Comparison of TOA and BOA LW Radiation Fluxes Inferred From Ground-Based Sensors, A-Train Satellite Observations and ERA Reanalyzes at the High Arctic Station Eureka Over the 2002–2020 Period

Yann Blanchard, Jacques Pelon, Christopher Cox, Julien Delanoë, Edwin Eloranta, Taniel Uttal

JGR Atmospheres

RESEARCH ARTICLE
10.1029/2020JD033615

Special Section:
The Arctic: An AGU Joint Special Collection

Comparison of TOA and BOA LW Radiation Fluxes Inferred From Ground-Based Sensors, A-Train Satellite Observations and ERA Reanalyzes at the High Arctic Station Eureka Over the 2002–2020 Period

Yann Blanchard[1], Jacques Pelon[2], Christopher J. Cox[3,4], Julien Delanoë[2], Edwin W. Eloranta[5], and Taniel Uttal[4]

[1]Centre pour l'Étude et la Simulation du Climat à l'Échelle Régionale (ESCER), Université du Québec à Montréal, Montréal, Canada, [2]Laboratoire Atmosphères, Milieux, Observations Spatiales, SU-UVSQ-CNRS, Paris, France, [3]Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, National Oceanic and Atmospheric Administration (NOAA), Boulder, CO, USA, [4]Physical Sciences Division, NOAA, Boulder, CO, USA, [5]Space Science and Engineering Center, University of Wisconsin-Madison, Madison, WI, USA

Abstract This study focuses on the accuracy of longwave radiation flux retrievals at the top and bottom of the atmosphere at Eureka station, Canada, in the high Arctic. We report comparisons between seven products derived from (a) calculations based on a combination of ground-based and space-based lidar and radar observations, (b) standard radiometric observations from the CERES sensor, (c) direct observations at the surface from a broadband radiation station, and (d) the ERA-Interim and ERA5 reanalyzes. Statistical, independent analyses are first performed to look at recurring bias and trends in fluxes at Top and Bottom of the Atmosphere (TOA, BOA). The analysis is further refined by comparing fluxes derived from coincident observations decomposed by scene types. Results show that radiative transfer calculations using ground-based lidar-radar profiles derived at Eureka agree well with TOA LW fluxes observed by CERES and with BOA LW fluxes reference. CloudSat-CALIPSO also shows good agreement with calculations from ground-based sensor observations, with a relatively small bias. This bias is shown to be largely due to low and thick cloud occurrences that the satellites are insensitive to owing to attenuation from clouds above and surface clutter. These conditions of opaque low clouds, cause an even more pronounced bias for CERES BOA flux calculation in winter, due to the deficit of low clouds identified by MODIS. ERA-I and ERA5 fluxes behave differently, the large positive bias observed with ERA-I is much reduced in ERA5. ERA5 is closer to reference observations due to better behavior of low and mid-level clouds and surface temperature.

Plain Language Summary Satellite and reanalysis data sets are widely used for climate and process studies in the Arctic in order to complement sparse ground-based measurements. This study compares ground-based observations of Arctic clouds and longwave fluxes at a Canadian High Arctic station with satellite and reanalysis products. Both statistical and coincident analyses show a good top of the atmosphere agreement, but reveal biases in surface fluxes that are due to the underestimation of the occurrence of low and thick clouds, frequent in the Arctic. The results allow for an evaluation of flux product uncertainties and an assessment of their limitations. The outcomes of this study can be applied over the entire Arctic region and can inform the instrumentation choices at various polar ground-based sites.

Key Points:

- Top-of-atmosphere flux observations, model calculations, and reanalyzes agree within the measurement uncertainty at a High Arctic Station
- Large bottom-of-Atmosphere biases are observed among data sets (reduced for ERA5) and are mainly due to incorrect vertical representation of low opaque clouds
- Ground-based active sensors are essential to complement space-based cloud observations in order to better understand Arctic climate change

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

Y. Blanchard,
yann.blanchard@usherbrooke.ca

Citation:

Blanchard, Y., Pelon, J., Cox, C. J., Delanoë, J., Eloranta, E. W., & Uttal, T. (2021). Comparison of TOA and BOA LW radiation fluxes inferred from ground-based sensors, A-Train satellite observations and ERA reanalyzes at the High Arctic station Eureka over the 2002–2020 period. *Journal of Geophysical Research: Atmospheres*, *126*, e2020JD033615. https://doi.org/10.1029/2020JD033615

Received 28 JUL 2020
Accepted 29 JAN 2021

Author Contributions:

Conceptualization: Yann Blanchard, Jacques Pelon
Formal analysis: Yann Blanchard, Christopher J. Cox, Julien Delanoë
Funding acquisition: Jacques Pelon
Investigation: Yann Blanchard, Jacques Pelon
Methodology: Yann Blanchard, Jacques Pelon

1. Introduction

Interest in the Arctic climate has increased as the effects of global warming have begun to manifest in the region over several decades (IPCC, 2013). These manifestations include increases in surface temperature that are larger than those observed at lower latitudes (Comiso & Hall, 2014; McBean et al., 2005); significant decreases in sea-ice extent and thickness (Lang et al., 2017; Palm et al., 2010; Serreze & Barry, 2011); and changes in Arctic cloud cycle and interactions (Abe et al., 2016; Liu et al., 2012; Sedlar et al., 2011). Large scale meteorological dynamical forcings on a more fragile sea-ice interface impact surface energy budgets and modify ice properties. Transport of aerosols (Ancellet et al., 2014, Igel et al., 2017; Rahn, 1981) and

Project Administration: Jacques Pelon
Resources: Jacques Pelon
Software: Christopher J. Cox
Supervision: Jacques Pelon
Visualization: Yann Blanchard
Writing – original draft: Jacques Pelon, Christopher J. Cox, Julien Delanoë, Edwin W. Eloranta, Taniel Uttal
Writing – review & editing: Jacques Pelon, Christopher J. Cox, Julien Delanoë, Taniel Uttal

larger water vapor intrusions from lower latitudes (Boisvert et al., 2015; Doyle et al., 2011; Liu et al., 2018) can affect cloud cycle and precipitation, as well as cloud radiative effects (Cox et al., 2015). Cloud radiation, especially from low-level clouds is a key component of the energy budget at the surface (English et al., 2015; Serreze & Barry, 2014; Sedlar et al., 2012, 2011; Shupe et al., 2013), and such clouds are directly impacted by the aforementioned atmospheric variability. A better understanding of feedbacks controlling Arctic change and the need for improved models (English et al., 2015; Kay et al., 2016; Li & Xu, 2020) emphasize the need to better constrain models with observations. To achieve this, particular attention must be given to the autumn (SON)-winter (DJF)-spring (MAM) period, during which transport of warmer and moister air masses from mid-latitudes may enhance sea-ice decline (Graham et al., 2017), including through modulation of longwave and shortwave cloud effects (Cox et al., 2016).

There are only a small number of surface land stations in the pan-Arctic region dedicated to atmospheric research (Uttal et al., 2016), and only a subset of these regularly make measurements using active instrumentation such as radar and lidar, which are necessary to retrieve cloud properties with vertical resolution. These retrieved profiles are valuable for understanding the vertical distribution and properties of cloud layers necessary to accurately model radiative transfer through the atmosphere (Shupe et al., 2013; Shupe, Tjernstrom, et al., 2015; Shupe, Turner, et al., 2015). The stations are located over land and many are coastal. Thus, data may be subject to spatial heterogeneity characteristic of such environments (e.g., orographic effects, specific atmospheric or ocean circulation flows, and variable surface reflectivity) and so may not be represented at the regional scale (Eastman & Warren, 2010; Shupe et al., 2011). New stations (drifting buoys) are being implemented over the Arctic ocean (Mariage et al., 2017; Provost et al., 2015), that should bring new information on aerosol and cloud profiles as well as the surface radiation budget (SRB), together with providing regional support to characterize SRB in combination with space observations. Estimation and direct observation of LW fluxes over the Arctic Ocean have been made possible from space-borne sensors for more than two decades, for example, from the Advanced Very High-Resolution Radiometer Polar Pathfinder Extended (Key et al., 2016) and from Clouds and the Earth's Radiant Energy System-Energy Balanced and Filled (CERES-EBAF) (Loeb et al., 2009). Those instruments provide data sets over a large swath that help to understand the energy budget in the Arctic and to better address the role of clouds. The advent of polar-orbiting satellite active sensors, with the success of CALIPSO/CloudSat missions (Stephens et al., 2018), allows for a more precise estimation of the regional Arctic cloud cover, quantification of cloud type vertical distribution, and inference of radiative fluxes at the regional scale (Kay & L'Ecuyer, 2013; Kay et al., 2016). The upcoming EarthCARE mission is designed to pursue and reinforce this progress through a continued instrumental synergy (Illingworth et al., 2015). However, while satellites provide the spatial coverage lacking from the surface stations, they do not directly observe the surface radiation budget and so must be validated.

Extensive characterization of Arctic SRB, therefore, necessitates a combination of the ground-based and satellite retrievals and a more accurate evaluation of all biases through comprehensive intercomparisons between observations. Previous work emphasized that the use of passive instruments (e.g., MODIS) alone is insufficient because of the underestimation of cloud fraction in DJF and SON (Blanchard et al., 2014; Liu et al., 2010, hereinafter B14). As cloud products from the satellite are commonly used to contribute to atmospheric reanalysis and to compute cloud radiative forcing, errors in cloud detection or biases in cloud products, as shown in Liu and Key (2016), may lead to errors in flux calculations. Consequently, B14 concluded that spaceborne lidar-radar synergy is essential for a complete representation of the cloud vertical profile, but that both surface and space observations are needed to reduce biases in all observations. Near-surface clouds are frequent in the central Arctic (Mariage et al., 2017; Uttal et al., 2002). Below about 1 km in altitude, space-based radar observations are inhibited by ground clutter (Palerme et al., 2019). Conversely, lidar sensitivity may be limited below clouds by attenuation, enhanced in presence of supercooled layers at cloud top. B14 found that the characterization of low clouds, as well as boundary layer events (composed of aerosols and/or precipitating ice crystals), are two of the principal challenges for spaceborne observations and the determination of radiation fluxes at the surface.

Longwave (LW) radiation is an important component of the energy budget in the Arctic and is indeed the only radiative flux during the polar night. LW is additionally particularly sensitive to the profile of atmospheric and cloud properties and therefore products such as CERES-EBAF are very sensitive to errors

in the profiles from which the calculations of the fluxes are made. In this study, we focus on the retrieval of LW radiation fluxes both at the top of the atmosphere (TOA) and at the bottom of atmosphere (BOA), as observed from the ground and from space and simulated from radiative transfer models over Eureka, Nunavut, Canada (80°N, 86°W). Eureka is a high-Arctic surface observatory with 5 years of overlapping measurements from the necessary instrumentation and is representative of a particularly dry region of the Arctic (Cox et al., 2012) where clouds are distributed over a wider range of heights than other locations (Shupe et al., 2011). The instrumentation and the time series of records at the Eureka site have allowed a significant number of studies related to climate (Cox et al., 2012; Lesins et al., 2010), comparisons of cloud cover (de Boer, 2009; Shupe, 2011; Shupe et al., 2011, B14) and downwelling fluxes (Cox et al., 2012). According to this latter study, the yearly average downwelling LW cloud radiative effect (difference of cloudy and clear air downwelling fluxes) at the BOA at Eureka is about 27 W/m^2. This work aims to analyze radiative flux comparisons following an approach similar to the one developed in B14. Namely, we perform two main analyses on a statistical basis using independent and coincident observations involving vertical profiles and cloud retrievals from synergistic use of lidar and radar data. We interpret results generally to conclude the applicability to the performance of the products under particular atmospheric regimes.

We first present upwelling and downwelling LW fluxes derived from satellite and surface observations, including calculations using cloud profile measurements at Eureka. All observations are compared to the ERA-Interim and ERA5 reanalyzes of the European Center for Medium Weather Forecast (ECMWF). Comparisons of seasonal variations of fluxes from statistical analyses and cloud vertical distribution based on independent data sets are detailed in Section 3. On the basis of the coincident data, a comparison of flux distributions and their differences are then discussed in Section 4. Finally, we discuss the results of the comparisons and identify biases and limitations.

2. Description of Observation Site and Data Sets

The focus surface observation sites are the Zero Altitude PEARL Auxiliary Laboratory (0PAL) and the Surface and Atmospheric Flux, Irradiance and Radiation Extension (SAFIRE), both part of the Polar Environment Atmospheric Research Laboratory (PEARL; Fogal et al., 2013) in Eureka, Nunavut, Canada, which is one of the high-latitude stations of the Network for the Detection of Atmospheric Composition Change (NDACC, http://www.ndaccdemo.org/stations/eureka-canada). It is also part of the IASOA network (International Arctic Systems for Observing the Atmosphere; Uttal et al., 2016), and, located at SAFIRE, a World Radiation Monitoring Center Baseline Surface Radiation Network (WCRP-BSRN, http://bsrn.awi.de/) station during the study period, 2007 and 2011 (Dreimel et al., 2018). Ground-truthing of satellite studies is one of the principal objectives of BSRN. Note that while the radiosonde launch facility, as well as the lidar and radar instruments, are located close to sea level at 0PAL, SAFIRE is located ~5 km northeast at 85 m (asl). Eureka is situated in the northernmost part of the Canadian Arctic Archipelago, a region having complex topography and variable surface type. However, despite this heterogeneity, the station offers a critical mass of observations and because of its latitude also a high frequency of polar satellite overpasses, enabling a larger number of coincident samples to be analyzed (B14). Figure 1 shows a map of Ellesmere and Axel Heiberg islands with the location of the Eureka station marked on the western coast of Ellesmere (a), the location of 0PAL and SAFIRE (b), as well as a photograph of the BSRN station highlighting flat, open area chosen for siting SAFIRE (c).

Kovacs and McCormick (2005) suggest that for cloud-comparison purposes a length scale of a few 10 km and a time scale of a few minutes is sufficient for identifying coincident observations. Based on this recommendation, we will define 25 km from Eureka as the maximum distance for the current study, which is also similar to the grid size of ERA-I, ERA5, and CERES. The region where spaceborne observations are analyzed is also shown in Figure 1.

CERES and MODIS have collected measurements since 1999 and 2002 onboard TERRA and AQUA, respectively. CloudSat and CALIPSO (hereafter referred to as C-C) were launched in 2006 and are part of the constellation of satellites formed with AQUA (A-Train). CloudSat has made measurements only during the daytime since 2011 due to a battery anomaly and the production of CloudSat 2B-FLXHR-LIDAR products (hereafter C-C-FLX) was discontinued at that time. Two releases of this product (R04 and R05) are however

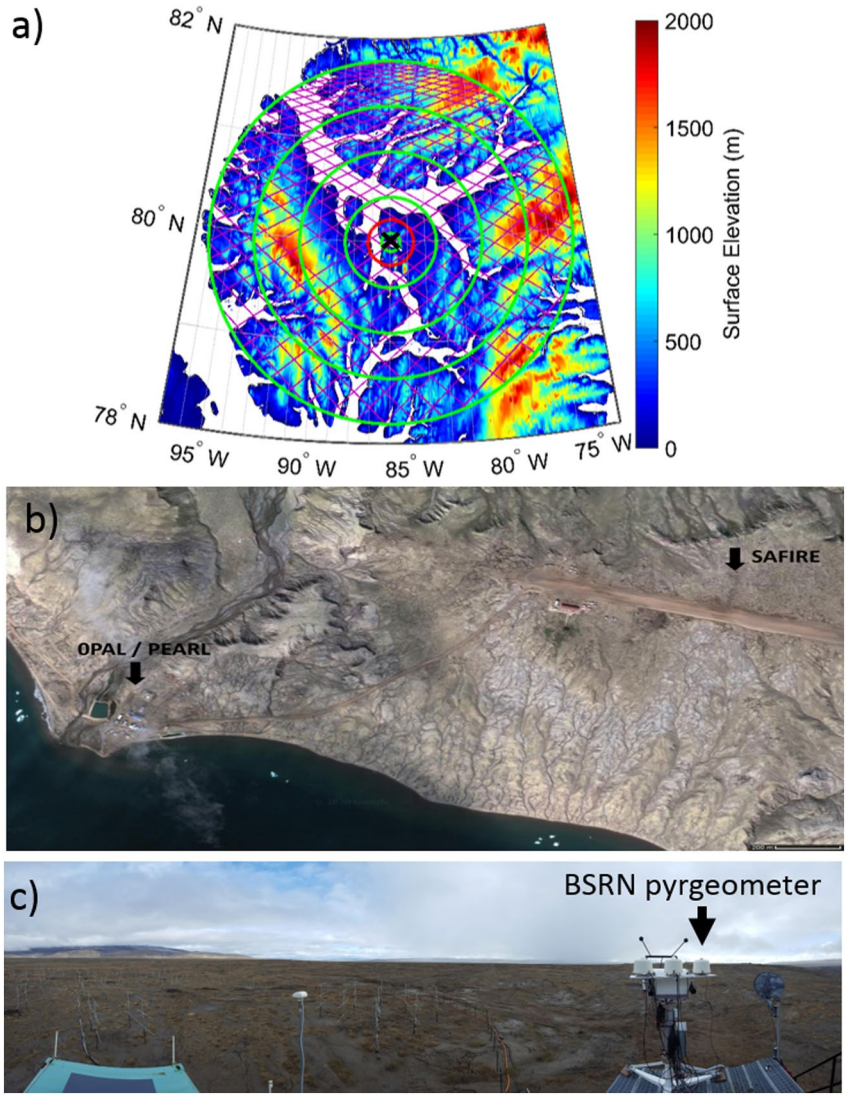

Figure 1. (a) A-Train tracks (in magenta) close to the Eureka station (black cross) during January 2007 superposed over the digital elevation model Global 30 arc s elevation data set (GTOPO30) used for CALIPSO data analysis. The green concentric circles (radius of 10, 25, 50, 100, 150, and 200 km) denote the area of this study. The 25 (red) circle delimits domains where surface orography and heterogeneity are minimized. (b) The location of the two active sensors at the Eureka 0PAL site and the radiation site (SAFIRE) (c) Panoramic view of the Eureka radiation site (SAFIRE) where the pyrgeometer is located. SAFIRE, Surface and Atmospheric Flux, Irradiance and Radiation Extension.

available from 2006 and almost over the same period. Although considering the whole period from 2002 to 2020 for the overall statistical analysis, the period of detailed analysis on coincident observations is limited to the overlap period spanning from June 2006 to May 2010, due to the availability of radiation products.

Table 1 summarizes the main characteristics of the data sets and methods used in this study.

2.1. Fluxes From Ground-Based Observations

Profiles of cloud properties are regularly measured above Eureka from combined radar and lidar measurements (B14). Here, we use these data as references for heights of cloud layers over the site. TOA and BOA fluxes based on the lidar and radar measurements at Eureka (hereafter "EUR-LR") were calculated using the Streamer radiative transfer model (RTM) (Key & Schweiger, 1988). The input parameters include atmospheric profiles (interpolated from twice-daily radiosonde measurements), aerosol optical depth from

Table 1
Satellite, Ground-Based, and Reanalysis Data Sets and Methods Used in This Study

Name	CERES	C-C-FLX	ERA-I/ERA5	EUR-LR	BSRN
Long name	CERES_SSF_Aqua-XTRK_Edition4A	CloudSat 2B-FLXHR-LIDAR	ECMWF ERA-Interim/ERA5	EUREKA-LIDAR-RADAR	Baseline Surface Radiation Network
Version	Edition 4A	Release 04 and 05			
Temporal resolution	2–4 s	0.16 s	3 h	3 min	1 min
Vertical resolution	N/A	240 m	60 levels (ERA-I); 137 levels (ERA5)	30 m	N/A
Footprint	20 × 20 km	1.4 × 1.7 km	0.125 ° × 0.125 °	N/A	N/A
Cloud properties	MODIS Collection 5 cloud products	CloudSat and CALIPSO	Reanalysis	From radar-lidar synergy	N/A
TOA fluxes	Observed	BugsRad RTM	Reanalysis	Streamer RTM	
BOA fluxes	Longwave Model B	BugsRad RTM	Reanalysis	Streamer RTM	Observed
References	Wielicki et al. (1996); Loeb et al. (2018)	Henderson et al. (2013)	Dee et al. (2011); Hersbach and Dee (2020)	Donovan and van Lammeren (2001); Shupe (2007)	McArthur, (2005); Driemel et al. (2018)

BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

the Eureka sunphotometer (part of AERONET, https://aeronet.gsfc.nasa.gov/), and cloud layer information (type of layer, altitude, layer optical depth, and mean effective radius) from ground-based lidar and radar (as detailed in Blanchard et al., 2017). Cloud type was derived from a multisensor classifier (Shupe, 2007) and particle sizes were retrieved from the ratio of radar and lidar backscatter cross-section (Eloranta et al., 2007) using default processing on the web site http://hsrl.ssec.wisc.edu. The purpose of this product is to be a comparable analog to the CloudSat and CALIPSO products, but from the perspective of the surface.

The LW flux data from the BSRN station are 1-min averages based on 1 Hz samples collected by a shaded Eppley pyrgeometer mounted on a sun tracker (Driemel et al., 2018). Grachev et al. (2018) reported on the intercomparability of the BSRN LW for an overlap period with another radiometer ~700 m east of the BSRN station at Eureka and found a negligible bias (~1 W/m^2) and with a standard deviation of 10.5 W/m^2 in the differences of hourly means.

The BSRN pyrgeometer was maintained approximately at daily intervals. In cold climates, this maintenance includes manual removal of ice from the sensor windows, which commonly occurs. The pyrgeometer was ventilated, which helps maintain temperature stability and mitigate the formation of ice. Unfortunately, despite these procedures, icing of the window frequently occurred on the pyrgeometer throughout the study period. Because the specific postprocessing procedures used on the data archived at BSRN are undocumented, we began with the raw data set and conducted our own quality control, including implementing the procedures recommended by Long and Shi (2008) as well as visual screening for signs of icing. The signal from the iced window is similar to the signal from clouds, making it difficult to identify. For an upward-facing LW measurement, the bias caused by the ice is generally positive and is large when the sky is clear and small when the sky is cloudy. Manual removal of ice by the technicians causes a change in the signal that is very fast compared to natural variability and this non-physical signal is easily identifiable, as is the decrease in radiance following the growth curve of the developing ice that precedes the cleaning backward in time. By identifying and removing these features, the visual screening likely removed most of the ice that occurred when the sky was radiatively clear and the bias was large, but the subsequent absence of the radiatively clear time periods in the record produces a climatological bias in the monthly means. Monthly means are only used in this study for qualitative purposes, so it was more important to have a representative estimate than a direct measurement for these periods. We, therefore, filled the gaps from the data removed because of icing with a calculation of the clear-sky downwelling LW following Long and Turner, 2008, which is based on Brutsaert's equation and requires only the radiometric measurements and collocated meteorology. Time periods that use these estimates are not incorporated into the validation analysis of this study.

In Section 3.2, we discuss the differences between the raw and cleaned data set, respectively named BSRN unfiltered and filtered.

The subset of observations coincident with the satellite overpasses that are used for comparison (Section 4) received further scrutiny individually, including analysis of logbook records, radar and lidar data, meteorology, and the other radiometric data in order to identify and remove additional suspect data that remained. This procedure had the added benefit of being well-suited to identify times where comparisons were likely to be influenced by cloud cover that was within the ~160° effective FOV of the pyrgeometer at times when the skies directly over Eureka were clear. Supporting S1 presents results from this data screening.

2.2. Fluxes From Satellite

Radiation measurements from the CERES instrument on AQUA and TERRA provide a direct observation of the upwelling TOA radiances (Wielicki et al., 1996), that is converted into fluxes using angular distribution models that provide a stable time series (Loeb et al., 2012). Specific comparisons of the CERES-EBAF TOA fluxes (Loeb et al., 2009) have been performed at high latitudes over the Arctic (Huang, Dong, Xi, Dolinar, Stanfield, & Qiu 2017; Kay & L'Ecuyer, 2013). CERES provides access to a long and homogeneous radiation database and has been used in numerous analyses. Based on these considerations, the AQUA CERES-SSF TOA fluxes (Ed4A) were taken as a reference for comparisons in the present study. The LW CERES surface (BOA) fluxes are achieved using cloud properties derived from MODIS and following the LW Model B algorithm described in Gupta et al. (1992). We used the AQUA CERES-SSF surface fluxes (Ed4A) in this study. They have been compared to other observations and validated against surface radiation measurements (Gupta et al., 2010; Kratz et al., 2020) for mid and low latitudes. In the Single Satellite Footprint (SSF) product Level 2, adding MODIS cloud retrievals and Goddard Meteorological Assimilation Office (GMAO) atmospheric profiles, the BOA fluxes at the surface are also available in a CERES field of view, about 20 × 20 km. CERES BOA fluxes are used as part of the Arctic Observation and Reanalysis Integrated System (ArORIS) gathering several data sets for climate studies in this region (Christensen et al., 2016). Initial comparisons over Greenland showed small dispersion (Christensen et al., 2016) confirmed by further studies over the whole Arctic (Huang, Dong, Xi, Dolinar, Stanfield, & Qiu 2017).

The CLOUDSAT-2B-FLXHR-LIDAR (hereafter named as C-C-FLX) products provide a direct estimation of TOA and BOA fluxes consistent with the liquid and ice water content and the cloud vertical profiles obtained from CloudSat, CALIOP, and MODIS measurements, using atmospheric profiles from ECMWF (Henderson et al., 2013; L'Ecuyer et al., 2008). In releases prior to R05, the TOA and BOA flux amounts were both defined as the first and the last non-zero value of FU and FD parameters (respectively upwelling and downwelling flux). In the last available version (R05), a new TOA variable is included to better represent atmospheric profiles above 25 km in calculations (Henderson & L'Ecuyer, 2020). In this study, we use the R05 products (http://www.cloudsat.cira.colostate.edu/data-products/level-2b/2b-flxhr-lidar), publicly available since March 2020, and discuss differences with the previous version R04. As the CLOUDSAT-2B-FLXHR-LIDAR R05 product is expected to be less prone to atmospheric biases due to cloud phase, well identified by lidar (Hu et al., 2009), we considered it to better represent clouds in flux calculations.

The C-C-FLX (R05) time series starts in June 2006 and ends in August 2010 (April 2011 for R04) while CERES on AQUA begins in July 2002. Both products are expected to differ due to factors related to cloud vertical distribution (Matus & L'Ecuyer, 2017). Whereas CERES is based on MODIS data inversion, C-C-FLX input is a direct retrieval of vertical profiles from CloudSat and CALIPSO active sensors. The cloud profiles are usually better constrained with active instruments, even if some biases remain (Chan & Comiso, 2011). Moreover, it has been shown that C-C misses some low clouds (B14). A second issue may be due to the spatial distribution as C-C-FLX fluxes are given at the radar footprints (1.4 km) along a track whereas the CERES grid is 20 × 20 km, somewhat smoothing spatial variability. For both data sets, we will discuss in Sections 4.1.1 and 4.2.1 the representativeness of taking the nearest pixels to the station versus an average of all the values located at less than 25 km from the station.

In the section related to cloud vertical profiles (Section 3.3), we also employed the raDAR-liDAR (DARDAR) mask (version 1.1.4) because it uses a comprehensive synergy of CALIOP and CloudSat observations

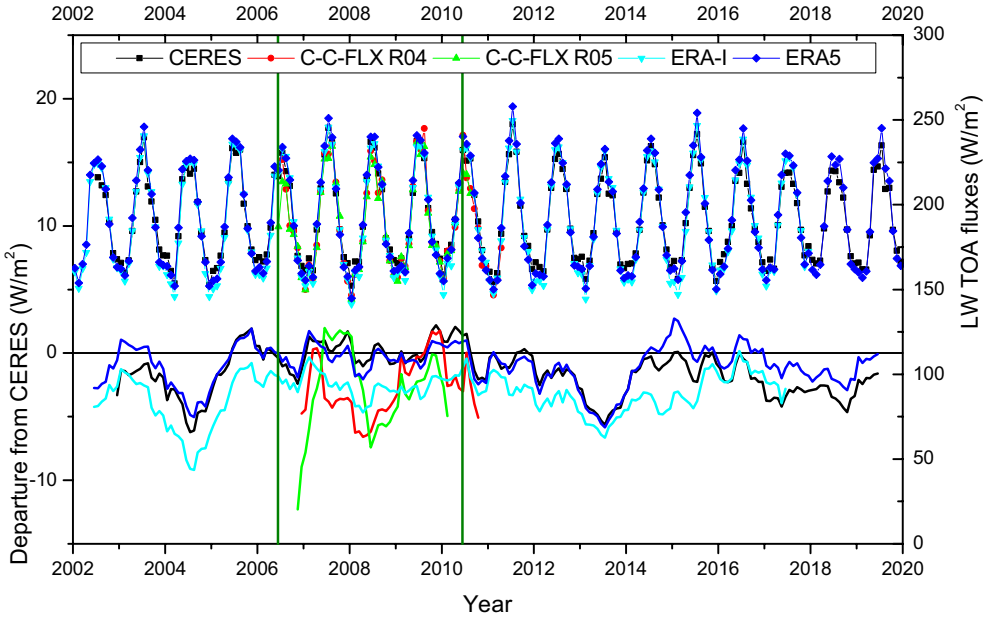

Figure 2. Monthly variation of TOA LW fluxes (right side scale) measured and derived from satellite (CERES in black and CloudSat-2B-FLXHR-LIDAR R04 (in red), R05 (in green), and re-analysis (ERA-I in cyan and ERA5 in blue) at Eureka from 2002 to 2020. Twelve-months moving average departures (left hand scale) from 4-years (June 2006–May 2010) CERES average are shown at the middle of the 12 months in the bottommost graph. The common observation period of this study is bounded by vertical green lines. CERES, Clouds and the Earth's Radiant Energy System; TOA, Top and Bottom of the Atmosphere.

(Ceccaldi et al., 2013; Delanoë & Hogan, 2008, 2010), differently than the CloudSat product, at a smaller vertical resolution of 60 m.

2.3. Fluxes From Re-Analysis

The ECMWF ERA-Interim (hereafter ERA-I) project is based on meteorological reanalysis that was assimilated from various data sets (Dee et al., 2011). ECMWF Integrated Forecast System uses four-dimensional variational data assimilation (4DVar). The ECMWF most advanced reanalysis product, ERA5, was recently released and provides several improvements compared to ERA-I, as detailed by Hersbach and Dee (2020), and uses a more advanced 4DVar assimilation scheme, and higher vertical (137 vs. 60 levels) and horizontal resolutions (31 vs. 79 km). In this study, we used the same spatial (0.125° × 0.125°, which is obtained by linear interpolation of the native grid and represents ~14 × 2 km at the latitude of Eureka) and temporal resolution (3 h) for both ECMWF products. This implies that for coincident comparison, the re-analysis products are 60 min later and 45 min earlier with respect to A-Train overpasses near 11:00 and 15:45 UTC, respectively.

3. Statistics From Independent Data Sets

In this section, as well as in the next section dealing with coincident measurements, we will first analyze TOA LW fluxes followed by downwelling fluxes at the surface.

3.1. TOA Fluxes

LW TOA monthly fluxes from CERES observations, CALIPSO-CloudSat flux calculations, and reanalysis from ERA5 are shown in Figure 2 for a period that extends from 2002 to 2020.

Monthly variations are similar amongst the data sets, minimum values being observed in DJF and maximum values in summer (JJA), with the range of the annual cycle being about 70–80 W/m^2, depending on

Table 2

Annual and Seasonal Variations of LW TOA Fluxes for CERES, C-C-FLX (R04 and R05), ERA-I, and ERA5 Over the Whole Data Set Period (a) and Coincident Period (b, July 2006 to May 2010) Based on Monthly Means

TOA	Time period	Annual mean		DJF		MAM		JJA		SON		ONDJFM	
		Mean (σ)	Trend (W/m² / year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)
a) STATISTICAL ANALYSIS OVER THE WHOLE OBSERVATION PERIOD													
CERES	07/2002–12/2019	191.2 (2.0)	−0.1 (0.1)	164.3 (2.8)	0.1 (0.1)	187.6 (3.0)	0.1 (0.2)	226.3 (6.6)	−0.5 (0.3)	186.9 (2.0)	−0.1 (0.1)	169.3 (1.8)	0.0 (0.1)
C-C-FLX R04	07/2006–04/2011	189.4 (2.4)		158.5 (4.8)		184.0 (4.0)		226.4 (8.5)		184.7 (2.8)		165.6 (3.2)	
C-C-FLX R05	06/2006–08/2010	191.5 (4.0)		158.9 (5.2)		184.4 (4.4)		220.9 (12.0)		188.7 (10.1)		166.5 (2.4)	
ERA-I	01/2002–12/2017	189.4 (2.0)	0.1 (0.1)	155.9 (3.0)	−0.1 (0.2)	184.6 (2.9)	0.3 (0.1)	229.9 (4.7)	0.1 (0.3)	187.0 (2.5)	0.0 (0.1)	163.5 (2.4)	0.1 (0.1)
ERA5	01/2002–12/2019	191.7 (1.8)	0.0 (0.1)	159.2 (2.7)	0.0 (0.1)	189.2 (3.0)	0.1 (0.1)	232.5 (5.8)	0.0 (0.3)	185.8 (2.6)	−0.1 (0.1)	165.1 (2.1)	0.0 (0.1)
b) STATISTICAL ANALYSIS OVER THE COMMON OBSERVATION PERIOD													
	Time period	Annual mean		DJF		MAM		JJA		SON		ONDJFM	
		Mean (σ)	Minus CERES	Mean (σ)	Minus CERES	Mean (σ)	Minus CERES	Mean (σ)	Minus CERES	Mean (σ)	Minus CERES	Mean (σ)	Minus CERES
CERES	06/2006–05/2010	192.7 (2.1)		164.8 (2.5)		187.7 (4.5)		230.7 (6.4)		187.4 (1.1)		169.5 (2.3)	
C-C-FLX R04	06/2006–05/2010	188.8 (4.3)	−3.8 (2.5)	159.3 (5.1)	−5.6 (3.1)	184.0 (4.0)	−3.7 (3.3)	227.4 (9.4)	−4.0 (5.9)	185.2 (2.9)	−2.2 (3.4)	166.1 (3.4)	−3.4 (1.2)
C-C-FLX R05	06/2006–05/2010	189.3 (6.2)	−4.7 (3.2)	158.9 (5.2)	−5.4 (2.4)	184.4 (4.4)	−3.3 (3.2)	222.2 (13.5)	−8.5 (7.8)	188.7 (10.1)	−0.9 (6.2)	166.5 (2.4)	−3.0 (0.6)
ERA-I	06/2006–05/2010	190.1 (0.6)	−2.6 (1.5)	156.4 (3.1)	−8.4 (1.1)	184.1 (2.6)	−3.6 (2.1)	231.9 (4.5)	1.2 (2.4)	187.9 (2.5)	0.5 (1.6)	163.3 (1.8)	−6.1 (1.4)
ERA5	06/2006–05/2010	192.3 (1.2)	−0.4 (0.9)	159.3 (3.3)	−5.5 (0.9)	188.7 (3.1)	1.0 (1.5)	234.9 (6.1)	4.2 (1.0)	186.2 (2.5)	−1.2 (1.5)	165.0 (2.2)	−4.4 (1.1)

Note. Linear trends are bolded when considered significant (more than 2 sigma). Color shading is representative of the difference with the reference (red when they are above CERES (darker) by more than 2 W/m^2, blue below −2 W/m^2, green is within 2 W/m^2).

CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

the data set. ERA5 and CERES are available over the full period from 2002 to 2020, whereas C-C-FLX R05 is limited to a period of 4 years (5 years for R04), as mentioned in the previous section. Plots in the lower graph of Figure 2 show the departure of 12-months moving average for each data set from CERES 2006–2010

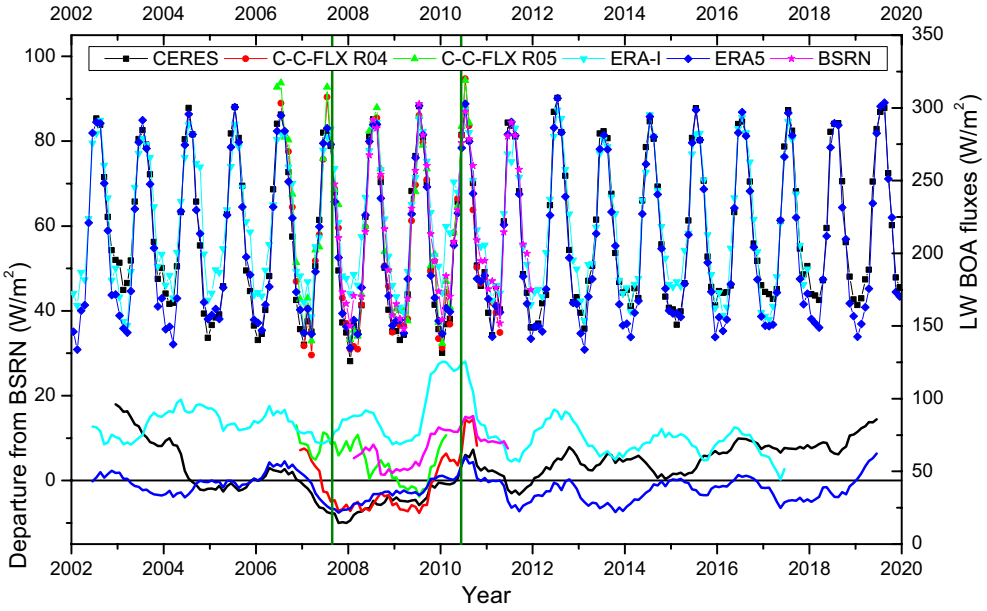

Figure 3. Monthly variation of BOA LW fluxes (right side scale) measured and derived from the ground (filtered BSRN), from satellite (CERES and CloudSat-2B-FLXHR-LIDAR R04 and R05) and reanalysis (ERA-I and ERA5) at Eureka from 2002 to 2020. Black dashed lines represent estimated trends on CERES annual minimum. 12-months moving average departures (left-hand scale) from 33-months (2007/09–2010/05) BSRN average are shown at the middle of the 12 months in the bottommost graph. The common observation period of this study is bounded by vertical green lines. BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System.

multi-year mean, which is the common observation period. It is seen that C-C-FLX R05 data are biased low as compared to CERES data 12-months moving average ranging from −12 to 3 W/m^2, with a mean value of −2 W/m^2. R04 shows a negative mean bias of −3 W/m2 with smaller interannual variability. Note that the 12-months moving average from ERA5 is generally in good agreement with those from CERES, except for a small difference of about +2 W/m^2 since 2014. Two periods in the whole CERES and ERA sequence were significantly different from the 2006 to 2010 average with departure larger than the overall standard deviation of 2.0 W/m^2. They occurred in 2004 and 2013 for both CERES and ERA5 data sets, before and after the reference period (June 2006–May 2010).

Average values, seasonal variations, trends, and standard deviations are reported in Table 2 for both 18 and 4-year periods. Departures from CERES, considered here as the reference for TOA LW fluxes, are shown in the second part of Table2. ERA-I statistics are also included to be discussed along with newly available ERA5.

The bottom part of Table 2b shows that the differences during the selected intensive common observation period, where both C-C-FLX releases are about 6 W/m^2 smaller than CERES in DJF and less than 4 W/m^2 the rest of the seasons. The difference of 4.5 W/m^2 in JJA between C-C-FLX R04 and R05 may be explained by more data available in R05 in 2006 JJA, but also by the longwave land emissivity that varies by surface type in R05 (Henderson & L'Ecuyer, 2020). The statistics over the whole observation period show that ERA-I has slightly smaller TOA LW fluxes than CERES but that ERA5 is similar to CERES and closer to the annual means. Both ERA data sets are close to CERES, except in DJF where they are closer to C-C-FLX R04, showing an 8.4 W/m^2 and 5.5 W/m^2 deficit relative to CERES (Table 2b). The use of ERA5 over ERA-I (with developments in model physics, core dynamics, assimilation system, higher spatial and temporal resolution) leads to an increase in LW fluxes by several W/m^2 and significantly reduces the bias with CERES, except in JJA where the bias is increased. SON shows the opposite behavior with decreased fluxes in ERA5 and degraded agreement. Note that the warm ground temperature bias in ERA reanalyzes (Wang et al., 2019) is not relevant in our study because weather observations at Eureka are assimilated (as seen in Section 3.4).

Table 3

Annual and Seasonal Variations of LW BOA Fluxes for CERES, C-C-FLX (R04 and R05), ERA-I, ERA5, and BSRN Over the Whole Data Set Period (a) and Coincident Period (b, September 2007 to May 2010) Based on Monthly Means

BOA	Time period	Annual mean		DJF		MAM		JJA		SON		ONDJFM	
		Mean (σ)	Trend (W/m² / year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)	Mean (σ)	Trend (W/m² /year)
a) STATISTICAL ANALYSIS OVER THE WHOLE OBSERVATION PERIOD													
BSRN filtered	09/2007–12/2011	216.1 (3.7)		168.9 (11.0)		189.5 (10.0)		282.4 (1.8)		219.0 (4.9)		178.9 (7.6)	
BSRN unfiltered	09/2007–12/2011	214.8 (3.7)		166.1 (8.4)		188.8 (10.0)		282.4 (1.8)		217.2 (5.0)		176.2 (6.5)	
C-C-FLX R04	07/2006–04/2011	205.9 (6.5)		152.7 (12.8)		182.0 (12.9)		286.0 (9.2)		211.0 (10.1)		165.1 (7.1)	
C-C-FLX R05	07/2006–08/2010	211.4 (7.0)		155.5 (7.9)		185.8 (11.7)		289.9 (13.0)		227.1 (12.7)		171.2 (4.1)	
CERES	07/2002–12/2019	211.2 (6.1)	0.5 (0.3)	160.2 (14.5)	0.7 (0.7)	192.1 (7.1)	0.1 (0.4)	286.0 (4.4)	0.3 (0.2)	208.1 (8.1)	0.3 (0.4)	170.1 (12.1)	0.6 (0.6)
ERA-I	01/2002–12/2017	220.0 (6.0)	−0.4 (0.3)	174.9 (9.3)	−0.9 (0.5)	201.4 (10.8)	−0.2 (0.6)	279.4 (5.2)	0.1 (0.3)	224.4 (7.8)	**−1.1 (0.3)**	186.8 (7.6)	−0.7 (0.4)
ERA5	01/2002–12/2019	206.5 (3.6)	0.0 (0.2)	153.5 (5.4)	−0.2 (0.3)	185.0 (4.7)	0.1 (0.2)	284.1 (4.9)	0.1 (0.2)	203.6 (6.6)	−0.1 (0.3)	164.0 (4.2)	−0.1 (0.2)
b) STATISTICAL ANALYSIS OVER THE COMMON OBSERVATION PERIOD													
		Annual mean		DJF		MAM		JJA		SON		ONDJFM	
	Time period	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN
filtered	09/2007–05/2010	205.8 (16.2)		163.9 (6.1)		192.4 (10.0)		280.9 (1.1)		217.5 (6.4)		175.8 (5.2)	
BSRN unfiltered	09/2007–05/2010	204.7 (16.1)	−1.1 (0.2)	162.5 (5.4)	−1.4 (0.7)	192.0 (9.4)	−0.4 (0.7)	280.9 (1.1)	0.0 (0.0)	215.3 (6.1)	−2.2 (0.6)	173.6 (4.7)	−2.1 (0.6)
C-C-FLX R04	09/2007–05/2010	195.9 (18.7)	−8.3 (2.2)	146.6 (5.3)	−17.3 (6.8)	183.0 (15.7)	−3.1 (3.1)	279.5 (9.7)	−1.4 (8.7)	209.4 (2.1)	−8.2 (7.0)	160.7 (3.4)	−13.5 (3.9)
C-C-FLX R05	09/2007–05/2010	203.1 (15.2)	−2.4 (4.4)	153.7 (8.5)	−9.7 (9.3)	187.1 (14.0)	0.4 (2.4)	281.1 (15.3)	0.2 (14.2)	223.6 (13.1)	−0.1 (8.7)	169.5 (3.0)	−4.7 (5.6)
CERES	09/2007–05/2010	193.9 (20.0)	−10.8 (4.0)	142.8 (2.1)	−21.2 (6.3)	187.8 (10.5)	0.8 (3.8)	284.9 (3.3)	4.0 (2.3)	198.2 (3.6)	−19.3 (2.9)	154.1 (2.6)	−21.3 (4.2)
ERA-I	09/2007–05/2010	218.6 (15.1)	13.8 (5.9)	181.6 (12.0)	17.6 (6.1)	212.6 (21.0)	25.4 (12.5)	279.3 (4.4)	−1.6 (3.3)	225.9 (4.4)	8.4 (2.3)	191.1 (11.4)	16.0 (6.0)
ERA5	09/2007–05/2010	196.4 (18.4)	−8.1 (2.7)	150.8 (4.0)	−13.2 (5.9)	188.3 (8.2)	2.2 (6.9)	283.8 (1.3)	2.9 (0.2)	198.9 (1.6)	−18.6 (4.8)	159.9 (3.1)	−15.1 (4.2)

Note. Linear trends are bolded when considered significant (more than 2 sigma). Colors indicate differences with BSRN fluxes taken as a reference (red when they are above BSRN by more than 2 W/m^2, blue below −2 W/m^2, green in between). Darker colors are used above ± 10 W/m^2.

BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

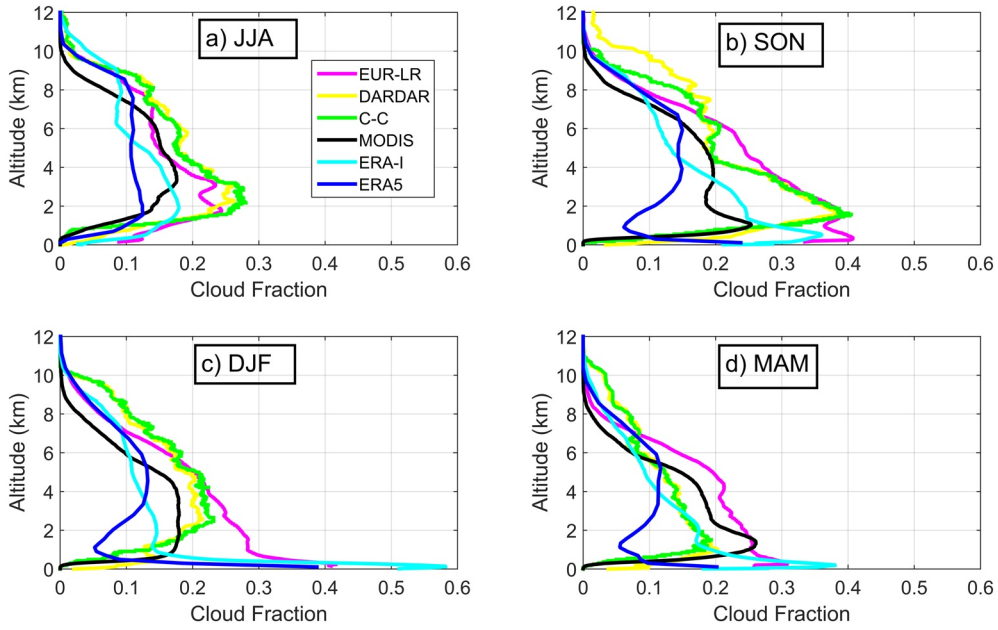

Figure 4. Cumulated seasonal vertical cloud distribution between June 2006 and May 2010 all the independent data sets at less than 25 km from the station.

CERES, ERA-I, and ERA5 do not show significant trends in LW TOA fluxes in the last 18 years, neither on average nor on a seasonal analysis. C-C-FLX annual data show a much larger variability between 2006 and 2010, as compared to the three other data sets that have a larger footprint.

To summarize, TOA averages over the whole period appear to be in good agreement for all data sets (within about 5 W/m^2). Only small differences (CERES can be larger by about 6–8 W/m^2 than C-C-FLX, ERA5, and ERA-I in DJF) are observed on seasonal TOA fluxes, and these differences are consistent with previous work (Loeb et al., 2018).

3.2. BOA Fluxes

Figure 3 presents the comparison of the LW BOA downwelling fluxes from CERES, ERA-I, and ERA5 over the full 2002–2020 period, and along with the C-C-FLX, the reference ground-based broad-band radiation data set over a more restricted period of time. Table 3 gives the yearly and seasonal average values as well as the estimated trends.

Note that the BSRN screening method described in Section 2.1 and supporting S1 helps to remove suspicious measurements (184,572 cases representing about 8% of the whole data set). Cox et al., (2020) looked at the impact of filtering and they concluded that while instantaneous biases are large, biases in the long-term means from ice are relatively small. The difference between the two BSRN data sets (unfiltered and filtered) is a combination of positive biases from ice in the version containing iced data and negative biases caused by the rejection of iced data under clear skies. This causes an under-representation of clear sky and then higher seasonal LW (up to 2.8 W/m^2 in DJF, see Table 3a for the difference between unfiltered and filtered BSRN). Therefore, those values presented in this section are not expected to be used for climatological analysis, but as a point of comparison with other data sets.

Figure 3 shows a good agreement between all LW BOA fluxes for the JJA, where annual cycle maxima among data sets vary within 5 W/m^2, except for C-C-FLX R05 between 2006 and 2010, where it exceeds 20 W/m^2. Large differences between the data sets are observable in DJF with divergence at minimum as large as 30 W/m^2 in 2003. Variation of DJF fluxes for CERES shows a V-shape trend (see Figure 3). The average decrease is about 30 W/m^2 from 2002 to the middle of the period analyzed for example, between 2008 and 2009, and an increase after (representing about 20 W/m^2). This increase in CERES fluxes may

be due to the increase in low cloud DJF temperature or changes in low cloud fraction over this part of the Arctic. This trend is however not seen on ERA5 (and ERA-I) data sets as was evidenced in meteorological trends observed in the Arctic (Graham, Cohen, et al., 2019; Jun et al., 2016). This V-shape trend observed in DJF over the 18 years tends to reduce the overall trend as reported in Table 3. Some residual low frequency (about 10 years) modulation is apparent, in particular during the DJF months, with a small peak-to-peak amplitude (about 3 W/m^2). BSRN data tend to agree with the uniformity of LW BOA fluxes measured in JJA at Eureka (standard deviation of 1.8 W/m^2 between 2007 and 2011, see Table 3), and the low values (~165 W/m^2) measured in the DJF from 2008 to 2010.

From Table 3 (see also Figure 3), it is seen that CERES statistics over the common period is about 10 W/m^2 smaller than BSRN averages. C-C-FLX is also smaller over this period but R05 shows reduced bias in all seasons compared to R04. Major changes were made in C-C-FLX R05 to improve the representation of cloud properties (cloud detection, supercooled liquid, and ice clouds microphysical properties) along with updated data ingested. The annual difference is still lower than BSRN by 2.4 W/m^2 over the common observation period and is mainly attributed to differences observed in DJF when differences are about 10 W/m^2, consistent with the aforementioned sampling limitations in the BSRN during the icing season. C-C-FLX R04 shows higher bias during the polar night (ONDJFM) with peak differences being observed between C-C-FLX R04 and BSRN in DJF that reaches −17.3 W/m^2.

Table 4
Seasonal LW TOA Average Fluxes for Coincident CERES and Other Retrievals for the Period Spanning From July 2006 to May 2010

TOA	Total		DJF		MAM		JJA		SON	
# of cases	249		56		67		47		79	
	Mean (σ)	Minus CERES < 25 km	Mean (σ)	Minus CERES < 25 km	Mean (σ)	Minus CERES < 25 km	Mean (σ)	Minus CERES < 25 km	Mean (σ)	Minus CERES < 25 km
CERES < 25km	192.1 (30.4)		164.0 (13.1)		189.7 (21.5)		239.9 (16.1)		185.6 (16.9)	
CERES nearest	192.2 (30.6)	0.1	163.7 (13.5)	−0.2	189.9 (21.6)	0.2	239.9 (17.0)	0.0	185.8 (16.9)	0.3
EUR-LR	191.1 (31.4)	−1.0	163.6 (16.7)	−0.4	187.2 (21.0)	−2.5	237.1 (19.2)	−2.8	186.6 (22.1)	1.0
C-C-FLX R04 < 25km	188.3 (31.4)	−3.8	158.7 (12.6)	−5.3	187.4 (22.5)	−2.2	237.8 (14.4)	−2.2	180.5 (17.4)	−5.1
C-C-FLX R04 nearest	187.9 (32.3)	−4.1	156.9 (13.1)	−7.1	186.8 (22.9)	−2.8	238.4 (16.0)	−1.5	180.9 (18.1)	−4.6
ERA-I	187.4 (32.0)	−4.7	156.9 (15.4)	−7.1	183.8 (24.3)	−5.9	237.2 (12.1)	−2.7	182.4 (17.2)	−3.2
ERA5	191.3 (34.7)	−0.7	160.1 (13.7)	−3.8	191.5 (24.5)	1.8	246.2 (14.5)	6.3	180.6 (20.6)	−4.9
# of cases	221		46		65		47		63	
	Mean (σ)	Minus CERES 25 km	Mean (σ)	Minus CERES 25 km	Mean (σ)	Minus CERES 25 km	Mean (σ)	Minus CERES 25 km	Mean (σ)	Minus CERES 25 km
C-C-FLX R05 < 25km	190.1 (31.7)	−4.3	157.9 (14.1)	5.6	187.3 (22.5)	−2.4	236.0 (14.5)	−3.9	182.1 (17.6)	−5.7
C-C-FLX R05 nearest	189.4 (33.2)	−5.0	155.4 (15.2)	−8.1	186.5 (23.4)	−3.2	237.3 (16.4)	−2.6	181.4 (17.9)	−6.4

Note. C-C-FLX R05 is put aside at the end of the table due to smaller data points compared. Standard deviations are in brackets. Colors are reported as in Table 2.

CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

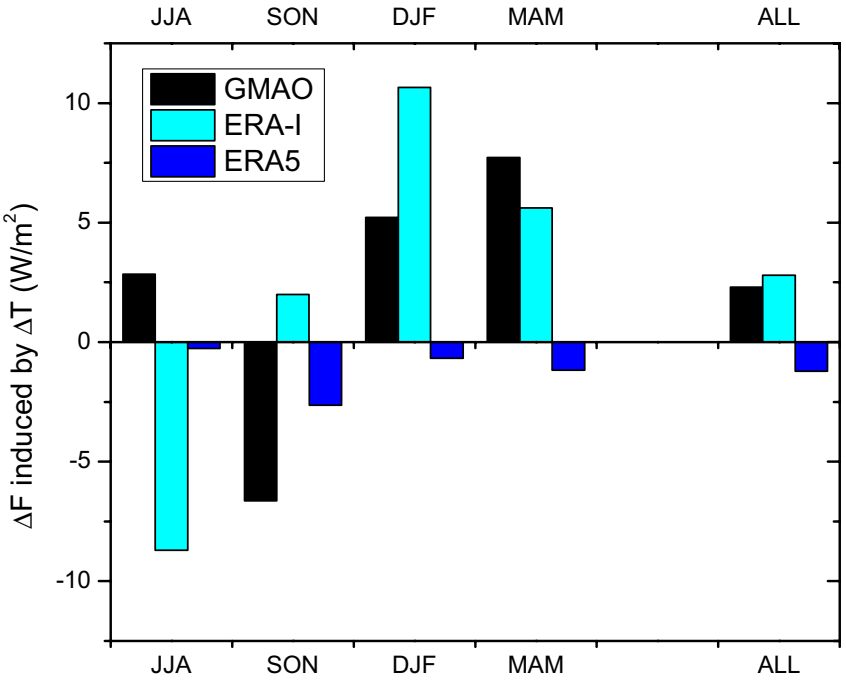

Figure 5. Induced flux differences based on mean seasonal temperature differences between June 2006 and May 2010 for the three reanalysis data sets used in flux calculations.

The reanalysis averages reported in Table 3 show interesting features with a correction of fluxes in ERA5, that reduces annual fluxes by 14 W/m^2 on average with respect to ERA-I. This reduction in LW downwelling fluxes occurs almost all year, except in June and July. As for TOA comparison, several reasons could explain the better performance of ERA5, for example, a more detailed data assimilation system with higher vertical resolution and better surface and radiation models. Although ERA5 is closer to BSRN than ERA-I, it still underestimates LW BOA by ~15 W/m^2 in ONDJFM, relatively to BSRN measurements. As seen from Figure 3, ERA5 and BSRN are in good agreement over the period of minimal DJF fluxes. Although in good agreement with CERES in JJA, ERA5 does not show DJF flux increases seen by CERES.

As a first conclusion on these seasonal average analyses of BOA fluxes (Table 3b), CERES and C-C-FLX R04 averages derived from space observations and ERA5 reanalysis are in rather good agreement although about 10 W/m^2 smaller than BSRN. They can even be larger than −20 W/m^2 during the polar night (ONDJFM). C-C-FLX R05 shows an improvement in all seasons with reduced differences compared to R04. Conversely, ERA-I appears to be biased high with respect to all observations except in JJA, where all results are in agreement within 4 W/m^2. The observed bias of ERA-I is coherent with previous analyses, where an over-estimation of low-cloud cover causes higher LW BOA in DJF, whereas ERA-I LW is subject to a dry bias in JJA (Chernokulsky & Mokhov, 2012; Huang, Dong, Xi, Dolinar, & Stanfield, 2017; Lenaerts et al., 2017; Zib et al., 2012; Zygmuntowska et al., 2012). Although relatively few studies with ERA5 evaluation in the Arctic are available, it seems that several biases of ERA-I are better addressed in ERA5, in terms of representation of temperature and humidity profiles and wind speed near the surface (Betts et al., 2019; Graham, Cohen, et al., 2019; Graham, Hudson, et al., 2019). However, it is not clear if a low cloud fraction is better represented.

In order to further analyze the origin of these differences, we come back to cloud vertical information as it was identified in B14 as a source of difference in sensitivities.

3.3. Cloud Vertical Structure and Type at Eureka

The abilities of each data set to correctly resolve cloud vertical distribution are explored in this section, as the latter plays a major role in radiation budget and flux calculations. In Figure 4, we compare the cumulat-

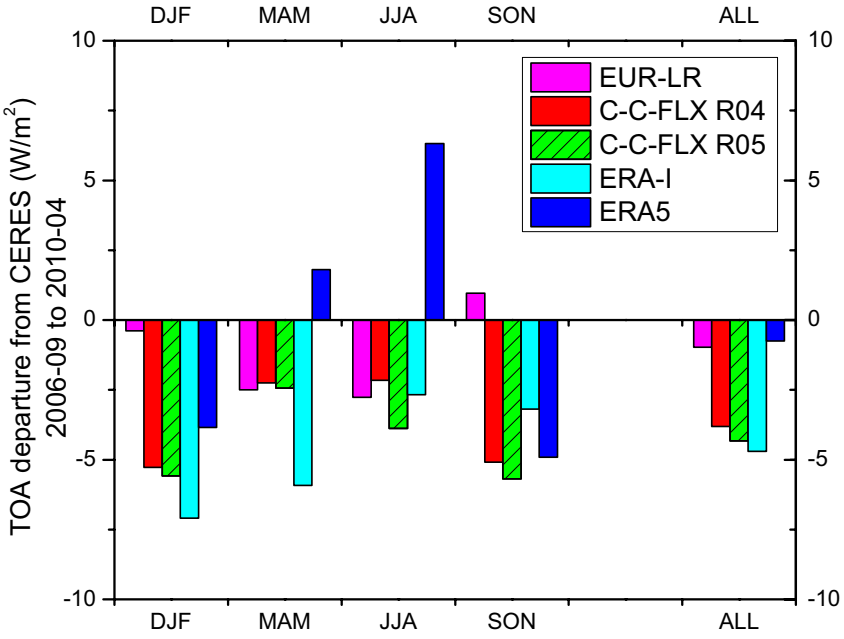

Figure 6. LW TOA departure from CERES as reported in Table 4. C-C-FLX R05 comparison with CERES is based on only 221 points (249 for R04 and other data sets). CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

ed seasonal vertical cloud distribution for all available data between June 2006 and May 2010, to foster the discussion of seasonal biases previously identified.

Over June 2006 to May 2010 period, the number of vertical profiles was 329,204 for EUR-LR (1 profile every 3 min), 57,976 for MODIS on AQUA (taking all the pixels whose center is less than 25 km from Eureka), 16,927 for DARDAR (481 overpasses; this data set, DARDAR-MASK-v1.1.4 is based on CloudSat and CALIPSO synergy, as described in Delanoë & Hogan, 2010 and Ceccaldi et al., 2013), 17,024 for CloudSat-CLDCLASS-LIDAR R05 (labeled as C-C in Figure 4) and 5,844 reanalyzes for ERA-I and ERA5. Note that differences in time sampling due to time of overpasses and instrument maintenance and failure are present in Figure 4 but are considered not to be a major driver in differences discussed below.

DARDAR, C-C, and EUR-LR have similar vertical distributions of cloud layers above 3 km. As detailed in B14 and Liu et al. (2017), low (<3 km) and very low (<1 km) clouds are difficult to address from space, and this is confirmed here from DARDAR, C-C, and MODIS for which cloud fractions are much lower than EUR-LR ground-based observations below 2 km, as evidenced in Figure 4. DARDAR and C-C are close in all seasons, except for high clouds ($z > 8$ km) in SON and DJF (that may be due to the use of a better vertical resolution in DARDAR, which is sampled at CALIOP vertical resolution). DARDAR and C-C give close results although DARDAR gives a higher amount of ice clouds and less mixed-phase clouds (see Appendix A). We find that the cloud fraction in reanalysis is generally biased low below 8 km, which is consistent with the findings of Liu and Key (2016) for a larger region of the Arctic.

In general, there is good agreement in the vertical profiles of cloud fraction excepting MODIS, ERA5, and ERA-I, which are systematically smaller than the other data sets (and this is particularly the case for ERA5). In MAM, although slightly smaller than ground-based observations, cloud fractions from all sources agree above 5 km, but significant discrepancies are observed below this altitude and even more below 2 km, in agreement with previous findings from B14.

Low clouds detected by EUR-LR are more frequent in all seasons but JJA, especially in DJF and MAM when the difference with the other data sets is larger (Figures 4c and 4d). An occurrence peak is observed by EUR-LR near the surface (within the first 500–1,000 m) during DJF (and MAM), only captured by ERA5 and ERA-I. MODIS strongly underestimates low clouds in JJA, and, to a lesser extent, in other seasons. B14 also

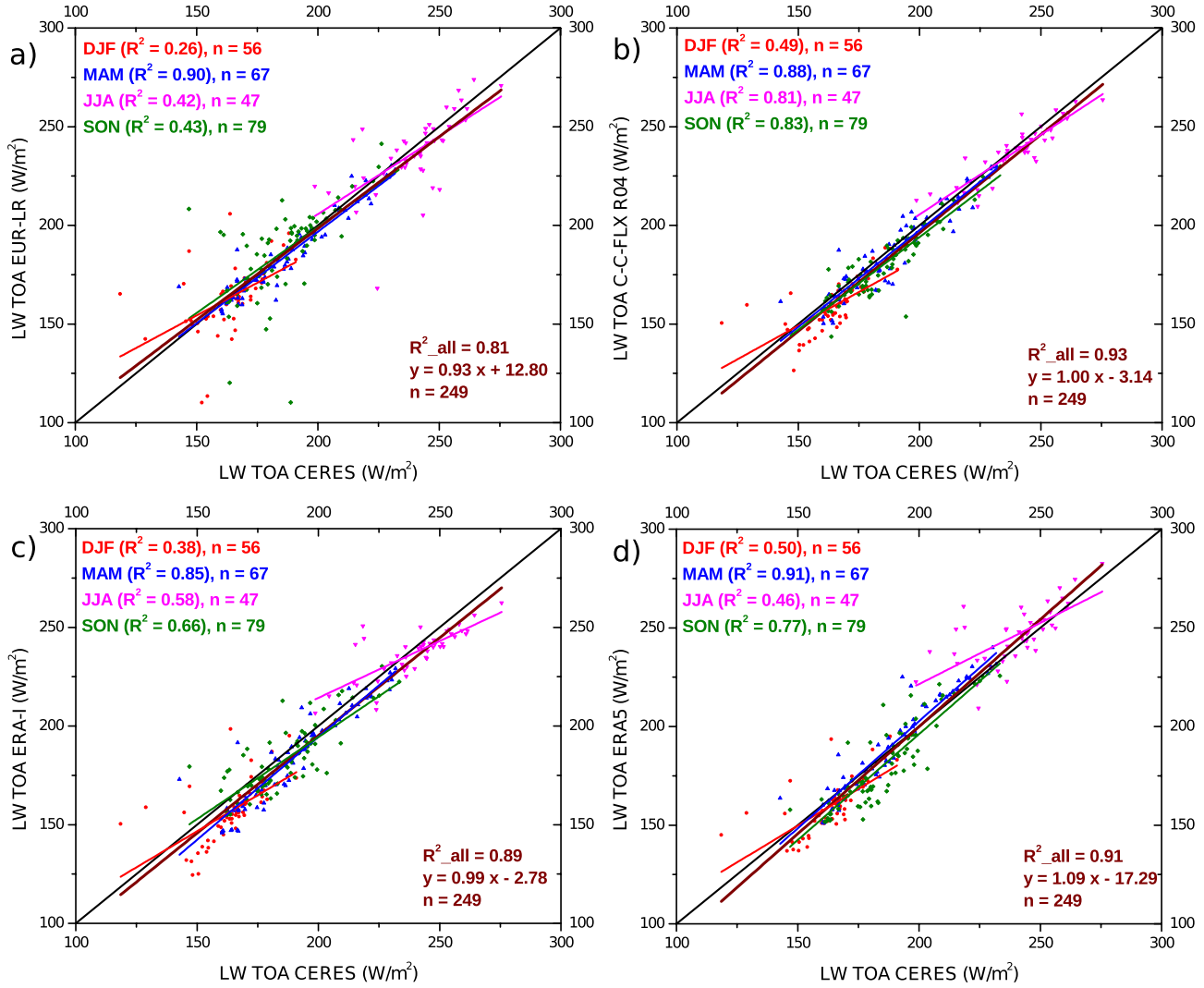

Figure 7. Seasonal TOA upwelling LW fluxes for CERES and EUR-LR (a), C-C-FLX R04 (b), ERA-I (c), and ERA5 (d) for 249 measurements at the same time, between July 2006 and May 2010. CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

showed that MODIS underestimates cloud fraction during DJF (from October to February). This is a known issue (Liu et al., 2010). MODIS cloud fraction biases vary with season, as the complete darkness during the polar night prevents the use of visible channels for cloud retrievals and implies the use of a nighttime cloud detection algorithm (Liu et al., 2004). MODIS distribution peaks at 1 km in MAM, when the temperature inversion is the strongest, about 10°C on average. ERA-I misses a large fraction of clouds below 8 km in all seasons but shows a large increase in the near surface cloud fraction from September to May. This was also discussed by Zygmuntowska et al. (2012) and Zib et al. (2012). Compared to ERA-I, ERA5 better captures mid-level clouds, which are mainly ice clouds, but largely misses low clouds especially between 500 and 1,000 m, where near-surface temperature inversion usually occurs.

In DJF, satellite observations and analyses dramatically lack low level clouds (between 0 and 3 km), where most Arctic clouds occur (Shupe et al., 2011). This is partly compensated in ERA-I and ERA5 by a peak of near surface clouds. In all seasons DARDAR, C-C, and MODIS lack low clouds below 1 km. Spaceborne radar detection suffers from surface contamination echo, and the lidar detection efficiency is decreased by attenuation in liquid water clouds. In most seasons, EUR-LR is missing some high clouds, due to the attenuation of lidar signal in opaque clouds and due to decreasing radar sensitivity with range. The better agreement (above 7 km) is obtained in JJA and the larger dispersion in this altitude range is observed in SON.

Table 5
Statistics of TOA Differences Between EUR-LR, C-C-FLX R04, and R05, ERA-I, ERA5, and CERES Based on Scatter Plots (Figure 7) and Gaussian Fit (Figure S4)

	EUR-LR	C-C-FLX R04	C-C-FLX R05	ERA-I	ERA5
R^2	0.81	0.93	0.94	0.89	0.91
RMSE (W/m^2)	13.9	8.9	8.8	11.7	10.8
Mean difference (W/m^2)	−1.3	−4.0	−4.6	−7.4	−2.7
σ (¼ of full width at $1/e^2$) (W/m^2)	6.9	6.7	6.8	7.2	8.6
Number of outliers (> 3 σ)	26	5	4	14	8

Note. That a smaller number of points were available for R05 (see Table 4).
CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

Large differences in BOA fluxes, in Table 3b, were observed during DJF, which is the period when significant mismatch appears in cloud vertical distribution. This deficit of low clouds lowers BOA fluxes calculated from the satellite data set, compared to BSRN. The smaller contribution of MODIS low clouds at this season can explain the lower LW BOA value of CERES compared to C-C-FLX R05. This is the opposite for MAM. The excess of near surface clouds for ERA-I in DJF and to a lower extent in MAM, combined with a warmer temperature profile compared to radiosondes (see Section 3.4), can also explain the over-estimation of downwelling fluxes for ERA-I. For ERA5, the cloud deficit in the whole column leads to an under-estimation of LW downwelling fluxes.

3.4. Atmospheric and Surface Temperature at Eureka

The quality of the satellite-derived LW fluxes depends on cloud profiles but also on the correct representation of the surface and atmospheric meteorological parameters in the calculations. We discuss here the importance of the temperature profile (including surface temperature) and the possible enhancement of the biases shown in the previous section.

The temperature profiles from reanalyzes that are used in TOA and BOA LW flux calculations are compared with radiosondes launched daily at the Eureka weather station for the June 2006–May 05 period. The temperature differences between reanalyzes and radiosondes (see Figure S2, bottom axis) show a relatively good agreement from 900 mb up to the tropopause, with a temperature difference less than 1K, similar to Graham, Hudson, and Maturilli (2019). The resulting plots (shown in S2) of the seasonal radiosonde temperature vertical profiles (top axis) show a persistent temperature inversion layer from September to May at a pressure level close to 900 hPa. This inversion layer is stronger in DJF with an average temperature difference of 10K. The presence of the inversion stratifies the boundary layer and inhibits the vertical transport of sensible and latent heats and the mixing with the free troposphere (Lesins et al., 2010).

All three reanalyzes show larger differences close to the surface, especially during wintertime when the inversion layer is stronger. In DJF (Figure S2c), both ECMWF reanalyzes overestimate the temperature at 1,000 hPa by 5°K. This is larger than what was observed by Graham, Hudson, and Maturilli (2019) who compared reanalyzes with radiosondes over the Arctic Ocean. On the contrary, GMAO underestimates the temperature in the inversion layer but gets a more accurate temperature at 1,000 hPa. This distinct behavior is similar but less pronounced in MAM and SON. In the presence of near-surface clouds, the radiative temperature is likely to be biased. The joint effects of the general underestimation of low clouds by satellite data sets and misrepresentation of their temperature are responsible for biases in both TOA and BOA LW fluxes.

The surface temperature used in flux calculations adds another source of uncertainty. Here we compare surface temperature from reanalyzes with the 2-m temperature measured at the weather station (Figure S3a). Except for ERA5, there is a general overestimation of 2-m temperature from September to May in reanalyzes, as observed over Arctic Sea ice in Wang et al. (2019). Contrary to the latter study, ERA5 shows a better agreement for all seasons with a difference of less than 1°K. To quantify the effect of surface temperature differences (ΔT), we look at the induced TOA flux differences (ΔF), following Equation 1 from the derivative of Stefan-Boltzmann's law:

$$\frac{\Delta F}{F} = 4\frac{\Delta Ts}{Ts} \tag{1}$$

Where F is the seasonal average flux (we used TOA fluxes measured from CERES, as shown in Table 2b), Ts is the seasonal average surface temperature mesured at the station (Figure S3a) and ΔTs are the mean temperature difference shown in Figure S3b.

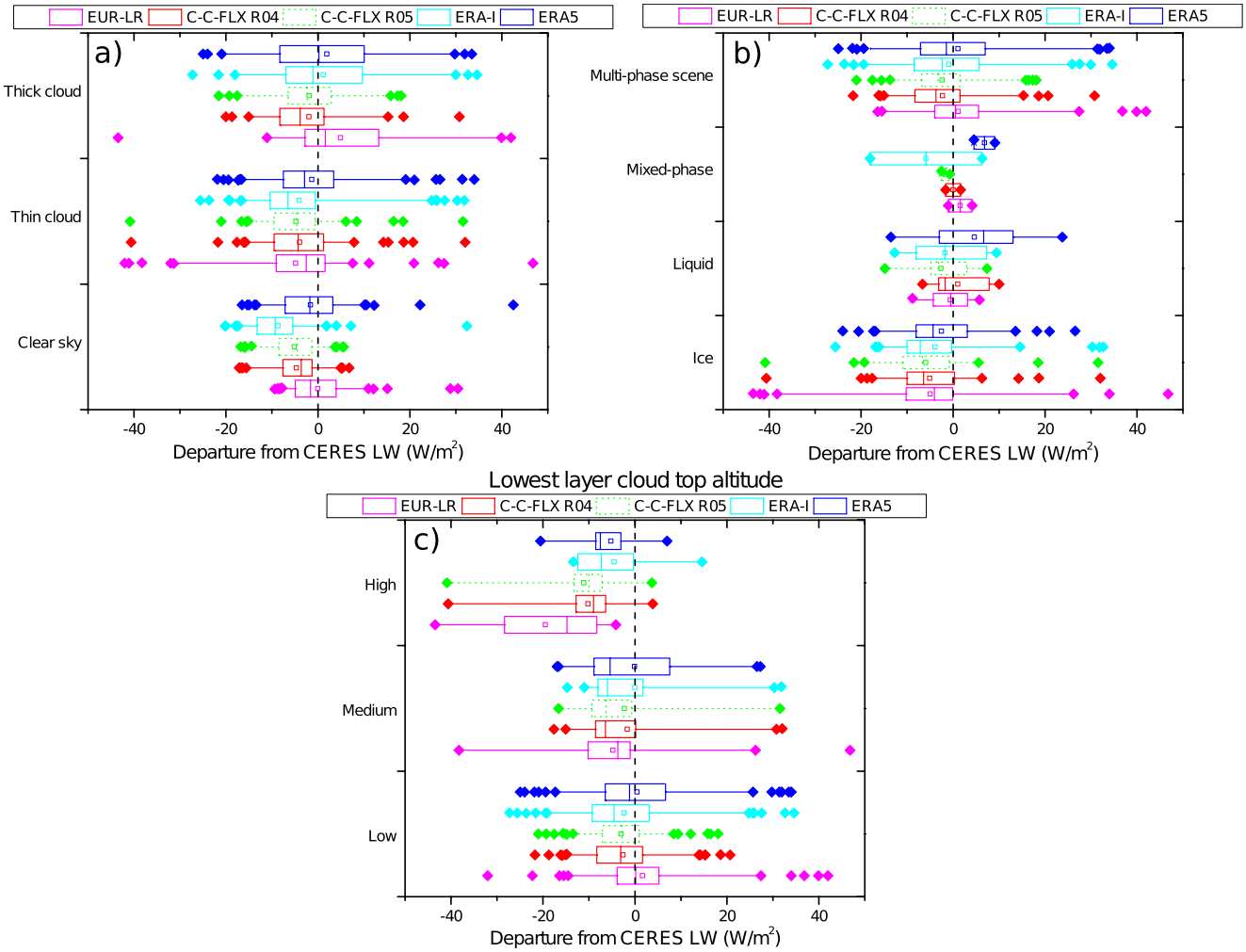

Figure 8. Difference of TOA LW fluxes between DATASET (EUR-LR, C-C-FLX R04, R05, ERA-I, and ERA5) minus CERES depending on (a) total visible optical depth, (b) phase of cloud layers, and (c) top height of the lowest layer. Boxes correspond to 25%, median, and 75% values, thin bars show 5% and 95%, and squares are used to show the mean. Outliers are also reported as colored diamonds. C-C-FLX R05 is reported in the dashed line because fewer cases were used in the comparison. CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

The induced flux differences vary in average between −1 and 3 W/m^2, depending on data sets (Figure 5). The seasonal variations are different amongst reanalyzes and can be as large as 11 W/m^2 for ERA-I, especially in DJF when the surface temperature is overestimated and the specific humidity close to the surface is underestimated (not shown). ERA5 surface temperature is however in better agreement with the station in all seasons, which causes reduced flux biases. As shown in Figure 5, the upwelling radiative effects of differences in surface temperatures are shown to be nonneglectable, especially during clear sky events, as shown in Section 4.2.3. These temperature differences also impact the determination of downward longwave flux, for example in the case of low level clouds that are in thermal equilibrium with the surface. The uncertainties are in the same order of magnitude, whereas they are somewhat reduced in clear air as the temperature profiles are contributing as a whole through weighting functions.

In this section, we discussed cloud profiles, atmospheric and surface temperature differences that contribute to the observed flux biases amongst data sets. To go further in the comparison of fluxes and to remove sampling bias, we looked to coincident data sets following the B14 approach.

Table 6
As for Table 4, Seasonal Variation of LW BOA for BSRN, EUR-LR, CERES, C-C-FLX (R04 and R05), ERA-I, and ERA5 for the Period From September 2007 to May 2010

BOA	Total		DJF		MAM		JJA		SON	
# of points	149		29		48		19		53	
	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN
BSRN Filtered	205.4 (50.0)		170.3 (37.4)		188.1 (40.4)		272.7 (39.2)		216.0 (40.3)	
EUR-LR	206.2 (50.8)	0.9	171.4 (36.9)	1.0	189.8 (42.0)	1.7	273.4 (37.2)	0.7	216.2 (43.0)	0.1
CERES < 25km	194.8 (52.7)	−10.5	151.0 (32.7)	−19.3	188.7 (44.5)	0.7	281.9 (35.0)	9.1	193.1 (35.3)	−23.0
CERES nearest	195.7 (53.3)	−9.7	151.1 (34.9)	−19.3	190.4 (43.8)	2.4	280.4 (39.1)	7.6	194.5 (37.7)	−21.5
C-C-FLX R04 < 25km	201.4 (51.4)	−4.0	160.2 (38.4)	−10.1	189.4 (41.5)	1.3	272.6 (35.2)	−0.1	209.1 (41.7)	−6.9
C-C-FLX R04 nearest	200.3 (52.3)	−5.1	159.0 (40.2)	−11.4	186.9 (42.8)	−1.1	269.8 (37.6)	−2.9	210.0 (42.2)	−6.0
ERA-I	222.6 (47.2)	17.3	191.9 (44.9)	21.5	215.4 (42.7)	27.3	279.1 (38.2)	6.4	225.8 (36.8)	9.8
ERA5	200.4 (47.3)	−4.9	159.8 (33.5)	−10.5	190.6 (33.0)	2.6	279.4 (27.8)	6.6	203.3 (34.9)	−12.8
# of points	130		24		48		19		39	
	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN	Mean (σ)	Minus BSRN
C-C-FLX R05 < 25km	209.6 (50.7)	1.3	169.3 (36.0)	0.5	193.7 (42.4)	5.7	270.8 (38.8)	−2.0	224.0 (39.7)	−2.0
C-C-FLX R05 nearest	207.7 (52.8)	−0.6	164.8 (37.4)	−4.0	192.1 (43.0)	4.0	269.2 (37.6)	−3.6	223.3 (45.2)	−2.7

Note. Standard deviations are in brackets. Colors are used as in Table 3.
BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

4. Coincident Measurements From Independent Data Sets

In this section, we will focus on TOA and BOA flux analyses for the subsets of coincident observations in space and time. We consider here the data sets that directly provide radiative fluxes (CERES and C-C-FLX R04 and R05) and data sets for which we have calculated fluxes (EUR-LR) using Streamer RTM. Note that ERA-I and ERA5 data points are not strictly coincident with A-Train overpasses, but as they are within approximately 1 h, we included them in the analysis. For TOA we compared all data sets CERES, C-C-FLX R04, and R05, EUR-LR, ERA-I, and ERA5 keeping CERES as a reference, whereas for BOA, the same data sets were considered with BSRN acting as the reference. In this section, the fluxes of each individual overpass within 25 km from the station are extracted and selected depending on the availability of other data sets. The mean fluxes (<25 km) and the nearest fluxes are used to discuss the spatial heterogeneity in Section 4.1.2, Tables 4 and 6. Then we keep the mean fluxes for the rest of the comparison.

4.1. TOA Fluxes

4.1.1. Mean Seasonal TOA Fluxes

Here, we will analyze the evolution of seasonal fluxes at TOA for the different data sets with respect to CERES and discuss correlations and histograms of spread. We will further study differences by type of scene. This analysis includes 249 coincident samples, seasonally distributed as DJF = 56; MAM = 67; JJA = 47; SON = 79. Mean seasonal fluxes from coincident measurements highlight any systematic bias between data sets. Table 4 summarizes the average values determined for all the seasons and the annual mean.

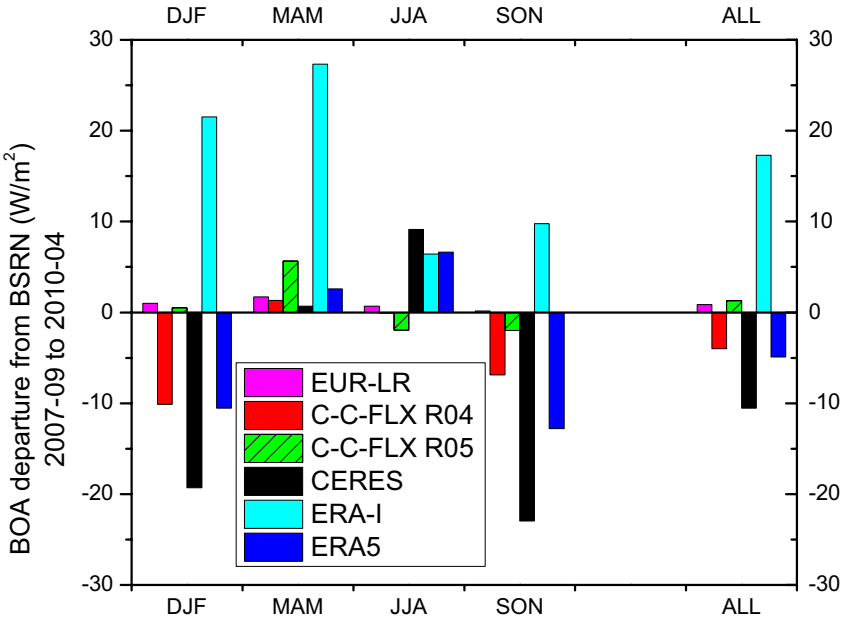

Figure 9. LW BOA departure from BSRN for coincident measurements as reported in Table 6. C-C-FLX R05 bars are dashed because the comparison with BSRN is based on only 130 points. BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere.

Table 4 allows for clarification of conclusions drawn from Table 2. There is a good agreement between all data sets within 5 W/m^2 for annual LW TOA, with values, in general, larger for CERES, the bias being higher during wintertime and, to a lesser extent, in SON. Streamer calculations, based on EUR-LR retrievals, agree well with CERES, except in MAM and JJA, where they are smaller (but by less than 3 W/m^2). C-C-FLX R05 shows a larger deficit than R04 in all seasons with a total bias of −4.3 W/m^2. Several changes in R05 could explain the difference with R04, namely longwave land emissivity and cloud properties (Henderson & L'Ecuyer, 2020).

We propose to take a closer look at those differences in terms of seasonal differences, depending on key parameters.

4.1.2. Seasonal Flux Differences and Spatial Heterogeneity

To assess the role of spatial heterogeneity, we compared the nearest and 25 km circle-mean value of each coincident measurement from CERES and C-C-FLX R04 and R05. Figure 6 displays LW TOA departures from CERES (<25 km) as reported in Table 4. It shows a systematic underestimation (mean annual bias of −3.1 W/m^2 with a standard deviation of 2.1 W/m^2) of LW for all data sets compared to CERES. Part of this difference could be explained by the spatial sampling over a heterogeneous and steep terrain (such as shown for Eureka in Figure 1), where surface temperatures are hard to precisely account for at the different pixel sizes among data set (CERES: 20 × 20 km, C-C-FLX: 1 km, ERA: 14 × 2 km). In DJF, this effect is expected to be smaller because the region is ubiquitously snow and ice-covered, limiting heterogeneity in surface emissivity.

The best agreement is seen from comparisons between CERES and the ground-based EUR-LR flux calculations with an absolute difference being maximum in MAM and JJA (about 2.5 W/m^2). Heterogeneity has a small impact for CERES (less than 0.3 W/m^2 because CERES measurements are to a certain extent smoothed in the 20 × 20 km pixel) and slightly higher for C-C-FLX (up to 1.8 W/m^2 for R04 and 2.5 W/m^2 for R05 in DJF).

During this time of year, thick liquid and mixed-phase clouds obstruct higher clouds, from a ground-based perspective (Figure 4). It further underscores the importance of the cloud vertical distribution in increasing the accuracy of the radiative transfer calculations. As in Table 2, we see in Table 4 and Figure 6 that the

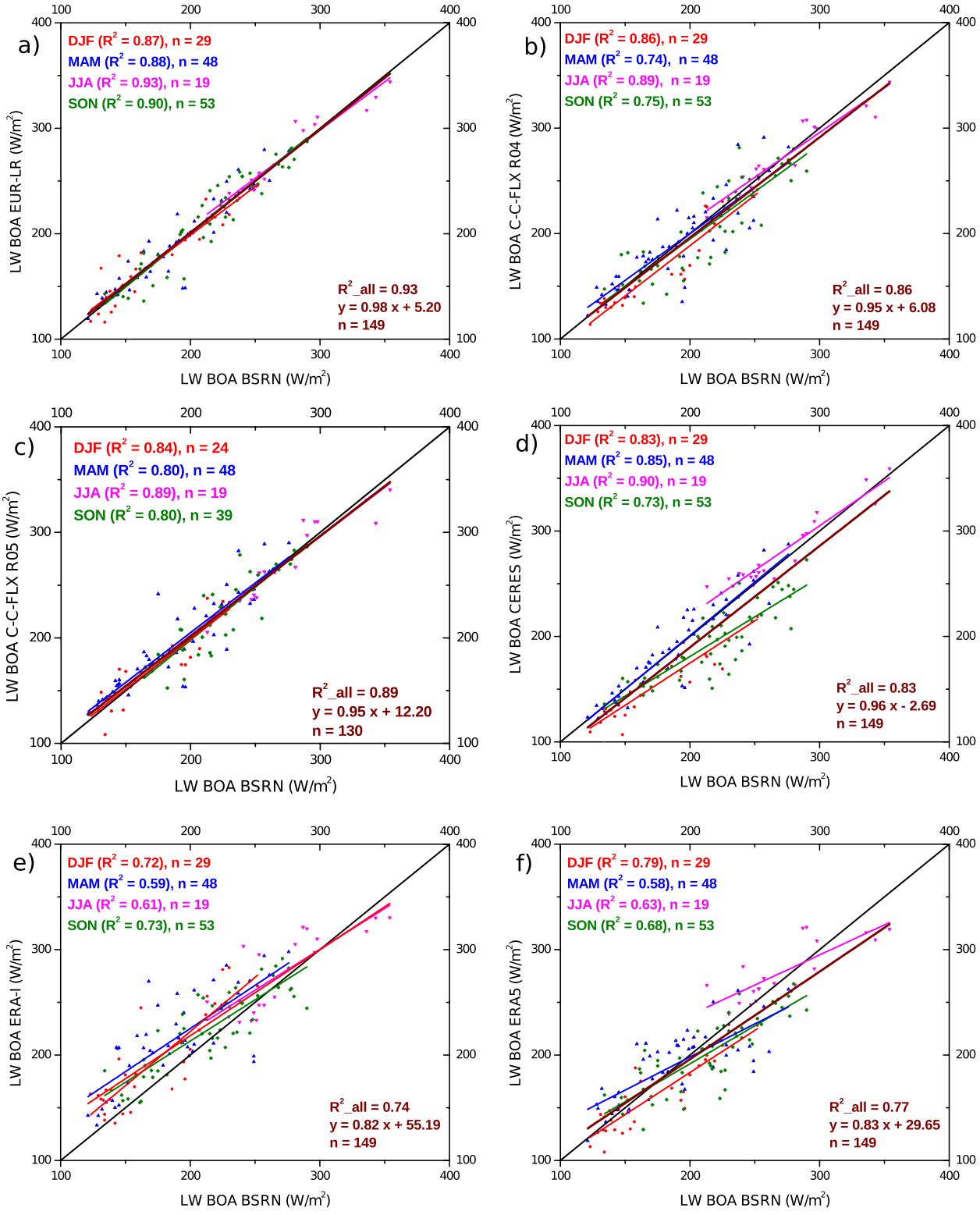

Figure 10. Seasonal BOA downwelling LW fluxes for BRSN and EUR-LR (a) C-C-FLX R04 (b), C-C-FLX R05 (c), CERES (d), ERA-I (e), and ERA5 (f) for the 149 measurements (except for R05) at the same time, between September 2007 and May 2010. BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System.

Table 7
Statistics of BOA Differences Between EUR-LR, C-C-FLX R04, and R05, CERES, ERA-I, ERA5, and BSRN Based on Scatter Plots (Figure 10) and Gaussian Fit (Figure S5)

	EUR-LR	C-C-FLX R04	C-C-FLX R05	CERES	ERA-I	ERA5
R^2	0.93	0.86	0.89	0.83	0.74	0.77
RMSE (W/m^2)	13.4	19.9	17.0	24.0	30.8	24.5
Mean difference (W/m^2)	0.8	−0.5	2.8	−6.8	2.6	−4.7
σ (¼ of full width at $1/e^2$) (W/m^2)	9.9	16.1	12.4	21.1	10.6	24.9
Number of outliers (>3 σ)	5	1	6	3	5	0

Note. A smaller number of points were available for R05 (see Table 6).
BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

C-C-FLX (R04 and R05) and CERES differences are negative for all seasons. A good agreement is however obtained (about −3.8 W/m^2) between C-C-FLX R04 and CERES. It is comparable to the one obtained in the statistical study although results for DJF and SON are degraded (see Figure 7). Compared with Table 2, ERA5 bias with CERES is still small on average but is degraded in JJA and fall. This could be due to the temporal sampling of 3-h re-analysis in seasons when atmospheric properties can rapidly change. ERA-I is behaving slightly differently with larger departures observed in DJF and MAM, consistently with Table 2. Overall the sampling effect does not appear as a first-order reason that can explain the differences in TOA between data sets.

Figure 7 shows scatter plots of the coincident retrieved LW TOA fluxes for EUR-LR, C-C-FLX R04, ERA-I, and ERA5 and histograms of their differences with respect to the CERES data set. Note that C-C-FLX R05 plots are not shown here as the conclusions are similar to R04 and the number of points is smaller (see Table 5 where results are reported). In the scatter plots we have identified both seasonal and overall correlation coefficients. In the histograms and Table 5, we have identified the biases one-fourth of the full width at $1/e^2$ of the maximum (e.g., one standard deviation—σ—of a Gaussian distribution), and the number of points outside three σ to indicate the outliers.

Looking at EUR-LR plots and histograms (S3), a few large seasonal outliers are evidenced in Figure 7a, except in MAM which corresponds to a smaller amount of high clouds (see Figure 4). These outliers are homogeneously distributed below and above the mean. The poorest correlations occurred in DJF, SON, and JJA with differences up to 80 W/m^2 as one can see from the outliers of the histogram, which lead to higher RMSE compared to other data sets. In the case of opaque clouds, ground-based instruments are not able to correctly resolve the vertical profiles of cloud fraction, particle size, and extinction at upper levels due to transmission losses. As a result, the mean cloud temperature is set too high and this causes an overestimation of LW TOA. Another critical scenario is the presence of high clouds, sometimes above opaque clouds. Due to the decreasing radar sensitivity with range and the fact that the lidar signal can be totally attenuated in opaque clouds, it is likely to miss those high clouds and then overestimates LW TOA, with a bias that depends on cloud layer optical depth.

The good agreement noticed before between C-C-FLX R04 (and R05) and CERES correspond to good correlation slopes (Figure 7b) with a weaker number of outliers than for EUR-LR. Differences are larger in DJF and in SON, where the slope is smaller than 1 and the outliers are more numerous, especially with large positives. In the histograms, DJF and SON are indeed characterized by a larger bias with respect to CERES (about −5 W/m^2). For these two seasons, a rather broad dispersion is observed with a few outliers at +30 W/m^2. In DJF and MAM, the difference shows a secondary peak at −10 W/m^2. This is not statistically significant, but it can be due to the under-estimation of low ice clouds if one looks at cloud vertical fraction with respect to EUR-LR, but the overall number of points remains small for that to be significant.

ERA-I appears slightly biased low by about 7.4 W/m^2 on average, and the number of outliers is large (Table 5). In Figure 7c, however, their distribution is different from the EUR-LR one with a large number of positive values creating a secondary peak at +30 W/m^2. Finally, we find that ERA5 LW TOA is on average

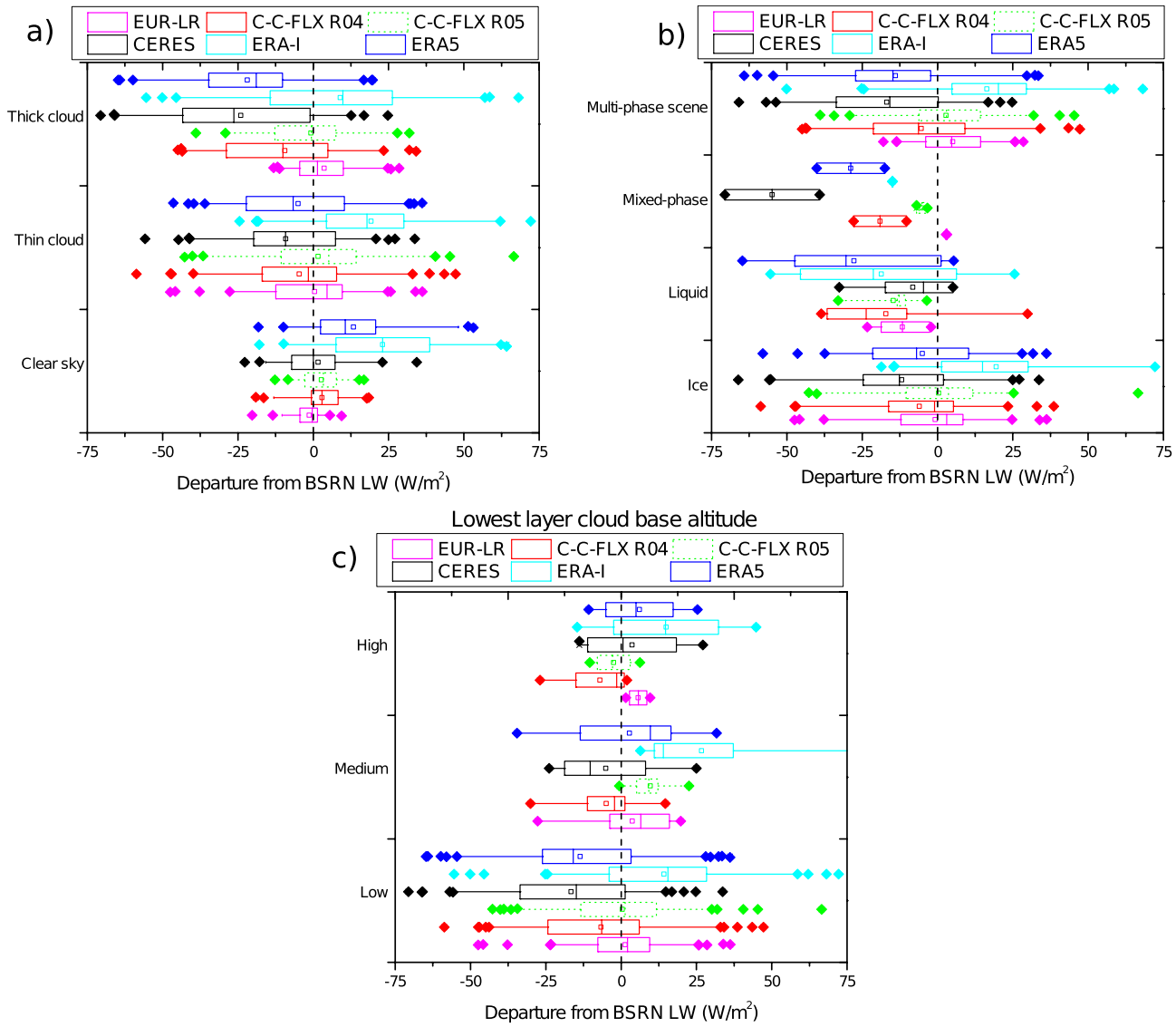

Figure 11. Difference of BOA LW fluxes depending on the total visible optical depth (a), phase of cloud layers (b), and bottom height of the lowest layer (c). Boxes correspond to 25%, median, and 75% values, thin bars show 5% and 95%, and squares are used to show the mean. Outliers are also reported as colored diamonds. BOA, Bottom of the Atmosphere.

in relatively better agreement if we consider all points with a global correlation of $R^2 = 0.91$, close to what is obtained for ERA-I. Reanalysis seasonal scatter plots and histograms (Figures 7c and 7d) highlight opposite patterns in winter and JJA, with a much smaller correlation. It can be in part explained by the fact that ERA5 overestimates cloud cover (especially low cloud) in the Arctic in wintertime, in a way similar to (but less than) the one already shown for ERA-I (see Figure 4), and consistent with previous work, done with ERA-I (Chernokulsky & Mokhov, 2012; Zygmuntowska et al., 2012). The modest correlation in JJA ($R^2 = 0.46$) may be linked to the fact that ECMWF reanalyzes underestimate by half the liquid water content of JJA clouds (Huang, Dong, Xi, Dolinar, & Stanfield, 2017; Zygmuntowska et al., 2012). This will be further discussed in the next subsection.

4.1.3. TOA Differences Decomposed by Cloud Optical Depth, Type of Scenes and Height of Lower Layer

To validate the hypothesis of Section 4.1.2, that most differences are due to clouds, we now plot the LW differences depending on the total visible optical depth (Figure 8a), the phase of cloud layers (Figure 8b),

Table 8
Findings for Each Data Set

	Cloud vertical distribution	TOA	BOA
BSRN	N/A	N/A	Used as a reference
EUR-LR	Detects a large number of hydrometeors close to the ground in winter. Is not able to detect high features above opaque clouds (in MAM and JJA).	Relatively good agreement (about 1 W/m^2). Issues when high clouds are not detected	Very good agreement with BSRN (bias less than 1 W/m^2)
C-C-FLXR04	Misses a significant amount of low ice clouds in winter. Good agreement with ground-based above 2 km.	Always smaller than CERES (about −4 W/m^2). Could be due to footprint compared to CERES (20 × 20 km), which smooths TOA fluxes and differences in surface temperature.	Overall bias close to −5 W/m^2. Better than CERES. Bigger bias in winter (−10 W/m^2)
C-C-FLX R05	Same as C-C-FLX R04	Smaller than CERES with larger bias compared to R04, probably due to different surface emissivity	Good agreement with BSRN (bias less than 1.5 W/m^2). TToo many mixed-phase are detected (in spring).
CERES	Misses clouds in all seasons, but this is more dramatic in winter over a snow surface.	Used as a reference	The overall bias of −10 W/m^2). Differences are higher in winter (−20 W/m^2) and autumn due to a wrong detection of low clouds.
ERA-I	Underestimates cloud fraction by a factor of 2. This is somewhat compensated by an overestimation of clouds very close to the surface.	Bias for clear sky due to a coarse temperature profile that misses temperature inversion.	Biggest overall bias (+20 W/m^2) larger in winter, as caused by a poor reanalysis of cloud vertical distribution and temperature profiles. Issues with water, ice clouds, and clear sky.
ERA5	Similar to ERA-I, but showing a lower amount of low clouds.	In good agreement with CERES, except in summer due to inaccurate liquid water content	Overall bias close to −5 W/m^2. Larger bias in winter (−10 W/m^2). Clear sky bias is still present (13 W/m^2) but reduced compared to ERA-I. Low liquid and ice clouds are the main sources of errors

BSRN, Baseline Surface Radiation Network; BOA, Bottom of the Atmosphere; CERES, Clouds and the Earth's Radiant Energy System; TOA, Top of the Atmosphere.

and the top altitude of the lowest layer (Figure 8c). All data sets plotted in solid line are compared to CERES with the same number of cases (249) whereas C-C-FLX R05 is shown in green dashed line because only 221 cases were available.

Note that the type of scene classification (either clear, thin, or thick clouds), phase of clouds (liquid water, ice, and mixed-phase or multiple phase scene), and top height of the lower layer are based on EUR-LR observations and therefore depends on instrument sensitivity and can be biased in the case of opaque clouds and very thin clouds. Thick/Thin clouds threshold is set to a total visible optical depth of 2. Multiple phase scene indicates that layers with different phase are present in the column.

Figure 8a confirms the relatively good agreement for TOA LW for clear-sky scenes. With a decreased departure from −8.7 to −2.6 W/m^2, ERA5 reanalyzes of the clear sky are improved compared to ERA-I, where surface emissivity, surface, and air temperature, humidity, or atmospheric absorption have been identified as a possible source of discrepancies (Graham, Hudson, et al., 2019). The comparison between both C-C-FLX data sets shows degraded statistics for R05 relative to R04, especially for the clear sky (median bias of −3.6 W/m^2 for R04 and −5.5 W/m^2 for R05). This could be due to changes in R05 implementation of longwave land emissivity, which is relatively complex to parametrize in heterogeneous and steep terrain like Eureka. There is a warm bias for EUR-LR due to the presence of thick clouds (COD > 2) when the lidar signal is extinguished and the cloud layer top altitude is not precisely found. Therefore, the EUR-LR cloud layer is wrongly positioned (too low, too warm). TOA departures based on cloud type are fairly similar amongst data sets. EUR-LR fails to get correct LW TOA when high clouds are present, due to a decrease in radar sensitivity for small particles (as discussed in Grenier et al., 2009 and Blanchard et al., 2017).

Ice layers are very frequent and cause a large spread in TOA differences. There are very few liquid-phase clouds only (7) and mixed-phase only (2) cases. Figure 8c shows that all types of clouds are mainly biased

low with respect to CERES TOA measurements, with an emphasis on high clouds. This is rather surprising for C-C-FLX R04 and R05, because of the high sensitivity of lidar to cirrus clouds as evidenced in the high C-C-FLX cloud occurrence reported in Figure 4. A possible reason could be inaccurate estimations of ice water content and microphysics in flux calculations, but this evaluation is beyond the scope of this study. It has also to be noted that some additional discrepancies could occur due to the temperature inversion layer which could be badly captured with GMAO or ERA5 coarse vertical resolutions.

Overall two main issues are confirmed here: The bias in high clouds for EUR-LR, and the bias in clear air identified for ERA-I is now corrected in ERA5.

4.2. BOA Fluxes

4.2.1. Mean Seasonal BOA Fluxes

Between September 2007 and May 2010, both active instruments and BSRN sensor were operational at Eureka. Repeating the same methodology as in Section 4.1, we first discussed annual and seasonal statistics.

From Table 6, we can see that there is a wider span of annual LW BOA averages amongst data sets, from 194.8 (CERES) to 222.6 W/m^2 (ERA-I). We restate that the lack of availability of ground-based measurements during the 2006 and 2008 JJA and icing screening from BSRN can induce sampling seasonal effects. Therefore, those values are not expected to be used for climatological analysis. We see that the spatial sampling effect of CERES and C-C-FLX fluxes (labeled as 25 km and nearest) is relatively small compared to the difference with BSRN and can be explained by a mixture of cloud edges or transition with a clear sky. Differences remain high and comparable for all fluxes excepting CERES. The dispersion on average annual values are of the order of 5 W/m^2, excepting CERES and ERA-I data, but those on seasonal values can be about twice larger in winter when the number of cases is reduced to 29. One can see that differences are larger than for JJA (about 6 W/m^2) when the number of points is even more reduced (19). In all cases, standard deviations remain high, and residual uncertainties on average values (standard deviation divided by the square root of the number of points) are 4 W/m^2 (annual average) to 8 W/m^2 (JJA). These values have to be kept in mind in the discussion of observed differences.

BSRN and EUR-LR agree well (within 1.7 W/m^2 over all seasons and better than 1 W/m^2 on average), confirming the high level of confidence of the combination of active measurements with the Streamer simulations. BSRN field-of-view angular integration and EUR-LR time integration are also contributing to this agreement.

While CERES and C-C-FLX R04 were in good agreement in Table 3, the coincident comparison showed that differences can be up to 16 W/m^2 in autumn and about 6.5 W/m^2 in annual mean (CERES being biased low with respect to C-C-FLX R04 by about 5 W/m^2). C-C-FLX R05 shows reduced bias in all seasons except in MAM and JJA. The better representation of ice and mixed-phase clouds in R05 could explain this improvement and this hypothesis will be discussed in the next section. Satellite observations are lower than BSRN in all seasons and more particularly in DJF (-10 to -20 W/m^2), except for C-C-FLX R05, while ERA-I is systematically much higher than BSRN for all seasons (between 6.4 and 27 W/m^2) and more than 15 W/m^2 on average, which is consistent with the overestimation of cloud fraction at low altitude (Zib et al., 2012). The several modifications implemented for ERA5 have a significant impact as it decreases BOA LW fluxes by 22 W/m^2 with respect to ERA-I, and even more in winter. We found that ERA5 is in general in much better agreement with other data sets.

Figure 9 is reporting the differences observed in Table 6. It evidences that the largest differences are observed for CERES and ERA-I, with annual biases of -11 and $+17$ W/m^2, respectively. Largest differences in autumn and winter (and MAM for ERA-I), when the cloud vertical distribution was divergent. The difference observed between MAM and JJA and the two other seasons between C-C-FLX and CERES is statistically meaningful. A closer look at those differences will help to understand the biases.

4.2.2. Seasonal Flux Differences

Looking at coincident fluxes in a way similar to TOA analysis helps to identify systematic seasonal, methodological or instrumental biases compared to BSRN reference. We must be aware however that due to the

footprint of satellite observations, it may be possible that although the separation distance is kept small, ground-based active instruments may not be looking at the same cloud. The reduced number of cases makes the multi-parameter analysis more difficult in terms of quantification. Figure 10 is reporting (as in Figure 7 for TOA fluxes) one-to-one flux comparisons and histograms of flux differences, however, evidence of significant differences. Table 7 summarizes the main parameters reported in histograms.

It is apparent from the histograms of BOA flux differences given in Figure 10 that all comparisons show very large dispersions except between EUR-LR and BSRN. No significant bias and very few outliers (close to ± 40 W/m^2, as evidenced from the narrower distribution) is observed in this last case. The correlation between EUR-LR and BSRN is indeed high ($R^2 = 0.93$, and $\sigma = 9.9$ W/m^2 and RMSE being the lowest between the data sets), but a comparison of individual coincident times can be off by up to 50 W/m^2. Those outliers are likely explained by the fact that the effective spatial resolution of active instruments after time averaging remains small compared to the BSRN pyranometer/pyrgeometer, located 2.3 km away from the station, which measures hemispheric (160°) LW fluxes.

The overall distributions are widely spread in almost all other cases from -60 to $+60$ W/m^2. Results between C-C-FLX (R04 and R05) and BSRN show a dispersion of seasonal differences rather contained, limiting the overall bias, slightly better for R05. Figure 10c shows a smaller dispersion of C-C-FLX R05 bias versus BSRN, as compared to CERES. This confirms the advantage of lidar-radar synergy from space. But there are still high seasonal variations, especially in MAM, when C-C are missing clouds below 5 km (Figure 4). Differences between R04 and R05 are minor but an overall better agreement is found for R05 (RMSE reduced by 2.9 W/m^2) and especially in winter (the R04 bias of -10 W/m^2 is reduced to $+0.5$ W/m^2 for R05). Some improvements in the R05 algorithm regarding ice and mixed-phase clouds could explain those differences.

CERES and ERA5 show the wider spreads with a dispersion (2 sigmas) of about 50 W/m^2. The mean biases vary from about -20 W/m^2 in winter and autumn to about $+10$ W/m^2 in JJA for CERES. In those winter and autumn seasons, MODIS is missing almost half of the low and mid clouds (see Figure 4). ERA5 is also biased low in winter and autumn compared to BSRN and larger in JJA. This result is in agreement with what was reported in Zib et al. (2012) for Ny-Ålesund and Barrow, but using ERA-Interim. Note that those stations are, however, coastal and their cloud fraction variation is different from the one at Eureka (Shupe, 2011; Shupe et al., 2011). ERA5 corrects the BOA LW positive bias for ERA-I in winter explained by the overestimation of very low cloud in reanalyzes (as reported in several studys, see Huang, Dong, Xi, Dolinar, & Stanfield, 2017; Zib et al., 2012; Zygmuntowska et al., 2012) as seen in Figure 4. C-C-FLX and ERA5 average values show biases in winter remain high (larger than -10 W/m^2) with broad distributions.

4.2.3. BOA Differences Decomposed by Type of Scenes

In this subsection, we discuss the overall relative difference of BOA fluxes to BSRN measurements as a function of the scene type as supported by ground-based lidar/radar observations.

Clear sky events (Figure 11a) are well captured by CERES, C-C-FLX, and EUR-LR while the mean bias is about $+13$ W/m^2 for ERA5. For opaque clouds (COD > 2), CERES and C-C-FLX are biased low because they miss correct cloud base heights as identified in Figure 4. Such clouds are expected to be mid- and low-level clouds. Indeed, high clouds that were missed by EUR-LR above opaque clouds (in the discussion about TOA) don't have a significant impact when looking at downwelling fluxes. However, it is confirmed that all other (CERES, ERA5, and to a smaller extend CC-FLX R05) LW BOA fluxes are biased low for low clouds by more than 20 W/m^2 with respect to BSRN BOA measurements, which appears to be the main driver of biases. The direct comparison of C-C-FLX releases (see Figure S6) confirms that R05 significantly reduces the strong biases identified for low and thick clouds. Figure 11b shows that there is a negative bias for ice layers for all data sets considered here except EUR-LR, as evidenced in Table 6. Mixed-phase (supercooled) clouds appear to be challenging for CERES as previously emphasized (Matus & L'Ecuyer, 2017). This also appears to be the case for ERA5 and ERA-I, but there are rather few mixed-phase cases here to draw any definitive conclusion (Figure 11c). As mentioned for TOA, ECMWF reanalyzes are still struggling to get the water content of liquid clouds correct. This remains true for ERA5 as was the case for ERA-I (Zygmuntowska et al., 2012).

4.3. Summary

Further to B14, we have applied in this paper the approach that was laid out for cloud occurrence to the analysis of LW radiation budget at the TOA and BOA at Eureka station. The statistical analyses are enforced by an approach of separating statistical independent analysis and coincident confrontation of observations constrained by scene types. This approach controls for some sampling and observational biases that affect the analysis, and the horizontal heterogeneity was found to be a small factor. The results indicate that there is rather good agreement in TOA fluxes (within a few W/m^2), but considerably less agreement in the arguably more important BOA fluxes. The main findings are summarized in Table 8.

For the TOA, we used CERES as a reference. From the statistical independent analysis, we found results comparable to what has been previously obtained, with the good agreement (better than 5 W/m^2) between data sets and low biases.

Observations of broadband longwave radiation using a surface passed pyrgeometer as part of BSRN were used as a reference for BOA analysis. A careful examination of each coincident case was undertaken to improve the quality and confidence in the measurements incorporated into the analysis. Comparison with fluxes determined using Streamer code using inputs from ground-based lidar-radar vertical profiles of cloud properties and meteorological data gave a very high agreement (with a standard deviation of less than 5 W/m^2), comparable to the agreement obtained for TOA fluxes. This is a remarkable result in the comparison of the BOA fluxes, for which deviation among data set is much larger, as with satellite and reanalysis data. Low opaque clouds in wintertime are found to be the most challenging to detect for passive radiometry due to the small temperature difference with the underlying snow surface. Those clouds are not well identified by CERES, as MODIS underestimates cloud fraction especially in winter and autumn (Figure 4). This remains an issue for active sensors as well, although to a smaller extent, but ground clutter, smaller droplets for optically thick water clouds for the radar, and transmission decrease for the lidar are significant issues limiting overall performance. Recent reanalyzes ERA5 are improved, as differences from references are reduced compared to ERA-Interim. Some bias, however, is persistent for clear sky and cloud vertical profile, which shows the need for improving model resolutions.

Statistical and coincident analysis revealed comparable agreement in TOA with biases smaller than 5 W/m^2 for all observations and analyses with respect to CERES observations. No obvious trend was found on the statistical data set. Narrow distributions are observed for satellite observations, but a larger dispersion is seen on analyses, with a larger number of outliers for ERA5. The difference observed appears to be mainly due to high clouds. Their occurrence is slightly smaller for CERES at higher altitudes. This may be due to the fact that the altitude attribution is underestimated by MODIS.

The results for BOA fluxes show more differences. Ground-based lidar-radar inputs to radiative transfer code (streamer) give at the same time unbiased fluxes with the lower dispersion with respect to BSRN reference. All other (CC-FLX, CERES, and ERA5) show biases ranging from 1 (C-C-FLX R05) to −10 W/m^2 (CERES and ERA5) analyzed as due to poor representation of low (mixed-phase) cloud properties (liquid water content). ERA5 corrects the very large positive longwave bias of +17 W/m^2 observed with ERA-I, but cloud distribution remains biased with respect to ground-based observations. Further improvements thus remain to be done in both the retrieved fluxes from observations and analyses to better address the liquid water content of complex Arctic low-level clouds observed in cold seasons.

It is important to remind that Eureka has a lower frequency of low level (usually mixed-phase) clouds and more high clouds compared to other arctic locations (Shupe et al., 2011). Given that low, thick clouds are the source of some of the biases identified in this paper, one might expect that errors associated with them are more influential at other locations where such clouds are also more prevalent.

5. Conclusions

Existing TOA flux observing and modeling strategies are in good agreement and seem sufficient. BOA fluxes on the other hand are more problematic and while there is an agreement between the ground-based broadband observations and ground-based radar-lidar retrievals, these are only for infrequent, single observatory

sites, and model and satellite methodologies to characterize BOA fluxes are still insufficient for monitoring or characterizing the Arctic system.

It is essential for future operations that active sensors at ground-based sites be operated in polar regions to complement space observations in order to correctly identify cloud vertical profiles. Without integration of the ground-based, ongoing reference data sets into observing strategies it seems unlikely that space-based observations or model projections will be able to independently measure or calculate the BOA fluxes that are a critical component for characterizing and monitoring the extreme environmental changes occurring in the Arctic environment.

Appendix A: Annual Cloud-Type Vertical Distribution

Cloud vertical distribution and cloud type are indeed key parameters in flux calculations. We here compare input vertical profiles from satellite and ground/based measurements and reanalysis between June 2006 and May 2010 (green lines in Figure 2). In this study, as a conclusion from B14, the EUR-LR is considered to be the reference for the low-level clouds, whereas space radar-lidar data are considered as such for upper-level data (>6 km).

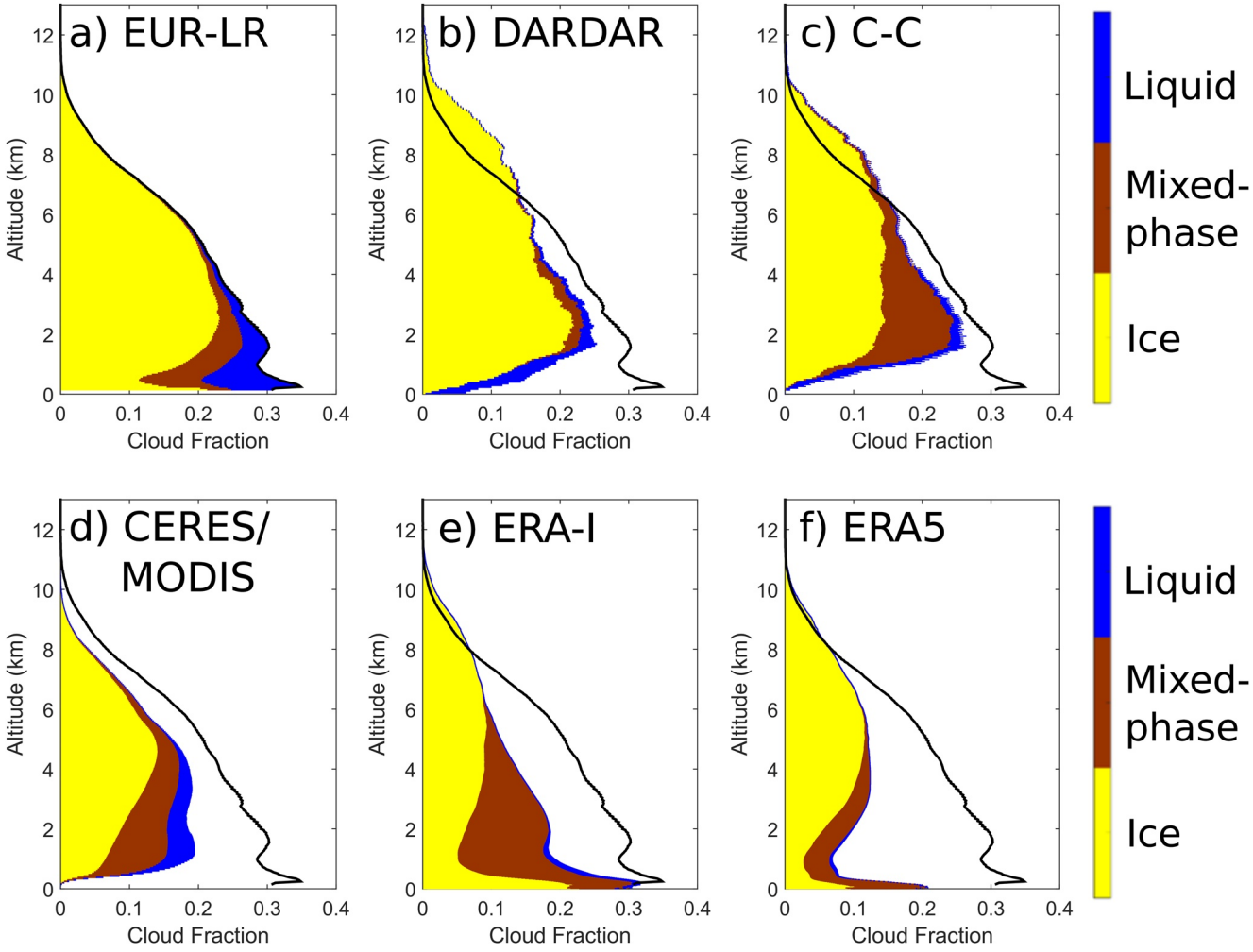

Figure A1. The cumulated vertical cloud-type distribution between June 2006 and May 2010 for all the independent data sets (a) EUR-LR; (b) DARDAR; (c) CloudSat 2B-CLDCLASS-LIDAR R05 (C-C); (d) MODIS; (e) ERA-I; (f) ERA5 at less than 25 km from the station as compared to ground-based observations at EUREKA (solid line) given as the total cloud occurrence in Figure 4. We have reported EUR-LR total cloud fraction from Figure A1a in all other Figures A1b, A1c, A1d, A1e, and A1f.

In Figure A1, DARDAR and C-C cloud vertical distributions are very close to EUR-LR above 2 km. MODIS is close to DARDAR, although MODIS is slightly more biased below 4 km. As detailed in B14, very low clouds are difficult to address from space, and this is confirmed here from DARDAR, C-C and MODIS for which cloud fractions are much lower than EUR-LR observations below 2 km. DARDAR and C-C are close although DARDAR gives a higher amount of ice clouds and less mixed-phase clouds. The finer vertical resolution for DARDAR (60 m) compared to C-C (240 m) might explain this difference as C-C would not be able to distinguish different water phases within one radar gate. ERA-I misses a large fraction of cloud below 8 km (Liu & Key, 2016). ERA5 appears to have a bias larger than ERA-I, and the fraction of mixed-phased clouds are observed to be much smaller. All behaviors are however rather similar, with more or less important bias in the vertical cloud fraction but significant biases below 2 km.

Data Availability Statement

Acknowledgments also go to ECMWF for the ERA5 and ERA-I reanalyzes and to the NASA CERES team for the production of the CERES data. NASA CERES–SSF-TOA and SSF-Surface products are available at http://ceres.larc.nasa.gov/. The BSRN data set can be accessed from World Radiation Monitoring Center at http://www.bsrn. awi.de/. Links to access data and a list of filenames used in this study can be found in Table S1.

Acknowledgments
Acknowledgments are due to the Canadian Network for the Detection of Atmospheric Change (CANDAC), Study of Environmental Arctic Change (SEARCH), and Environment Canada for their operational support as well as radiosondes and observations data at Eureka. The authors would like to acknowledge the support from the NOAA Arctic Research Program on processing ground-based data. The authors are grateful to Kenneth Moran from NOAA and Edwin Eloranta from the University of Wisconsin-Madison for providing MMCR and AHSRL data, respectively. The authors thank the CALIPSO team at NASA Langley Research Center (https://www-calipso.larc.nasa.gov/), the ICARE data Center in Lille/France (https://www.icare.univ-lille.fr/) and the CloudSat team at Colorado State University for the availability of the level-2 data (http://www.cloudsat.cira.colostate.edu/data-products/). This research was supported by the Center National d'Etudes Spatiales (CNES).

References

Abe, M., Nozawa, T., Ogura, T., & Takata, K. (2016). Effect of retreating sea ice on Arctic cloud cover in simulated recent global warming. *Atmospheric Chemistry and Physics*, *16*(22), 14343–14356. https://doi.org/10.5194/acp-16-14343-2016

Ancellet, G., Pelon, J., Blanchard, Y., Quennehen, B., Bazureau, A., Law, K. S., & Schwarzenboeck, A. (2014). Transport of aerosol to the Arctic: Analysis of CALIOP and French aircraft data during the spring 2008 POLARCAT campaign. *Atmospheric Chemistry and Physics*, *14*(16), 8235–8254. https://doi.org/10.5194/acp-14-8235-2014

Betts, A. K., Chan, D. Z., & Desjardins, R. L. (2019). Near-surface biases in ERA5 over the Canadian Prairies. *Frontiers in Environmental Science*, *7*. https://doi.org/10.3389/fenvs.2019.00129

Blanchard, Y., Pelon, J., Eloranta, E. W., Moran, K. P., Delanoë, J., & Sèze, G. (2014). A synergistic analysis of cloud cover and vertical distribution from A-Train and ground-based sensors over the high Arctic station Eureka from 2006 to 2010. *Journal of Applied Meteorology and Climatology*, *53*(11), 2553–2570. https://doi.org/10.1175/jamc-d-14-0021.1

Blanchard, Y., Royer, A., O'Neill, N. T., Turner, D. D., & Eloranta, E. W. (2017). Thin ice clouds in the Arctic: Cloud optical depth and particle size retrieved from ground-based thermal infrared radiometry. *Atmospheric Measurement Techniques*, *10*(6), 2129–2147. https://doi.org/10.5194/amt-10-2129-2017

Boisvert, L. N., & Stroeve, J. C. (2015). The Arctic is becoming warmer and wetter as revealed by the atmospheric infrared sounder. *Geophysical Research Letters*, *42*(11), 4439–4446. https://doi.org/10.1002/2015gl063775

Ceccaldi, M., Delanoë, J., Hogan, R. J., Pounder, N. L., Protat, A., & Pelon, J. (2013). From CloudSat-CALIPSO to EarthCare: Evolution of the DARDAR cloud classification and its comparison to airborne radar-Lidar observations. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7962–7981. https://doi.org/10.1002/jgrd.50579

Chan, M. A., & Comiso, J. C. (2011). Cloud features detected by MODIS but not by CloudSat and CALIOP. *Geophysical Research Letters*, *38*(24). https://doi.org/10.1029/2011gl050063

Chernokulsky, A., & Mokhov, I. I. (2012). Climatology of total cloudiness in the Arctic: An intercomparison of observations and reanalyzes. *Advances in Meteorology*, *2012*, 1–15. https://doi.org/10.1155/2012/542093

Christensen, M. W., Behrangi, A., L'ecuyer, T. S., Wood, N. B., Lebsock, M. D., & Stephens, G. L. (2016). Arctic observation and reanalysis integrated system: A new data product for validation and climate study. *Bulletin of the American Meteorological Society*. *97*(6), 907–916. https://doi.org/10.1175/BAMS-D-14-00273.1

Comiso, J. C., & Hall, D. K. (2014). Climate trends in the Arctic as observed from space. *WIREs Climate Change*, *5*(3), 389–409. https://doi.org/10.1002/wcc.277

Cox, C. J., Morris, S. M., Uttal, T., Burgener, R., Hall, E., Kutchenreiter, M., et al. (2020). The de-icing comparison experiment (D-ICE): A study of broadband radiometric measurements under icing conditions in the Arctic, *Atmospheric Measurements Techniques*, *14*(2), 1205–1224. https://doi.org/10.5194/amt-2020-397

Cox, C. J., Uttal, T., Long, C. N., Shupe, M. D., Stone, R. S., & Starkweather, S. (2016). The role of springtime Arctic clouds in determining autumn sea ice extent. *Journal of Climate*, *29*(18), 6581–6596. https://doi.org/10.1175/jcli-d-16-0136.1

Cox, C. J., Walden, V. P., & Rowe, P. M. (2012). A comparison of the atmospheric conditions at Eureka, Canada, and Barrow, Alaska (2006–2008). *Journal of Geophysical Research*, *117*(D12). https://doi.org/10.1029/2011jd017164

Cox, C. J., Walden, V. P., Rowe, P. M., & Shupe, M. D. (2015). Humidity trends imply increased sensitivity to clouds in a warming Arctic. *Nature Communications*, *6*(1). https://doi.org/10.1038/ncomms10117

de Boer, G., Eloranta, E. W., & Shupe, M. D. (2009). Arctic mixed-phase stratiform cloud properties from multiple years of surface-based measurements at two high-latitude locations. *Journal of the Atmospheric Sciences*, *66*(9), 2874–2887. https://doi.org/10.1175/2009jas3029.1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

Delanoë, J., & Hogan, R. J. (2008). A variational scheme for retrieving ice cloud properties from combined radar, Lidar, and infrared radiometer, *Journal of Geophysical Research*, *113*, D07204. https://doi.org/10.1029/2007JD009000

Delanoë, J., & Hogan, R. J. (2010). Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Journal of Geophysical Research*, *115*. https://doi.org/10.1029/2009jd012346

Donovan, D. P., & van Lammeren, A. C. A. P. (2001). Cloud effective particle size and water content profile retrievals using combined Lidar and radar observations: 1. Theory and examples. *Journal of Geophysical Research*, *106*(D21), 27425–27448. https://doi.org/10.1029/2001JD900243

Doyle, J. G., Lesins, G., Thackray, C. P., Perro, C., Nott, G. J., Duck, T. J., et al. (2011). Water vapor intrusions into the high Arctic during winter. *Geophysical Research Letters*, *38*(12). https://doi.org/10.1029/2011gl047493

Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., et al. (2018). Baseline surface radiation network (BSRN): Structure and data description (1992–2017), *Earth System Science Data*, *10*, 1491–1501. https://doi.org/10.5194/essd-10-1491-2018

Eastman, R., & Warren, S. G. (2010). Interannual variations of Arctic cloud types in relation to sea ice. *Journal of Climate*, *23*(15), 4216–4232. https://doi.org/10.1175/2010jcli3492.1

Eloranta, E. W., Uttal, T., & Shupe, M. (2007). Cloud particle size measurements in Arctic clouds using Lidar and radar data. Presented at the *2007 IEEE International Geoscience and Remote Sensing Symposium*. https://doi.org/10.1109/igarss.2007.4423292

English, J. M., Gettelman, A., & Henderson, G. R. (2015). Arctic radiative fluxes: Present-day biases and future projections in CMIP5 models. *Journal of Climate*, *28*(15), 6019–6038. https://doi.org/10.1175/jcli-d-14-00801.1

Fogal, P. F., Leblanc, L. M., & Drummond, J. R. (2013). The polar environment atmospheric research laboratory (PEARL): Sounding the atmosphere at 80°North. *Arctic*, *66*(3). https://doi.org/10.14430/arctic4321

Grachev, A. A., Persson, P. O. G., Uttal, T., Akish, E. A., Cox, C. J., Morris, S. M., et al. (2018). Seasonal and latitudinal variations of surface fluxes at two Arctic terrestrial sites. *Climate Dynamics*, *51*(5–6), 1793–1818. https://doi.org/10.1007/s00382-017-3983-4

Graham, R. M., Cohen, L., Petty, A. A., Boisvert, L. N., Rinke, A., Hudson, S. R., et al. (2017). Increasing frequency and duration of Arctic winter warming events. *Geophysical Research Letters*, *44*(13), 6974–6983. https://doi.org/10.1002/2017gl073395

Graham, R. M., Cohen, L., Ritzhaupt, N., Segger, B., Graversen, R. G., Rinke, A., et al. (2019). Evaluation of six atmospheric reanalyzes over Arctic sea ice from winter to early summer. *Journal of Climate*, *32*(14), 4121–4143. https://doi.org/10.1175/jcli-d-18-0643.1

Graham, R. M., Hudson, S. R., & Maturilli, M. (2019). Improved performance of ERA5 in Arctic gateway relative to four global atmospheric reanalyzes. *Geophysical Research Letters*, *46*(11), 6138–6147. https://doi.org/10.1029/2019gl082781

Grenier, P., Blanchet, J. P., & Muñoz-Alpizar, R. (2009). Study of polar thin ice clouds and aerosols seen by CloudSat and CALIPSO during midwinter 2007. *Journal of Geophysical Research*, *114*(D9). https://doi.org/10.1029/2008jd010927

Gupta, S. K., Darnell, W. L., & Wilber, A. C. (1992). A parameterization for longwave surface radiation from satellite data: Recent improvements. *Journal of Applied Meteorology*, *31*(12), 1361–1367. https://doi.org/10.1175/1520-0450(1992)0312.0.co;2

Gupta, S. K., Kratz, D. P., Stackhouse, P. W., Wilber, A. C., Zhang, T., & Sothcott, V. E. (2010). Improvement of surface longwave flux algorithms used in CERES processing. *Journal of Applied Meteorology and Climatology*, *49*(7), 1579–1589. https://doi.org/10.1175/2010jamc2463.1

Henderson, D. S., & L'Ecuyer, T. (2020). *CloudSat level 2B fluxes and heating rates with lidar [2B-FLXHR-LIDAR] process description and interface control document*. Retrieved from http://www.cloudsat.cira.colostate.edu/sites/default/files/products/files/2B-FLXHR-LIDAR_PDICD.P1_R05.rev0_.pdf

Henderson, D. S., L'Ecuyer, T., Stephens, G., Partain, P., & Sekiguchi, M. (2013). A multisensor perspective on the radiative impacts of clouds and aerosols. *Journal of Applied Meteorology and Climatology*, *52*(4), 853–871. https://doi.org/10.1175/jamc-d-12-025.1

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*: 1999– 2049. https://doi.org/10.1002/qj.3803

Hu, Y., Winker, D., Vaughan, M., Lin, B., Omar, A., Trepte, C., et al. (2009). CALIPSO/CALIOP cloud phase discrimination algorithm. *Journal of Atmospheric and Oceanic Technology*, *26*(11), 2293–2309. https://doi.org/10.1175/2009jtecha1280.1

Huang, Y., Dong, X., Xi, B., Dolinar, E. K., & Stanfield, R. E. (2017). The footprints of 16 year trends of Arctic springtime cloud and radiation properties on September sea ice retreat. *Journal of Geophysical Research: Atmospheres*, *122*(4), 2179–2193. https://doi.org/10.1002/2016jd026020

Huang, Y., Dong, X., Xi, B., Dolinar, E. K., Stanfield, R. E., & Qiu, S. (2017). Quantifying the uncertainties of reanalyzed Arctic cloud and radiation properties using satellite surface observations. *Journal of Climate*, *30*(19), 8007–8029. https://doi.org/10.1175/jcli-d-16-0722.1

Igel, A. L., Ekman, A. M. L., Leck, C., Tjernström, M., Savre, J., & Sedlar, J. (2017). The free troposphere as a potential source of arctic boundary layer aerosol particles. *Geophysical Research Letter*, *44*, 7053–7060. https://doi.org/10.1002/2017GL073808

Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., et al. (2015). The EarthCARE satellite: The next step forward in global measurements of clouds, aerosols, precipitation, and radiation. *Bulletin of the American Meteorological Society*, *96*(8), 1311–1332. https://doi.org/10.1175/bams-d-12-00227.1

IPCC (2013). Climate change 2013: The physical science basis. In Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Eds.), *Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on climate change*. Cambridge University Press.

Jun, S.-Y., Ho, C.-H., Jeong, J.-H., Choi, Y.-S., & Kim, B.-M. (2016). Recent changes in winter Arctic clouds and their relationships with sea ice and atmospheric conditions. *Tellus A: Dynamic Meteorology and Oceanography*, *68*(1), 29130. https://doi.org/10.3402/tellusa.v68.29130

Kay, J. E., & L'ecuyer, T. (2013). Observational constraints on Arctic Ocean clouds and radiative fluxes during the early 21st century. *Journal of Geophysical Research: Atmospheres*, *118*(13), 7219–7236. https://doi.org/10.1002/jgrd.50489

Kay, J. E., L'Ecuyer, T., Chepfer, H., Loeb, N., Morrison, A., & Cesana, G. (2016). Recent advances in Arctic cloud and climate research. *Current Climate Change Report*, *2*(4), 159–169. https://doi.org/10.1007/s40641-016-0051-9

Key, J., Wang, X., Liu, Y., Dworak, R. & Letterly, A. (2016). The AVHRR polar pathfinder climate data records. *Remote Sensing*. *8*, 167. https://doi.org/10.3390/rs8030167

Key, J. R., & Schweiger, A. J. (1998). Tools for atmospheric radiative transfer: Streamer and FluxNet. *Computers & Geosciences*, *24*(5), 443–451. https://doi.org/10.1016/s0098-3004(97)00130-1

Kovacs, T., & McCormick, P. (2005). *Cloud-aerosol Lidar and infrared pathfinder satellite observations (CALIPSO) quidpro quo validation plan*. Retrieved from http://calipsovalidation.hamptonu.edu/QPQ_plan062206.html

Kratz, D. P., Gupta, S. K., Wilber, A. C., & Sothcott, V. E. (2020). Validation of the CERES edition-4A surface-only flux algorithms. *Journal of Applied Meteorology and Climatology*, *59*(2), 281–295. https://doi.org/10.1175/jamc-d-19-0068.1

Lang, A., Yang, S., & Kaas, E. (2017). Sea ice thickness and recent Arctic warming. *Geophysical Research Letters*, *44*(1), 409–418. https://doi.org/10.1002/2016gl071274

L'Ecuyer, T. S., Wood, N. B., Haladay, T., Stephens, G. L., & Stackhouse, P. W. (2008). Impact of clouds on atmospheric heating based on the R04 CloudSat fluxes and heating rates data set. *Journal of Geophysical Research*, *113*. https://doi.org/10.1029/2008jd009951

Lenaerts, J. T. M., Van Tricht, K., Lhermitte, S., & L'ecuyer, T. S. (2017). Polar clouds and radiation in satellite observations, reanalyzes, and climate models. *Geophysical Research Letters*, *44*(7), 3355–3364. https://doi.org/10.1002/2016gl072242

Lesins, G., Duck, T. J., & Drummond, J. R. (2010). Climate trends at Eureka in the Canadian high arctic. *Atmosphere-Ocean*, *48*(2), 59–80. https://doi.org/10.3137/ao1103.2010

Li, Z., & Xu, K. M. (2020). Arctic clouds simulated by a multiscale modeling framework and comparisons with observations and conventional GCMs. *Journal of Geophysical Research: Atmospheres*, *125*(1). https://doi.org/10.1029/2019jd030522

Liu, Y., Ackerman, S. A., Maddux, B. C., Key, J. R., & Frey, R. A. (2010). Errors in cloud detection over the Arctic using a satellite imager and implications for observing feedback mechanisms. *Journal of Climate*, *23*(7), 1894–1907. https://doi.org/10.1175/2009jcli3386.1

Liu, Y., & Key, J. R. (2016). Assessment of Arctic cloud cover anomalies in atmospheric reanalysis products using satellite data. *Journal of Climate*, *29*(17), 6065–6083. https://doi.org/10.1175/jcli-d-15-0861.1

Liu, Y., Key, J. R., Frey, R. A., Ackerman, S. A., & Menzel, W. P. (2004). Nighttime polar cloud detection with MODIS. *Remote Sensing of Environment*, *92*(2), 181–194. https://doi.org/10.1016/j.rse.2004.06.004

Liu, Y., Key, J. R., Liu, Z., Wang, X., & Vavrus, S. J. (2012). A cloudier Arctic expected with diminishing sea ice. *Geophysical Research Letters*, *39*(5). https://doi.org/10.1029/2012gl051251

Liu, Y., Key, J. R., Vavrus, S., & Woods, C. (2018). Time evolution of the cloud response to moisture intrusions into the Arctic during winter. *Journal of Climate*, *31*(22), 9389–9405. https://doi.org/10.1175/jcli-d-17-0896.1

Liu, Y., Shupe, M. D., Wang, Z., & Mace, G. (2017). Cloud vertical distribution from combined surface and space radar-Lidar observations at two Arctic atmospheric observatories. *Atmospheric Chemistry and Physics*, *17*(9), 5973–5989. https://doi.org/10.5194/acp-17-5973-2017

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., et al. (2018). Clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 data product. *Journal of Climate*, *31*(2), 895–918. https://doi.org/10.1175/jcli-d-17-0208.1

Loeb, N. G., Kato, S., Su, W., Wong, T., Rose, F. G., Doelling, D. R., et al. (2012). Advances in understanding top-of-atmosphere radiation variability from satellite observations. *Surveys in Geophysics*, *33*, 359–385. https://doi.org/10.1007/s10712-012-9175-1

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., et al. (2009). Toward optimal closure of the earth's top-of-atmosphere radiation budget. *Journal of Climate*, *22*(3), 748–766. https://doi.org/10.1175/2008jcli2637.1

Long, C. N., & Shi, Y. (2008). An automated quality assessment and control algorithm for surface radiation measuredments. *The Open Atmospheric Science Journal*, *2*, 23–37. https://doi.org/10.2174/1874282300802010023

Long, C. N., & Turner, D. D. (2008). A method for continuous estimation of clear-sky downwelling longwave radiative flux developed using ARM surface measurements. *Journal of Geophysical Research*, *113*, D18206. https://doi.org/10.1029/2008JD009936

Mariage, V., Pelon, J., Blouzon, F., Victori, S., Geyskens, N., Amarouche, N., et al. (2017). IAOOS microlidar-on-buoy development and first atmospheric observations obtained during 2014 and 2015 arctic drifts. *Optics Express*, *25*(4), A73. https://doi.org/10.1364/oe.25.000a73

Matus, A. V., & L'ecuyer, T. S. (2017). The role of cloud phase in Earth's radiation budget. *Journal of Geophysical Research: Atmospheres*, *122*(5), 2559–2578. https://doi.org/10.1002/2016jd025951

McArthur, L. J. B. (2005). *Baseline surface radiation netword (BSRN)*. Operations manual version 2.1, WCRP-121, WMO/TD-No. 1274

McBean, G., Alekseev, G., Chen, D., Foerland, E., Fyfe, J., Groisman, P. Y., et al. (2005). Arctic climate: Past and present. In *Arctic climate impact assessment Scientific Report* (pp. 21–60). Cambridge University Press. Retrieved from http://www.acia.uaf.edu/PDFs/ACIA_Science_Chapters_Final/ACIA_Ch02_Final.pdf

Palerme, C., Claud, C., Wood, N. B., L'ecuyer, T., & Genthon, C. (2019). How does ground clutter affect CloudSat snowfall retrievals over ice sheets? *IEEE Geosciences Remote Sensing Letters*, *16*(3), 342–346. https://doi.org/10.1109/lgrs.2018.2875007

Palm, S. P., Strey, S. T., Spinhirne, J., & Markus, T. (2010). Influence of Arctic sea ice extent on polar cloud fraction and vertical structure and implications for regional climate. *Journal of Geophysical Research*, *115*(D21). https://doi.org/10.1029/2010jd013900

Provost, C., Pelon, J., Sennéchael, N., Calzas, M., Blouzon, F., et al. (2015). IAOOS (Ice-Atmosphere-Arctic Ocean Observing System, 2011–2019). Mercator Ocean Quarterly Newsletter, *Mercator Ocean, Special Issue with ICE-ARC* (pp.13–15).

Rahn, K. A. (1981). Relative importances of north America and Eurasia as sources of arctic aerosol. *Atmospheric Environment*, *15*(8), 1447–1455. https://doi.org/10.1016/0004-6981(81)90351-6

Sedlar, J., Shupe, M. D., & Tjernström, M. (2012). On the relationship between thermodynamic structure and cloud top, and its climate significance in the Arctic. *Journal of Climate*, *25*(7), 2374–2393. https://doi.org/10.1175/jcli-d-11-00186.1

Sedlar, J., Tjernström, M., Mauritsen, T., Shupe, M. D., Brooks, I. M., Persson, P. O. G., et al. (2011). A transitioning Arctic surface energy budget: The impacts of solar zenith angle, surface albedo and cloud radiative forcing. *Climate Dynamics*, *37*(7–8), 1643–1660. https://doi.org/10.1007/s00382-010-0937-5

Serreze, M. C., & Barry, R. G. (2011). Processes and impacts of Arctic amplification: A research synthesis. *Global and Planetary Change*, *77*(1–2), 85–96. https://doi.org/10.1016/j.gloplacha.2011.03.004

Serreze, M. C., & Barry, R. G. (2014). *The Arctic climate system*. Cambridge University Press. https://doi.org/10.1017/cbo9781139583817

Shupe, M. D. (2007). A ground-based multisensor cloud phase classifier. *Geophysical Research Letters*, *34*(22). https://doi.org/10.1029/2007gl031008

Shupe, M. D. (2011). Clouds at Arctic atmospheric observatories. Part II: Thermodynamic phase characteristics. *Journal of Applied Meteorology and Climatology*, *50*(3), 645–661. https://doi.org/10.1175/2010jamc2468.1

Shupe, M. D., Persson, P. O. G., Brooks, I. M., Tjernström, M., Sedlar, J., Mauritsen, T., et al. (2013). Cloud and boundary layer interactions over the Arctic sea ice in late summer. *Atmospheric Chemistry and Physics*, *13*(18), 9379–9399. https://doi.org/10.5194/acp-13-9379-2013

Shupe, M. D., Tjernstrom, M., & Persson, P. O. G. (2015). Challenge of Arctic clouds and their implications for surface radiation (in "State of the Climate in 2014"). *Bulletin of American Meteorological Society*, *96*(7), S130–S131.

Shupe, M. D., Turner, D. D., Zwink, A., Thieman, M. M., Mlawer, E. J., & Shippert, T. (2015). Deriving Arctic cloud microphysics at Barrow, Alaska: Algorithms, results, and radiative closure. *Journal of Applied Meteorology and Climatology*, *54*(7), 1675–1689. https://doi.org/10.1175/jamc-d-15-0054.1

Shupe, M. D., Walden, V. P., Eloranta, E., Uttal, T., Campbell, J. R., Starkweather, S. M., & Shiobara, M. (2011). Clouds at Arctic atmospheric observatories. Part I: Occurrence and macrophysical properties. *Journal of Applied Meteorology and Climatology*, *50*(3), 626–644. https://doi.org/10.1175/2010jamc2467.1

Stephens, G., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., et al. (2018). CloudSat and CALIPSO within the A-Train: Ten years of actively observing the earth system. *Bulletin of the American Meteorological Society*, *99*(3), 569–581. https://doi.org/10.1175/bams-d-16-0324.1

Uttal, T., Curry, J. A., Mcphee, M. G., Perovich, D. K., Moritz, R. E., Maslanik, J. A., et al. (2002). Surface heat budget of the Arctic Ocean. *Bulletin of the American Meteorological Society*, *83*(2), 255–275. https://doi.org/10.1175/1520-0477(2002)0832.3.co;2

Uttal, T., Starkweather, S., Drummond, J. R., Vihma, T., Makshtas, A. P., Darby, L. S., et al. (2016). International Arctic systems for observing the atmosphere: An international polar year legacy consortium. *Bulletin of the American Meteorological Society*, *97*(6), 1033–1056. https://doi.org/10.1175/bams-d-14-00145.1

Wang, C., Graham, R. M., Wang, K., Gerland, S., & Granskog, M. A. (2019). Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: Effects on sea ice thermodynamics and evolution. *The Cryosphere*, *13*(6), 1661–1679. https://doi.org/10.5194/tc-13-1661-2019

Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee, R. B., Louis Smith, G., & Cooper, J. E. (1996). Clouds and the earth's radiant energy system (CERES): An earth observing system experiment. *Bulletin of the Ameroican Meteorological Society*, *77*(5), 853–868. https://doi.org/10.1175/1520-0477(1996)0772.0.co;2

Zib, B. J., Dong, X., Xi, B., & Kennedy, A. (2012). Evaluation and intercomparison of cloud fraction and radiative fluxes in recent reanalyzes over the Arctic using BSRN surface observations. *Journal of Climate*, *25*(7), 2291–2305. https://doi.org/10.1175/jcli-d-11-00147.1

Zygmuntowska, M., Mauritsen, T., Quaas, J., & Kaleschke, L. (2012). Arctic clouds and surface radiation—A critical comparison of satellite retrievals and the ERA-interim reanalysis. *Atmospheric Chemistry and Physics*, *12*(14), 6667–6677. https://doi.org/10.5194/acp-12-6667-2012