AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2016JD026443

Key Points:
- CERES SYN1deg flux estimates over sea ice and snow show smallest errors but suffer from underestimated sea ice albedo in early summer
- The AVHRR-based cloud optical thickness in CLARA data is clearly overestimated over snow and sea ice, leading to issues in flux estimation
- Cloud fraction in GEWEX SRB differs significantly from CERES and CLARA over snow and sea ice

Supporting Information:
- Supporting Information S1

Correspondence to:
A. Riihelä,
aku.riihela@fmi.fi

Citation:
Riihelä, A., J. R. Key, J. F. Meirink, P. Kuipers Munneke, T. Palo, and K.-G. Karlsson (2017), An intercomparison and validation of satellite-based surface radiative energy flux estimates over the Arctic, *J. Geophys. Res. Atmos.*, *122*, 4829–4848, doi:10.1002/2016JD026443.

Received 2 JAN 2017
Accepted 10 APR 2017
Accepted article online 13 APR 2017
Published online 5 MAY 2017

An intercomparison and validation of satellite-based surface radiative energy flux estimates over the Arctic

Aku Riihelä[1], Jeffrey R. Key[2], Jan Fokke Meirink[3], Peter Kuipers Munneke[4], Timo Palo[5], and Karl-Göran Karlsson[6]

[1]Finnish Meteorological Institute, Helsinki, Finland, [2]National Oceanic and Atmospheric Administration, Madison, Wisconsin, USA, [3]Royal Netherlands Meteorological Institute, De Bilt, Netherlands, [4]Institute for Marine and Atmospheric Research, Utrecht University, Utrecht, Netherlands, [5]University of Tartu, Tartu, Estonia, [6]Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

Abstract Accurate determination of radiative energy fluxes over the Arctic is of crucial importance for understanding atmosphere-surface interactions, melt and refreezing cycles of the snow and ice cover, and the role of the Arctic in the global energy budget. Satellite-based estimates can provide comprehensive spatiotemporal coverage, but the accuracy and comparability of the existing data sets must be ascertained to facilitate their use. Here we compare radiative flux estimates from Clouds and the Earth's Radiant Energy System (CERES) Synoptic 1-degree (SYN1deg)/Energy Balanced and Filled, Global Energy and Water Cycle Experiment (GEWEX) surface energy budget, and our own experimental FluxNet / Satellite Application Facility on Climate Monitoring cLoud, Albedo and RAdiation (CLARA) data against in situ observations over Arctic sea ice and the Greenland Ice Sheet during summer of 2007. In general, CERES SYN1deg flux estimates agree best with in situ measurements, although with two particular limitations: (1) over sea ice the upwelling shortwave flux in CERES SYN1deg appears to be underestimated because of an underestimated surface albedo and (2) the CERES SYN1deg upwelling longwave flux over sea ice saturates during midsummer. The Advanced Very High Resolution Radiometer-based GEWEX and FluxNet-CLARA flux estimates generally show a larger range in retrieval errors relative to CERES, with contrasting tendencies relative to each other. The largest source of retrieval error in the FluxNet-CLARA downwelling shortwave flux is shown to be an overestimated cloud optical thickness. The results illustrate that satellite-based flux estimates over the Arctic are not yet homogeneous and that further efforts are necessary to investigate the differences in the surface and cloud properties which lead to disagreements in flux retrievals.

Plain Language Summary The amounts of solar and thermal radiative energies toward and away from the Earth's surface over the Arctic are the main driver of, e.g., the annual melt and freezing cycle of ice and snow. In this paper, we investigate how state-of-the-art satellite-based data sets of these radiative energy fluxes agree, or disagree, with each other and reference measurements made in the Arctic Ocean and on the Greenland Ice Sheet. We found that different data sets have not only individual strengths but also individual weaknesses which should be improved upon to enable a cohesive investigation of the radiative energy balance at annual and decadal time scales.

1. Introduction

The Earth's surface radiative energy budget (SRB) consists of the downwelling and upwelling shortwave and longwave fluxes. The SRB is a key factor regulating the climate of our planet, particularly for the polar regions. Over the Arctic, insolation, cloudiness, and temperature vary greatly over the course of a year, affecting the SRB and triggering melt and freeze events in the region's cryosphere. The Arctic SRB is driven by complex surface-atmosphere interactions and feedbacks, such as the surface (ice) albedo feedback [e.g., *Wetherald and Manabe*, 1975; *Curry et al.*, 1996] and the cloud radiative forcing effect [*Stephens*, 2005; *Letterly et al.*, 2016; *Wang and Key*, 2005].

The Arctic cryosphere is undergoing severe changes; air temperature is increasing [*Chylek et al.*, 2009], sea ice extent and volume are in decline [e.g., *Rothrock et al.*, 1999; *Stroeve et al.*, 2007; *Comiso et al.*, 2008], the albedo of the sea ice zone is decreasing [*Riihelä et al.*, 2013; *Pistone et al.*, 2014], the spring snow cover of the Northern Hemisphere is shrinking [*Derksen and Brown*, 2012], and Arctic permafrost shows signs of decay [e.g., *Zhang et al.*, 2005]. Exploration of the connections between these changes and changes in Arctic SRB

requires accurate characterization of the flux components with comprehensive spatiotemporal coverage at decadal time scales. Satellite remote sensing is the most cost-effective tool to achieve this goal.

The determination of radiative fluxes from satellite observations is a complex process, requiring robust algorithms to map both atmospheric and surface properties, as well as sophisticated radiative transfer calculations based on such inputs. Present efforts on the validation of satellite-based surface flux estimates focus strongly on lower latitudes, with only a few studies that focus on the Arctic [*Maslanik et al.*, 2001; *Chiacchio et al.*, 2002; *Liu et al.*, 2005; *Kay and L'Ecuyer*, 2013; *Dong et al.*, 2016; *Christensen et al.*, 2016]. Validation and intercomparison studies over Arctic sea ice are very rare, owing partly to the scarcity of in situ flux measurements. The goal of this paper is to address this deficiency through the following approach.

First, we present an Arctic-wide intercomparison study between the current state-of-the-art flux data sets: the Clouds and the Earth's Radiant Energy System (CERES) Synoptic 1-degree (SYN1deg)/Energy Balanced and Filled (EBAF) [*Wielicki et al.*, 1996; *Loeb et al.*, 2009; *Rutan et al.*, 2015] and the Global Energy and Water Cycle Experiment (GEWEX) SRB [*Stackhouse et al.*, 2011]. Second, to complement the comparative analysis, we have utilized a fast neural network-based radiative transfer calculator called FluxNet [*Key and Schweiger*, 1998] to calculate radiative fluxes based on the Satellite Application Facility on Climate Monitoring (CM SAF) CLARA-A2 (CM SAF cLoud, Albedo and RAdiation data set, Advanced Very High Resolution Radiometer (AVHRR) release 2 [*Karlsson et al.*, 2016]) daily mean cloud properties augmented with surface properties data from the AVHRR Polar Pathfinder-Extended (APP-x) data set [*Key et al.*, 2016]. This allows us to conduct a sensitivity analysis on the retrieved surface flux dependency on cloud and surface properties from AVHRR and CERES/MODIS. Third, we have validated the satellite-based surface fluxes against quality-monitored in situ measurements made at sites in the Arctic Ocean and on the Greenland Ice Sheet (GrIS) during polar summer 2007. While both CERES and GEWEX data sets have been validated against large amounts of in situ measurements [*Rutan et al.*, 2015; *Zhang et al.*, 2015], we focus on areas that have received little attention and are traditionally considered very challenging for flux retrievals.

The paper is structured as follows. First we introduce the existing data sets and the FluxNet-CLARA data set created for the purposes of this intercomparison. We then review the validation and intercomparison methods. The results are structured from large to small features, beginning with a general Arctic-wide intercomparison to place the similarities and differences in perspective, before proceeding to validation of the flux estimates with in situ observations. This is followed by a sensitivity study on the impact of cloud characterization uncertainties on the flux estimates. The results are then discussed against the backdrop of existing literature, before drawing conclusions and recommendations for future work.

2. Data Sets

The terminology surrounding the surface radiative energy budget is heterogeneous; therefore, a brief summary of the primary terms and acronyms used throughout this paper is in order. The solar radiation flux reaching Earth's surface is henceforth called the downwelling shortwave flux, or SWD. The solar radiation reflected by the surface is called the upwelling shortwave flux, or SWU. The thermal radiation of the atmosphere toward the surface (and measured at the surface) is called the downwelling longwave flux, or LWD. Finally, the thermal radiation emitted from Earth's surface toward the atmosphere is called the upwelling longwave flux, or LWU.

2.1. CERES

We have used two data sets from the CERES data set family. Where possible, we use the well-validated Energy Balanced and Filled (EBAF, ed2.8) data. However, the EBAF fluxes are not available at the daily mean time scale, which is required by our validation reference data spanning a single polar summer period. Therefore, for the validation part of our study we use the Synoptic 1-degree daily fluxes (SYN1deg, ed. 3A). The main difference between the data sets is that the surface fluxes in EBAF have been adjusted (through nudging of surface and atmospheric properties) to achieve consistency in the top-of-atmosphere fluxes between CERES and the SYN1deg monthly mean. All CERES flux calculations are based on the Fu-Liou correlated-k 2/4 stream radiative transfer model [*Fu and Liou*, 1992]. All CERES data are provided to users on a 1° equal-angle grid. The diminishing grid cell area of an equal-angle grid close to the poles is compensated by using a coarser grid in the processing, with results replicated to the 1° final grid [*Wielicki et al.*, 1998].

Table 1. Principal Attributes of the Satellite-Based Data Used in the Study

Data Set	Shortwave/Longwave Spectrum (μm)	Remote Sensing Instrument(s) Used as Input	Radiative Transfer Scheme	Source of Cloud Properties	Coverage
CERES EBAF/SYN1deg	0–5/5–100	CERES, MODIS on Terra/Aqua\ (also GEO imagers for low latitudes)	SW and LW: Langley Fu-Liou [*Fu and Liou*, 1992; *Kato et al.*, 1999]	MODIS [*Minnis et al.*, 2011]	2000–2016
GEWEX SRB	0.2–4/4–100	AVHRR	SW: *Pinker and Laszlo* [1992] LW: *Fu et al.* [1997]	ISCCP DX [*Rossow and Schiffer*, 1999]	1983–2007
FluxNet-CLARA	0.3–4/4.03–500	AVHRR	SW and LW: FluxNet/Streamer [*Key and Schweiger*, 1998]	Primarily CLARA-A2 [*Karlsson et al.*, 2016]	1982–2015 (CLARA) and 1982–2016 (APP-x)

Daily mean CERES SYN fluxes are averages of temporally interpolated hourly fluxes. Temporal coverage of CERES is 2000–2016. The principal attributes and coverage of the CERES and other data sets are summed in Table 1.

2.2. GEWEX

For GEWEX SRB we used release 3.0 data. Like CERES products, GEWEX SRB is based on temporally averaged cloud properties (1 h for CERES/MODIS; 3 h for GEWEX), using a correction to normalize the data with the full diurnal cycle. The GEWEX fluxes are based on radiative transfer calculations for a plane-parallel absorbing/scattering atmosphere using the delta-Eddington approximation [*Pinker and Laszlo*, 1992]. Again, similarly to CERES, the GEWEX fluxes are provided to users at a 1° equal-angle grid, but internal flux calculations at a grid which coarsens toward the poles [*Zhang et al.*, 2015]. Temporal coverage of GEWEX is 1983–2007.

2.3. CLARA

For the CLARA data record, we used the second release (CLARA-A2—described in detail by *Karlsson et al.* [2016]) compiled from historic AVHRR data covering the time period 1982–2015. Original AVHRR cloud products at ~4 km horizontal resolution and from all available NOAA and Metop satellites are aggregated into daily means in a 25 × 25 km^2 resolution grid covering the Arctic region. The cloud cover, cloud phase, and cloud top temperature products (CFC, CPH, and CTT) are defined for all months, but the other cloud property products (i.e., cloud optical thickness (COT) and particle effective radius (r_{eff}) for liquid and ice clouds) are only available during the Northern Hemisphere summer (more specifically, for solar zenith angles below 75°). Additional details on the CLARA-A2 cloud properties are available in the supporting information.

2.4. APP-x

The AVHRR Polar Pathfinder-Extended (APP-x) product suite provides a comprehensive set of variables for studies of trends and interactions in polar climate. APP-x is derived from the AVHRR Polar Pathfinder (APP) product, which is a fundamental climate data record composed of channel reflectances, brightness temperatures, and viewing and illumination geometry. APP-x is a thematic climate data record that provides information on some surface characteristics, cloud properties, and radiative fluxes. Both data sets cover the Arctic and Antarctic over the period 1982 through the present, using twice-daily composites at a 25 × 25 km^2 spatial resolution [*Key et al.*, 2016]. A more detailed description of the APP-x parameters used in this study is also available in the supporting information.

2.5. In Situ

Our reference data for validating the flux retrievals over the Arctic come from two sources: one in the Arctic Ocean and one on the Greenland Ice Sheet. Both cover the polar summer of 2007. Our Arctic Ocean in situ flux measurements are from the Tara drifting ice camp, where continuous radiative flux measurements were made between April and October 2007 [*Vihma et al.*, 2008]. The ice camp drifted between 88.5°N and 85.9°N during the summer; the drift is accounted for in the satellite data collocation. The expedition coincided with a record-low sea ice extent season, making this time period particularly interesting.

The fluxes were measured with pairs of Eppley Precision Spectral Pyranometer/Precision Infrared Radiometer (PIR) pyranometers/pyrgeometers which were monitored and maintained daily for riming or instrument tilt.

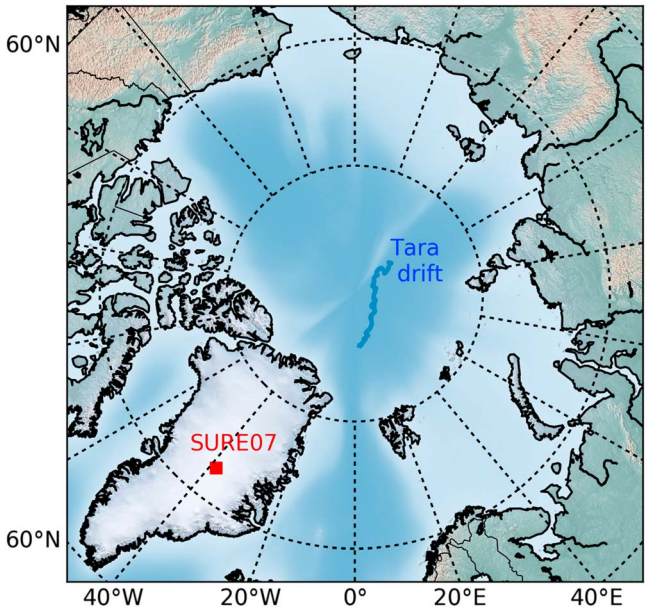

Figure 1. Locations of the drifting Tara ice camp (blue) and the SURE07 campaign (red) during the summer of 2007. (Note that sizes of location markers do not represent measured areas; all in situ measurements were at surface level.)

The Tara flux measurements have an estimated uncertainty of 3% for SW and 5% for LW [*Vihma et al.*, 2008, *Riihelä et al.*, 2010]. LWU measurements over midsummer were subjected to substantial solar heating and were thus excluded from analysis for quality assurance.

A concurrent measurement campaign (SURE07) took place at the Summit Geophysical Observatory on top of the Greenland Ice Sheet (3209 m above sea level, 72°34′N 38°28′W) between 8 June and 20 July 2007 [*Kuipers Munneke et al.*, 2009]. Continuous flux measurements were performed during June–July. For SWD and SWU, Kipp and Zonen CM21 pyranometers were used. The upward-looking sensor was actively ventilated and continuously checked for riming. LWD and LWU were measured by using ventilated Eppley PIR pyrgeometers. Instrument tilt was regularly measured and corrected to within 0.1°. Hourly values were computed from 1 min instantaneous observations and have an estimated uncertainty of $<$2%. Locations of the Tara drift (23.3. – 19.9.2007) and the SURE07 campaign site are shown in Figure 1.

3. Methods

3.1. FluxNet-CLARA Calculations

The use of a comprehensive radiative transfer model for the calculation of radiative fluxes with large satellite data sets is computationally prohibitive. Here we use an empirical parameterization of a radiative transfer model that is based on a neural network. The tool, called FluxNet, is a fast neural network-based implementation of the radiative transfer model Streamer for computing surface fluxes [*Key and Schweiger*, 1998]. Streamer is a very flexible, medium spectral resolution, plane-parallel model. Streamer can compute fluxes or radiances and is suitable for simulating radiative processes at the surface or within the atmosphere. FluxNet computes only surface and top-of-atmosphere fluxes but is 100 to 10,000 times faster than Streamer with little loss in accuracy. Root-mean-square differences between FluxNet and Streamer are approximately 11 W m^{-2} for downwelling shortwave and 7 W m^{-2} for downwelling longwave with near-zero biases.

Daily means of the CLARA-A2 cloud properties and 14:00 local solar time APP-x surface properties were the primary FluxNet inputs, as listed in Table 2. The total precipitable water (TPW) in each grid cell was not available over the Arctic from satellite observations over our full period of interest, and thus, ERA-Interim TPW was used.

All inputs were processed at a 25 km resolution EASE-2 Arctic projection (native to APP-x and CLARA-A2), which is also the resolution of the output fluxes. As cloud microphysical parameters are often not retrievable for sparse cloud situations, grid cells where the daily mean cloud fraction (CFC) was less than 16% were

Table 2. FluxNet Inputs and Their Sources

FluxNet Input	Acronym [Unit, if Any]	Source
Surface skin temperature	T_{surf} [K]	APP-x data (14:00 LST)
Surface emissivity	e	Constant 0.988
Surface albedo, broadband	α	APP-x data (14:00 LST)
Liquid cloud particle effective radius	$r_{eff\text{-}liq}$ [μm]	CLARA-A2 daily mean
Ice cloud particle effective radius	$r_{eff\text{-}ice}$ [μm]	CLARA-A2 daily mean
Cloud water/ice content	LWC/IWC [g/m^3]	Constant 0.1 LWC and 0.05 IWC
Cloud optical thickness	COT	CLARA-A2 daily mean, nudged by APP-x
Cloud top temperature	CTT [K]	CLARA-A2 daily mean
Cloud fraction	CFC	CLARA-A2 daily mean
Aerosol optical depth	AOD	Constant 0.1
Total column ozone	O3 [DU]	Constant 350
Total column water amount	TPW [g/m^2]	ERA-Interim reanalysis

classified as clear-sky for the radiative flux calculation and their cloud properties were not used. Grid cells classified as cloudy but with missing cloud parameters were discarded.

For each grid cell and day, hourly solar zenith angles were determined from latitude and longitude data by using the Pyorbital software package. Clear-sky grid cells' hourly calculations were based on FluxNet standard clear-sky input parameters. For grid cells classified as cloudy, FluxNet was run twice for each hour, once using the daily mean cloud properties for liquid clouds, and again for ice clouds. The resulting liquid and ice cloud output fluxes were averaged by using the CLARA-A2 daily mean cloud phase (i.e., relative fraction of liquid clouds) as weight. Conceptually, our approach was to "freeze" the atmosphere and clouds in the daily mean state, while the Sun traverses over each grid cell, finally averaging the fluxes into daily means according to cloud-type frequency.

We will show in section 5 that our approach does not affect the quality of the resulting flux estimates significantly. Some inputs were set to constant values because of lacking comprehensive input data, as noted in Table 2. However, literature supports our choices of constants. Arctic aerosol optical depth (AOD) typically varies seasonally between 0.05 and 0.15 [*Tomasi et al.*, 2007]. Our primary interest is in snow and ice, whose emissivities typically lie in the 0.98–0.99 range [e.g., *Salisbury and D'Aria*, 1994], although we acknowledge that emissivity does vary somewhat with wavelength, snow grain size and shape, and viewing geometry [*Hori et al.*, 2006]. Pan-Arctic ozone atmospheric content from the Ozone Measurement Instrument and the National Aeronautics and Space Administration (NASA) Modern-Era Retrospective Analysis for Research and Applications reanalysis range seasonally between 300 and 450 Dobson units [*NASA Ozone watch*, 2016]. Furthermore, the impact of ozone on the shortwave and longwave radiative fluxes is minor, justifying our use of a realistic constant. Cloud liquid water content (LWC) and ice water content (IWC) are set to slightly larger values than typically measured over Arctic sea ice [*Shupe et al.*, 2005; *Shupe et al.*, 2006] as a compromise between sea ice and high-latitude land areas. Again, spatiotemporal variation in both LWC and IWC will impact the daily FluxNet retrievals, but longer-term means should remain realistic.

Both CLARA-A2 and APP-x contain cloud optical thickness estimates based on the algorithm by *Nakajima and King* [1990]. Our early processing tests showed that basing the FluxNet calculations on APP-x COT overestimate SWD over sea ice, whereas the CLARA-A2 COT produced an underestimated SWD. To achieve a reasonable quality in FluxNet-CLARA SWD, we thus chose to base our FluxNet calculations on CLARA-A2 ice and liquid cloud optical thickness (COT) that has been "nudged" toward APP-x by half of their difference (APP-x total COT was weighted by relative fraction of liquid clouds (CPH) in CLARA-A2). The magnitude of the correction is admittedly arbitrary and based solely on the assessment of relative radiative flux retrieval performance in our testing. We used the logarithmic daily mean COT from CLARA as it conserves radiative fluxes better than the arithmetic mean; CERES SYN1deg cloud optical thicknesses are also based on logarithmic averaging.

3.2. Intercomparison and Validation Methods

Having reprojected the CERES and GEWEX data sets to the EASE-2 grid native to CLARA-A2 and APP-x, our first method of intercomparison is to simply provide difference maps of the radiative flux component estimates. It should be kept in mind, however, that the projection and resolution of the CERES and GEWEX

data sets were not designed for polar studies; true near-pole longitudinal resolution is actually substantially less than 1°, meaning that their data will be replicated considerably when reprojecting to the North Pole-centric EASE-2 projection at 25 km resolution. Thus, when comparing to the AVHRR global area coverage-based FluxNet-CLARA estimates whose native resolution is considerably closer to the EASE-2 grid resolution, some sampling-based disagreement in the zonal means of the highest latitudes is expected even for equally performing retrieval algorithms. While zonal mean intercomparisons of the data sets are useful and we do present them here for 70°N–90°N, this constraint must be acknowledged in addition to the ever-present potential for mutually canceling errors improving agreement in zonal mean studies.

We validate the flux component estimates against quality-controlled in situ reference measurements by using root-mean-square error (RMSE), as well as the mean flux difference and its standard deviation as primary quality indicators. Validation of the satellite-based flux estimates requires that the estimates be matched to the in situ data spatially, temporally, and spectrally and that the representativeness of the small-scale in situ measurement be assessed against the coarser satellite estimate.

We achieve the first two requirements by nearest-neighbor collocation of satellite data with in situ measurements averaged over each day. The in situ flux measurements were made with industry-standard Kipp and Zonen and Eppley pyranometers/pyrgeometers whose wavebands differ slightly from the satellite estimate definitions; in situ pyranometers are sensitive to shortwave radiation between 300 and 2800 nm, whereas the shortwave satellite estimates in this study are generally defined between 300 and 4000 nm. However, the calibration of the instruments takes place against reference radiometers that are sensitive to the full shortwave radiation band. The resulting impact on flux estimates over snow may therefore be assumed to be negligible.

The comparison of satellite-based flux estimates, which have large area-integrated footprints, with in situ reference measurements is potentially problematic in cases where atmospheric and surface conditions vary substantially between the in situ and satellite scales. This representativeness issue is expected to be negligible for the validation at Summit Greenland, where the snow cover is uniform and free from melt in 2007. Representativeness of atmospheric conditions over GrIS is more difficult to assess. *Dong et al.* [2008] found good agreement between spatially averaged stratus cloud properties from the Moderate Resolution Imaging Spectroradiometer (MODIS) and surface observations at a snow-free low-latitude site, but the validity of their result for high-latitude snow and ice sites is unknown.

Similarly, representativeness of the validation over Arctic Ocean is challenging to define. On sea ice, surface conditions may vary even over short spatial scales. The surface albedo in particular decreases substantially between snow-covered ice, bare ice, melt ponded ice, and open water [*Perovich and Polashenski*, 2012]. The issue is ameliorated for the Tara validation by the fact that the ice camp drifted mainly in high-concentration multiyear ice, where spatial albedo variation is smaller.

We may also draw an analogue between Tara measurements and the Surface Heat Budget of the Arctic Ocean campaign, where research aircraft were concurrently deployed over the ice camp to compare local and aircraft measurements of cloud, surface, and radiation properties. Both *Doelling et al.* [2001] and *Benner et al.* [2001] argue that representativeness between in situ and satellite measurements for radiative fluxes over Arctic sea ice is good, provided that in situ surface albedo is representative of the larger area. As stated above, this is assumed to be the case at Tara. However, while this analogue points toward good comparability between satellite-based and in situ data, the true representativeness of Tara versus the satellite data sets cannot be quantified for lack of intermediate-resolution data.

Another issue to consider is the impact of reprojecting CERES and GEWEX to a finer-resolution EASE-2 grid for analysis, replicating their data to fill the grid. To investigate, we duplicated both the zonal-mean comparison for June and the Tara validation below by using GEWEX and CERES data at their original 1° resolution, coarsening the FluxNet-CLARA estimates through a linearly weighted radial resampling algorithm. The algorithm resampled FluxNet-CLARA data to the CERES grid by averaging all FluxNet-CLARA flux samples within 75 km of each CERES grid cell center, each sample being given a weight inversely proportional to its distance from the CERES grid cell center. This approach results in a more realistic coarsening of the FluxNet-CLARA data to CERES resolution. We used the standard implementation of the algorithm in the pyresample library.

The results showed that the reprojection has an impact between 5 and 10 W/m^2 in the June zonal mean shortwave fluxes between 80 and 90°N (compared to the analysis at the 25 km grid) but less than 3 W/m^2 in the other cases. The impact on Tara validation RMSE was between 2 and 5 W/m^2, depending on flux component. The 1° resolution zonal mean comparison result table and Tara validation figures are available in Table S1 and Figures S1–S4 in the supporting information.

4. Intercomparison and Validation Results

4.1. Intercomparison

Figure 2 shows a spatially resolved example (June 2007 monthly mean) of the downwelling short- and longwave fluxes (SWD/LWD) from the reprojected data sets. While GEWEX SRB and CERES EBAF SWD fluxes have many common features, they also show disagreements, particularly over sea ice and GrIS. The FluxNet-CLARA SWD flux calculated in this study often falls between the two, with a notable exception over the central part of GrIS. The downwelling longwave fluxes are similar between CERES EBAF and GEWEX SRB over GrIS, yet markedly different over the sea ice zone. FluxNet-CLARA LWD, which is based on the AVHRR-like GEWEX SRB, resembles CERES EBAF closely over sea ice, implying that cloud treatment over sea ice differs between GEWEX SRB and CLARA-A2 more than between CERES EBAF and CLARA-A2. Over land, the FluxNet-CLARA LWD is clearly larger than the other two data sets.

This example, while illustrative, showcases only a single month. We next evaluate the whole GEWEX-CERES overlap period of March 2000 to December 2007 by comparing 10° zonal mean fluxes from 70–80°N and 80–90°N. It should be kept in mind that the zonal bands chosen reflect different surface conditions; 70–80°N is a mixture of land/snow surfaces, open water, and the marginal areas of the sea ice zone, while 80–90°N is the inner part of the Arctic Ocean, with continuous, typically high-concentration sea ice cover with its distinctive seasonal atmospheric and surface cycle [*Perovich and Polashenski*, 2012].

The results are shown in Figure 3. Despite the smoothing effect of zonal mean usage, clear differences between CERES and GEWEX can be seen in nearly all flux components. Most notable differences are in the summer upwelling shortwave flux, pointing to different surface albedo treatments, and in the winter downwelling longwave flux, pointing to different winter cloud property retrievals.

To better assess the summer flux differences, we calculated the zonal mean differences and their standard deviations in the flux components in Table 3 for all 2000–2007 June months. The period 2000–2007 was chosen because it is the longest overlap period between GEWEX SRB and CERES SYN1deg. June was chosen as the example month because it is the month of largest fluxes over the Arctic, highlighting the differences between data sets. Choosing a different summer month changes the shortwave comparison results; in July, most SWD/SWU zonal mean flux differences are smaller both absolutely and relatively. However, the 10° zonal mean differences also hide mutually canceling larger local scale differences within the zone. For example, in July, CERES EBAF SWD is substantially larger than GEWEX SRB over Ellesmere Island and North Greenland, whereas GEWEX SRB SWD is larger over the sea ice close to the Pole. Counterparts of Table 3 for July and August are included in Tables S2 and S3.

In June, the shortwave zonal mean flux differences between GEWEX SRB and CERES EBAF are clearly larger than for longwave fluxes, exceeding 20 W/m^2 for both zonal bands and both up- and downwelling components.

Similarly, differences in shortwave fluxes between FluxNet-CLARA and CERES EBAF in June are larger than longwave fluxes. The disagreement in SWD and SWU is especially pronounced over the sea ice zone (80–90°N), exceeding 40 W/m^2. In general, the shortwave fluxes in FluxNet-CLARA over the highest latitudes are consistently the lowest of the three (Figure 3) and with more spatial variability (Figure 2).

The June shortwave flux differences between GEWEX and FluxNet-CLARA are less pronounced but, interestingly, the longwave fluxes disagree substantially even though both data sets are based on AVHRR observations. The seasonal cycle in LWD over 80–90°N in GEWEX is subdued compared to CERES. LWU in FluxNet-CLARA between 70 and 80°N is clearly largest of the data sets, likely related to the simplification of using a constant emissivity for all surfaces.

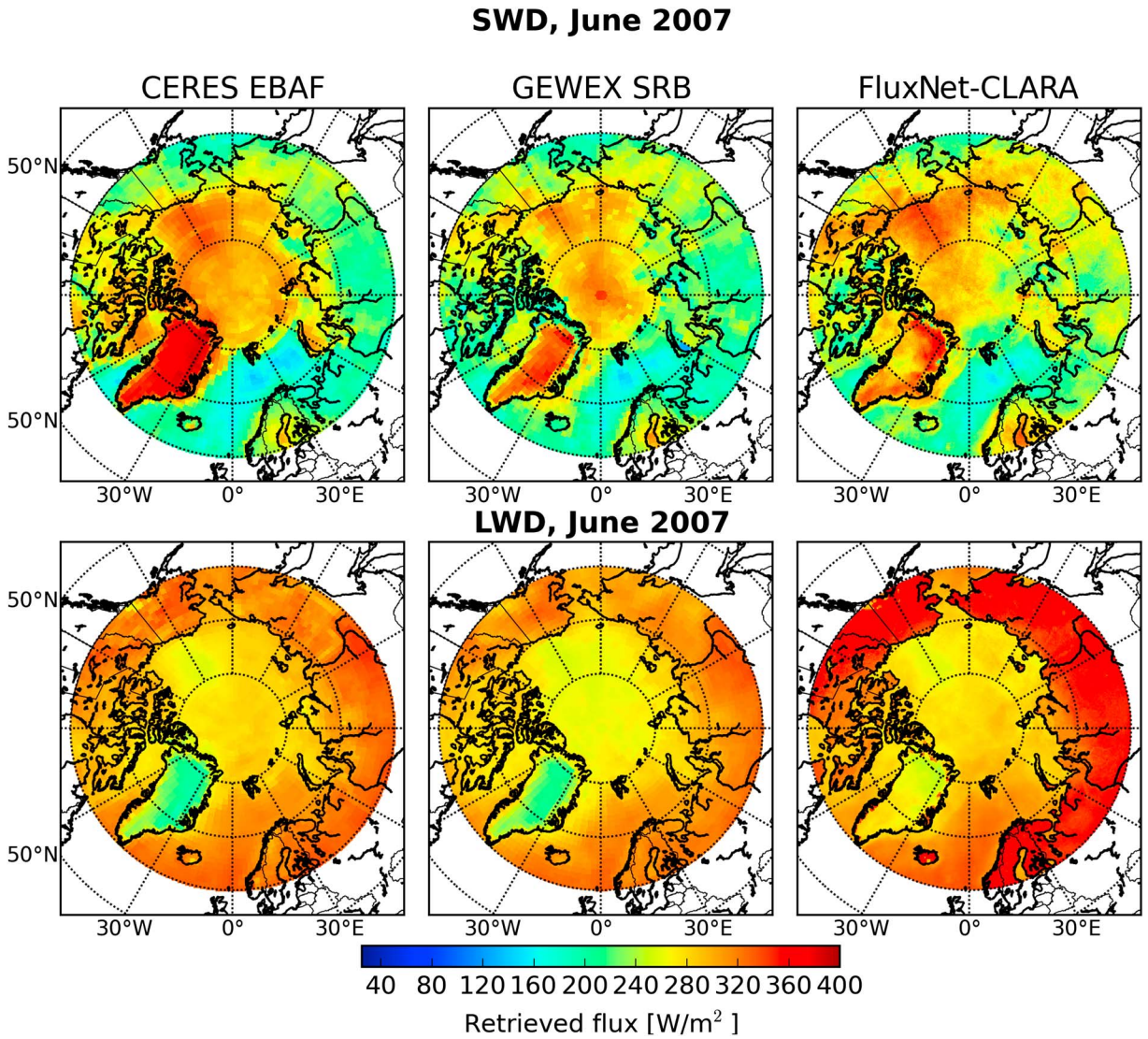

Figure 2. Monthly means of (top row) downwelling shortwave radiation (SWD) and (bottom row) downwelling longwave radiation (LWD) for June 2007 over the Arctic from the (left column) CERES EBAF (edition 3A), (middle column) GEWEX SRB, and (right column) FluxNet-CLARA data sets. All data sets have been reprojected to a common Arctic EASE-2 grid with a latitude cut-off at 60°N.

Our results indicate that the three flux estimates agree only partially. Thus, it is instructive to study the differences in the context of quality-monitored in situ observations, while keeping in mind the spatial representativeness issues discussed in section 3.

4.2. Tara Validation

We begin by comparing daily mean flux estimates over summer 2007 against in situ observations made at the Tara ice camp. The results are shown by flux component in Figures 4–7. The period of comparison is 1 April to 30 September 2007. As CERES EBAF flux data do not contain daily means, we now use CERES SYN1deg daily means instead. As shown in Figure 3, the two data sets are usually fairly closely aligned.

For SWD over Tara (Figure 4), CERES appears both least biased and having the smallest scatter. Both GEWEX SRB and FluxNet-CLARA are prone to large day-to-day variation in retrieval accuracy, with GEWEX SRB exhibiting a positive bias throughout the summer and FluxNet-CLARA, in turn, showing a negative bias in early summer. For both CERES and GEWEX it should be kept in mind that their original data are very coarse over the highest latitudes where Tara drifted, thus contributing uncertainty into the comparability of the

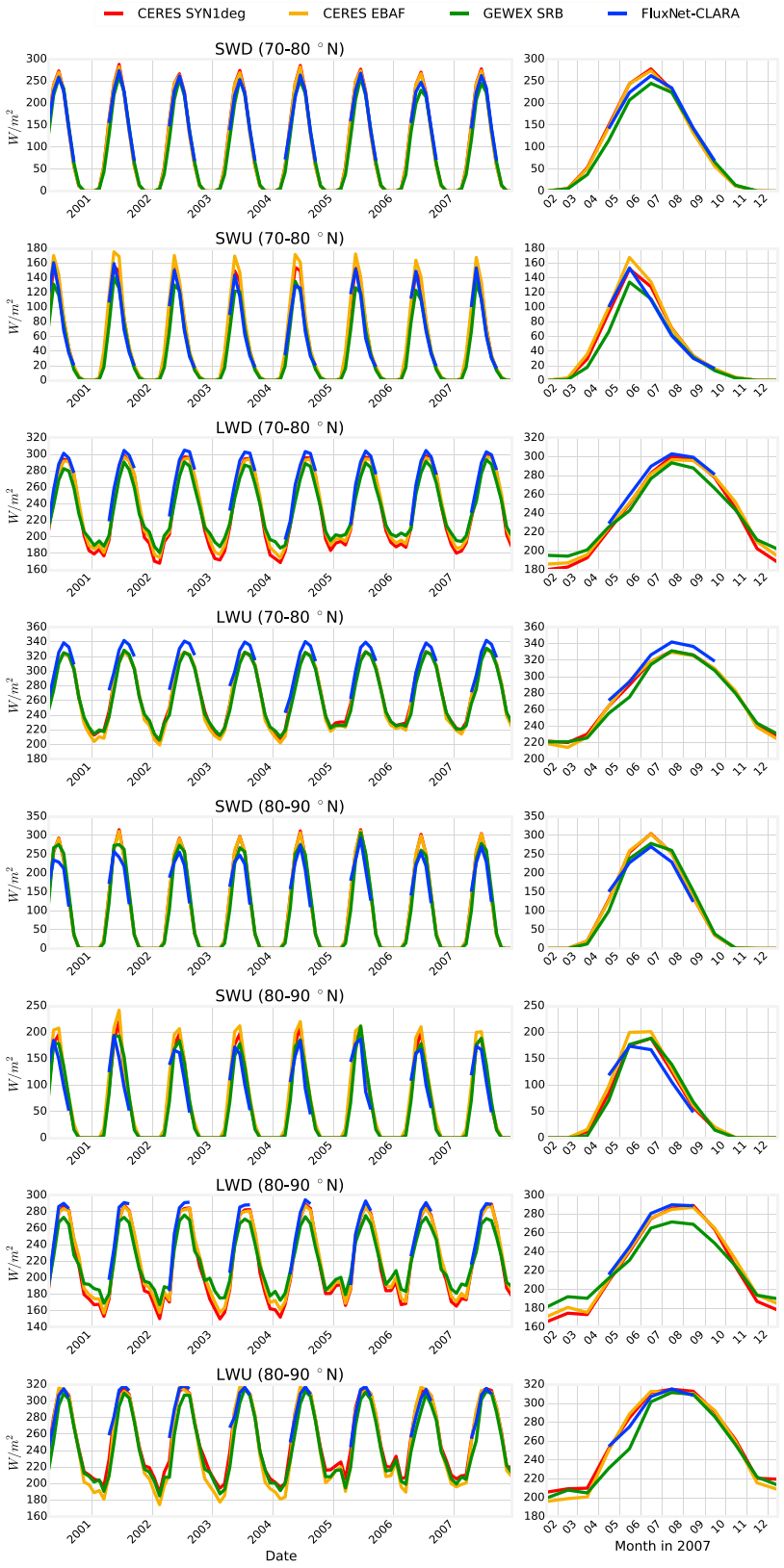

Figure 3. Monthly mean radiative flux estimates between March 2000 and December 2007 from CERES SYN1deg/EBAF, GEWEX SRB, and FluxNet-CLARA averaged for the area between (top) 70–80°N and (bottom) 80–90°N. The subplots on the right show a magnified view of the flux estimates for 2007.

Table 3. Mean Differences in Surface Radiative Energy Budget Components From CERES, GEWEX, and FluxNet-CLARA, Calculated From June Monthly Mean Estimates Between 2000 and 2007[a]

Mean Difference (SD) of SRB Component [W/m^2]	CERES EBAF-GEWEX SRB	CERES EBAF-FluxNet-CLARA	GEWEX SRB-FluxNet-CLARA
Latitude Band 70–80°N, June 2000–2007			
SWD	21.65 (10.05)	11.63 (6.25)	−10.03 (6.60)
SWU	28.33 (9.85)	27.82 (8.28)	−0.51 (4.48)
LWD	7.66 (2.57)	−9.71 (1.40)	−17.37 (2.74)
LWU	5.65 (3.23)	−10.83 (3.83)	−16.48 (2.43)
Latitude Band 80–90°N, June 2000–2007			
SWD	24.40 (11.29)	43.31 (16.11)	18.91 (14.66)
SWU	26.78 (14.96)	48.16 (20.18)	21.37 (11.66)
LWD	10.37 (3.53)	−6.72 (3.58)	−17.09 (1.60)
LWU	17.60 (7.7)	3.78 (8.03)	−13.82 (5.14)

[a]Differences and their standard deviations (in parentheses) calculated for latitude bands 70–80°N and 80–90°N.

measurements. Also, validation against late September SWD measurements should not be given high significance because of the very low flux density.

Figure 5 shows the validation for SWU. The patterns understandably follow those of SWD, although here CERES also appears to underestimate the in situ flux, with the RMSE being comparable between the three estimates. For explaining the SWU differences, it is critical to both examine the surface albedo treatment in the three data sets and assess the impact of differing spatial resolution in the albedo of the in situ site and the satellite estimates. Figure 8a shows the surface albedo of the satellite estimates over Tara versus the in situ measurements. The surface albedo of the satellite data sets shown here is calculated from the all-sky

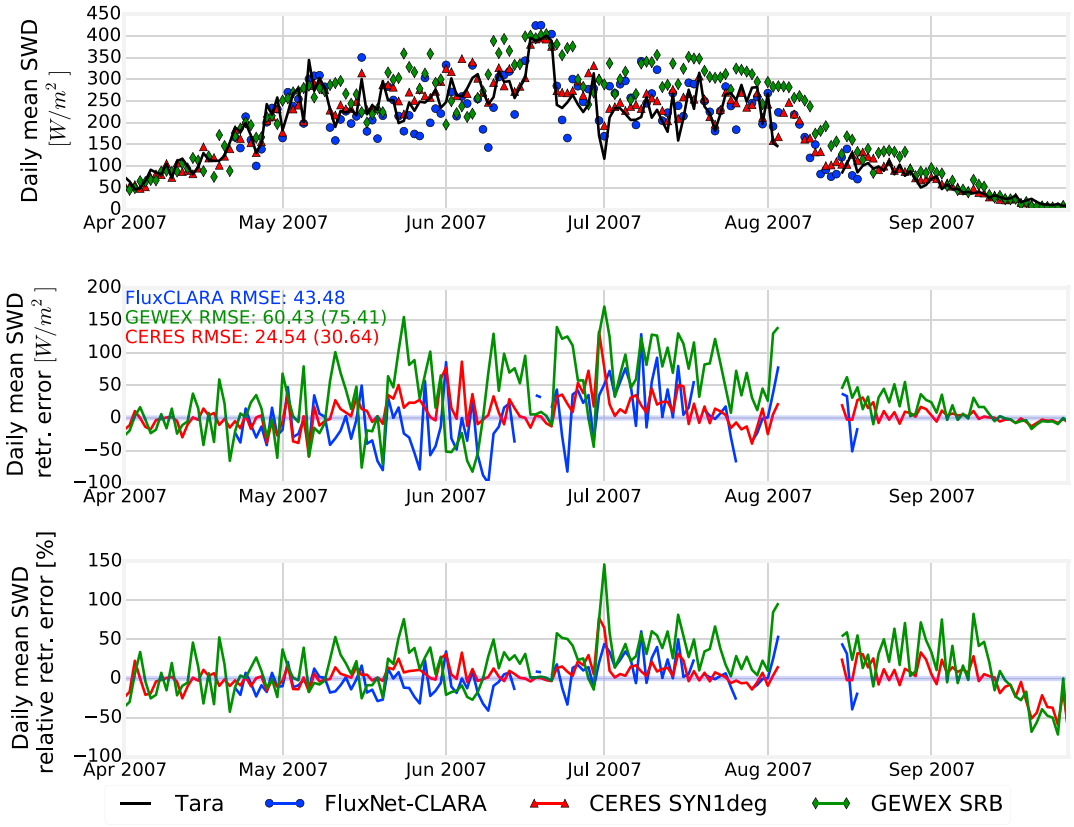

Figure 4. Validation of daily mean SWD flux estimates from CERES SYN1deg, GEWEX SRB, and FluxNet-CLARA against in situ SWD observations made at Tara ice camp during polar summer 2007. The RMSE in parentheses is for a subset corresponding to the FluxNet-CLARA temporal coverage. (top) Retrieved and measured daily mean fluxes. (middle) Satellite-based retrieval errors versus in situ. (bottom) Relative retrieval errors.

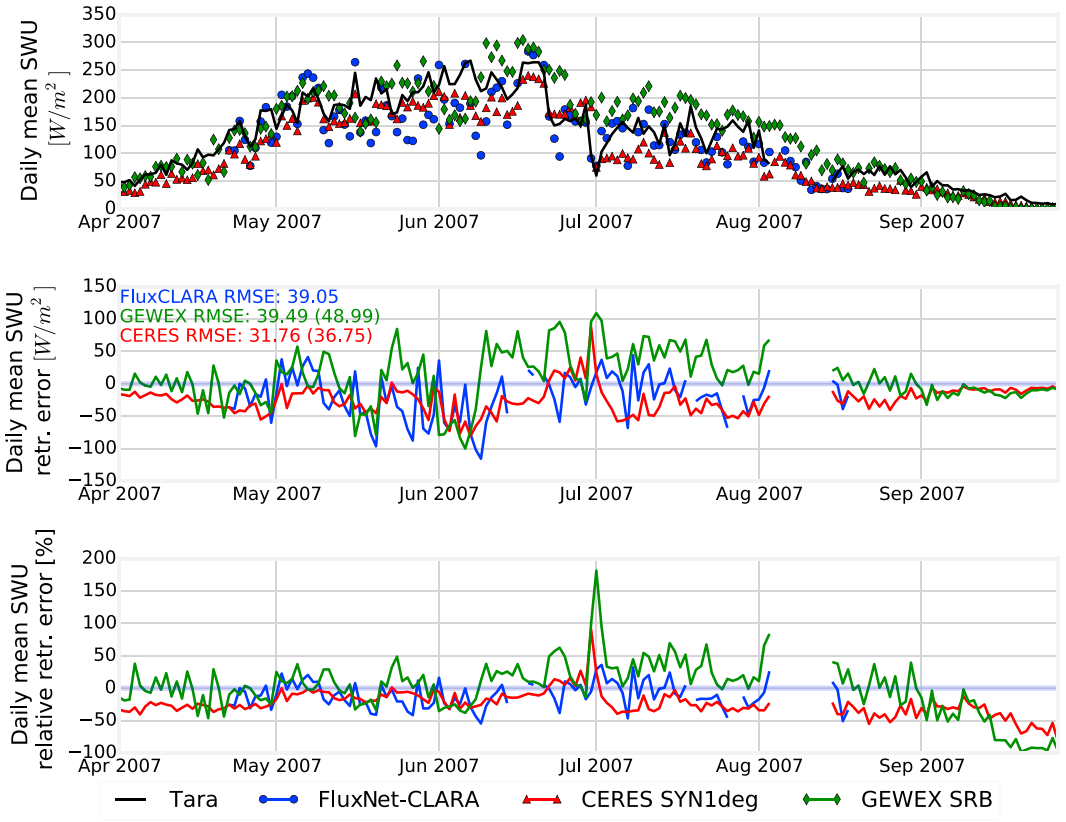

Figure 5. Validation of daily mean SWU flux estimates from CERES SYN1deg, GEWEX SRB, and FluxNet-CLARA against in situ SWU observations made at Tara ice camp during polar summer 2007. The RMSE in parentheses is for a subset corresponding to the FluxNet-CLARA temporal coverage. (top) Retrieved and measured daily mean fluxes. (middle) Satellite-based retrieval errors versus in situ. (bottom) Relative retrieval errors.

up- and downwelling shortwave fluxes. The surface albedo of FluxNet-CLARA, which is from the APP-x data set, and the albedo of GEWEX SRB are relatively well aligned despite the difference in their spatial resolution. In contrast, the albedo of CERES SYN1deg is substantially lower. In particular, the early summer albedo in CERES SYN1deg appears negatively biased, as the snow cover of the sea ice zone has been shown to have a large albedo of ~0.7–0.8 prior to the onset of summer melt in May–June at both large and small scales [*Perovich et al.*, 1998, *Riihelä et al.*, 2013]. The Tara albedo measurements prior to May 12 may also exhibit artificial variation [*Vihma et al.*, 2008].

Figure 6 shows the validation results over Tara for LWD. Both CERES and FluxNet-CLARA slightly overestimate LWD, whereas GEWEX SRB shows a slight underestimation. All three estimates are generally with 10% of the in situ measurement.

The validation for LWU (Figure 7) must be considered inconclusive because of the exclusion of most midsummer measurements. The available data indicate that all three estimates perform with similar accuracy, although CERES LWU appears to saturate at approximately 315 W/m^2 during midsummer, while both GEWEX SRB and FluxNet-CLARA retain a more realistic intraseasonal variation.

As the downwelling fluxes are primarily a function of local cloud properties, we compared CERES SYN1deg daily mean cloud property retrievals (based on MODIS) to the FluxNet-CLARA daily mean cloud properties from the AVHRR-based CLARA-A2 data set. The comparison is shown in Figure 8. Cloud optical thickness (COT; Figure 8) is clearly larger in CLARA-A2 than in CERES/Moderate Resolution Imaging Spectroradiometer (MODIS) during early and midsummer. Total CLARA-A2 cloud optical thickness (COT) is calculated as the weighed sum of liquid and ice cloud COT. On the other hand, the effective droplet radius of liquid clouds is very similar between the two (Figure 8c), yet the effective ice crystal radius of ice clouds is substantially larger and more variable in CERES/MODIS compared to that in CLARA-A2 (Figure 8d).

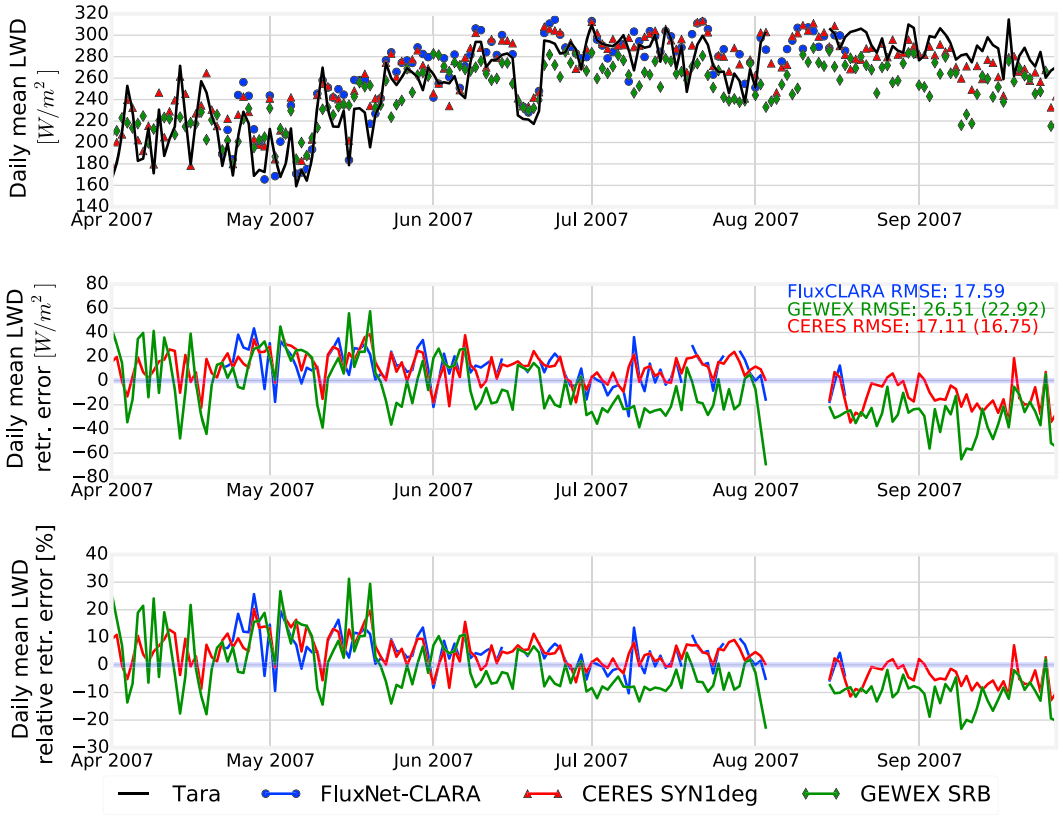

Figure 6. Validation of daily mean LWD flux estimates from CERES SYN1deg, GEWEX SRB, and FluxNet-CLARA against in situ LWD observations made at Tara ice camp during polar summer 2007. The RMSE in parentheses is for a subset corresponding to the FluxNet-CLARA temporal coverage. (top) Retrieved and measured daily mean fluxes. (middle) Satellite-based retrieval errors versus in situ. (bottom) Relative retrieval errors.

4.3. SURE07 Validation

The results of the validation of the flux components against SURE07 in situ measurements are shown in Figure 9. Similar to the results over sea ice, CERES SYN1deg generally follows the in situ measurements closely, with GEWEX SRB and FluxNet-CLARA displaying larger variation in retrieval accuracy. In contrast to the sea ice results, however, here FluxNet-CLARA clearly underestimates the shortwave fluxes and overestimates longwave fluxes. Also, GEWEX and CERES shortwave fluxes seem to be more closely aligned over the ice sheet than over sea ice, although the limited study period must be taken into account.

What causes the increased retrieval errors in FluxNet-CLARA over the ice sheet? To investigate, we compared cloud and surface properties between FluxNet-CLARA and CERES SYN1deg and GEWEX SRB for surface albedo and cloud fraction only. The comparison is shown in Figure 10. Surface albedo between the data sets is highly similar and cannot explain the differences in flux estimation. The comparison of cloud fraction reveals that CERES and FluxNet-CLARA are alike, while GEWEX SRB cloud fraction is very low consistently throughout the study period (secondary axis in Figure 10a). Contrasting with this, retrieved COT between CERES and FluxNet-CLARA is very different, with FluxNet-CLARA COT often being more than 4 times larger and substantially more variable. It should also be kept in mind that the FluxNet-CLARA COT shown is the estimate based on CLARA-A2 but nudged toward APP-x, i.e., having ameliorated the largest overestimations.

Compared to the substantial disagreement in COT, effective radii of liquid and ice clouds differ far less but still notably (Figures 10c and 10d). The difference in ice crystal effective radius is on the same order of magnitude as over sea ice, while the difference in liquid cloud r_{eff} is now both larger and more variable.

The results raise some interesting questions. Clearly, the underestimated SWD/SWU and overestimated LWD fluxes in FluxNet-CLARA over the ice sheet can be linked to cloud microphysical properties. The very large

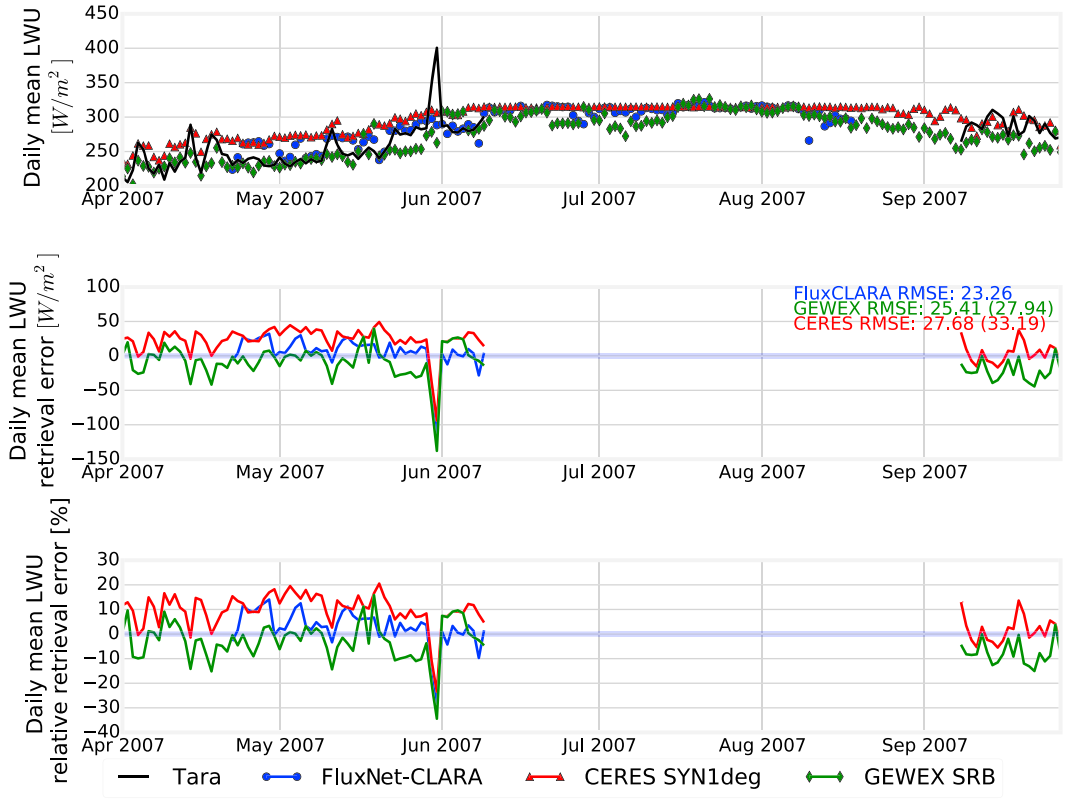

Figure 7. Validation of daily mean LWU flux estimates from CERES SYN1deg, GEWEX SRB, and Fluxnet-CLARA against in situ LWU observations made at Tara ice camp during polar summer 2007. LWU fluxes were not included at Tara between 9 June and 8 September. The RMSE in parentheses is for a subset corresponding to the FluxNet-CLARA temporal coverage. (top) Retrieved and measured daily mean fluxes. (middle) Satellite-based retrieval errors versus in situ. (bottom) Relative retrieval errors.

COT will both suppress SWD and enhance LWD, which is what the validation shows. Also, since cloud fraction between CERES and CLARA-A2 is very similar, incorrect division of clear-sky and cloudy-sky flux calculations cannot be the cause of the large FluxNet-CLARA biases. The large difference in estimated cloud fraction between GEWEX and CERES/FluxNet-CLARA given the reasonably good flux estimation results for GEWEX does also raise the question of how the very low cloudiness over the ice sheet in GEWEX is compensated for to produce generally nonbiased flux estimates. To answer some of these questions, we performed a sensitivity analysis by using CERES surface and cloud properties to drive the FluxNet calculations, its results being shown in the following section.

5. Sensitivity Study

The expected uncertainty in satellite retrievals of the surface radiation budget was assessed by *Key et al.* [1997]. They found that the retrievals are particularly sensitive to errors in certain cloud and surface parameters. In particular, surface albedo, skin temperature, and cloud optical thickness are the sources of largest uncertainty in net flux estimates. Also, because of the persistent cloudiness over the Arctic during summer, the total uncertainty tends to be weighted more by the cloud properties than surface properties.

Given the large discrepancies observed in this study between the cloud and surface properties from AVHRR versus those from CERES/MODIS, we conducted an experiment where FluxNet was run over the Tara site (i.e., Arctic Ocean) by using cloud and surface properties from CERES/MODIS, and then the inputs were changed, one by one, to those from CLARA-A2/APP-x to test the impact on retrieved fluxes. The various flux estimates were then compared to Tara in situ flux measurements by using RMSE as the quality metric. The results of the experiment are shown in Table 4. AOD, T_{surf}, O_3, emissivity, cloud water contents, and TPW inputs were retained from the original FluxNet run.

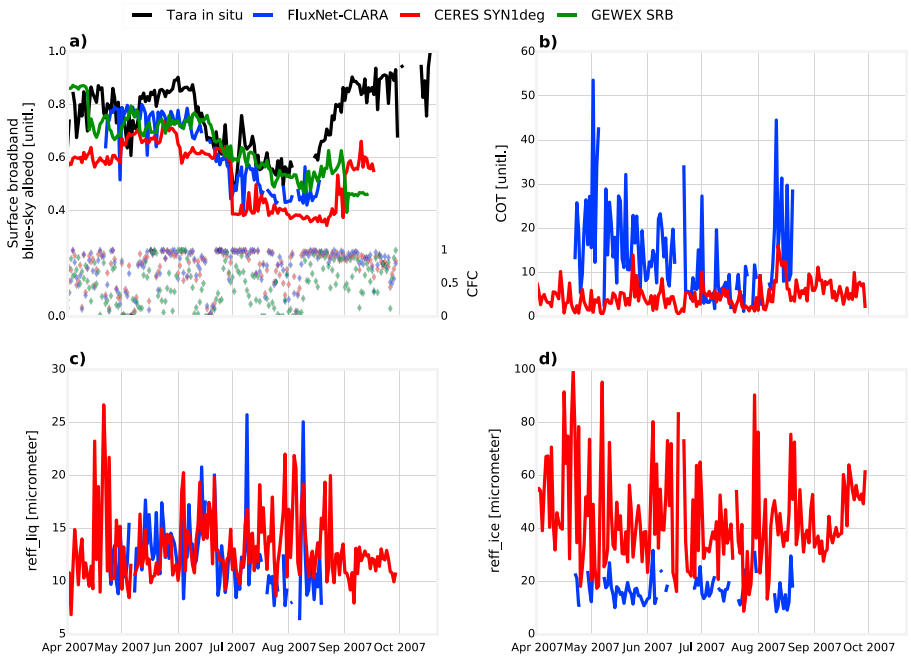

Figure 8. Surface and atmospheric properties over Tara ice camp as retrieved by FluxNet-CLARA/APP-x, CERES SYN1deg, and GEWEX SRB (albedo and cloud fraction only). (a) Lines indicate surface albedo from FluxNet-CLARA versus CERES SYN1deg and GEWEX SRB (all calculated from SW fluxes) and in situ measurements. The diamond markers indicate daily mean cloud fraction. (b) Cloud optical thickness. (c) Liquid cloud droplet effective radius in micrometers. (d) Ice cloud crystal effective radius in micrometers.

We observe that the retrieval quality of FluxNet-based flux estimates is comparable to the original CERES SYN1deg, if the cloud and surface properties in Table 4 are taken from CERES/MODIS. It is also clear that an overestimation of COT in AVHRR-based retrievals worsens the retrieval quality significantly and is the single greatest source of error in SWD and SWU. The magnitude of retrieval error increase is roughly in line with that shown by *Key et al.* [1997]. The results also imply that the CERES SYN1deg surface albedo over Tara is underestimated, as changing that to the APP-x albedo improves SWU RMSE significantly.

For the longwave fluxes, the COT overestimation remains the single largest source of uncertainty. While changing the effective radius of ice cloud crystal from CERES/MODIS to AVHRR appeared to improve slightly upon the original CERES SYN1deg quality, we note that RMSE differences of a few watts per square meter fall within the uncertainty envelope of FluxNet-Streamer itself and are thus too small to draw any conclusions from. However, the improvement in LWU RMSE against the original CERES is substantial and appears to originate from the slight overestimation tendency in CERES SYN1deg over early summer and secondarily from an apparent saturation of CERES LWU at approximately 315 W/m^2 over the midsummer.

6. Discussion

Our results show clear differences in both short- and longwave fluxes between CERES, GEWEX, and our experimental FluxNet-CLARA retrievals over the Arctic, resulting from substantial variation in the surface and cloud properties used as input for the flux calculations. The blue-sky surface albedo and cloud optical thickness in particular show the largest disagreement between CERES/MODIS and CLARA-A2/APP-x. Most cloud properties such as COT were not provided in the GEWEX SRB data set and thus were not evaluated here. However, the GEWEX SRB cloud fraction estimate was provided and was found to deviate strongly from both CLARA-A2 and CERES/MODIS for both the Arctic Ocean and Greenland Ice Sheet sites. This suggests that the cloud detection and characterization algorithms are likely quite different even between AVHRR-based data sets.

Based on our findings, it seems clear that the cloud optical thicknesses from AVHRR/CLARA data over the Arctic Ocean and GrIS are overestimated compared to corresponding MODIS retrievals. Challenges in

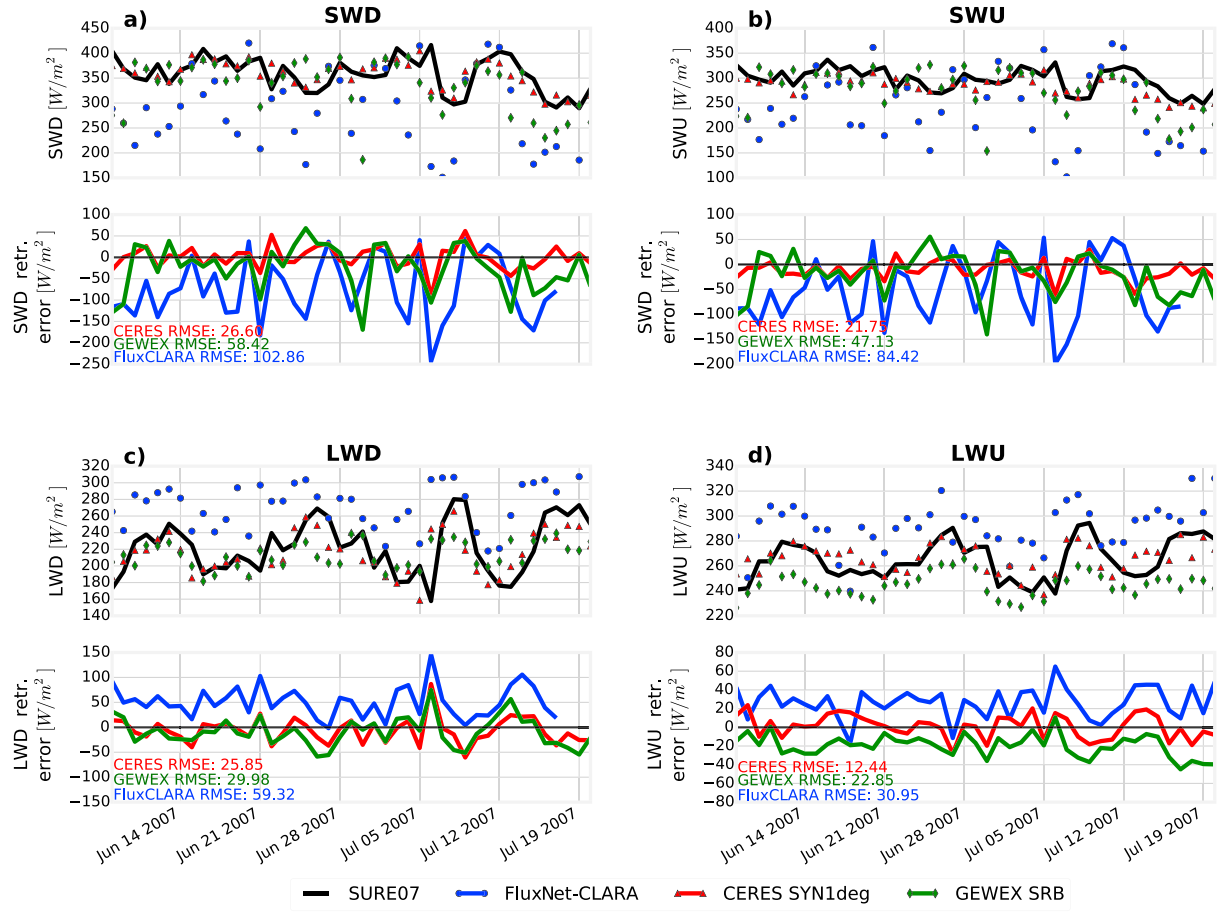

Figure 9. Daily mean flux retrieval errors at Summit station/SURE07 campaign. The red triangle indicates CERES SYN1deg, the green diamond indicates GEWEX SRB, and the blue circle indicates FluxNet-CLARA. In situ observations are marked with a black thick line. The lower subplots of each flux component show daily mean flux retrieval error of the estimates. (a) SWD. (b) SWU. (c) LWD. (d) LWU.

AVHRR-based COT determination over the Arctic Ocean have been reported previously by *Xiong et al.* [2002], and our data confirm those results. The problem is that in the AVHRR visible/near-infrared channels 1 (0.63 μm) and 2 (0.86 μm), thick clouds and ice-covered surfaces have comparable reflectances. Therefore, these bright surfaces are prone to being interpreted as thick clouds in Nakajima-King-type retrievals. On MODIS an additional channel (band 5 at 1.24 μm) is available, in which ice-covered surfaces have a much lower albedo and exhibit considerably more contrast with clouds. Yet the question of determining the correct range and seasonal cycle (if any) for COT over the Arctic Ocean remains elusive due to missing ground truth. A potential solution to ascertaining the valid range of COT over sea ice and improving the AVHRR algorithm could be to collocate CALIPSO/Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) lidar-based COT measurements [*Winker et al.*, 2009] to near-simultaneous AVHRR COT estimates to investigate the retrievals in detail. Despite potential issues with CALIOP COT retrievals of liquid clouds [*Winker et al.*, 2009], this prospect is not only interesting but also outside the scope of the current study. Lidar-based COT measurements cannot reliably characterize thick clouds; therefore, the inclusion of Cloudsat data is also advisable.

The results also indicate that the CERES/MODIS sea ice surface albedo is likely underestimated. While some of the difference can be explained by the different spatial resolutions of the data sets, it is well known that the multiyear sea ice zone, where Tara ice camp drifted, has a continuous and optically thick snow cover in the early spring and summer, having an albedo in excess of 0.7 [*Perovich et al.*, 2002]. Combined with the typically high sea ice concentration of the region poleward of 85°N, it is logical to conclude that even the large-scale surface albedo of the Tara area should be above 0.7 in the April–June period. Both GEWEX SRB and APP-x albedo are in that range, whereas CERES SYN1deg albedo varies between 0.6 and 0.7. Furthermore, when

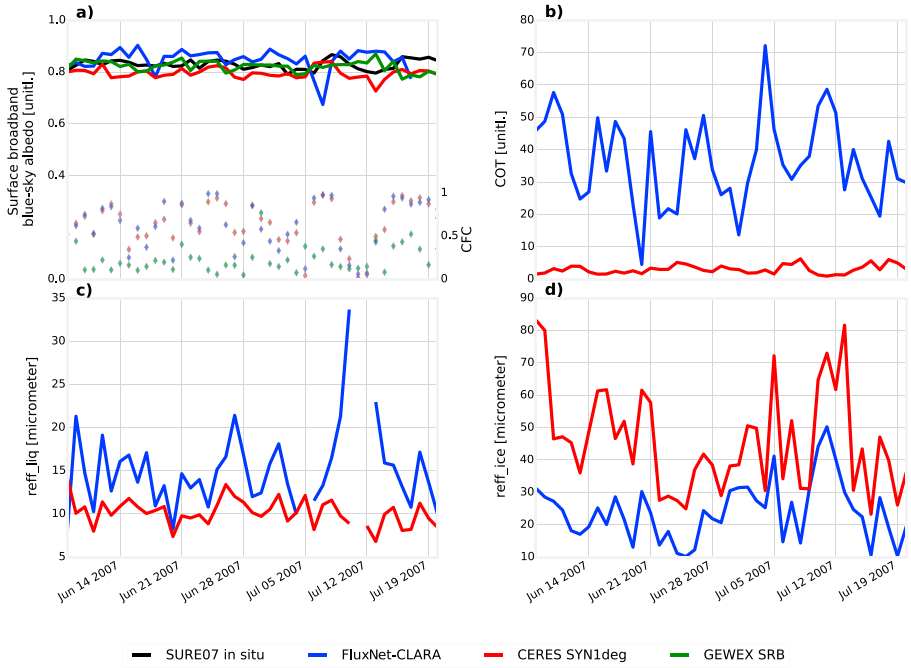

Figure 10. As in Figure 8, but over the Summit ice camp during the SURE07 campaign. (a) Surface albedo and cloud fraction. (b) Cloud optical thickness. (c) Liquid cloud droplet effective radius in micrometers. (d) Ice cloud crystal effective radius in micrometers.

CERES albedo was replaced with APP-x albedo in our sensitivity study, the agreement in SWU versus Tara in situ observations improved. It should be noted that CERES EBAF SWU is typically larger than CERES SYN1deg SWU poleward of 80°N for some years between 2000 and 2007 (Figure 3)—thus partially ameliorating the albedo underestimation.

The temporal coverage of this study was limited to a single summer season by choice to ensure that the in situ data were from a continuously monitored high-quality source. Some prior long-term validation studies on surface radiative fluxes have been performed [*Liu et al.*, 2005, *Dong et al.*, 2010, *Zhang et al.*, 2013] on a limited number of Arctic locations, using primarily high-latitude Baseline Surface Radiation Network [*Ohmura et al.*, 1998] sites. The most recent longer term in situ measurement data sets from programs like the Programme for Monitoring the Greenland Ice Sheet [*van As et al.*, 2011] do not match the joint CERES-GEWEX overlap period of 2000–2007 and were thus inapplicable. Also, the in situ data needed to contain all four flux components in order to be a valid reference for a full SRB study, invalidating some measurement sites in the Arctic. Based on our findings, there is a clear need for long-term multidata set

Table 4. Sensitivity of FluxNet-Based Flux Estimates to Using CLARA/APP-x Input Versus Using CERES SYN1deg Input for Surface and Cloud Properties Over Tara Ice Camp[a]

Parameter From CLARA/APP-x	SWD RMSE/Tara	SWU RMSE/Tara	LWD RMSE/Tara	LWU RMSE/Tara
ALB	38.04	**23.95**	14.46	23.83
COT and CPH	**50.00**	**65.37**	**17.01**	24.91
CFC	34.41	40.94	15.07	23.77
CTT	33.45	41.07	15.01	24.05
Reff-liq	33.88	41.49	14.89	24.60
Reff-ice	34.68	42.65	**13.79**	24.28
FluxNet—all inputs from CERES	33.81	40.97	14.67	**23.81**
Original CERES SYN1deg	**30.64**	36.75	16.75	**33.19**

[a]Each row indicates the FluxNet input variable taken from CLARA-A2/APP-x, all others being taken from CERES SYN1deg. The columns indicate RMSE calculated against Tara in situ measurements for the study period; the bottom row indicates the CERES SYN1deg RMSE against Tara in situ measurements. For all calculated RMSE, $N = 111$ days (valid CLARA coverage). All RMSE expressed in W/m^2. Lowest/highest RMSE indicated in bold font.

intercomparison studies over the Arctic in the future, although the limited in situ data availability has to be acknowledged.

The extraterrestrial solar constant used for GEWEX processing was 1367 W/m^2, whereas CERES publications report using 1361 W/m^2 [*Loeb et al.*, 2009]. The solar constant used in FluxNet is based on that in the Streamer model, which is 1354.2 W/m^2 based on the mean Earth-Sun distance for the spectral band of 0.28–4.0 μm. (It would be 1367 W/m^2 for a spectral range of 0.2–5 μm.) These differences in the solar constant could conceivably influence the SWD retrievals based on solar wave band observations. However, because the surface radiative energy fluxes depend on a complex set of cloud, surface, and atmospheric properties, differences in solar constant are not likely to translate into equally large differences in SWD. Also, flux differences between the data sets in this study over the Arctic exceed any differences in the solar constant by a wide margin, making it unlikely that this difference plays a crucial role.

Similarly, the differences in shortwave and longwave band definitions between the data sets (Table 1) are minor and likely of consequence. *Li et al.* [2010] critiqued several current radiative transfer models for typically defining the shortwave band only up to 4 μm, missing or misplacing the 3–4 W/m^2 of shortwave (solar) energy existing beyond that limit. However, the CERES and GEWEX radiative transfer algorithms account for this energy [*Rose and Charlock*, 2002]. Also, the observed differences in shortwave fluxes between data sets and in situ measurements in this study were typically much larger. We therefore conclude that while this effect may have some minor impact, it cannot explain the results.

As the spatial resolutions of CERES and GEWEX are considerably coarser than the FluxNet inputs CLARA-A2 and APP-x, one may question if the flux differences are not simply effects of different spatial scales. To investigate, we recomputed the validation of FluxNet-CLARA against Tara measurements, smoothing the FluxNet-CLARA estimates with radial weight resampling by using a radius of 100 km. The RMSE against in situ improved by 9 W/m^2 in SWD and by 5 W/m^2 in SWU, whereas the longwave flux RMSEs changed by less than 1 W/m^2. The higher dependence on spatial scale in shortwave fluxes is logical, given the broader spatial dynamic in SWD over the Arctic Ocean (Figure 2). Still, the scale dependence does not appear strong enough to account for the majority of the differences found in this study. This logic is also supported by the results of our sensitivity study, where FluxNet performance was comparable to CERES against Tara measurements using the CERES 1° mean cloud/surface properties as input (Table 4).

The close agreement between original CERES and FluxNet estimates given CERES/MODIS cloud and surface parameters as FluxNet input also justifies our processing approach (of using the daily mean atmospheric and surface state with hourly traversing Sun over each grid cell). Based on the validation and intercomparison results, the use of a constant emissivity clearly affects the quality of LWU estimate, but our other constant parameters seem to be sufficiently well fitted to prevailing conditions. The FluxNet estimates were not computed for the polar winter period, where challenges in cloud identification and characterization could impact flux estimation significantly [e.g., *Schweiger et al.*, 1999; *Jun et al.*, 2016; *Karlsson et al.*, 2016].

7. Conclusions

We have evaluated the surface radiative energy budget components from CERES SYN/EBAF and GEWEX SRB against each other, in situ observations from the Arctic Ocean (Tara ice camp), and Greenland Ice Sheet (SURE07 campaign), as well as the experimental FluxNet-CLARA data set created for this study. Along with the flux components, we have also compared the surface and cloud properties central to flux estimation. The main temporal focus of the study was in the polar summer 2007. The main findings of this study are as follows:

1. CERES SYN1deg fluxes show the smallest root-mean-square error against in situ fluxes over both the sea ice and the ice sheet. The improvement over AVHRR-based fluxes is logical, given the enhanced spectral information available for MODIS cloud retrievals, as well as the high calibration accuracy of CERES itself. GEWEX SRB and FluxNet-CLARA flux retrievals performed comparably well. Differences in spatial resolution of the compared data can explain some of the differences, but not their majority.
2. The CERES SYN1deg surface albedo over sea ice is likely underestimated. Flux estimates computed here with APP-x albedo and CERES/MODIS clouds showed a better agreement with Tara in situ measurements than the original CERES data. Given how the upwelling shortwave flux in CERES EBAF data is typically

higher than CERES SYN1deg over sea ice, the underestimation in EBAF fluxes is likely ameliorated. APP-x and GEWEX SRB albedo were found to be of comparable magnitude.

3. Cloud optical thickness (COT) in CLARA-A2 over the sea ice and the ice sheet is substantially overestimated, while cloud fraction and the effective radius of water droplets in liquid clouds agreed very well with CERES/MODIS retrievals. Ice crystal effective radius in CLARA-A2 was consistently smaller than CERES/MODIS for both sites. The overestimated COT causes the FluxNet-CLARA downwelling shortwave/longwave fluxes to be underestimated/overestimated, respectively.
4. CERES SYN1deg upwelling longwave flux over sea ice appears to saturate to (does not exceed) ~315 W/m^2 during midsummer. Given that neither FluxNet-CLARA nor GEWEX SRB shows a similar behavior, the lack of variability in CERES SYN1deg is likely an issue in the data generation rather than a physical phenomenon. Given that near-surface air temperatures at Tara were stable at close to 0°C for the midsummer [*Vihma et al.*, 2008], the actual LWU of the period was possibly quite close to the 315 W/m^2 level, suggesting low retrieval errors for all data sets. However, this remains conjecture due to the lack of in situ reference data. FluxNet-CLARA upwelling longwave fluxes were clearly overestimated over land areas due to simplifications made during processing, e.g., the assumption of a fixed surface emissivity.

The differences in flux estimates also showed a substantial spatial variability over the Arctic. Maps comparing the data sets for April–September 2007 are in the supporting information.

This study has shown that the state-of-the-art radiative flux estimates over the Arctic have clear differences, driven by their widely varying cloud and surface properties. In particular, the correct estimation of cloud optical thickness over snow and ice from AVHRR observations remains the single largest obstacle in creating high-quality SRB flux estimates from CLARA-A2/APP-x data with the FluxNet radiative transfer calculator. Also, the question of determining the correct COT over sea ice and snow has not been studied sufficiently to date. We propose the use of lidar cloud retrievals from, e.g., CALIPSO/CALIOP, as a potential source of unbiased, independent reference data for thin clouds, against which improvements in COT retrievals can be made. Inclusion of, e.g., Cloudsat data, as a reference for thick clouds is also advised.

Further, while the results here are illustrative, the spatial and temporal scopes were limited to a single year. The paucity of quality-controlled in situ flux measurement data over sea ice is a limiting factor for expanded studies on this topic, but we nevertheless recommend further investigations and efforts to harmonize the flux estimates of the various data sets and improve the shortcomings identified in this study.

Acknowledgments

The work of A.R. has been funded by the Academy of Finland research project 287399. K.G.K. and J.F.M. acknowledge support by the EUMETSAT CM SAF project. P.K.M. acknowledges funding from Utrecht University and support from C. H. Reijmer and M. R. van den Broeke (Utrecht University). CERES data are from https://ceres.larc.nasa.gov. GEWEX data are from the NASA Langley Research Center Atmospheric Sciences Data Center: https://eosweb.larc.nasa.gov/project/srb/srb_table. CLARA-A2 data are from https://wui.cmsaf.eu. APP-x data are from ftp://data.ncdc.noaa.gov/cdr/appx/. In situ and FluxNet-CLARA data are available from the authors. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision.

References

Benner, T. C., J. A. Curry, and J. O. Pinto (2001), Radiative transfer in the summertime Arctic, *J. Geophys. Res.*, *106*, 15,173–15,183, doi:10.1029/2000JD900422.

Chiacchio, M., J. Francis, and P. Stackhouse Jr. (2002), Evaluation of methods to estimate the surface downwelling longwave flux during Arctic winter, *J. Appl. Meteorol.*, *41*(3), 306–318.

Christensen, M. W., A. Behrangi, T. LEcuyer, N. B. Wood, M. D. Lebsock, and G. L. Stephens (2016), Arctic observation and reanalysis integrated system: A new data product for validation and climate study, *Bull. Am. Meteorol. Soc.*, *97*, 907–915.

Chylek, P., C. K. Folland, G. Lesins, M. K. Dubey, and M. Wang (2009), Arctic air temperature change amplification and the Atlantic Multidecadal Oscillation, *Geophys. Res. Lett.*, *36*, L14801, doi:10.1029/2009GL038777.

Comiso, J. C., C. L. Parkinson, R. Gersten, and L. Stock (2008), Accelerated decline in the Arctic sea ice cover, *Geophys. Res. Lett.*, *35*, L01703, doi:10.1029/2007GL031972.

Curry, J. A., J. L. Schramm, W. B. Rossow, and D. Randall (1996), Overview of Arctic cloud and radiation characteristics, *J. Clim.*, *9*(8), 1731–1764.

Derksen, C., and R. Brown (2012), Spring snow cover extent reductions in the 2008–2012 period exceeding climate model projections, *Geophys. Res. Lett.*, *39*, L19504, doi:10.1029/2012GL053387.

Doelling, D. R., P. Minnis, D. A. Spangenberg, V. Chakrapani, A. Mahesh, S. K. Pope, and F. P. Valero (2001), Cloud radiative forcing at the top of the atmosphere during FIRE ACE derived from AVHRR data, *J. Geophys. Res.*, *106*, 15.

Dong, X., P. Minnis, B. Xi, S. Sun-Mack, and Y. Chen (2008), Comparison of CERES-MODIS stratus cloud properties with ground-based measurements at the DOE ARM Southern Great Plains site, *J. Geophys. Res.*, *113*, D03204, doi:10.1029/2007JD008438.

Dong, X., B. Xi, K. Crosby, C. N. Long, R. S. Stone, and M. D. Shupe (2010), A 10 year climatology of Arctic cloud fraction and radiative forcing at Barrow, Alaska, *J. Geophys. Res.*, *115*, D17212, doi:10.1029/2009JD013489.

Dong, X., B. Xi, S. Qiu, P. Minnis, S. Sun-Mack, and F. Rose (2016), A radiation closure study of Arctic stratus cloud microphysical properties using the collocated satellite-surface data and Fu-Liou radiative transfer model, *J. Geophys. Res. Atmos.*, *121*, 10,175–10,198, doi:10.1002/2016JD025255.

Fu, Q., and K. N. Liou (1992), On the correlated k-distribution method for radiative transfer in nonhomogeneous atmospheres, *J. Atmos. Sci.*, *49*(22), 2139–2156.

Fu, Q., K. N. Liou, M. C. Cribb, T. P. Charlock, and A. Grossman (1997), Multiple scattering parameterization in thermal infrared radiative transfer, *J. Atmos. Sci.*, *54*(24), 2799–2812.

Hori, M., et al. (2006), In-situ measured spectral directional emissivity of snow and ice in the 8–14 μm atmospheric window, *Remote Sens. Environ.*, *100*(4), 486–502.

Jun, S. Y., C. H. Ho, J. H. Jeong, Y. S. Choi, and B. M. Kim (2016), Recent changes in winter Arctic clouds and their relationships with sea ice and atmospheric conditions, *Tellus A*, *68*, 29,130, doi:10.3402/tellusa.v68.29130.

Karlsson, K.-G., et al. (2016), The second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data, Submitted to ACP, 2016.

Kato, S., T. P. Ackerman, J. H. Mather, and E. E. Clothiaux (1999), The k-distribution method and correlated-k approximation for a shortwave radiative transfer model, *J. Quant. Spectrosc. Radiat. Transfer*, *62*(1), 109–121.

Kay, J. E., and T. L'Ecuyer (2013), Observational constraints on Arctic Ocean clouds and radiative fluxes during the early 21st century, *J. Geophys. Res. Atmos.*, *118*, 7219–7236, doi:10.1002/jgrd.50489.

Key, J., and A. J. Schweiger (1998), Tools for atmospheric radiative transfer: Streamer and FluxNet, *Comput. Geosci.*, *24*(5), 443–451.

Key, J., X. Wang, Y. Liu, R. Dworak, and A. Letterly (2016), The AVHRR Polar Pathfinder climate data records, *Remote Sens.*, *8*(3), 167, doi:10.3390/rs8030167.

Key, J. R., A. J. Schweiger, and R. S. Stone (1997), Expected uncertainty in satellite-derived estimates of the surface radiation budget at high latitudes, *J. Geophys. Res.*, *102*, 15,837.

Kuipers Munneke, P., M. R. Van den Broeke, C. H. Reijmer, M. M. Helsen, W. Boot, M. Schneebeli, and K. Steffen (2009), The role of radiation penetration in the energy budget of the snowpack at Summit, Greenland, *Cryosphere*, *3*(2), 155–165.

Letterly, A., J. Key, and Y. Liu (2016), The influence of winter cloud on summer sea ice in the Arctic, 1983–2013, *J. Geophys. Res. Atmos.*, *121*, 2178–2187, doi:10.1002/2015JD024316.

Li, J., C. L. Curry, Z. Sun, and F. Zhang (2010), Overlap of solar and infrared spectra and the shortwave radiative effect of methane, *J. Atmos. Sci.*, *67*(7), 2372–2389.

Liu, J., J. A. Curry, W. B. Rossow, J. R. Key, and X. Wang (2005), Comparison of surface radiative flux data sets over the Arctic Ocean, *J. Geophys. Res.*, *110*, C02015, doi:10.1029/2004JC002381.

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong (2009), Toward optimal closure of the Earth's top-of-atmosphere radiation budget, *J. Clim.*, *22*(3), 748–766.

Maslanik, J., J. Key, C. W. Fowler, T. Nguyen, and X. Wang (2001), Spatial and temporal variability of satellite-derived cloud and surface characteristics during FIRE-ACE, *J. Geophys. Res.*, *106*(15), 233.

Minnis, P., et al. (2011), CERES edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data—Part I: Algorithms, *IEEE Trans. Geosci. Remote Sens.*, *49*(11), 4374–4400.

Nakajima, T., and M. S. King (1990), Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory, *J. Atmos. Sci.*, *47*, 1878–1893.

NASA Ozone Watch (2016), https://ozonewatch.gsfc.nasa.gov/meteorology/NH.html [Accessed on Dec 29, 2016].

Ohmura, A., et al. (1998), Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research, *Bull. Am. Meteorol. Soc.*, *79*(10), 2115–2136.

Perovich, D. K., C. S. Roesler, and W. S. Pegau (1998), Variability in Arctic sea ice optical properties, *J. Geophys. Res.*, *103*, 1193–1208.

Perovich, D. K., T. C. Grenfell, B. Light, and P. V. Hobbs (2002), Seasonal evolution of the albedo of multiyear Arctic sea ice, *J. Geophys. Res.*, *107*(C10), 8044, doi:10.1029/2000JC000438.

Perovich, D. K., and C. Polashenski (2012), Albedo evolution of seasonal Arctic sea ice, *Geophys. Res. Lett.*, *39*, L08501, doi:10.1029/2012GL051432.

Pinker, R. T., and I. Laszlo (1992), Modeling surface solar irradiance for satellite applications on a global scale, *J. Appl. Meteorol.*, *31*(2), 194–211.

Pistone, K., I. Eisenman, and V. Ramanathan (2014), Observational determination of albedo decrease caused by vanishing Arctic sea ice, *Proc. Natl. Acad. Sci. U.S.A.*, *111*(9), 3322–3326.

Riihelä, A., V. Laine, T. Manninen, T. Palo, and T. Vihma (2010), Validation of the Climate-SAF surface broadband albedo product: Comparisons with in situ observations over Greenland and the ice-covered Arctic Ocean, *Remote Sens. Environ.*, *114*(11), 2779–2790.

Riihelä, A., T. Manninen, and V. Laine (2013), Observed changes in the albedo of the Arctic sea-ice zone for the period 1982–2009, *Nat. Clim. Change*, *3*(10), 895–898.

Rose, F., and T. P. Charlock (2002), New Fu-Liou code tested with ARM Raman lidar aerosols and CERES in pre-CALIPSO sensitivity study, in *11th Conf. on Atmospheric Radiation, Am. Meteor. Soc.*, P.4.8., Ogden, Utah. [Available at http://ams.confex.com/ams/pdfpapers/42757.pdf.]

Rutan, D. A., S. Kato, D. R. Doelling, F. G. Rose, L. T. Nguyen, T. E. Caldwell, and N. G. Loeb (2015), *J. Atmos. Oceanic Technol.*, *32*, 1121–1143, doi:10.1175/JTECH-D-14-00165.1.

Rossow, W. B., and R. A. Schiffer (1999), Advances in understanding clouds from ISCCP, *Bull. Am. Meteorol. Soc.*, *80*(11), 2261.

Rothrock, D. A., Y. Yu, and G. A. Maykut (1999), Thinning of the Arctic sea-ice cover, *Geophys. Res. Lett.*, *26*, 3469–3472.

Salisbury, J. W., and D. M. D'Aria (1994), Emissivity of terrestrial materials in the 3–5 μm atmospheric window, *Remote Sens. Environ.*, *47*(3), 345–361.

Schweiger, A. J., R. W. Lindsay, J. R. Key, and J. A. Francis (1999), Arctic clouds in multiyear satellite data sets, *Geophys. Res. Lett.*, *26*, 1845–1848.

Shupe, M. D., T. Uttal, and S. Y. Matrosov (2005), Arctic cloud microphysics retrievals from surface-based remote sensors at SHEBA, *J. Appl. Meteorol.*, *44*(10), 1544–1562.

Shupe, M. D., S. Y. Matrosov, and T. Uttal (2006), Arctic mixed-phase cloud properties derived from surface-based sensors at SHEBA, *J. Atmos. Sci.*, *63*(2), 697–711.

Stackhouse, P. W., Jr., S. K. Gupta, S. J. Cox, T. Zhang, J. C. Mikovitz, and L. M. Hinkelman (2011), The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset, *GEWEX News*, *21*(1), 10–12.

Stephens, G. L. (2005), Cloud feedbacks in the climate system: A critical review, *J. Clim.*, *18*(2), 237–273, doi:10.1175/JCLI-3243.1.

Stroeve, J., M. M. Holland, W. Meier, T. Scambos, and M. Serreze (2007), Arctic sea ice decline: Faster than forecast, *Geophys. Res. Lett.*, *34*, L09501, doi:10.1029/2007GL029703.

Tomasi, C., et al. (2007), Aerosols in polar regions: A historical overview based on optical depth and in situ observations. *J. Geophys. Res.*, *112*, D16205, doi:10.1029/2007JD008432.

van As, D., et al. (2011), Programme for monitoring of the Greenland Ice Sheet (PROMICE): First temperature and ablation record, *Geol. Surv. Denmark Greenland Bull.*, *23*, 73–76.

Vihma, T., J. Jaagus, E. Jakobson, and T. Palo (2008), Meteorological conditions in the Arctic Ocean in spring and summer 2007 as recorded on the drifting ice station Tara, *Geophys. Res. Lett.*, *35*, L18706, doi:10.1029/2008GL034681.

Wang, X., and J. Key (2005), Arctic surface, cloud, and radiation properties based on the AVHRR Polar Pathfinder data set. Part I: Spatial and temporal characteristics, *J. Clim.*, *18*(14), 2558–2574.

Wetherald, R. T., and S. Manabe (1975), The effects of changing the solar constant on the climate of a general circulation model, *J. Atmos. Sci.*, *32*(11), 2044–2059.

Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. Louis Smith, and J. E. Cooper (1996), Clouds and the Earth's Radiant Energy System (CERES): An Earth observing system experiment, *Bull. Am. Meteorol. Soc.*, *77*(5), 853–868.

Wielicki, B. A., et al. (1998), Clouds and the Earth's Radiant Energy System (CERES): Algorithm overview, *IEEE Trans. Geosci. Remote Sens.*, *36*(4), 1127.

Winker, D. M., M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young (2009), Overview of the CALIPSO mission and CALIOP data processing algorithms, *J. Atmos. Oceanic Technol.*, *26*(11), 2310–2323.

Xiong, X., D. Lubin, W. Li, and K. Stamnes (2002), A critical examination of satellite cloud retrieval from AVHRR in the Arctic using SHEBA data, *J. Appl. Meteorol.*, *41*(12), 1195–1209.

Zhang, T., et al. (2005), Spatial and temporal variability in active layer thickness over the Russian Arctic drainage basin, *J. Geophys. Res.*, *110*, D16101, doi:10.1029/2004JD005642.

Zhang, T., P. W. Stackhouse, S. K. Gupta, S. J. Cox, J. C. Mikovitz, and L. M. Hinkelman (2013), The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: A systematic quality control, production and application approach, *J. Quant. Spectrosc. Radiat. Transfer*, *122*, 127–140.

Zhang, T., P. W. Stackhouse, S. K. Gupta, S. J. Cox, and J. C. Mikovitz (2015), The validation of the GEWEX SRB surface longwave flux data products using BSRN measurements, *J. Quant. Spectrosc. Radiat. Transfer*, *150*, 134–147.