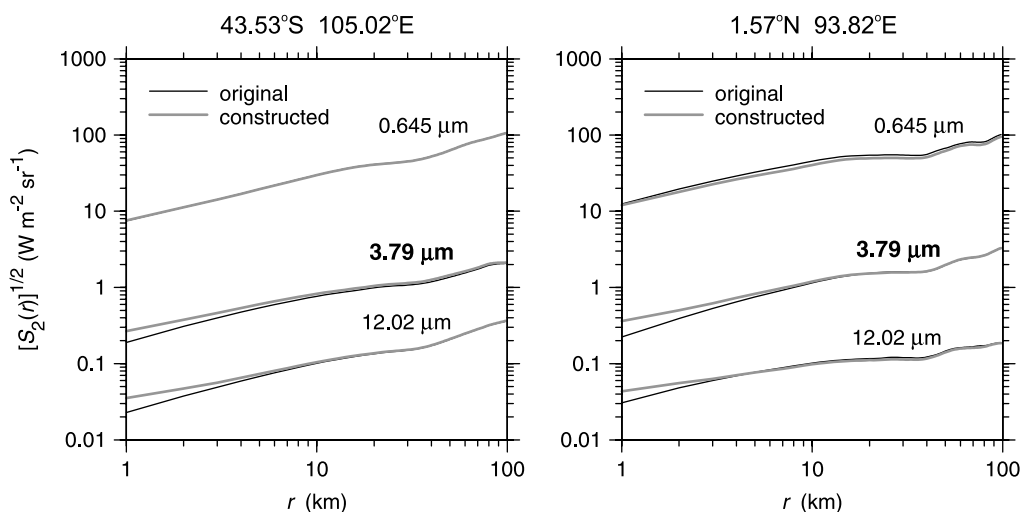

Figure 13. Square root of the ensemble average 1D structure function (along the 135 km axis) of the original and reconstructed images shown in Figure 12. While the 0.645 and 12.02 μm channels were used to perform the reconstruction, the 3.79 μm channel was not. The 3.79 μm curves were shifted up one decade to facilitate easy viewing.

Figure 14 shows zonal plots of domain-average, cloud-only, $\langle\tau\rangle$, $\langle h\rangle$ and effective values of liquid-droplet and ice-crystal sizes for the orbit on 25 January 2007 between 60°S and 60°N. The size of each domain was 21 × 40 km^2. Results are similar for other orbits, domain sizes and minor alterations to the combination of channels used in Eq. (1). Mean bias errors (MBEs) are for the full algorithm using four channels. As expected from the goodness-of-fit results reported in Figure 4, $\langle\tau\rangle$ and $\langle h\rangle$ were captured extremely well with almost no bias. Near 35°N, however, the algorithm had obvious difficulty with ground snow and cloud. This can be ameliorated by limiting the construction algorithm

to fill columns identified as cloudy by MODIS, but if cloud screening errors stem from MODIS retrievals then any improvement might be relative to those retrievals only.

The same can be said for mean droplet effective radius, where the majority of mean values were between 10 and 20 μm with associated MBEs less than ±2 μm. For mean ice-crystal effective radius, overall bias was again almost nil and its MBEs relative to mean values resembled the other fields. For those instances where MBEs were relatively large, however, fractional areas of ice cloud tended to be very small. Hence, ice radius across many searchable segments of the RXS was a sparse variable and subject to poor sampling.
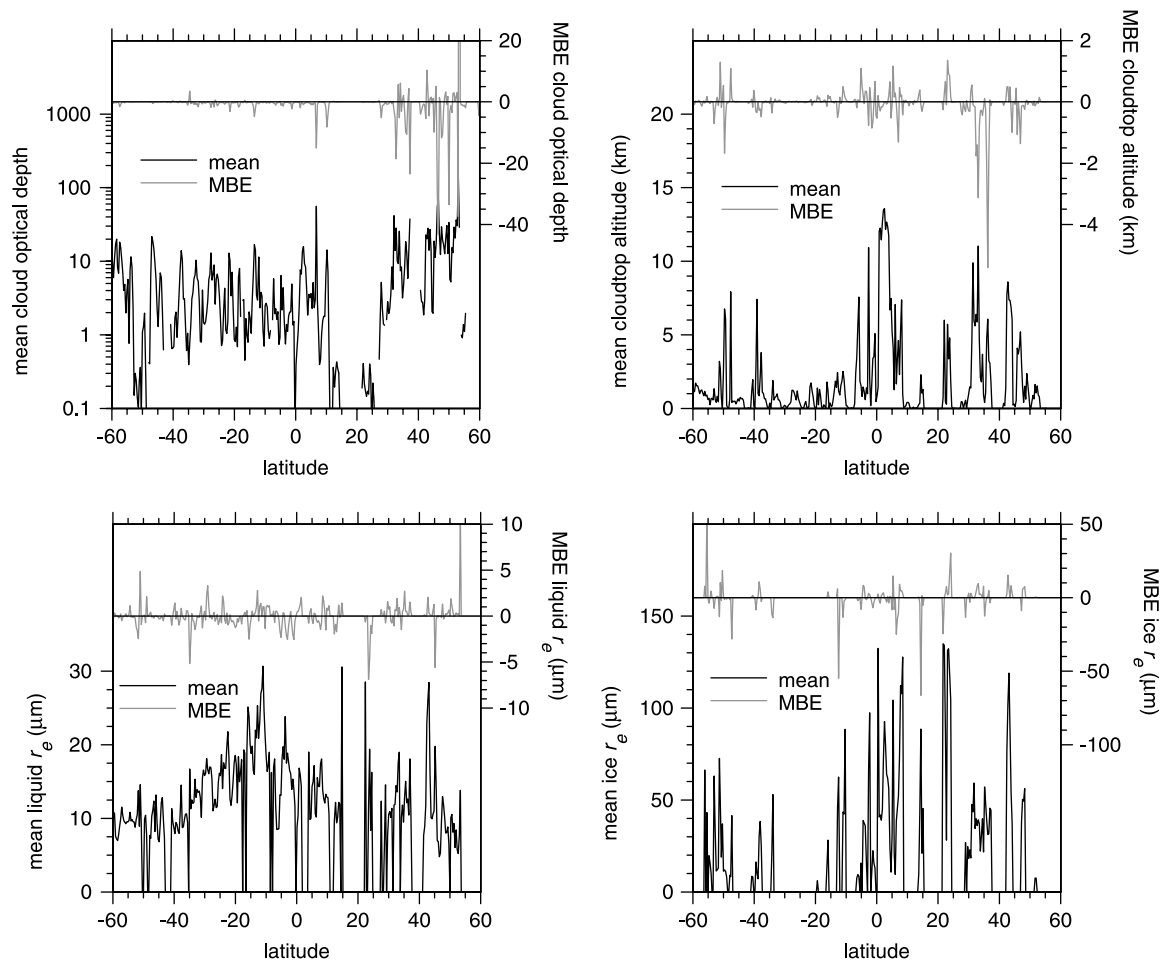
Figure 14. Domain-average mean cloud properties retrieved directly from MODIS imagery and the mean bias error (MBE) associated with their reconstructed counterparts based on Eq. (1) using the same channels as in Figure 12. Domains were $21 \times 40\,\text{km}^2$ and came from [60°S, 60°N] for an orbit on 25 January 2007.

While this naturally generates random error, it can also lead to some large MBEs on account of small sample sizes.

4.5. Comparison with CERES TOA fluxes

As a final assessment, domains were constructed so as to match 20 km CERES footprints, 1D radiative transfer (RT) codes were applied to them and estimated TOA fluxes were compared with corresponding values inferred from CERES radiances and angular direction models (ADMs) (Loeb *et al.*, 2005, 2006). Figure 15 shows CERES and 1D RT results as well as μ_0 as functions of θ_{lat} between 60°S and 60°N for the Sun-up period of 25 January 2007 (CloudSat granule 3960). Fields were constructed by the full algorithm using all four channels listed in Figure 4 and domain sizes were $21 \times 40\,\text{km}^2$ (i.e. two successive CERES pixels). With the aid of the scatter plots on the right, it is clear that the construction algorithm produced quite reliable fields, as correlation coefficients between CERES and the 1D RT are 0.89 for all SW points and 0.93 for all LW points. Some notable discrepancies, however, deserve comment.

Firstly, this orbit came up the Indian Ocean and over the India–Bangladesh coast. The underestimation of reflected SW at TOA by the 1D model in the area marked A in Figure 15 is due to the neglect of substantial aerosol loading coming off the continent: Kolkata and Balasore, India were both reporting light winds from the north and haze. This

aerosol was evident in CALIPSO data. Future studies for EarthCARE will naturally include aerosols.

Secondly, between about 30°N and 37°N (Himalayan Plateau) the 1D model seriously overestimated and underestimated reflected SW in fairly rapid succession. MODIS cloud optical depths were between 10 and 50 and total cloud fractions were ~0.5. For this area, visible and near-infrared surface albedos reported in the CERES dataset were typically 0.15 and 0.3, respectively, thus implying no snow, which corroborates with MODIS's 0.645 μm imagery. It is not clear at this stage why estimates of TOA fluxes became erratic, as cloud-property reconstructions agreed well with MODIS inferences. Possible explanations include poor MODIS cloud retrievals or poor CERES ADM-derived estimates of TOA SW flux, given that this portion of Earth is very mountainous.

The areas indicated by B in Figure 15 exhibit extreme underestimates of OLR by the 1D model. Cloudless skies prevailed and reflected SW values for the actual and constructed domains were in excellent agreement. The source of the problem was poor surface (and possibly lower atmospheric) temperatures T_s from the Goddard database (Bloom *et al.*, 2005). A similar situation existed on 19 April 2007 over the Rajasthan Desert, but the case shown here was easier to analyze due to several meteorological stations in the vicinity. Figure 16 shows actual and reconstructed images corresponding to channels 0.645 and 12.02 μm. One sees
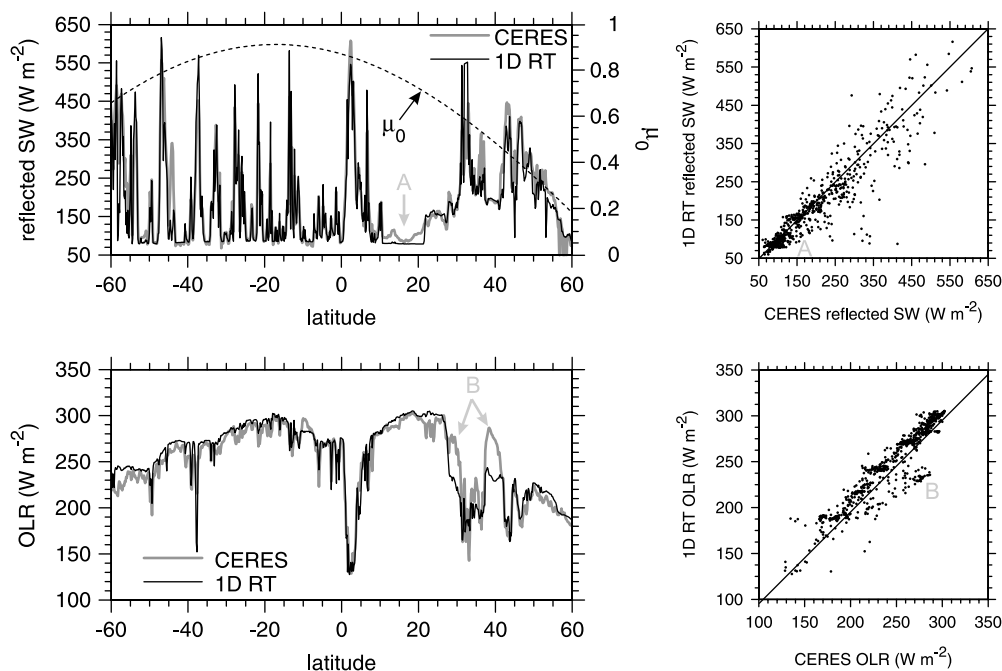
Figure 15. The left plots show CERES TOA reflected SW and outgoing LW fluxes at TOA as functions of latitude for domains measuring $21 \times 40\,km^2$ for an orbit on 25 January 2007 (same as used to produce Figure 14). Also shown are results from 1D radiative transfer models (1D RT) applied to constructed domains as well as the cosine of solar zenith angle μ_0. The plots on the right show CERES and 1D RT data from plots on the left plotted against each other.
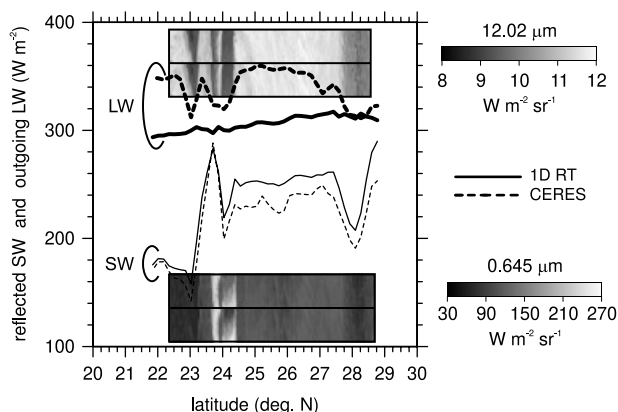

Figure 16. Paired images show original MODIS radiances (lower image for each pair) for the Rajasthan Desert and the constructed counterpart using Eq. (1) with channels 0.62–0.67 μm, 2.105–2.155 μm, 8.4–8.7 μm and 11.77–12.27 μm. Dashed lines represent CERES TOA SW and LW fluxes along the centre of the images while solid lines represent the corresponding values from 1D radiative transfer models applied to the constructed domain using MODIS surface albedos and temperature and moisture profiles from a forecast model.

immediately that the construction algorithm had no trouble in reconstructing TOA radiances. Indeed, surface albedo information passed to the 1D SW model was sufficiently accurate so as to model reflected SW to within $20\,W\,m^{-2}$ of CERES for most of the stretch. Often, however, CERES OLR was near $350\,W\,m^{-2}$ whereas the 1D LW model gave typically $300\,W\,m^{-2}$.

At 23.0°N, 70.5°E along the RXS ground track, the National Oceanic and Atmospheric Administration (NOAA) reported $T_s \simeq 303\,K$, which was used in the 1D model. Surface observations close to the overpass at Rajkot, India (22.3°N 70.8°E) reported a T_s value of ~313 K. Assuming that $OLR = \epsilon \sigma T_s^4$, where ϵ is effective broadband emissivity, use of Rajkot's $T_s \simeq 313\,K$ and

CERES's $OLR = 350\,W\,m^{-2}$ gives $\epsilon \simeq 0.64$. Using this value for ϵ with $T_s \simeq 303\,K$ gives $OLR = 302\,W\,m^{-2}$, about $50\,W\,m^{-2}$ less than CERES. This, therefore, raises a cautionary note about ancillary data that are outside the purview of the satellite mission and radiative closure experiments: retrievals can appear to be poor for reasons that have little or nothing to do with them, in this case, initialization with poor surface temperatures.

It is interesting to redo the comparison between model and CERES data for scenes constructed from just the infrared channels; 8.55 and 12.02 μm. Results were generally so similar to those shown in Figure 15 that they have been plotted in Figure 17 as model results using four channels versus those using two channels. The correlation for reflected SW is 0.952, while for LW it is 0.997. In other words, the only impact on TOA fluxes worth mentioning that would result from losing the two SW channels is in the SW for the most reflective clouds; the impact on OLR would be negligible.

As a final point, Figure 18 shows cumulative frequency distributions of SW and LW TOA flux errors, defined here as model − observed, for $21 \times 40\,km^2$ domains within [60°S, 60°N] for eight orbits, five from January 2007 and three from April 2007. Depending on orbit, estimated reflected SW TOA broadband fluxes were within $\pm 10\,W\,m^{-2}$ of CERES values (i.e. EarthCARE's goal for each domain) for between about 25% and 45% of the time. Corresponding rates for OLR were about 45%–75%.

Also from Figure 18, reflected SW and OLR overall bias (median) errors are $-2.9\,W\,m^{-2}$ and $0.4\,W\,m^{-2}$, with corresponding interquartile ranges of $\sim 40\,W\,m^{-2}$ and $\sim 15\,W\,m^{-2}$. For just eight partial orbits, these values appear to be quite acceptable given that at this point in the study little to no effort has been put into details involving ice-cloud optical properties, aerosols, surface properties and synergistic retrieval of cloud properties, as the main concern was documentation and demonstration of the construction
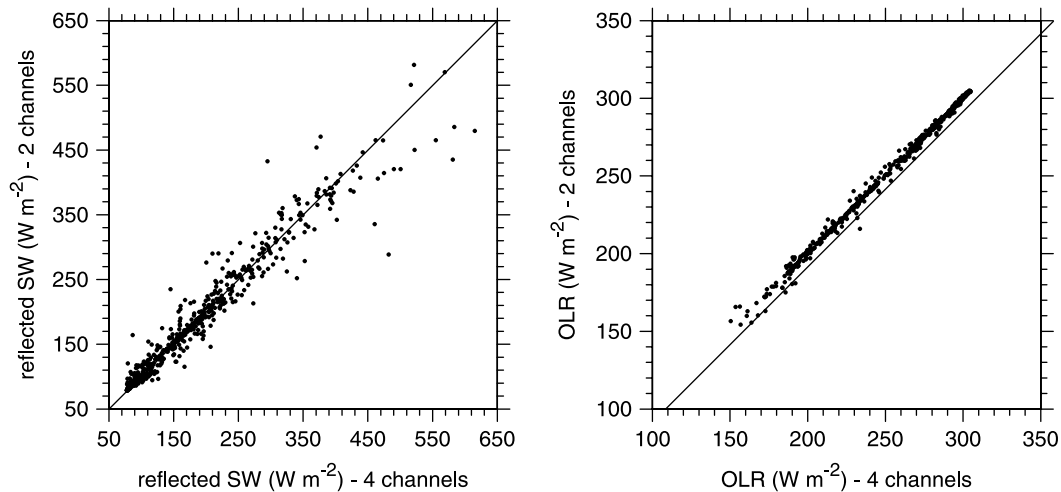
Figure 17. For the same domains as in Figure 15, reflected SW and OLR at the TOA predicted by the 1D radiative transfer models for constructed fields using four MODIS channels (0.62–0.67 μm, 2.105–2.155 μm, 8.4–8.7 μm and 11.77–12.27 μm) plotted against the use of two channels (0.62–0.67 μm and 11.77–12.27 μm).
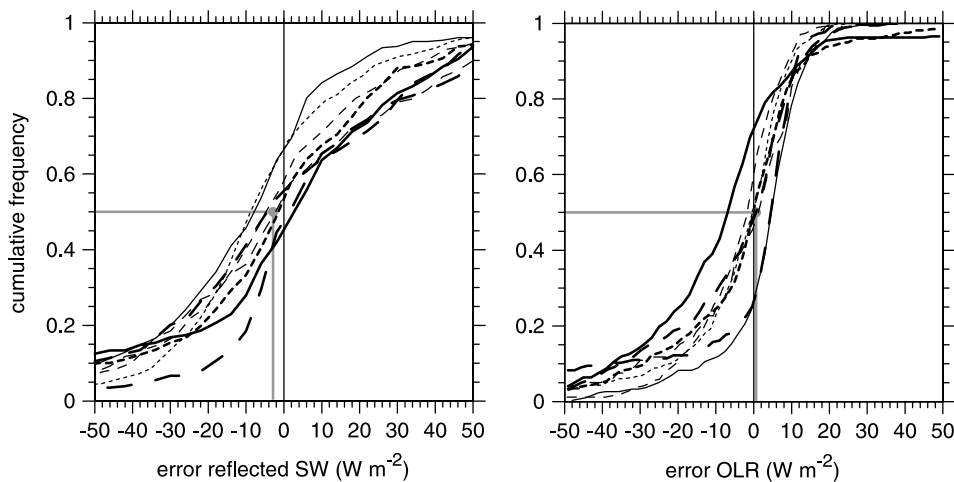

Figure 18. Cumulative frequency distributions of errors in reflected SW and OLR at the TOA for [60°S, 60°N] sections of eight orbits. Errors were defined as 1D RT values minus CERES values. Shaded regions indicate errors less than 10 W m^{-2}. Grey lines indicate overall medians (approximately equal to mean bias errors given near-symmetric distributions).

algorithm. Once these issues are addressed and 3D RT is used, these biases are expected to decrease and the fraction of cases with errors less than ±10 W m^{-2} to increase.

5. Summary and recommendations

Although the price tags for active-sensor cloud missions, like CloudSat and EarthCARE, are high, so too are the knowledge returns. Conventionally, cloud properties have been and will continue to be inferred from data collected by relatively affordable passive sensors. Today's satellite-based active sensors do not scan across-track and so provide only a narrow, essentially 2D, nadir cross-section of data. Covering this cross-section, and extending well beyond it in the across-track direction, are data from a co-registered passive imager. Bringing all data together allows for synergistic active–passive retrieval of a 2D vertical cross-section of cloud and aerosol properties (referred to here as the retrieved cross-section or RXS). The algorithm presented here seeks to extend the synergistic process through construction of 3D domains of cloud, aerosol and surface properties around the RXS. The focus here was on clouds.

The construction algorithm's main purpose is to aid verification of EarthCARE's RXS data via continuous radiative closure using coincidental broadband SW and LW radiances or fluxes measured at TOA. To do this requires a 3D map of cloud (plus aerosol and surface) properties, as well as ancillary meteorological data. This is because (1) the baseline configuration of the broadband radiometer's (BBR) footprint is ∼10 km, which is much wider than the RXS at ∼1 km, and (2) even if the BBR's footprint were comparable to that of the RXS, radiation flows laterally and so horizontal transport can be expected to be non-negligible.

The algorithm seeks to fill columns associated with an off-RXS pixel by performing a limited search along the RXS for the closest RXS pixel with passive radiances matching closely those of the pixel the column of which is to be filled. Once a suitable *donor* is identified, its associated column is designated by proxy to represent the column 'beneath' the off-nadir *recipient* pixel. This is repeated for an entire orbit until all pixels out to a specified distance are filled, thereby creating a 3D swath of cloud, aerosol and surface properties. One can then perform either 1D or 3D radiative transfer (RT) calculations on 3D domains sectioned off from the full 3D swath. For this introductory study, only

1D RT was considered; a companion article explores the use of 3D RT. The goal of this article was to demonstrate the credibility of the construction algorithm using data from A-train satellites, the only data that resemble those expected from EarthCARE.

It was shown that the algorithm has little difficulty in *reconstructing* MODIS imagery for channels used and not used to perform the reconstruction. Usually distance between recipient and donor columns was less than 30 km and often less than 5 km. This shows that there is usually more than enough passive radiance information close by along the RXS. Additionally, the algorithm had little difficulty reconstructing MODIS-inferred values of cloud optical depth, particle effective sizes, and cloud-top altitude.

It was demonstrated that the algorithm performs best when the cost function used to find the proxy RXS column uses a visible and an infrared window channel. This is reasonable, as the visible and infrared channels work simultaneously to secure optical depth (cloud geometric thickness) and cloud-top height, respectively (Rossow and Schiffer, 1999). Including more channels in the cost function helps only slightly (and might even begin to degrade performance). It should be noted that additional attributes beyond individual pixel radiances, such as the mean of 3×3 pixels surrounding the centre pixel (Gabriel *et al.*, 2008), were tried but did not help; in fact, they often hindered. These were considered in an attempt to acknowledge 3D RT effects. At 1 km resolution, however, too much useful information is likely lost and too much extraneous information might be given too much weight. Nevertheless, more work should be devoted to testing before dispensing entirely with this idea.

The ability of the construction algorithm to produce 3D cloud fields can be tested, to some extent, using A-train data and attempting to reconstruct the RXS itself. This was done by applying the algorithm as usual but barring computation of the cost function for $\pm n$ pixels either side of the one to be filled. This mimics filling columns that are n pixels away from the RXS in the across-track direction. Results were shown for a complex, multilayer deep tropical cloud system. Reconstruction was very satisfactory out to distances of at least 20 km. As clouds become more planar and homogeneous, with decreasing numbers of distinct layers, performance improves.

As a final test, 1D radiative transfer models were applied to constructed fields and estimated TOA broadband fluxes were compared with corresponding CERES values for domains that measured 21 km wide in the across-track and 40 km in the along-track direction. Generally speaking, model estimates tracked CERES's values extremely well for the eight partial orbits considered. Typically ~35% of modelled domain-averaged reflected solar fluxes were within $\pm 10\,\mathrm{W\,m}^{-2}$ of CERES. For outgoing longwave radiation this number jumped to ~65%. These tests also pointed to the need for ancillary meteorological data, beyond the purview of the satellite missions, to provide reliable estimates of surface optical properties and profiles of atmospheric state.

Several steps are required to round out assessment of the construction algorithm, and we are at various stages of taking them. Firstly, aerosols and a wider range of surface conditions (e.g. snow and ice) must be addressed. Secondly, end-to-end simulations, including full radiative closure experiments, that use data from cloud models should be performed as they will enable the partition of errors due to active–passive retrievals, scene construction, RT modelling and ancillary data. Thirdly, 3D solar and infrared radiative transfer models should be explored, as differences between 1D and 3D results could be significant for certain cloud types.

Acknowledgement

References

Barker HW, Liu D. 1995. Inferring optical depth of broken, shallow clouds from Landsat imagery. *J. Climate* **8**: 2620–2630.

Barker HW, Stephens GL, Partain PT, Bergman JW, Bonnel B, Campana K, Clothiaux EE, Clough S, Cusack S, Delamere J, Evans KF, Fouquart Y, Freidenreich S, Galin V, Hou Y, Kato S, Li J, Mlawer E, Morcrette J.-J, O'Hirok W, Räisänen P, Ramaswamy V, Ritter B, Rozanov E, Schlesinger M, Shibata K, Sporyshev P, Sun Z, Wendisch M, Wood N, Yang F. 2003. Assessing 1D atmospheric solar radiative transfer models: Interpretation and handling of unresolved clouds, *J. Climate* **16**: 2676–2699.

Bloom SA, *et al.* 2005. '*Documentation and validation of the Goddard Earth Observing System (GEOS) data assimilation system version-4*', NASA Technical Report Series on Global Modeling and Data Assimilation NASA/TM-2005-104606, Vol 181, 26pp.

Cess RD, Harrison EF, Minnis P, Barkstrom BR, Ramanathan V, Kwon TY. 1992. Interpretation of seasonal cloud–climate interactions using Earth radiation budget experiment data. *J. Geophys. Res. Atmos.* **97**(D7): 7613–7617.

Davis AB, Marshak A. 2004. Photon propagation in heterogeneous optical media with spatial correlations: Enhanced mean-free-paths and wider-than-exponential free-path distributions. *J. Quant. Spectrosc. Radiat. Transfer* **84**: 3–34.

Davis AB, Marshak A, Cahalan RF, Wiscombe WJ. 1997. The Landsat scale break in stratocumulus as a three-dimensional radiative transfer effect: Implications for cloud remote sensing. *J. Atmos. Sci.* **54**: 241–260.

Delanoë J, Hogan RJ. 2010. Combined CloudSat–CALIPSO-MODIS retrievals of the properties of ice clouds. *J. Geophys. Res.* **115**: D00H29. DOI:10.1029/2009JD012346.

Dobbie JS, Li J, Chylek P. 1999. Two and four stream optical properties for water clouds at solar wavelengths. *J. Geophys. Res.* **104**: 2067–2079.

Donovan DP, Voors RH, van Zadelhoff GJ, Acaretta JR. 2008. '*ECSIM Model and Algorithms Document*', Technical Report ECSIM-KNMI-TEC-MAD01-R. KNMI: De Bilt, Netherlands.

ESA. 2001. '*The Five Candidate Earth Explorer Missions: EarthCARE – Earth Clouds, Aerosols and Radiation Explorer*', ESA SP-1257(1), September 2001, 130 pp. Prepared by J. P. V. Poiares Baptista.

Fu Q. 1996. An accurate parameterization of the solar radiative properties of cirrus clouds for climate models. *J. Climate* **9**: 2058–2082.

Fu Q, Yang P, Sun WB. 1998. An accurate parameterization of the infrared radiative properties of cirrus clouds for climate models. *J. Climate* **11**: 2223–2237.

Gabriel P, Barker HW, O'Brien D, Ferlay N, Stephens GL. 2008. Statistical approaches to plane-parallel error identification and retrievals of optical and microphysical properties in 3D clouds Part 1: Bayesian inference. *J. Geophys. Res.* **114**: D06207. DOI:10.1029/2008JD011005.

Kato S, Rose FG, Charlock TP. 2005. Computation of domain-averaged irradiance using satellite-derived cloud properties. *J. Atmos. Oceanic Technol.* **22**: 146–164.

Kato S, Sun-Mack S, Miller WF, Rose FG, Chen Y, Minnis P, Wielicki BA. 2010. Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles. *J. Geophys. Res.* **115**: D00H28. DOI:10.1029/2009JD012277.

L'Ecuyer T. 2007. CloudSat project: Level 2 fluxes and heating rates product process description and interface control document. Available at http://cloudsat.cira.colostate.edu/ICD/2B-FLXHR/2B-FLXHR_PDICD_5.1.pdf.

Li J. 2002. Accounting for unresolved clouds in a 1D infrared radiative transfer code. Part I: Solution for radiative transfer, including cloud scattering and overlap. *J. Atmos. Sci.* **59**: 3302–3320.

Lindner TH, Li J. 2000. Parameterization of the optical properties for water clouds in the infrared. *J. Climate* **13**: 1797–1805.

Loeb NG, Kato S, Loukachine K, Smith NM. 2005. 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. Oceanic Technol.* **22**: 338−351.

Loeb NG, Kato S, Loukachine K, Smith NM. 2006. 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. Oceanic Technol.* **24**: 564−584.

Marshak A, Davis AB. 2005a. *3D Radiative Transfer in Cloudy Atmospheres*. Springer: Heidelberg; 686 pp.

Marshak A, Davis AB. 2005b. 'Scale-by-scale analysis and fractal cloud models'. In *3D Radiative Transfer in Cloudy Atmospheres*, Marshak A, Davis AB (eds). Springer: Heidelberg; 686 pp.

Minnis P, Trepte QZ, Sun-Mack S, Chen Y, Doelling DR, Young DF, Spangenberg DA, Miller WF, Wielicki BA, Brown RR, Gibson SC, Geier EB. 2008. Cloud detection in non-polar regions for CERES using TRMM VIRS and Terra and Aqua MODIS data. *IEEE Trans. Geosci. Remote Sensing.* **46**: 3857−3884.

Minnis P, Sun-Mack S, Young DF, Heck PW, Garber DP, Chen Y, Spangenberg DA, Arduini RF, Trepte QZ, Smith WL Jr, Ayers JK, Gibson SC, Miller WF, Chakrapani V, Takano Y, Liou K-N, Xie Y. 2010a. CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data, Part I: Algorithms. *IEEE Trans. Geosci. Remote Sensing.* In Press.

Minnis P, Sun-Mack S, Chen Y, Khaiyer MM, Yi Y, Ayers JK, Brown RR, Dong X, Gibson SC, Heck PW, Lin B, Nordeen ML, Nguyen L, Palikonda R, Smith WL Jr, Spangenberg DA, Trepte QZ, Xi B. 2010b. CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data, Part II: Examples of average results and comparisons with other data. *IEEE Trans. Geosci. Remote Sensing.* In Press.

Monin AS, Yaglom AM. 1975. *Statistical Fluid Mechanics, Volume 2.* MIT Press: Boston, MA.

Nastrom GD, Gage KS. 1985. A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.* **43**: 950−960.

Oreopoulos L, Mlawer E. 2010. The Continual Intercomparison of Radiation Codes (CIRC): Assessing anew the quality of GCM radiation algorithms. *Bull. Am. Meteorol. Soc.* **91**: 305−310.

Ramanathan V. 1987. The role of Earth radiation budget studies in climate and general circulation research. *J. Geophys. Res. Atmos.* **92**: 4075−4095.

Rossow WB, Schiffer RA. 1999. Advances in understanding clouds from ISCCP. *Bull. Am. Meteorol. Soc.* **11**: 2261−2287.

Rutan D, Rose F, Roman M, Manalo-Smith N, Schaaf C, Charlock T. 2009. 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.

Sassen K, Wang Z. 2008. Classifying clouds around the globe with the CloudSat radar: 1-year of results. *Geophys. Res. Lett.* **35**: DOI:10.1029/2007GL032591.

Scinocca JF, McFarlane NA, Lazare M, Li J. 2008. The CCCma third generation AGCM and its extension into the middle atmosphere. *Atmos. Chem. Phys.* **8**: 7055−7074.

Stephens GL, Gabriel PM, Tsay S-C. 1991. Statistical radiative transfer in one-dimensional media and its application to the terrestrial atmosphere. *Trans. Theory Stat. Phys.* **20**: 139−175.

Stephens GL, Austin RT, Benedetti A, Boain R, Durden SL, Illingworth A, Mace G, Miller SD, Mitrescu C, O'Conner E, Rossow WB, Sassen K, Vane DG, Wang Z. 2002. The CloudSat mission and the EPS constellation: A new dimension of space-based observations of clouds and precipitation. *Bull. Am. Meteorol. Soc.* **83**: 1771−1790.

Wielicki BA, Barkstrom BR, Harrison EF, Lee RB III, Smith LG, Cooper JE. 1996. Mission to Planet Earth: Role of clouds and radiation in climate. *Bull. Am. Meteorol. Soc.* **77**: 853−868.

Winker DM, Hunt WH, McGill MJ. 2007. Initial performance assessment of CALIOP. *Geophys. Res. Lett.* **34**: DOI:10.1029/2007GL030135.

World Climate Research Programme (WCRP). 1994. Report of International Workshop In: *Cloud−radiation interactions and their parameterization in climate models.* 18−20 Oct 1993, *WCRP-86, WMO/TD-No. 648.* 147 pp. WCRP, Camp Springs, MD.