

RESEARCH ARTICLE

10.1029/2017JD027377

Key Points:

- Component and process-based model evaluation is performed using surface energy budget observations at Summit, Greenland
- The surface temperature response to radiative forcing is underestimated in the evaluated models
- Cross-model comparisons reveal the importance of representing radiative and thermal properties of the snowpack in models

Correspondence to:

N. B. Miller,
millernb@colorado.edu

Citation:

Miller, N. B., Shupe, M. D., Lenaerts, J. T. M., Kay, J. E., de Boer, G., & Bennartz, R. (2018). Process-based model evaluation using surface energy budget observations in central Greenland. *Journal of Geophysical Research: Atmospheres*, 123, 4777–4796. <https://doi.org/10.1029/2017JD027377>

Received 28 JUN 2017

Accepted 4 APR 2018

Accepted article online 16 APR 2018

Published online 16 MAY 2018

Process-Based Model Evaluation Using Surface Energy Budget Observations in Central Greenland

Nathaniel B. Miller^{1,2} , **Matthew D. Shupe**^{1,2} , **Jan T. M. Lenaerts**³ , **Jennifer E. Kay**^{1,3} , **Gijs de Boer**^{1,2} , and **Ralf Bennartz**^{4,5} ¹Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, CO, USA,²NOAA Earth System Research Laboratory, Boulder, CO, USA, ³Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, CO, USA, ⁴Department of Earth and Environment Sciences, Vanderbilt University, Nashville, TN, USA, ⁵Space Science and Engineering Center, UW-Madison, Madison, WI, USA

Abstract Energy exchange at the Greenland ice sheet surface governs surface temperature variability, a factor critical for representing surface melt. Physical processes link driving forces to subsequent surface energy budget responses, including radiative, turbulent, and ground heat fluxes, and ultimately control surface temperature evolution. A reanalysis product (ERA-Interim, ERA-I), operational model (Climate Forecast System version 2, CFSv2), and climate model (Community Earth System Model, CESM) are evaluated using a comprehensive set of surface energy budget observations and process-based relationships obtained at Summit, Greenland. Simulated downwelling longwave radiation is underestimated, which is linked to a deficiency of liquid-bearing clouds. Lower than observed surface albedo, especially in ERA-I, compensates for summer deficiencies in downwelling longwave radiation. In winter, such deficiencies are compensated by an overestimation of the sensible heat flux. Process-based relationships convey that all three models underestimate the response of surface temperature to changes in radiative forcing, primarily due to an overactive ground heat flux response in ERA-I, turbulent heat fluxes in CFSv2, and sensible heat flux in CESM. Cross comparison of three distinct models indicates that the ground heat flux response for ERA-I, CFSv2, and CESM is too high, too low, and comparatively accurate, respectively, signifying the benefit of using an advanced representation of snow properties. Relatively small biases in CESM surface albedo suggest that advances in the representation of cloud microphysics result in more realistic radiative forcing. These results provide insight into model strengths and deficiencies, indicating the importance of representing physical processes when portraying cloud impacts on surface temperature variability.

1. Introduction

Accurate representation of surface energy exchange is critical in modeling the Greenland Ice Sheet (GIS) surface temperature. In the context of increased summer GIS surface temperatures in a warming climate (Hanna et al., 2008), fluctuations in the surface energy budget (SEB) are becoming more consequential to melt duration, extent, and the amount of energy available to melt snow and ice. Satellite observations have indicated two major melt events, unprecedented since the start of the observational period, and increasing GIS surface temperatures since 2000 (Hall et al., 2013). The difference between snow accumulation and surface meltwater runoff ultimately determines the GIS surface mass balance (SMB). In recent years an overall decrease in the SMB has contributed to global sea level rise (McMillan et al., 2016) driven by increased meltwater runoff (van den Broeke et al., 2016). Projections of future increase in snow accumulation are smaller than the increase in surface runoff (Fettweis et al., 2013), indicating a sustained decrease in SMB and suggesting that additional GIS mass loss will occur in the future. Increases of melt water discharge could weaken the Atlantic meridional overturning circulation (Rahmstorf et al., 2015) and have moderate to severe societal implications for coastal populations throughout the globe (Hauer et al., 2016; Intergovernmental Panel on Climate Change, 2013).

Modeling tools are useful to represent past, present, and future states of the Earth system and its components because they are spatially and temporally continuous and can represent many aspects of the climate system that cannot be observed. Here three types of modeling tools are discriminated. First, by using a static version of data assimilation techniques and model parameters, reanalysis products are a valuable tool for representing the past state of the climate system across many decades. Reanalysis products incorporate observational

data sets into the model framework to produce a spatially consistent and comprehensive portrait of most surface and atmospheric variables. The reanalyses rely on physical parameterizations and tend to have higher uncertainties in regions with less assimilation data. Second, operational weather prediction systems are used to provide accurate short-term forecasts, because they assimilate observations to get an initial atmospheric state; furthermore, the model physics are updated more frequently than a reanalysis in order to achieve the best possible forecast of the atmospheric state and surface properties. Third, global climate models (GCMs) are valuable to attribute past and present climate variability and change, and to provide future climate scenarios driven by emission scenarios. In addition, GCMs are able to estimate natural variability of the climate system because numerous realizations of the climate can be produced, thus enhancing our understanding of the significance of the observed changes. GCMs are fully coupled model frameworks between atmosphere, ocean (including sea ice), and land. In general, all three of these model perspectives have different strengths, yet in order to be effective analytical tools these models are expected to represent physical processes and their interactions.

Global climate models can be used to investigate if Greenland SMB trends are anthropogenically forced. The Community Earth System Model (CESM) suggests that the emergence of an anthropogenically derived positive SMB signal near Summit Station (72.6°N, 38.5°W; 3,211 m) is correlated with a negative SMB signal at the GIS periphery (Fyke et al., 2014). However, it is possible that the signal-to-noise ratio used to determine the emergence of the signal is influenced by an underestimation of cloud liquid over central Greenland resulting from excessive production of snow (McIlhattan et al., 2017) and an associated cold bias in the maximum summer 2-m temperatures (Kay, Bourdages et al., 2016). If representation of mixed-phase clouds were more realistic, surface temperatures would increase in response to cloud presence (Miller et al., 2015), thus decreasing GIS SMB (Van Tricht et al., 2016) and increasing the small amount of mass lost due to sublimation (Cullen et al., 2014). Hence, model deficiencies may limit our ability to represent the climate system and predict future changes.

Direct observations are an important foundation for model evaluation and development. Furthermore, reciprocity between models and observations serves to enhance our understanding of Arctic cloud-climate processes (Kay, L'Ecuyer et al., 2016). In the past, observations have been sparse in the Arctic (Serreze et al., 2000), but increased spatial coverage of satellite, ground-based observations, and reanalysis products have become more readily available (Christensen et al., 2016). Yet there are still sparse ground-based observations in the Arctic and notable uncertainties in estimates of surface temperatures from satellite (Shuman et al., 2014) and reanalysis products (Zhang et al., 2011) in central Greenland. Similarly, difficulties in reproducing diurnal cycles of near-surface temperature in stable atmospheric boundary layers, which occur frequently above the GIS, beleaguer model performance (Holtslag et al., 2013). Daily minimum temperatures affect the nighttime subsurface temperatures that precondition the snowpack in a way that influences the subsequent maximum surface temperature (Solomon et al., 2017). In addition, models may have compensating biases when simulating the various components of the SEB (Boeke & Taylor, 2016). Thus, to evaluate model surface temperature biases, it is advantageous to measure a comprehensive set of all SEB components.

Distilling key physical processes into SEB relationships derived from observations is a useful approach to assess model representation of these processes because a model inherently balances and accounts for all energy at the surface. Here a comprehensive set of measurements at Summit, Greenland is used to compare observed SEB components and processes to the European Centre for Medium Range Weather Forecast Interim Reanalysis Product, National Oceanic and Atmospheric Administration (NOAA) Climate Forecast System version 2, and the National Center for Atmospheric Research CESM. Process-based relationships (Miller et al., 2017), which relate radiative forcing terms to SEB responses, provide a unique perspective on how realistically the ice sheet/atmosphere interaction is represented. Essentially, turbulent and ground heat flux responses to changes in surface radiative forcing control the evolution of surface temperature. These observationally derived relationships are used to discern whether or not surface temperatures are realistically responding to changes in surface radiative forcing and pinpointing potential deficiencies in how modeled surface energy fluxes modulate changes in downwelling radiation. Using this process evaluation, it can be determined if an improved representation of cloud properties or other processes in a model could effectively improve the representation of surface temperature variability, which is critical for capturing extreme melt events that rarely occur at Summit Station (Nghiem et al., 2012) but occur more frequently at other parts of the GIS.

The unique process-based evaluation technique presented in this paper can reveal critical model shortcomings that will ultimately enable improvements in reanalyses, operational models, and GCMs, including the improved ability to represent climate feedbacks. The objective here is to outline the methodology and apply it to example models. Such an approach, when applied to many models with distinct physical parameterizations, can offer insight into the strengths and weaknesses of different approaches. For this specific application, the goal is to enable improved representation of atmosphere/ice sheet interactions. Our results strictly use model output to indicate areas of strengths and weaknesses when representing the surface temperature on the central GIS, beneficial for both model use and development.

2. Data Sets

A comprehensive observational data set is recorded at Summit Station to independently measure and derive each component of the SEB (Miller et al., 2017). Summit Station is located at 72.6°N and 38.5°W, with an elevation of 3,211 m resulting in dry and cold atmospheric conditions. At Summit Station there is a large range of possible surface temperatures throughout the year. Surface melt events have occurred twice in the last 130 years (Nghiem et al., 2012), a surface temperature of -68.8°C occurred on 23 March 2014 (Miller et al., 2017), and a late twentieth century climatological (1987–1999) average annual temperature of -29.4°C is reported by Shuman et al. (2001). The Sun is completely below the horizon from approximately 14 November to 27 January and at least partially above the horizon for a 24-hr period from 7 May to 6 August. The relative homogeneity of the snowpack allows for point observations to more accurately represent a larger domain by taking an average over a time window. The results of this study are specific to Summit Station, yet the physical processes and resultant relationships may be applicable to the greater dry-snow zone in central Greenland.

The SEB at Summit is a combination of radiative (Q), sensible heat (SH), latent heat (LH), and ground heat (G) fluxes, and balances according to the following equation:

$$0 = Q + SH + LH + G \quad (1)$$

All SEB components described here have the same convention, such that a positive value represents an energy flux toward the surface and negative value represents energy flux away from the surface. Any other contributions to the SEB not represented in equation (1) are exceedingly small over the central GIS and are assumed to be negligible. At other GIS locations melt is a more important factor in the SEB, but in general at Summit Station the melt term can be neglected because of the rarity of surface melt. Although Summit Station is climatically similar to the greater accumulation zone in central Greenland, the conclusions strictly apply to only the location of the observational data set.

2.1. Model Description

Accurate representation of surface temperatures and the physical processes that determine surface temperature variability are important in all types of models including reanalysis, operational models, and climate models. Thus, an example of each model type is evaluated, including ERA-Interim, Climate Forecast System version 2, and the CESM.

The European Center for Medium-Range Weather Forecasts provides 3-hourly forecasts initialized every 12 hr as part of its ERA-Interim (ERA-I) Reanalysis (Dee et al., 2011). The spatial resolution of ERA-I near Summit is approximately 80 km by 25 km, and the time period considered here spans January 2010 to May 2016. ERA-I model fields are interpolated to the Summit coordinates from the four model coordinates in closest proximity. The ERA-I elevation at Summit, calculated from the surface geopotential, is 3,172 m. Since the system is not coupled the ground heat flux is regarded as the residual of the other SEB terms, maintaining conservation of energy at the surface, which is critical to the modeling framework. Defining the ground heat flux as the residual of the other SEB terms assumes the effect of data assimilation on SEB closure is small and that there are no other notable surface energy fluxes besides those listed in equation (1). A method similar to the one described in Tjernström and Graversen (2009) is used to investigate the worst-case contribution of the data assimilation process to inaccuracies in estimated monthly mean G values and process-based analysis. Differences between forecasts 3 hr after initialization and forecasts 15 hr after initialization, 12 hr prior, indicate a maximum monthly difference of 0.8 W/m^2 and an average difference of 0.3 W/m^2 independent of season and environmental factors such as downwelling radiation.

The NOAA National Centers for Environmental Prediction created the Climate Forecast System (CFS) model for seasonal retrospective and operational forecasts. CFSv2 (Saha et al., 2014) output, spanning January 2011

to October 2016, is the operational extension of the Climate Forecast System Reanalysis (Saha et al., 2010) with updated coupled atmosphere-ocean-land model physics. The data are representative of instantaneous forecast values at 0, 6, 12, and 18 UTC. The spatial resolution is 0.2 by 0.2°, which is equivalent to a grid box of approximately 22.6 km by 6.9 km around Summit. CFSv2 model fields are interpolated to the Summit coordinates from the four model coordinates in closest proximity. The CFSv2 altitude at Summit, calculated from the surface geopotential height, is 3,184 m. Any small SEB residual is considered a part of the ground heat flux in order to maintain conservation of surface energy in the modeling framework.

The CESM is a fully coupled global climate model (Hurrell et al., 2013) maintained by the Climate and Global Dynamics Laboratory at the National Center for Atmospheric Research. This global climate model should generally reproduce the state of the current climate within its natural variability but is not expected to simulate the timing of observed surface and atmospheric states. In fact, a process-based evaluation framework is well suited for a free running model in that this perspective can provide some degree of insight without having to run the model for multiple years or ensemble members. The 5-year CESM data set shown here is a development run, composed of model improvements including a more accurate representation of liquid-bearing clouds by implementation of advanced microphysical schemes (Gettelman & Morrison, 2015). Improving the representation of clouds has been a major thrust for the development of the embedded Community Atmosphere Model 6, because in the publicly available version of CESM the lack of liquid-bearing clouds over Greenland was a major deficiency in the Community Atmosphere Model 5 (Kay, Bourdages et al., 2016), likely due to an overactive Wegener-Bergeron-Findeisen process (McIlhatten et al., 2017). In addition, CESM has an advanced representation of snow properties, compared to ERA-I and CFSv2, including the simulation of solar radiation penetration. The CESM firn layer in central Greenland has 12 vertical levels and has time-varying snow density and albedo.

The CESM data, output at 3-hourly intervals, are produced using fully coupled simulations, with interactive atmosphere-land-ocean components and greenhouse gas forcing of 393 ppmv for the period 2011–2014. A spatial resolution of 0.94° (latitude) by 1.25° (longitude) translates to a grid box size around Summit of 104 km × 43 km. The CESM model fields are reported using the grid box that includes Summit (73.04°N and 38.75°W). The CESM altitude at Summit, calculated from the surface geopotential, is 3,079 m.

2.2. Measurements

The radiative flux is separated into upwelling (\uparrow) and downwelling (\downarrow) components of the longwave (LW) and shortwave (SW) radiation, thusly

$$Q = LW\downarrow - LW\uparrow + SW\downarrow - SW\uparrow. \quad (2)$$

The upwelling and downwelling radiation components, measured by Kipp and Zonen CG4 pyrgeometers (LW) and Kipp and Zonen CM22 pyranometers (SW), are provided by the Swiss Federal Institute of Technology. The SH and LH fluxes are primarily calculated via the bulk aerodynamic method (Persson et al., 2002) and the two-level profile method (Steffen & DeMaria, 1996), respectively. Potential sources of uncertainty for the SH flux estimates include the stability correction functions, the LW-derived surface temperature values, and the assumed roughness length. Uncertainty in the LH flux estimates stem from uncertainty in the mixing ratio values and stability functions. The ground heat flux is estimated from subsurface temperature probes for a 1-yr period only, providing a complete annual cycle of estimates of all the SEB terms from July 2013 to June 2014. Measurements of the radiative components, SH, and LH predate the beginning of the ground heat flux measurements (1 July 2013), with start dates of January 2011, June 2011, and March 2012, respectively.

The SEB calculations and meteorological measurements (30-min temporal resolution) used as input are detailed in Miller et al. (2017) and can be accessed online from the Arctic Data Center (Shupe & Miller, 2016). Redundancy of direct measurements is crucial for deriving a closed SEB in a harsh Arctic environment where extremely cold temperatures create operational challenges. The National Oceanic and Atmospheric Administration's (NOAA) Global Monitoring Division provides quality controlled and daily serviced meteorological temperature data and radiative flux measurements. Cross comparisons between measurements are used to make LW radiative flux adjustments (Miller et al., 2015). Redundant meteorological data from the Closing the Isotope Balance at Summit experiment are used to fill in data gaps and provide comparable measurements and energy flux estimates. The 2-m temperature bias between the Closing the Isotope Balance at Summit and NOAA/Global Monitoring Division values is 0.72 K, attributable to variability in sensor location and deviation from the 2-m height due to snow accumulation.

Ground-based observations at Summit Station are used to assess model simulations over model grid box scales; hence, the observations, when compared to models with coarse spatial resolution (ERA-I and CESM), are smoothed over a 3-hr running window to remove local variability in a given parameter. The 3-hr window is the approximate amount of time it would take a typical parcel of air to advect across a model grid box of approximately 80 km. The relative homogeneity of the central GIS, compared to complexities associated with terrestrial sites (Foken, 2008), limits the sampling error when comparing point measurements to larger domains. A 3-hr running window is applied to G in all models, reducing the variability due to large measurement uncertainty of the ground heat flux.

In order to compare to models with 3-hr running averages (ERA-I and CESM), the 30-min observational components are averaged over the model integration period, requiring the observed values to be available for at least two thirds of the averaging period. In order to compare observations to CFSv2, which has higher spatial resolution and provides an instantaneous value every 6 hr, the observations are averaged within ± 30 min of the forecast time. As a result of different averaging approaches, the observational data vary modestly for comparison with each model, despite being derived from the same observational data set. Measurement uncertainty, as reported in Miller et al. (2017), is reduced when considering monthly mean SEB values, because the reduction is proportional to the square root of the sample size. The typical monthly uncertainties of measurements sampled every 3 hr are approximately 0.2 W/m^2 for most SEB components, except G , which is 0.7 W/m^2 . The monthly mean uncertainties of measurements sampled every 6 hr are approximately 2.5 times greater than the 3-hourly uncertainties.

Miller et al. (2017) found that the SEB at Summit is radiatively driven, such that the statistics over an annual cycle indicate that changes in the radiative forcing terms ($\text{LW}\downarrow + \text{net SW}$) are completely compensated by the response terms ($\text{SH} + \text{LH} + G - \text{LW}\uparrow$) in equations (1) and (2). The responses of individual SEB terms to radiative forcing are calculated by performing a linear fit using the sum of the forcing terms as the dependent variable and taking into account measurement uncertainty of all components (see Miller et al., 2017 for details). The resultant slope represents the response of a given term to changes in radiative forcing over a specified time period. Monthly relationships among parameters are examined here. The main drivers of variability in the forcing terms are clouds, especially liquid-bearing clouds, and/or insolation. Cloud presence has a positive radiative forcing effect throughout the entire year, due to high surface albedo year round, which limits the cloud SW cooling effect (Miller et al., 2015). Variability in surface albedo can also influence the forcing terms but at Summit Station the relatively constant surface albedo is less of a factor than variations in downwelling radiation. $\text{LW}\uparrow$ is considered to be a proxy for a surface temperature, because it is proportional to temperature to the fourth power. Thus, a positive response of $\text{LW}\uparrow$ to radiative forcing indicates warming of the surface. The responses of SH , LH and G modulate the degree to which $\text{LW}\uparrow$ responds to changes in radiative forcing.

Surface temperatures are estimated from the $\text{LW}\uparrow$ measurements using a greybody approximation (Persson et al., 2002), with an assumed surface emissivity of 0.985 (Warren, 1982). Surface albedo is calculated as the daily total $\text{SW}\uparrow$ divided by the daily total $\text{SW}\downarrow$, which reduces errors due to subdaily changes in snow metamorphism, possible tilting of sensors, and poor cosine response (van den Broeke et al., 2004). Liquid water path (LWP) is derived from a pair of microwave radiometers (MWR), similar to Turner et al. (2007), operated as part of the Integrated Characterization of Energy, Clouds, Atmospheric state, and Precipitation at Summit project (Shupe et al., 2013). During May–June 2014 one MWR was off-site for repairs; hence, there are no statistics available for LWP during these months.

3. Model Evaluation

To provide context for SEB processes, model surface temperatures are first evaluated before comparisons of SEB components and relationships are investigated to explore why the temperature biases exist in each of the model products. Modeled skin temperatures, hereafter considered to be the surface temperature, are compared to the observationally derived surface temperatures. ERA-I surface temperatures are higher than observed with a difference in the annual mean value of 3.9°C . Monthly statistics of surface temperature values (Figure 1) indicate that the ERA-I monthly averages are greater than the observed values due to large discrepancies in minimum temperatures. In contrast to ERA-I, both CFSv2 and CESM have individual months that have lower mean surface temperatures than observed. The annual mean CFSv2 surface temperature difference is 3.3°C higher than observed, although in summer (June–August) CFSv2 is 1.8°C colder than observed.

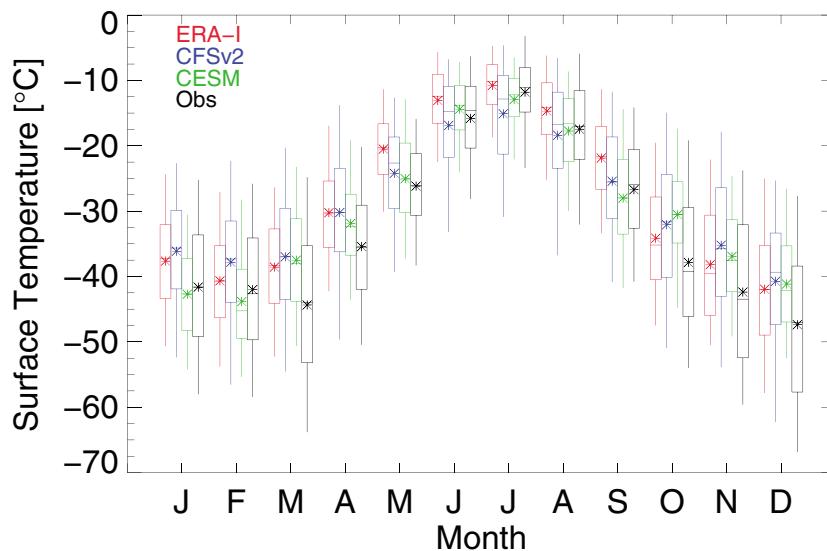


Figure 1. Monthly surface temperature distributions are represented by box-and-whisker plots (the box indicates the 25th and 75th percentiles, the whiskers indicate 5th and 95th percentiles, the middle line is the median and the * indicates the mean). ERA-I (red) distributions are from January 2010 to May 2016. CFSv2 (blue) distributions are from January 2011 to October 2016. CESM (green) distributions are from 5 years of climate model output. Observed (black) distributions are from January 2011 to June 2014. CFSv2 = Climate Forecast System version 2; CESM = Community Earth System Model; ERA-I = ERA-Interim.

The CESM summer temperature difference is effectively zero, and the annual mean surface temperature is 2.2°C higher than observed.

Differences between modeled and observed surface temperatures are not attributable to any one factor alone. Elevation differences in the models likely contribute to inaccuracies of surface and 2-m temperatures. Using a surface slope lapse rate of 7.1°C/km (Steffen & Box, 2001), the underestimation of elevation by ERA-I (-39 m), CFSv2 (-27 m), and CESM (-132 m) translates to an estimated overestimation of surface temperature by 0.28°C , 0.19°C , and 0.94°C , respectively. In many cases these biases are substantially smaller than the overall surface temperature overestimates. Surface temperature discrepancies vary between seasons and models due to model representation of the amount of radiation transferred to the surface and the associated responses of the turbulent and ground heat fluxes. In this section, all SEB components and associated process-based relationships are investigated for each model individually starting with the ERA-I.

3.1. ERA-I

To investigate the warm bias of ERA-I, comparisons of the individual radiative components are performed to characterize differences in the radiative balance at the surface. A subsampled data set contains matching sampling times from the year July 2013 to June 2014 where there exist quality controlled observations of each parameter. Sampling over longer time periods for specific parameters (ERA-I: January 2010 to May 2016 and Obs: January 2011 to June 2014) provides context on how deviations of the year-long data set result from interannual variability. In most cases, the data set that includes the longer time periods yields similar qualitative results to the data set with the matching sampling times for the 1-year period, thus the figures given here compare monthly statistics between the longer time periods.

Generally, $\text{LW}\downarrow$ is larger in the observations than in ERA-I (Figure 2a), with exceptions in April, May, and September, resulting in an annual mean ERA-I $\text{LW}\downarrow$ bias of -7.8 W/m^2 in the 1-year data and -4.6 W/m^2 for the extended data sets. Despite the underestimation of $\text{LW}\downarrow$ in ERA-I, the $\text{LW}\uparrow$ comparisons (Figure 2b) are consistent with the warm biases in Figure 1, indicating that differences in surface emissivity are not the primary reason for the surface temperature differences. The ERA-I annual mean $\text{SW}\downarrow$ deficiency is -7.6 W/m^2 in the 1-year data set and -7.8 W/m^2 in the extended data set (Figure 2c). Similar to the monthly $\text{SW}\downarrow$ values being too low in ERA-I, the $\text{SW}\uparrow$ values are lower in the model compared to observations, although the $\text{SW}\uparrow$ differences are greater (annual biases of 1 year: -17.3 W/m^2 ; extended data: -17.9 W/m^2). Because clouds generally serve to increase the $\text{LW}\downarrow$ and reduce the $\text{SW}\downarrow$, the similarities in underestimation of $\text{LW}\downarrow$ and $\text{SW}\downarrow$

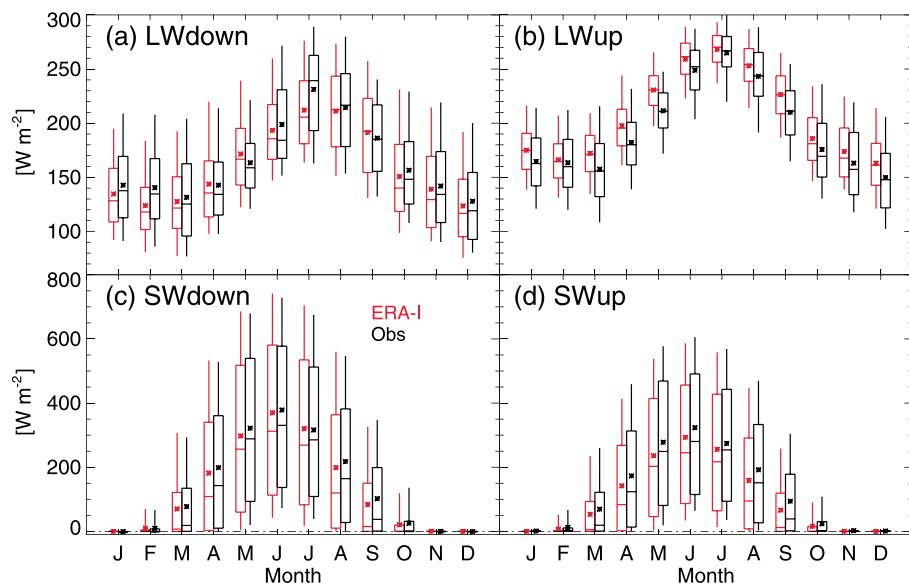


Figure 2. Monthly distributions of (a) $LW\downarrow$, (b) $LW\uparrow$, (c) $SW\downarrow$, and (d) $SW\uparrow$. ERA-I distributions are shown in red, and the observed (black) distributions are from January 2011 to June 2014, when available. SW = shortwave; LW = longwave; ERA-I = ERA-Interim.

suggest that the ERA-I clouds have optical properties with large solar back scattering and insufficient cloud LW optical thickness.

The turbulent and ground heat flux comparisons between ERA-I and the observations (Figure 3) indicate other possible contributions to surface temperature biases. The surface is warmed (cooled) by turbulent mixing of overlying warm (cold) air, deposition (sublimation), or conduction of subsurface heat toward (away from) the surface. Compared to the observations, ERA-I SH flux is biased high from September to March and has greater variability in all months, as indicated by the monthly interquartile range (Figure 3a). The ERA-I LH flux

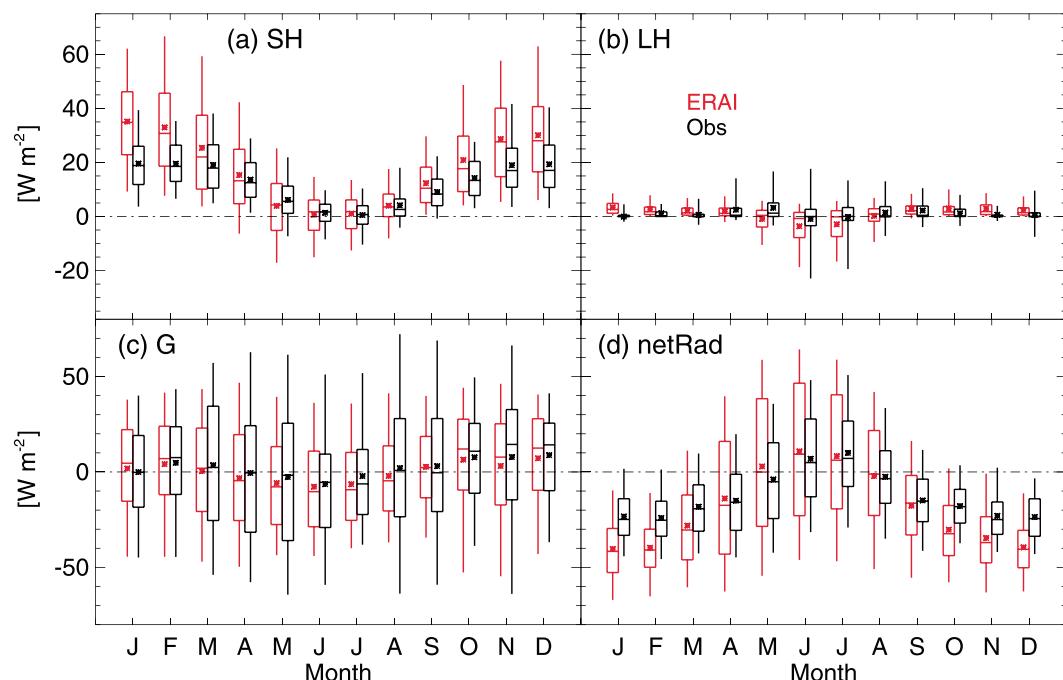


Figure 3. Monthly distributions of (a) SH, (b) LH, (c) G, and (d) net radiation. ERA-I distributions are shown in red, and the observed (black) distributions are from January 2011 to June 2014, when available. SH = sensible heat; LH = latent heat; ERA-I = ERA-Interim; G = ground heat.

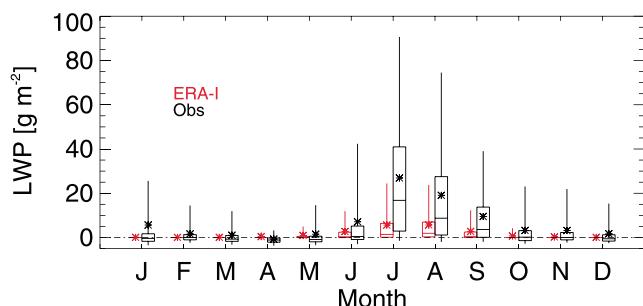


Figure 4. Monthly distributions of liquid water path. ERA-I distributions are shown in red, and the observed (black) distributions are from January 2011 to June 2014, when available. LWP = liquid water path; ERA-I = ERA-Interim.

At Summit Station, the presence of liquid-bearing clouds results in additional radiation reaching the surface, compared to clear-sky scenes (Miller et al., 2015). The amount of liquid water contained in overlying clouds, or the LWP, modulates the amount of downwelling radiation impinging on the surface, decreasing SW_{\downarrow} and increasing LW_{\downarrow} . Additional cloud properties that modulate the surface radiation are ice water path, cloud temperature, and cloud optical thickness in both the SW and LW spectrum. The surface temperature adjusts accordingly, resulting in new radiative balance at the surface, altering the temperature gradients between the overlying atmosphere, surface, and subsurface. Monthly ERA-I and MWR-based estimates (Figure 4) show a substantial underestimation of LWP by ERA-I throughout the annual cycle. This deficiency of LWP is consistent with the deficiency of LW_{\downarrow} in ERA-I in most months. The deficiency of SW_{\downarrow} could be explained by an overestimation of ice clouds in ERA-I, which is explored further in section 4.

Miller et al. (2017) report that variability in the SEB is primarily driven by the radiative forcing terms, which are a combination of LW_{\downarrow} plus net SW. At Summit the Sun is below the horizon from mid-November to early February. Figures 5a and 2a show that from November to February the ERA-I deficiencies in the LW_{\downarrow} component correspond to deficiencies in the LWP. The occurrence of liquid-bearing clouds in January 2014 was more frequent than during January 2011–2013 and, thus, the ERA-I January radiative forcing bias was -25.8 W/m^2 for the 1-year data compared to -4.5 W/m^2 for the extended data. Despite the observations indicating the monthly LWP being greater than ERA-I values, the ERA-I forcing terms are greater than the observations and have more variability from April to September (Figure 5a). The summer (June–August) ERA-I forcing bias is 13.5 W/m^2 and 8.4 W/m^2 for the 1-year and all data sets, respectively. Excess SW radiation absorbed by the surface in ERA-I, due to the surface albedo being too low (Figure 5b), results in an overestimation of the forcing terms during periods of large insolation. From March to September the average albedo in ERA-I is 0.79 compared to 0.87 in the observations. The observed values are similar to May–July broadband albedo values at Summit reported by Wright et al. (2014), which varied between approximately 0.82–0.90.

The underestimation of liquid-bearing clouds and surface albedo has compensating effects on surface temperature biases. The response terms, which comprise SH, LH, G, and LW_{\uparrow} components of the SEB, are dependent on the magnitude of the forcing terms. Hence, to investigate whether biases in the response terms are attributable to discrepancies in radiative forcing, a process-based perspective is utilized to indicate

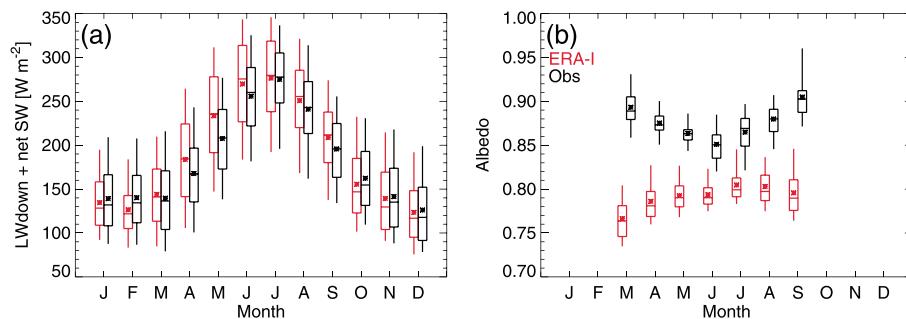


Figure 5. Monthly distributions of (a) LW_{\downarrow} + net SW and (b) surface albedo. ERA-I distributions are shown in red, and the observed (black) distributions are from January 2011 to June 2014, when available. SW = shortwave; LW = longwave; ERA-I = ERA-Interim.

is lower than observed from April to August and higher throughout the rest of the year (Figure 3b). Stronger net radiative cooling by ERA-I in the winter (Figure 3d) corresponds to an overestimation of the SH flux, in contrast to June–August when small biases in radiative forcing values in ERA-I result in a more accurate estimation of the SH flux. The annual ERA-I bias of G for the 1-year data set is -1.1 W/m^2 (Figure 3c). The high variability in the observed G values is in part due to the large uncertainty in estimating G . July, August, and November have the largest G biases, when the monthly mean ERA-I G values are lower than observed. January is the only month when ERA-I mean values of G are higher than observed. The biases in the turbulent and ground heat flux components could be due to difficulties in boundary layer parameterizations, subsurface processes, or the modeled radiative forcing occurring within a limited range.

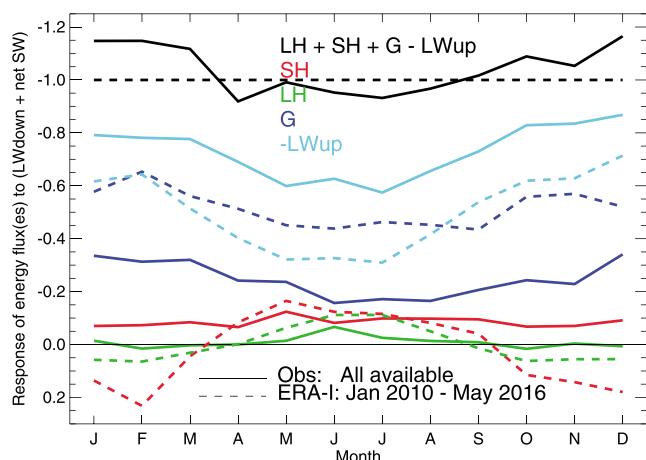


Figure 6. Annual cycle of monthly linear regression of responses to the forcing terms. ERA-I data (dashed) is from January 2010 to May 2016, and the observational data (solid) are representative of all available data for the given subset. SW = shortwave; LW = longwave; LH = latent heat; SH = sensible heat; G = ground heat; ERA-I = ERA-Interim.

mating the temperature gradient between the surface and the overlying air, and consequently overestimating the SH flux. The LH response is also positive from October to March, and the April–August LH response is similar to, although somewhat larger than, the observed response. The inferred ERA-I ground heat flux response is overestimated throughout the annual cycle, producing a response 33% greater than the observed response. The ground heat flux response serves to diminish the LW↑ response, which is 23% lower than what is observed by the broadband measurements. Generally, the bias in ERA-I surface albedo produces more radiative forcing but is compensated for by an overactive response of G, resulting in moderately higher surface temperatures.

3.2. CFSv2 Evaluation

Similar to ERA-I, CFSv2 LW↓ is underestimated and LW↑ is overestimated for most of the annual cycle (Figures 7a and 7b). Throughout this section the biases of the extended data period will be indicated in parenthesis to provide a multiyear perspective, if available. The LW↑ annual bias for CFSv2 is 8.6(9.9) W/m² for the 1-year (extended) data set, indicating that annual mean surface temperatures are too high. However,

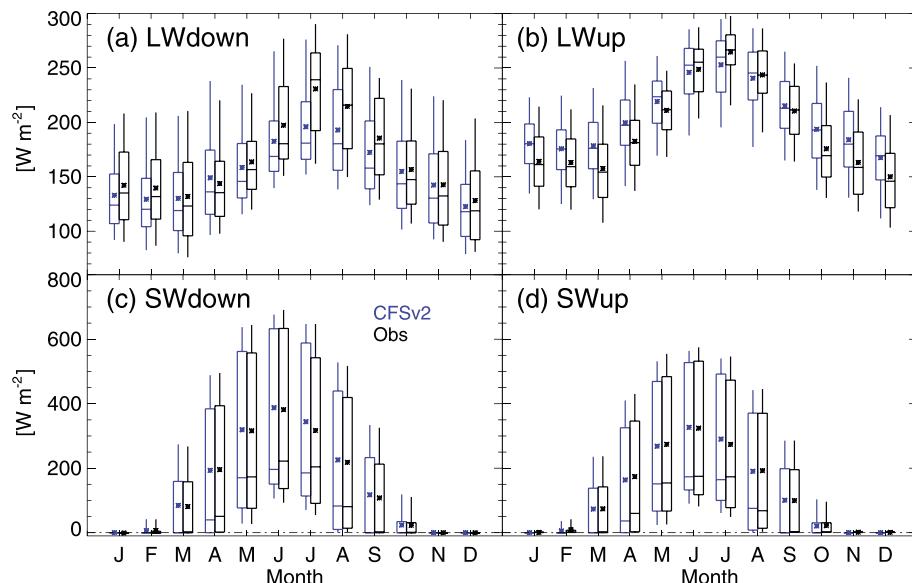


Figure 7. Monthly distributions of (a) LW↓, (b) LW↑, (c) SW↓, and (d) SW↑. Climate Forecast System version 2 distributions are shown in blue and the observed (black) distributions are from January 2011 to June 2014, when available. SW = shortwave; LW = longwave.

if changes in forcing elicit a realistic response. By defining G as the residual of the other SEB terms, the sum of the ERA-I response terms completely accounts (i.e., -1.0 slope) for any changes to the forcing terms in all months (Figure 6). The observations do not show an exact accounting (i.e., -1.0 slope) of the total response to changes in the observations (up to 15% differences in some months), indicating possible seasonal errors in the measurement of some of the SEB component responses.

Comparisons between ERA-I and observationally derived responses illuminate differences and similarities regarding the modulation of surface temperature responses by the sensible, latent, and ground heat fluxes. From April to September the ERA-I SH flux response is within 5% of the observed values (Figure 6), yet from October to March the modeled SH response discrepancy increases and changes sign, indicating an increase in SH for an increase in radiative forcing at the surface. Observationally, the winter SH flux is diminished for relatively small and large radiative forcing periods. ERA-I does have diminished SH flux for small radiative forcing but does not reproduce this behavior for large radiative forcing conditions. This discrepancy could be due to the model surface temperatures not warming enough during periods of large radiative forcing, thus overestimating the temperature gradient between the surface and the overlying air, and consequently overestimating the SH flux.

The LH response is also positive from October to March, and the April–August LH response is similar to, although somewhat larger than, the observed response. The inferred ERA-I ground heat flux response is overestimated throughout the annual cycle, producing a response 33% greater than the observed response.

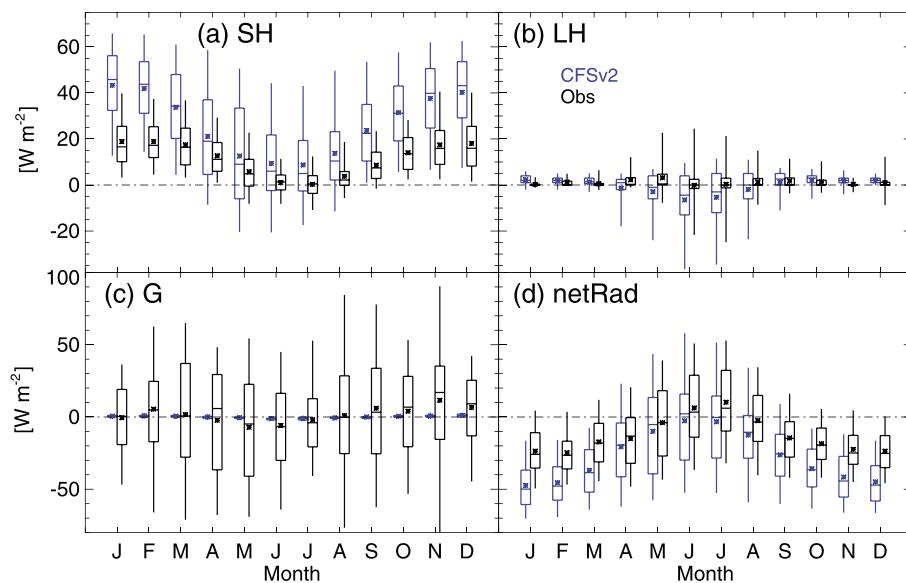


Figure 8. Monthly distributions of (a) SH, (b) LH, (c) G, and (d) net radiation. Climate Forecast System version 2 distributions are shown in blue, and the observed (black) distributions are from January 2011 to June 2014, when available. LH = latent heat; SH = sensible heat; G = ground heat.

during summer (June–August) the $LW\uparrow$ is lower than observed, with a bias of $-2.4(-5.8) \text{ W/m}^2$, indicating that the surface is too cold. $SW\downarrow$ has similar variability (interquartile range) and range (difference between the 5th and 95th percentiles) compared to observed values (Figure 7c), with a difference in the annual mean values of $3.3(4.8) \text{ W/m}^2$. Annually, CFSv2 underestimates $LW\downarrow$ by $-10.2(-9.4) \text{ W/m}^2$. The negative bias in $LW\downarrow$ and positive bias in $SW\downarrow$ suggests that CFSv2 under-represents liquid-bearing clouds. The annual difference between the observed and modeled $SW\uparrow$ is less than 1 W/m^2 (Figure 7d). The monthly mean net radiative fluxes in CFSv2 are less than the observed radiative balance in all months (Figure 8d) due to the combined deficiency in $LW\downarrow$ and excess of thermal emission.

An annual SH positive bias of $17.2(15.0) \text{ W/m}^2$, which manifests in all months, is a major contributor to enhanced warming of the surface in CFSv2 (Figure 8a). LH flux biases are smaller than those of the SH flux and lead to enhanced warming of the surface in the winter and more cooling during the summer (Figure 8b). The CFSv2 ground heat flux values and variability are extremely small, indicating heat flux into the snowpack is negligible.

The CFSv2 radiative forcing terms have an annual mean bias of $-6.0(-5.4) \text{ W/m}^2$ (Figure 9a). In contrast to ERA-I, there is a deficit of summer CFSv2 forcing despite similarities in enhanced SW fluxes compared to observations. The CFSv2 albedo is close to the observed albedo value (Figure 9b), with a bias of $-0.024(-0.025)$, indicating a slight overestimation of the proportion of SW radiation absorbed by the surface.

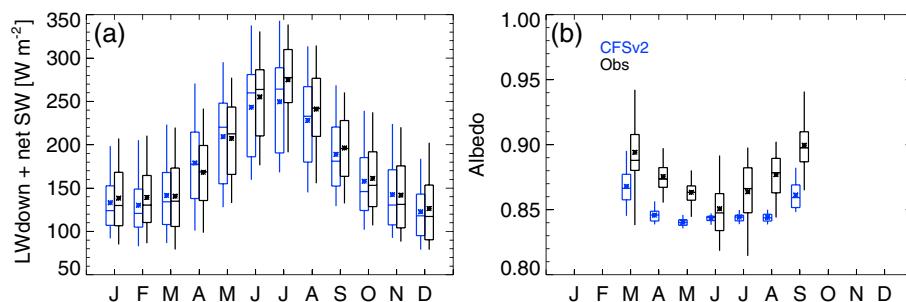


Figure 9. Monthly distributions of (a) $LW\downarrow + \text{net SW}$ and (b) surface albedo. CFSv2 distributions are shown in blue, and the observed (black) distributions are from January 2011 to June 2014, when available. SW = shortwave; LW = longwave; CFSv2 = Climate Forecast System version 2.

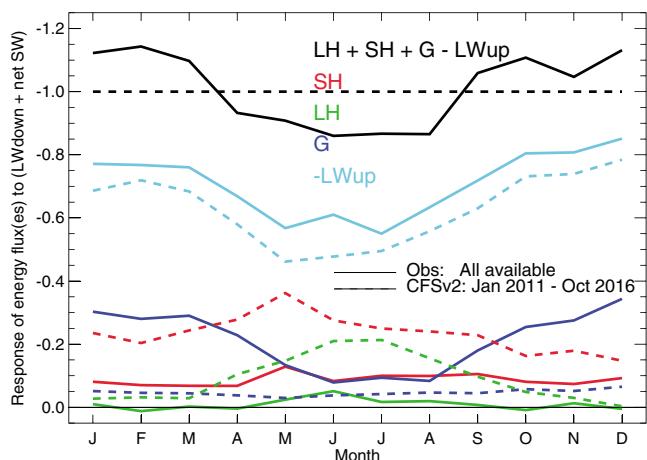


Figure 10. Annual cycle of monthly linear regression of responses to the forcing terms. CFSv2 data (dashed) is from January 2011 to October 2016, and the observational data (solid) are representative of all available data for the given subset. CFSv2 = Climate Forecast System version 2; SW = shortwave; LW = longwave; LH = latent heat; SH = sensible heat; G = ground heat.

but the results must be interpreted with caution, as the time period may not cover the full range of internal variability. A fully spun-up, transient twentieth century CESM2 simulation will be required to fully explore the performance of CESM2, output of which is currently not yet available. We aim to use this method for a detailed evaluation of the final CESM2 once such simulations are available. The processes-based evaluation approach allows for examining a shorter time period needed to discern model issues, making the technique useful for model evaluation during the development stage.

The CESM annual $LW\uparrow$ bias of 4.7 W/m^2 indicates that surface temperatures in CESM are generally higher than observed, although the model is actually slightly too cold in five winter and late summer months (Figure 11b). The largest $LW\uparrow$ deficit occurs in February (-6.8 W/m^2) and the largest surplus in October (20.8 W/m^2). The annual CESM $LW\downarrow$ bias is -10.9 W/m^2 , with only October producing slightly more $LW\downarrow$ than observed

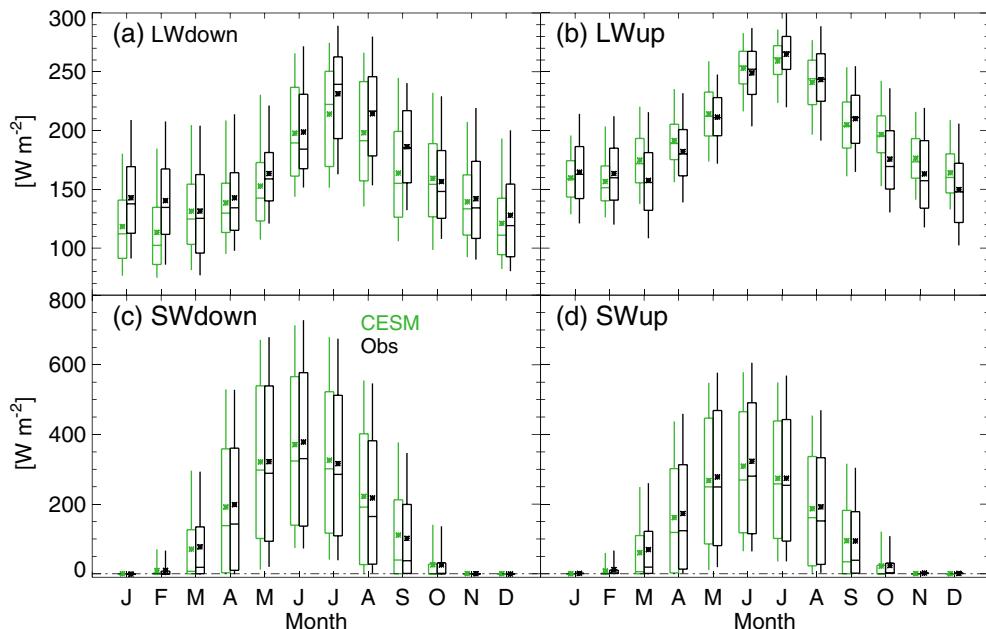


Figure 11. Monthly distributions of (a) $LW\downarrow$, (b) $LW\uparrow$, (c) $SW\downarrow$, and (d) $SW\uparrow$. Community Earth System Model distributions are shown in green, and the observed (black) distributions are from January 2011 to June 2014, when available. SW = shortwave; LW = longwave.

The CFSv2 responses of the sensible, latent, and ground heat fluxes all compensate for increases or decreases in the radiative forcing terms throughout the annual cycle (Figure 10). The LH and SH responses are greater than observed responses, averaging 8% and 15% higher responses for all available data, respectively. The monthly CFSv2 ground heat flux response ranges between 2% and 7%, much less than the observed range of 8–34%. The weak response of the ground heat flux is more than offset by the stronger than observed responses of the turbulent heat fluxes. The result is that the difference of annual mean CFSv2 and observed $LW\uparrow$ response is 8%, less than the difference of the equivalent ERA-I response (23%). Generally, the CFSv2 surface temperatures are too high despite deficits in the radiative forcing terms because the overactive turbulent heat flux response limits cooling of the surface.

3.3. CESM Evaluation

The 5-year CESM data set is a fully coupled simulation with interactive atmosphere-land-ocean components. This free-running model is not expected to reproduce actual synoptic conditions for a given time period. Hence, this subsection will only compare monthly statistics for the CESM results and the maximum data available for the corresponding observational data set. The data set is representative of current climate conditions,

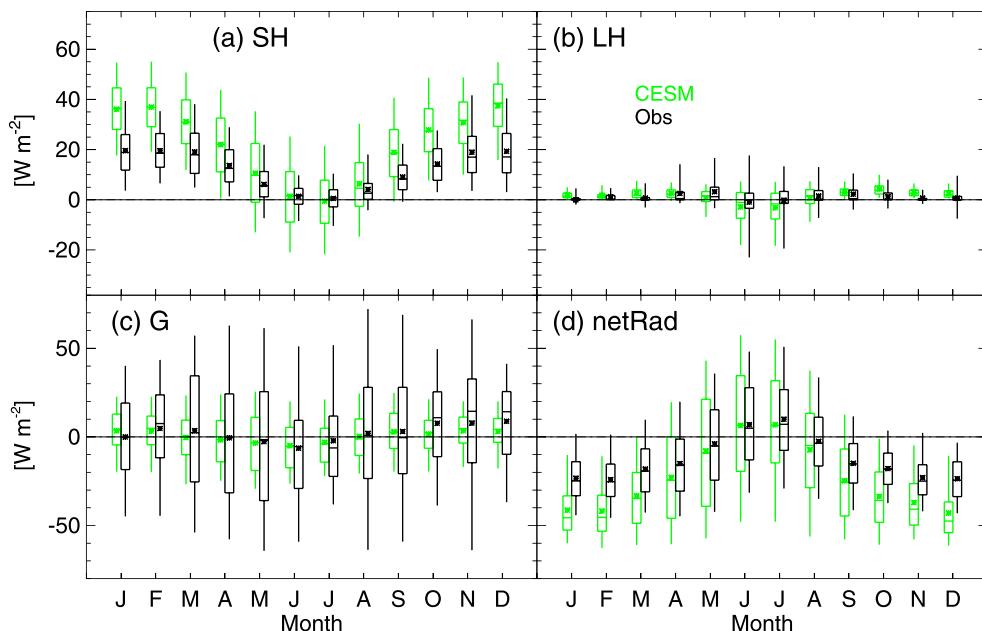


Figure 12. Monthly distributions of (a) SH, (b) LH, (c) G, and (d) net radiation. CESM distributions are shown in green and the observed (black) distributions are from January 2011 to June 2014, when available. LH = latent heat; SH = sensible heat; CESM = Community Earth System Model; G = ground heat.

(Figure 11a). It is noteworthy that while the monthly range of $LW\downarrow$ data is similar, CESM has narrower monthly ranges of $LW\uparrow$, indicating that CESM has lower surface temperature variability due to the enhanced modulation of surface radiative forcing by the turbulent and/or ground heat fluxes. The annual mean biases of the $SW\downarrow$ and $SW\uparrow$ components are 0.3 and -5.0 W/m^2 , respectively (Figures 11c and 11d). The CESM net radiative flux is biased low (-10.8 W/m^2) with smaller underestimates from May to August (Figure 12d), indicating enhanced near-surface stability in CESM from September to April compared to observations. The modeled monthly variability of net radiation is larger than observed in summer.

The CESM SH warms the surface more than is reported in the observations for all months except June and July, contributing to an annual bias of 9.4 W/m^2 (Figure 12a). In addition, in summer the monthly variability of the modeled SH is greater than the observed variability. The deficit in the CESM LH values from May to August is offset by the surplus throughout the rest of the year resulting in a mean annual bias of 0.4 W/m^2 (Figure 12b). The annual CESM G bias is -1.8 W/m^2 (Figure 12c), and monthly variability is substantially smaller than observed, although the qualitative annual cycle is captured quite well.

Instances of high LWP occur in CESM (Figure 13), indicating a stark improvement in the ability to model liquid-bearing clouds throughout the annual cycle, compared to the first version of CESM (Kay, Bourdages, et al., 2016). The 95th percentiles of the monthly CESM distributions (upper whiskers, Figure 13) are in many months greater than the observed values, while median values are typically smaller. This skewness indicates that the model may underestimate the frequency of occurrence of liquid-bearing clouds in a given month while those that do occur sometimes have a greater optical thickness. LWP biases correspond to biases in the down-welling radiation. For example, the overestimation of LWP in June by CESM relates to a surplus of $LW\downarrow$ and a deficit of $SW\downarrow$.

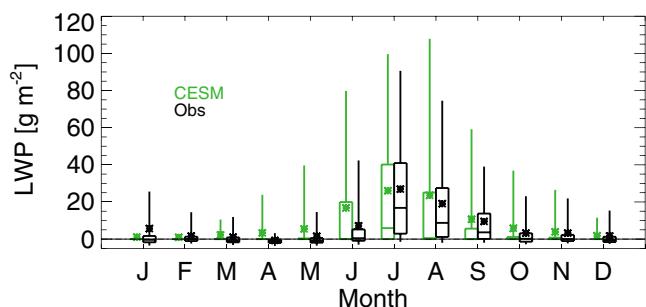


Figure 13. Monthly distributions of liquid water path. CESM distributions are shown in green, and the observed (black) distributions are from January 2011 to June 2014, when available. CESM = Community Earth System Model; LWP = liquid water path.

As a result of a more realistic representation of liquid-bearing clouds, summertime radiative forcing terms are generally more realistic than those from ERA-1 and CFSv2, while winter values are substantially underestimated (Figure 14a). Annually, the CESM forcing bias is -6.7 W/m^2 , from June to August (summer) the bias is -4.3 W/m^2 and from December to February (winter) it is -17.1 W/m^2 . On average, the albedo is 0.03 lower

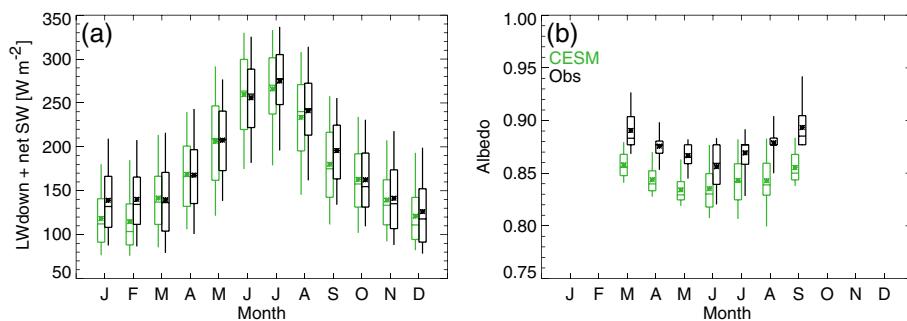


Figure 14. Monthly distributions of (a) $LW\downarrow + \text{net SW}$ and (b) surface albedo. CESM distributions are shown in green and the observed (black) distributions are from January 2011 to June 2014, when available. CESM = Community Earth System Model; SW = shortwave; LW = longwave.

in CESM compared to observations, indicating about a 3–4% surplus of the amount of SW radiation absorbed by the surface. Thus, the summer deficit in radiative forcing is largely due to the $LW\downarrow$ term, offset somewhat by the extra net SW radiation.

CESM differences in individual SEB components are, in part, due to interannual variability of the atmospheric properties, but consideration of the responses to radiative forcing isolates how well CESM captures the processes-based relationships, which are independent of the year-to-year variability of the forcing terms. The ground heat flux response in CESM (Figure 15) has an annual bias of 1%, although the stronger summer response (difference of −9%) offsets the weaker winter response (difference of 8%). Over the annual cycle the CESM SH and LH responses are greater than the observed responses by 12% and 1%, respectively. The CESM SH responses are greater than the observed responses in every month, in contrast to the biases of the monthly LH responses, which are slightly negative from September to March. The CESM $LW\uparrow$ response is 21% lower than the observed response, indicating that surface temperature variably, in response to changes in radiative forcing, is underestimated in CESM. The primary factor responsible for this overly weak $LW\uparrow$ response is the over active response of the SH flux. Generally, CESM best represents the radiative forcing, yet the response analysis demonstrates an incorrect partitioning of that energy at the surface.

4. Discussion and Summary

Observationally based estimates of all SEB components provide a comprehensive perspective to evaluate surface temperature biases in ERA-I, CFSv2, and CESM. All the energy flux at the surface must be accounted for; thus, misrepresentation of any SEB component will lead to misrepresentation in another and/or a warm

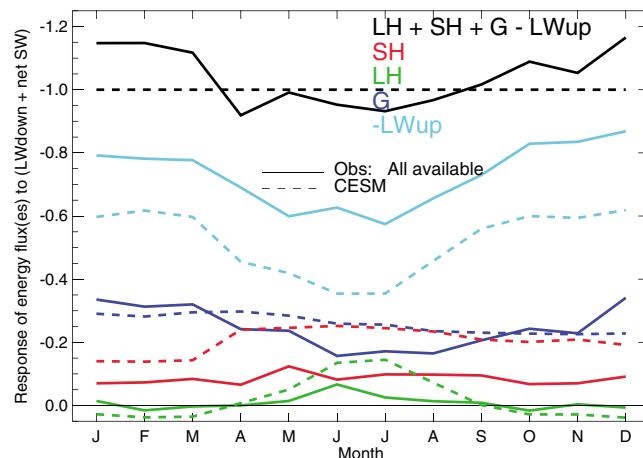


Figure 15. Annual cycle of monthly linear regression of responses to the forcing terms. CESM data (dashed) is from a 5-year data set, and the observational data (solid) are representative of all available data for the given subset. CESM = Community Earth System Model; SW = shortwave; LW = longwave; LH = latent heat; SH = sensible heat; G = ground heat.

Table 1
Seasonally Averaged Monthly Mean Biases for 1-Year Period Spanning July 2013 to June 2014 and (Extended Data Periods)

	ERA-I		CFSv2		CESM	
	Summer	Winter	Summer	Winter	Summer	Winter
T_{2m}	1.9(1.9)	1.7(3.5)	-0.5(-1.2)	4.0(4.4)	(-0.4)	(0.7)
$LW\downarrow$	-5.9(-9.2)	-16.8(-9.7)	-16.2(-23.8)	-12.7(-8.3)	(-11.5)	(-19.5)
$LW\uparrow$	7.7(7.7)	1.5(8.8)	-2.4(-5.8)	14.3(15.5)	(-1.4)	(0.8)
$SW\downarrow$	-8.8(-8.0)	N/A	10.7(13.5)	N/A	(2.3)	N/A
$SW\uparrow$	-28.9(-27.1)	N/A	1.4(5.8)	N/A	(-6.5)	N/A
SH	-0.2(-0.1)	16.1(13.2)	9.2(8.9)	28.9(23.2)	(0.4)	(17.3)
LH	-1.4(-2.2)	2.1(2.2)	-5.4(-5.1)	1.5(1.2)	(-1.9)	(1.5)
G	-3.0(-3.2)	1.4(-0.2)	1.2(1.3)	-3.1(-3.0)	(-0.4)	(-1.3)
Forcing	13.5(8.4)	-15.6(-6.9)	-7.3(-16.8)	-11.3(-6.0)	(-4.3)	(-17.1)

Note. T_{2m} biases are in units of Kelvin, and the SEB component biases are in units of W/m^2 . Summer spans June–August and winter spans December–February. Winter SW values are not applicable (N/A) due to negligible insolation. ERA-I = ERA-Interim; CFSv2 = Climate Forecast System version 2; CESM = Community Earth System Model; SW = shortwave; LW = longwave; LH = latent heat; SH = sensible heat; G = ground heat.

or cold bias in modeled surface temperature. $LW\uparrow$, which is proportional to surface temperature to the fourth power, is used as a proxy for surface temperature comparisons because this circumvents implicit assumptions involved in estimating surface temperature from LW-derived observations yet still captures the important role of surface temperature in the SEB. Temperature at a reference height of 2 m (T_{2m}) is highly correlated with surface temperature (Shuman et al., 2014) and provides a metric for how well the near-surface atmosphere is represented. Table 1 compares ERA-I, CFSv2, and CESM biases of 2-m temperature, radiative components, SH, LH, G, and forcing terms. To highlight how the biases differ throughout the annual cycle, seasonal averages of monthly mean biases are calculated for two radiatively distinct periods: summer (June–August) and winter (December–February).

The largest temperature biases are for the CFSv2 winter, indicating a $4^{\circ}C$ and $14 W/m^2$ warm bias (Table 1) of T_{2m} and $LW\uparrow$, respectively. In fact, all modeled periods shown in Table 1 indicate a higher than observed T_{2m} , except summer CFSv2 and CESM values. The weak to strong warm biases in most model periods, and weak ($< 1^{\circ}C$) cold bias in CFSv2 and CESM summer, are in contrast to the large underestimation of $LW\downarrow$ in all model periods. For all three models the winter $LW\downarrow$ deficit is offset by a large surplus of SH flux warming the surface, providing an explanation for why a lack of $LW\downarrow$ coincides with higher than observed surface temperatures. During summer CFSv2 maintains a positive SH bias, in contrast to ERA-I and CESM, which are within $1 W/m^2$ of observed values, indicating that ERA-I and CESM's summer $LW\downarrow$ deficit is compensated by other factors.

CESM has reasonable representation of summer $SW\downarrow$ radiation despite the deficit in $LW\downarrow$, suggesting a more accurate modeling of the cloud SW optical thickness compared to LW optical thickness. In summer CFSv2 has a deficit in $LW\downarrow$ and an associated surplus of $SW\downarrow$ (Table 1) most likely due to insufficient modeled clouds that emit downwelling radiation and reflect insolation. ERA-I's summer $SW\downarrow$ is too small despite the deficit of $LW\downarrow$, suggesting the modeled clouds have artificially large SW albedo and underestimate the LW cloud optical thickness. The $SW\downarrow$ bias should be partially offset by a corresponding $SW\uparrow$ bias, unless the surface albedo is inaccurate. Indeed, ERA-I surface albedo is 9% lower than observed values, indicating that the excess SW absorbed by the surface compensates for under-representation of radiative warming due to clouds. The lower than observed ERA-I albedo, and consequent underestimation of $SW\uparrow$, is partially offset by the positive summer $LW\uparrow$ bias.

Commonality of the LH flux seasonal biases across the three models indicates more summer sublimation (-1 to $-5 W/m^2$) and more winter deposition (1 to $2 W/m^2$) in each model. The amplified seasonal cycle of LH exchange acts to dampen seasonal changes of surface temperatures, although the magnitude of LH differences is limited by the relatively dry conditions at Summit. The seasonal biases of G are different in each model; ERA-I is negative in summer and near zero in winter, CFSv2 is slightly positive in summer and negative in winter, and CESM is slightly negative during both seasons and, as such, no cross-model relationship between monthly mean surface temperature biases and G biases is found.

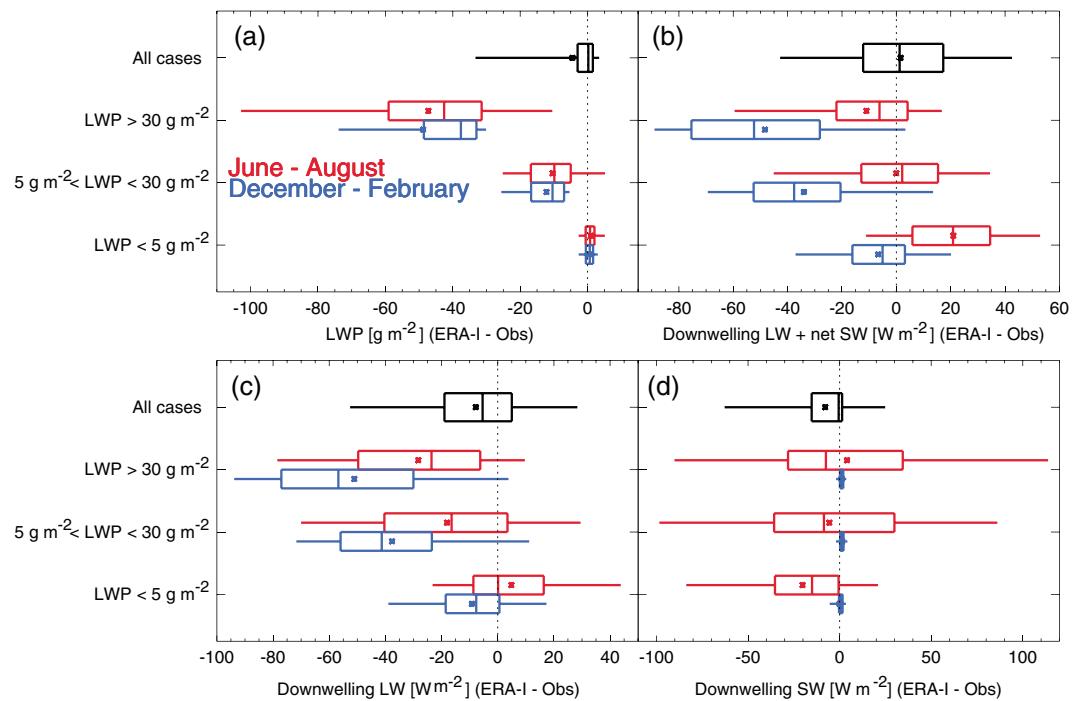


Figure 16. Distributions of ERA-I minus observed values during concurrent 3-hr periods. Distributions of (a) LWP, (b) $LW\downarrow + \text{net SW}$, (c) $LW\downarrow$, and (d) $SW\downarrow$ are shown for observations that indicate thick ($LWP > 30 \text{ g/m}^2$), thin ($5 \text{ g/m}^2 < LWP < 30 \text{ g/m}^2$) or the absence of ($LWP < 5 \text{ g/m}^2$) liquid-bearing clouds. Black statistics are for all available cases throughout the annual cycle, red statistics occur in summer, and blue statistics occur in winter. SW = shortwave; LW = longwave; LWP = liquid water path; ERA-I = ERA-Interim.

It might be expected that biases in forcing terms ($LW\downarrow + \text{net SW}$) would lead to appropriate warm or cold biases in surface temperatures. Yet, a strong correlation between biases in forcing terms and surface temperature is not evident. SH can provide sizable compensation to offset deficiencies in radiative forcing. Despite small monthly biases of LH and G, the response of these fluxes to changes in radiative forcing generally serves to weaken the diurnal cycle of the surface temperature. Diurnal variability of surface temperature is an important consideration, especially in summer when daily maximum temperatures are important for determining if the melting point is reached at a later time, although 4 times daily sampling or 3-hr temporal resolution data could miss short warming events.

Factors that would lead to downwelling radiation biases include inaccurate representation of mixed-phase clouds, cloud base height, or frequency of cloud occurrence (Shupe & Intrieri, 2004). At Summit, clouds are present approximately 85% of the time (Miller et al., 2015; Shupe et al., 2013), indicating the importance of the cloud microphysics and resultant cloud optical depth, the latter of which is primarily dependent on LWP and the physical depth of the ice cloud (Miller et al., 2015). Bennartz et al. (2013) estimates that clouds at Summit are opaque to LW radiation for LWP greater than 30 g/m^2 (thick) and that liquid-bearing clouds with LWP greater than 5 g/m^2 and less than 30 g/m^2 (thin) provide the maximum warming of the surface for periods with large solar elevation angles. While it is beyond the scope of the paper to do a full cloud evaluation, ERA-I results are used here as an example of how the SEB results can provide further insight into pathways for model improvement.

Summer and winter season differences between ERA-I values and observations of LWP and the resulting biases of downwelling radiation during cases of thick, thin, and no liquid-bearing clouds are shown in Figure 16. ERA-I underestimates LWP in winter and summer and has little to no liquid present when the observations also indicate a lack of liquid in the atmosphere (Figure 16a). Thus, the forcing terms are underestimated during instances of observed thick liquid-bearing clouds (Figure 16b) and in the winter when LWP biases enhance the downwelling LW biases (Figure 16c). The low surface albedo of ERA-I does compensate for the lack of thin liquid-bearing clouds in the summer to produce the smallest bias in the forcing terms for cases of thin LWP. When liquid-bearing clouds are not present, winter underestimates of radiative forcing may be driven

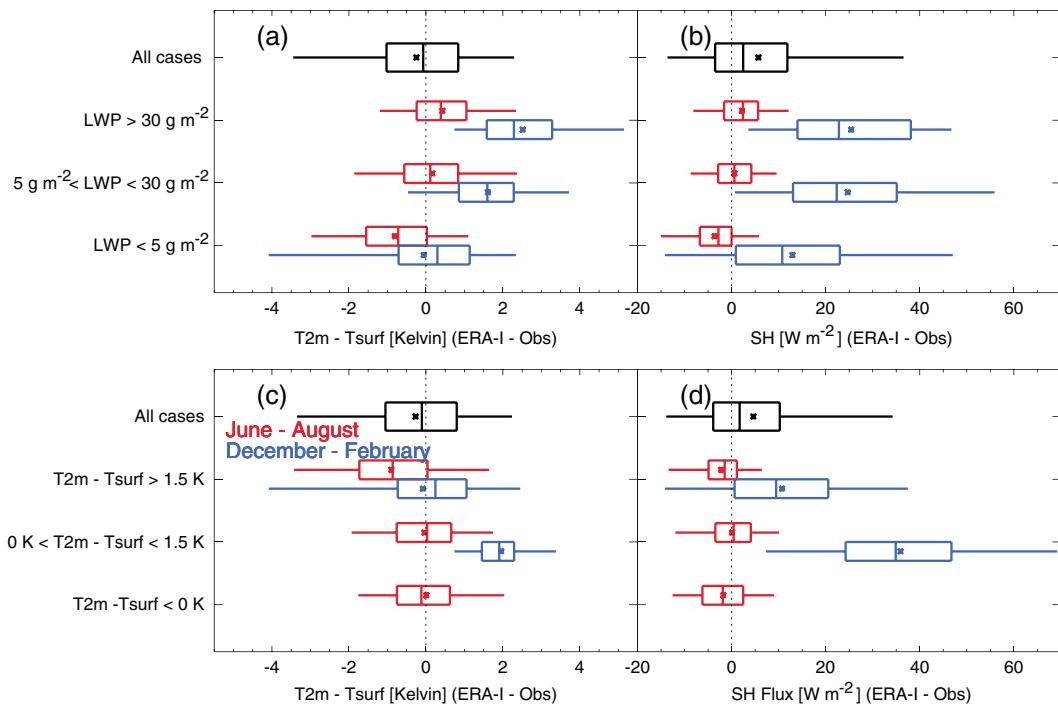


Figure 17. Distributions of ERA-I minus observed values during concurrent 3-hr periods. Distributions of (a) 2 m temperature minus surface temperature and (b) sensible heat flux are shown for observations that indicate thick ($LWP > 30 \text{ g/m}^2$), thin ($5 \text{ g/m}^2 < LWP < 30 \text{ g/m}^2$), or the absence of ($LWP < 5 \text{ g/m}^2$) liquid-bearing clouds. Distributions of (c) 2 m temperature minus surface temperature and (d) sensible heat flux are shown for observations that indicate highly stable ($T_{2m} - T_{surf} > 1.5 \text{ K}$), moderately stable ($0 \text{ K} < T_{2m} - T_{surf} < 1.5 \text{ K}$), or unstable ($T_{2m} - T_{surf} < 0 \text{ K}$) regimes as estimated from the difference in temperature between 2 m and the surface. Black statistics are for all available cases throughout the annual cycle, red statistics occur in summer, and blue statistics occur in winter. LWP = liquid water path; SH = sensible heat; ERA-I = ERA-Interim.

by deficiencies in representing ice clouds while summer overestimates are clearly linked to the underestimated surface albedo, which more than compensates for the underestimated SW↑. Figures 16c and 16d further suggest that summer ice clouds are too optically thick, reflecting too much downwelling SW and emitting too much downwelling LW.

The turbulent and ground heat fluxes respond to changes of the forcing terms through changes in surface temperature, which increases in the presence of liquid-bearing clouds. The surface temperature response affects the temperature gradient between the air above the ice sheet and the surface, which then controls the heat exchange via turbulence. The underestimation of winter LWP in ERA-I (Figure 16a) is associated with an overestimation of the temperature difference ($T_{2m} - T_{surf}$) between the overlying atmosphere and the surface (Figure 17a). In winter, the biases of the temperature differences are greatest for moderate observed stability ($0 \text{ K} < T_{2m} - T_{surf} < 1.5 \text{ K}$) and smaller biases occur during cases of high observed stability ($> 1.5 \text{ K}$, Figure 17c). By comparing Figures 17c and 17d it is evident that the strong winter biases in SH flux are a result of the strong near-surface stability influenced by deficient liquid clouds (Figure 17b). In summer, difficulties representing clouds are largely offset by the underestimated surface albedo, resulting in relatively smaller biases in near-surface stability (Figure 17c) and SH flux (Figure 17d). The largest summer biases occur for cases of high observed stability, which usually occur during the absence of liquid-bearing clouds. For these summer cloud-free conditions, ERA-I underestimates near-surface stability, likely again because of the overestimate of net radiative forcing due to the low surface albedo.

Building off of this foundational assessment of individual parameters, comparisons of processes-based relationships provide further insight into model performance. Table 2 summarizes the differences in modeled and observed responses to changes in the forcing terms during summer and winter. In all models there is an underestimation of the LW↑ response throughout the annual cycle, which is related to the model difficulties in representing the range of surface temperatures. Generally, the LW↑ difference is larger in the summer

Table 2Differences in Unitless Response (Modeled Response - Observed Response) to Changes in Forcing ($LW\downarrow + \text{net SW}$)

Response	ERA-I		CFSv2		CESM	
	Summer	Winter	Summer	Winter	Summer	Winter
$-LW\uparrow$	0.26(0.27)	0.16(0.16)	0.08(0.09)	0.02(0.07)	(0.23)	(0.20)
G	-0.29	-0.40	0.05	0.28	(-0.09)	(0.08)
SH	-0.04(-0.01)	0.38(0.26)	-0.21(-0.16)	-0.12(-0.11)	(-0.15)	(-0.08)
LH	0.02(-0.06)	0.04(0.06)	-0.11(-0.16)	-0.03(-0.02)	(-0.08)	(0.03)

Note. A negative value suggests that the model provides more compensation for a change in forcing. The primary numbers are for a 1-year period (July 2013 to June 2014) and the numbers in parenthesis are differences between the extended time periods, when available. ERA-I = ERA-Interim; CFSv2 = Climate Forecast System version 2; CESM = Community Earth System Model; SW = shortwave; LW = longwave; LH = latent heat; SH = sensible heat; G = ground heat.

compared to winter. ERA-I $LW\uparrow$ response is 26–27% too low in the summer, which coincides with a positive bias in the forcing terms (Table 1).

While all models underestimate the response of $LW\uparrow$ and surface temperature to radiative forcing, each has a unique set of reasons why. The G , SH , and LH responses modulate how effectively surface temperature (i.e., $LW\uparrow$) responds to the radiative influence of the forcing terms. ERA-I has a larger than observed response of G throughout the annual cycle, reducing the $LW\uparrow$ response. In contrast, CFSv2's G response is less than observed and CESM's G response is greater than observed in summer (−9%) and less than observed in winter (8%). During summer the ERA-ISH response is within 1–4% of observations, while in winter the positive SH response to changes in the forcing terms leads to 26–38% less compensation than observed. Thus, the difference in the ERA-I winter SH response offsets the difference in the G response. Both CFSv2 and CESM overestimate the SH response throughout the annual cycle, effectively limiting the surface temperature responses to changes in radiative forcing. The differences in the LH responses are comparatively small for the one year data set, where the biggest differences are in summer CFSv2 and summer CESM. The differences of the summer LH responses are larger when comparing the extended data sets, suggesting that if the observed LH responses in 2012 are more typical than the summer consisting of July–August 2013 and June 2014, then summer LH response plays a larger role in limiting modeled surface temperature variability.

Generally, the representation of the ground heat flux response is important because it provides the most compensation to changes in radiative forcing. CESM, a coupled climate model, provides the most realistic G response because it has the most advanced representation of snow properties and thus more dynamic surface and subsurface interactions. The CESM model uses a vertically resolved land surface model to represent snow compaction of the ice sheet firn and allow for changes in snow density based on temperature and wind speed (van Kampenhout et al., 2017). In contrast, ERA-I has only one snow layer with a constant density, resulting in strong coupling between the atmosphere and the GIS due to the snowpack being too thermally active (Dutra et al., 2015). CFSv2 is similar to ERA-I in that it uses a simple snow layer model, but in the case of CFSv2 the result is a small response of G . In the CFSv2 Noah land surface model the ground heat flux is estimated from the ratio of the temperature gradient across the depth of the snowpack to the snowpack depth, which at Summit is ≈ 1.3 m, resulting in very little heat conduction through this thick layer. Increased vertical resolution in the ERA-I or CFSv2 snowpack would enhance the accuracy of the ground heat flux response to changes in radiative forcing, as is evident in CESM. Although, simulating solar radiation penetration would likely increase the ground heat flux response and consequently decrease the already suppressed surface temperature response.

As had been shown here, representation of surface albedo is important because it directly affects the amount of radiative forcing at the surface by determining the amount of SW radiation absorbed by the surface. In ERA-I the parameter is set to a globally representative value for snow of about 0.8. CFSv2 surface albedo is about 0.84 with apparently some dependence on solar zenith angle in March and September. CESM has the most seasonal and monthly variability indicating an ability to respond to changes in solar zenith angle, cloud presence, and evolution of the snowpack (e.g., Gardner & Sharp, 2010); however, surface albedo is still biased somewhat low. Accurate representation of surface albedo in combination with representative downwelling radiation yields the most realistic modeled radiative forcing.

SH and LH are directly linked to boundary layer stability and the associated small-scale turbulent features. Implementation of scaling parameterizations of eddy diffusivity for stable boundary layers varies widely

across models (Holtslag et al., 2013). In addition to differences of surface layer parameterizations, discrepancies of modeled turbulent heat flux values could be the result of differences in low-level wind speed, vertical resolution of the models, and assumptions for roughness length values (Cassano et al., 2001). While a full assessment of the model boundary layer schemes is beyond the scope of this paper, it is clear that each of the models evaluated here has unique issues with how the turbulent heat fluxes respond to radiative forcing.

5. Conclusion

A comprehensive set of observations at Summit, Greenland is used to evaluate three different modeling approaches. A reanalysis, an operational forecast model, and a global climate model all produce an annual mean warm bias of surface temperature, generally due to difficulties in representing the coldest surface temperatures. Monthly statistics and seasonal biases of each SEB component help discern the major contributors to discrepancies of modeled surface temperatures. Model evaluation utilizing process-based relationships focuses on evaluating how realistically SH, LH, and G modulate changes in radiative forcing. This approach separates the evaluation of radiative forcing of the surface, mainly driven by variability in clouds and/or insolation, from evaluation of the surface energy exchange via ice sheet/atmosphere coupling.

Realistic representation of surface temperature variability at Summit, Greenland would support confidence in modeled temperatures across the central GIS, which would in turn give confidence in a model's ability to represent GIS melt extent. Ideally, modeled surface temperature biases would be small due to accurate representation of downwelling radiation, surface albedo, and realistic responses of the SEB terms to radiative forcing. Offsetting biases of various SEB components often combine to produce surface temperatures that compare decently with observations on monthly or annual time scales. The physical coupling of the GIS and the atmosphere is manifested in the ground heat flux response, which is an important modulating factor of the surface temperature response to radiative forcing. For ice sheet modeling it is advantageous to have a multilayer representation of snowpack, preferably with environmentally dependent snow properties. It is likely that an improved accounting of the ground heat flux variability would also improve the representation of SH, LH, and their associated responses to radiative forcing via a more realistic representation of surface temperature and stability within the boundary layer (i.e., colder surface temperatures could weaken turbulent mixing.) Thus, it is important to consider both the strengths and shortcomings of a reanalysis product when using it to evaluate free-running climate models that may have superior ice sheet/atmosphere coupling.

Accurately representing the observed physical processes is important for models of all types, even though they are forced and/or constrained in different ways. For example, it is imperative to understand the limitations associated with a reanalysis product before utilizing it to force another model or study a climate process. In Greenland, for example, a summer warm surface temperature bias would lead to an overestimation of past surface melt extent, an albedo that was too low would lessen a summer warming trend due to an increase in cloud cover, and near-surface stability that was too weak would enhance the exchange of water vapor at the surface. Climate models are also dependent on accurate representation of physical processes so that associated feedback mechanisms might be reasonably represented for a changing climate. One consequence of higher atmospheric temperatures in the future is that once surface temperatures reach the melting point the surface temperature no longer responds to changes in radiative forcing. When surface temperatures are limited to 0°C the modulating responses of LH, SH, and G are altered and ultimately control the extent to which changes in radiative forcing impact melt.

References

Bennartz, R., Shupe, M. D., Turner, D. D., Walden, V. P., Steffen, K., Cox, C. J., et al. (2013). July 2012 Greenland melt extent enhanced by low-level liquid clouds. *Nature*, 496, 83–86. <https://doi.org/10.1038/nature12002>

Boeke, R. C., & Taylor, P. C. (2016). Evaluation of the Arctic surface radiation budget in CMIP5 models. *Journal of Geophysical Research: Atmospheres*, 121, 8525–8548. <https://doi.org/10.1002/2016JD025099>

Cassano, J. J., Parish, T. R., & King, J. C. (2001). Evaluation of turbulent surface flux parameterizations for the stable surface layer over Halley, Antarctica. *Monthly Weather Review*, 129, 26–46. [https://doi.org/10.1175/1520-0493\(2001\)129<0026:EOTSFP>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0026:EOTSFP>2.0.CO;2)

Christensen, M. W., Behrangi, A., L'Ecuyer, T. S., Wood, N. B., Lebsack, 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), 907–915. <https://doi.org/10.1175/BAMS-D-14-00273.1>

Cullen, N. J., Mölg, T., Conway, J., & Steffen, K. (2014). Assessing the role of sublimation in the dry snow zone of the Greenland ice sheet in a warming world. *Journal of Geophysical Research: Atmospheres*, 119, 6563–6577. <https://doi.org/10.1002/2014JD021557>

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, 553–597. <https://doi.org/10.1002/qj.828>

Acknowledgments

This research was supported by the National Science Foundation grants PLR1303879 and PLR1314156. Gijs de Boer's work was supported by the U.S. Department of Energy's Atmospheric Systems Research program, under grant number DE-SC0013306. The surface energy budget product was derived from radiation measurements provided by the Swiss Federal Institute of Technology, and turbulent and conductive heat flux measurements are provided by David Noone (PLR1023574). The observational data set can be accessed online from the Arctic Data Center (Shupe & Miller, 2016). ERA-Interim results are provided by ECMWF (European Centre for Medium-Range Weather Forecasts, 2009). NCEP CFSv2 data were provided by the NOAA/OAR/ESRL Physical Science Division and can be obtained from <http://cfs.ncep.noaa.gov>. The specific CESM development output is available from <ftp://ftp1.esrl.noaa.gov/psd3/arctic/summit/cesm/>.

Dutra, E., Sandu, I., Balsamo, G., Beljaars, A., Freville, H., Vignon, E., & Brun, E. (2015). Understanding the ECMWF winter surface temperature biases over Antarctica (ECMWF Technical Memoranda No. 762). European Centre for Medium-Range Weather Forecasts.

European Centre for Medium-Range Weather Forecasts (2009). Era-interim project, research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. <https://doi.org/doi.org/10.5065/D6CR5RD9>

Fettweis, X., Franco, B., Tedesco, M., van Angelen, J. H., Lenaerts, J. T. M., van den Broeke, M. R., & Gallée, H. (2013). Estimating the Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model MAR. *The Cryosphere*, 7, 469–489. <https://doi.org/10.5194/tc-7-469-2013>

Foken, T. (2008). The energy balance closure problem: An overview. *Ecological Applications*, 140(6), 1351–1367.

Fyke, J. G., Vizcaíno, M., & Lipscomb, W. H. (2014). The pattern of anthropogenic signal emergence in Greenland ice sheet surface mass balance. *Geophysical Research Letters*, 41, 6002–6008. <https://doi.org/10.1002/2014GL060735>

Gardner, A. S., & Sharp, M. J. (2010). A review of snow and ice albedo and the development of a new physically based broadband albedo parameterization. *Journal of Geophysical Research*, 115, F01009. <https://doi.org/10.1029/2009JF001444>

Gettelman, A., & Morrison, H. (2015). Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes. *Journal of Climate*, 28, 1268–1287. <https://doi.org/10.1175/JCLI-D-14-00102.1>

Hall, D. K., Comiso, J. C., DiGirolamo, N. E., Shuman, C. A., Box, J. E., & Koenig, L. S. (2013). Variability in the surface temperature and melt extent of the Greenland ice sheet from MODIS. *Geophysical Research Letters*, 40, 2114–2120. <https://doi.org/10.1002/grl.50240>

Hanna, E., Huybrechts, P., Steffen, K., Cappelen, J., Huff, R., Shuman, C., et al. (2008). Increased runoff from melt from the Greenland ice sheet: A response to global warming. *Journal of Climate*, 21, 331–341. <https://doi.org/10.1175/2007JCLI1964.1>

Hauer, M. E., Evans, J. M., & Mishra, D. R. (2016). Millions projected to be at risk from sea-level rise in the continental United States. *Nature Climate Change*, 6, 691–695. <https://doi.org/10.1038/nclimate2961>

Holtslag, A. M. A., Svensson, G., Baas, P., Basu, S., Beare, B., Beljaars, A. C. M., et al. (2013). Stable atmospheric boundary layers and diurnal cycles: Challenges for weather and climate models. *Bulletin of the American Meteorological Society*, 94(11), 1691–1706. <https://doi.org/10.1175/BAMS-D-11-00187.1>

Hurrell, J., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The Community Earth System Model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, 94, 1339–1360. <https://doi.org/10.1175/BAMS-D-12-00121.1>

Intergovernmental Panel on Climate Change (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (1535 pp.). Cambridge, UK and New York: Cambridge University Press. <https://doi.org/10.1017/CBO9781107415324>

Kay, J. E., Bourdages, L., Miller, N. B., Morrison, A., Yettella, V., Chepfer, H., & Eaton, B. (2016). Evaluating and improving cloud phase in the Community Atmosphere Model version 5 using spaceborne lidar observations. *Journal of Geophysical Research: Atmospheres*, 121, 4162–4176. <https://doi.org/10.1002/2015JD024699>

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 Reports*, 2, 159–169. <https://doi.org/10.1007/s40641-016-0051-9>

McLennan, E. A., L'Ecuyer, T., & Miller, N. B. (2017). Observational evidence linking Arctic supercooled liquid cloud biases in CESM to snowfall processes. *Journal of Climate*, 30(12), 4477–4495.

McMillan, M., Leeson, A., Shepherd, A., Briggs, K., Armitage, T. W. K., Hogg, A., et al. (2016). A high-resolution record of Greenland mass balance. *Geophysical Research Letters*, 43, 7002–7010. <https://doi.org/10.1002/2016GL069666>

Miller, N. B., Shupe, M. D., Cox, C. J., Noone, D., Persson, P. O. G., & Steffen, K. (2017). Surface energy budget responses to radiative forcing at Summit, Greenland. *The Cryosphere*, 11(1), 497–516. <https://doi.org/10.5194/tc-11-497-2017>

Miller, N. B., Shupe, M. D., Cox, C. J., Walden, V. P., Turner, D. D., & Steffen, K. (2015). Cloud radiative forcing at Summit, Greenland. *Journal of Climate*, 28, 6267–6280. <https://doi.org/10.1175/JCLI-D-15-0076.1>

Nghiem, S. V., Hall, D. K., Mote, T. L., Tedesco, M., Albert, M. R., Keegan, K., et al. (2012). The extreme melt across the Greenland ice sheet in 2012. *Geophysical Research Letters*, 39, L20502. <https://doi.org/10.1029/2012GL053611>

Persson, P. O. G., Fairall, C. W., Andreas, E. L., Guest, P. S., & Perovich, D. K. (2002). Measurements near the Atmospheric Surface Flux Group tower at SHEBA: Near-surface conditions and surface energy budget. *Journal of Geophysical Research*, 107(C10), 8045. <https://doi.org/10.1029/2000JC000705>

Rahmstorf, S., Box, J. E., Feulner, G., Robinson, M. E. M. A., Rutherford, S., & Schaffernicht, E. J. (2015). Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nature Climate Change*, 5, 475–480. <https://doi.org/10.1038/nclimate2554>

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2010). The NCEP Climate Forecast System reanalysis. *Bulletin of the American Meteorological Society*, 91, 1015–1057.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., et al. (2014). The NCEP Climate Forecast System version 2. *Journal of Climate*, 27, 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>

Serreze, M. C., Walsh, J. E., III, F. S., Osterkamp, T., Dyrurgerov, M., Romanovsky, V., et al. (2000). Observational evidence of recent change in the northern high-latitude environment. *Climatic Change*, 46, 159–207. <https://doi.org/10.1023/A:1005504031923>

Shuman, C. A., Hall, D. K., DiGirolamo, N. E., Mefford, T. K., & Schnaubelt, M. J. (2001). A dozen years of temperature observations at the Summit: Central Greenland automatic weather stations 1987–99. *Journal of Applied Meteorology*, 40, 741–752.

Shuman, C. A., Hall, D. K., DiGirolamo, N. E., Mefford, T. K., & Schnaubelt, M. J. (2014). Comparison of near-surface air temperatures and MODIS ice-surface temperatures at Summit, Greenland (2008–13). *Journal of Applied Meteorology and Climatology*, 53, 2172–2180. <https://doi.org/10.1175/JAMC-D-14-00223.1>

Shupe, M. D., & Intrieri, J. M. (2004). Cloud radiative forcing of the Arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle. *Journal of Climate*, 17, 616–628.

Shupe, M. D., & Miller, N. B. (2016). Surface energy budget at Summit, Greenland. Arctic Data Center. <https://doi.org/10.18739/A2Z37J>

Shupe, M. D., Turner, D. D., Walden, V. P., Bennartz, R., Cadeddu, M. P., Castellani, B., et al. (2013). High and dry: New observations of tropospheric and cloud properties above the Greenland ice sheet. *Bulletin of the American Meteorological Society*, 94(C10), 169–186. <https://doi.org/10.1175/BAMS-D-11-00249.1>

Solomon, A., Shupe, M. D., & Miller, N. B. (2017). Cloud-atmospheric boundary layer-surface interactions on the Greenland ice sheet during the July 2012 extreme melt event. *Journal of Climate*, 30(9), 3237–3252. <https://doi.org/10.1175/JCLI-D-16-0071.1>

Steffen, K., & Box, J. (2001). Surface climatology of the Greenland ice sheet: Greenland climate network 1995–1999. *Journal of Geophysical Research*, 106(D24), 33,951–33,964.

Steffen, K., & DeMaria, T. (1996). Surface energy fluxes of Arctic winter sea ice in Barrow Strait. *Journal of Applied Meteorology*, 35, 2067–2079.

Tjernström, M., & Graversen, R. G. (2009). The vertical structure of the lower Arctic troposphere analysed from observations and the ERA-40 reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 135, 431–443. <https://doi.org/10.1002/qj.380>

Turner, D. D., Clough, S. A., Liljegren, J. C., Clothiaux, E. E., Cady-Pereira, K. E., & Gaustad, K. L. (2007). Retrieving liquid water path and precipitable water vapor from the Atmospheric Radiation Measurement (ARM) microwave radiometers. *IEEE Transactions on Geoscience and Remote Sensing*, 45(11), 3680–3690.

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg, W. J., et al. (2016). On the recent contribution of the Greenland ice sheet to sea level change. *The Cryosphere*, 10(5), 1933–1946. <https://doi.org/10.5194/tc-10-1933-2016>

van den Broeke, M., van As, D., Reijmer, C., & van de Wal, R. (2004). Assessing and improving the quality of unattended radiation observations in Antarctica. *Journal of Atmospheric and Oceanic Technology*, 21, 1417–1431.

van Kampenhout, L., Lenaerts, J. T. M., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., & van den Broeke, M. R. (2017). Improving the representation of polar snow and firn in the Community Earth System Model. *Journal of Advances in Modeling Earth Systems*, 9, 2583–2600. <https://doi.org/10.1002/2017MS000988>

Van Tricht, K., Lhermitte, S., Lenaerts, J., Gorodetskaya, I., L'Ecuyer, T., Noël, B., et al. (2016). Clouds enhance Greenland ice sheet meltwater runoff. *Nature Communications*, 7(10266), 1–9. <https://doi.org/10.1038/ncomms10266>

Warren, S. G. (1982). Optical properties of snow. *Reviews of Geophysics and Space Physics*, 20(1), 67–89.

Wright, P., Bergin, M., Dibb, J., Lefer, B., Domine, F., Carman, T., et al. (2014). Comparing MODIS daily snow albedo to spectral albedo field measurements in central Greenland. *Remote Sensing of Environment*, 140, 118–129. <https://doi.org/10.1016/j.rse.2013.08.044>

Zhang, Y., Seidel, D. J., Golaz, J., Deser, C., & Tomas, R. A. (2011). Climatological characteristics of Arctic and Antarctic surface-based inversions. *Journal of Climate*, 24, 5167–5186.